Topics in Current Chemistry 331

Zongwei Cai
Shuying Liu *Editors*

Applications of MALDI-TOF Spectroscopy

Springer

331
Topics in Current Chemistry

Editorial Board:

K.N. Houk, Los Angeles, CA, USA
C.A. Hunter, Sheffield, UK
M.J. Krische, Austin, TX, USA
J.-M. Lehn, Strasbourg, France
S.V. Ley, Cambridge, UK
M. Olivucci, Siena, Italy
J. Thiem, Hamburg, Germany
M. Venturi, Bologna, Italy
P. Vogel, Lausanne, Switzerland
C.-H. Wong, Taipei, Taiwan
H.N.C. Wong, Shatin, Hong Kong
H. Yamamoto, Chicago, IL, USA

For further volumes:
http://www.springer.com/series/128

Aims and Scope

The series Topics in Current Chemistry presents critical reviews of the present and future trends in modern chemical research. The scope of coverage includes all areas of chemical science including the interfaces with related disciplines such as biology, medicine and materials science.

The goal of each thematic volume is to give the non-specialist reader, whether at the university or in industry, a comprehensive overview of an area where new insights are emerging that are of interest to larger scientific audience.

Thus each review within the volume critically surveys one aspect of that topic and places it within the context of the volume as a whole. The most significant developments of the last 5 to 10 years should be presented. A description of the laboratory procedures involved is often useful to the reader. The coverage should not be exhaustive in data, but should rather be conceptual, concentrating on the methodological thinking that will allow the non-specialist reader to understand the information presented.

Discussion of possible future research directions in the area is welcome.

Review articles for the individual volumes are invited by the volume editors.

Readership: research chemists at universities or in industry, graduate students.

Zongwei Cai • Shuying Liu
Editors

Applications of MALDI-TOF Spectroscopy

With contributions by

D. Asakawa · B. Bernevic · Z. Cai · D. Debois · K. Demeure · E. De Pauw · E.B. Erba · X. Gao · Y. Guo · Z. He · N. Liu · S. Liu · M. Lu · L. Ma · S. Mädler · M. Przybylski · R.Z. Qi · L. Quinton · K. Schellander · N. Smargiasso · R.J. Sugrue · I. Susnea · K. Tang · B.-H. Tan · H. Wang · M. Wicke · H. Yang · W. Yu · R. Zenobi · Z. Zhao · T.A. Zimmerman

Springer

Editors
Zongwei Cai
Department of Chemistry
Hong Kong Baptist University
Hong Kong
China

Shuying Liu
Changchun Institute of Applied Chemistry
Chinese Academy of Sciences
Changchun
China

ISSN 0340-1022 ISSN 1436-5049 (electronic)
ISBN 978-3-642-35664-3 ISBN 978-3-642-35665-0 (eBook)
DOI 10.1007/978-3-642-35665-0
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012954392

© Springer-Verlag Berlin Heidelberg 2013
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)

Preface

The emergence of matrix-assisted laser desorption ionization (MALDI) and electrospray ionization-mass spectrometry (ESI-MS) was a revolutionary event in the history of mass spectrometry. Both techniques allow the ionization of large biomolecules, which has been a big challenge for more than 20 years of the last century. MALDI as a mature "soft ionization" technique has been widely applied in the different disciplines of science with great success since 1988, especially in the field of life sciences. In general, the merits of MALDI are suitability for solid samples, high sensitivity, easy sample handling, salt tolerance, time-saving determination, and it is simpler to interpret than ESI-MS due to the singly charged ions. MALDI is an excellent complement to ESI.

New matrices and related techniques have emerged endlessly, and the applications have significantly increased. However, the mechanism of MALDI still remains unclear. The ionization-desorption principle of MALDI-MS is well accepted, based on the co-crystallization of analytes with so-called matrices, which are light-absorbing small organic molecules. When activated by a laser, the matrix ionizes the analyte due to the proton transfer reaction. In this volume, several chapters are addressed.

MALDI is referred to as a "soft" ionization technique, because the spectrum shows mostly intact, singly charged ions for the analyte molecules. However, sometimes and in some cases, MALDI causes minimal fragmentation so it is not as "soft" as ESI, depending on the characteristics of analytes and the conditions of laser used. In-source decay (ISD) MS shows significant fragmentation under certain conditions, which may be useful for structure determination. In this volume, some topics involve the principles and the applications of ISD. Typically, MALDI is coupled with time-of-flight (TOF) analyzers with rather high resolution and a good cost performance, therefore MALDI-TOF-MS has become a nice combination in MS analysis.

We invited eight recognized mass spectrometrists to contribute their reviews from different aspects in the field of MALDI-MS. All of these authors have published excellent papers in the leading MS journals. The volume is divided roughly into three parts, though there is some cross-linking. The first part involves large biomolecular analysis, the second part is the application to small molecule analysis, and the third focusses on bioinformatics. In the first part, the hot topic of the study of noncovalent complexes of biomolecules by using MALDI-MS is

discussed. Zenobi's review describes the major applications of MALDI-MS to investigate noncovalent complexes of biomolecules and highlights the strengths and the limitations of the approaches to detect and quantify them. The second chapter is the application of MALDI-TOF-MS to proteome analysis using stain-free gel electrophoresis given by Przybylski's group. The approach may be a useful complement to staining techniques for mass spectrometric proteome analysis. The third chapter is MALDI-MS for nucleic acid analysis, given by Tang's group. The next two chapters involve the application of the MALDI-ISD method to protein/peptide disulfide determination and the discussion of the principle in detail. These chapters are contributed by Liu's and De Pauw's groups, respectively.

The second part involves the application of MALDI-MS to small molecules. Cai's review describes the application of MALDI-TOF-MS in the analysis of traditional Chinese medicines (TCMs). This is a very complicated system, which requires a long time to separate the components, so the MALDI technique may become one of the promising methods in TCM study for some cases. Guo reviews recent advances and applications of MALDI-TOF-MS analysis for small molecule biochemical, organic, and organometallic compounds.

Finally, Qi's article focuses on bioinformatic methods used for analysis of MALDI-MS data, which is also important for mass spectrometrists.

Here we would like to express our sincerest thanks to all contributors and colleagues from the Springer; without their efforts and encouragement the publication of this volume would have been impossible. Of course, we cannot collect together all aspects of MALDI-MS in a single volume of *Topics in Current Chemistry*, and we regret that some renowned mass spectrometrists especially in the MALDI-TOF field were unable to contribute to this volume. We welcome comments regarding possible oversights from our peers and colleagues which we hope to be able to include in future volumes.

Changchun Shuying Liu and Zongwei Cai
2013

Contents

Top Curr Chem (2013) 331: 1–36
DOI: 10.1007/128_2011_311
© Springer-Verlag Berlin Heidelberg 2012
Published online: 28 February 2012

MALDI-ToF Mass Spectrometry for Studying Noncovalent Complexes of Biomolecules

Stefanie Mädler, Elisabetta Boeri Erba, and Renato Zenobi

Abstract Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) has been demonstrated to be a valuable tool to investigate noncovalent interactions of biomolecules. The direct detection of noncovalent assemblies is often more troublesome than with electrospray ionization. Using dedicated sample preparation techniques and carefully optimized instrumental parameters, a number of biomolecule assemblies were successfully analyzed. For complexes dissociating under MALDI conditions, covalent stabilization with chemical cross-linking is a suitable alternative. Indirect methods allow the detection of noncovalent assemblies by monitoring the fading of binding partners or altered H/D exchange patterns.

Keywords Chemical crosslinking · First shot phenomenon · Intensity fading · MALDI matrix · SUPREX

Contents

S. Mädler, E. Boeri Erba and R. Zenobi (✉)
Department of Chemistry and Applied Biosciences, ETH Zurich, 8093 Zurich, Switzerland
e-mail: zenobi@org.chem.ethz.ch

Abbreviations

ACN	Acetonitrile
ADP	Adenosine-5′-diphosphate
AMNP	2-Amino-4-methyl-5-nitropyridine
AMP	Adenosine-5′-monophosphate
ANA	2-Aminonicotinic acid
AP	Atmospheric pressure
ATP	Adenosine-5′-triphosphate
ATT	6-Aza-2-thiothymine
CHCA	α-Cyano-4-hydroxycinnamic acid
DHAP	Dihydroxyacetophenone
DHB	Dihydroxybenzoic acid
DNA	Deoxyribonucleic acid
ESI	Electrospray ionization
FA	Ferulic acid
hERα LBD	Human estrogen receptor α ligand binding domain
HIV	Human immunodeficiency virus
HPA	3-Hydroxypicolinic acid
IR	Infrared
iTRAQ	Isobaric tag for relative and absolute quantitation
K_a	Association constant
K_d	Dissociation constant
LILBID	Laser induced liquid *beam* or *bead* ionization/desorption
MALDI	Matrix-assisted laser desorption/ionization
MCP	Microchannel plate detector
MS	Mass spectrometry
NHS	*N*-Hydroxysuccinimide
PNA	*p*-Nitroaniline
RNA	Ribonucleic acid
SA	Sinapinic acid
TFA	Trifluoroacetic acid
THAP	Trihydroxyacetophenone
THF	Tetrahydrofuran
ToF	Time-of-flight
TrpR	Tryptophan repressor
UV	Ultraviolet

1 Introduction

Biomolecules that are involved in key cellular processes such as reproduction, growth, and development can communicate via physical interactions. To achieve this, biomolecules specifically recognize their interaction partner. They generally

dock to their partners through short-range biophysical interactions, such as hydrogen bonds, van der Waals forces, and hydrophobic interactions. In other words, intracellular as well as cell–cell communication functions because transient noncovalent interactions are formed between biomolecules. Many such noncovalent interactions take place, including protein–protein, protein–ligand, protein–metal ion, protein–carbohydrate, protein–DNA, DNA–DNA, and DNA/RNA-drug interactions.

Protein–protein interactions are the kernels for the formation of multiprotein complexes, which consist of two or more noncovalently bound proteins. Such multiprotein complexes are often quite large "molecular machines" that perform an equally complex, specific biological function. For instance, the proteasomes are protein complexes for molecular degradation of damaged proteins in all eukaryotes and archaea and in some bacteria [1]. Noncovalent bonds between proteins and between protein and polysaccharides are involved in the recognition of antigens by antibodies [2]. Antibodies, gamma globulin proteins that exist in the blood and other bodily fluids of vertebrates, show a very high binding affinity and specificity for specific antigens. Interactions between proteins and small ligands such as drugs or nucleotides (e.g., adenosine-5′-triphosphate, ATP) are critical for keeping cells fully functional. For instance, ATP plays a fundamental role in signal transduction processes in eukaryotes and many prokaryotes [3]. Proteins can contain one or several metal ions as cofactor, which are generally crucial for proteins to function properly. Hemoglobin, myoglobin, and hemerythrin are important examples of metalloproteins that contain iron and bind oxygen [4]. Noncovalent interactions between proteins and DNA are of paramount importance for the storage and readout of genetic information in living organisms. In eukaryotes, segments of DNA double helix are coiled around a central core formed by eight proteins (histones) to form the so-called nucleosome [5]. Nucleosomes are defined as the basic unit of DNA packaging in eukaryotes. Finally, DNA–DNA interactions are omnipresent [6]: DNA generally forms the well-known DNA double strands that are held tightly together by hydrogen bonds.

A large variety of analytical techniques such as nuclear magnetic resonance, X-ray crystallography, ultracentrifugation, and spectroscopic techniques (fluorescence, circular dichroism, light scattering, or surface plasmon resonance) have been developed to study such interactions (see, for example [7]). Recently, mass spectrometry (MS) has emerged as a powerful tool to investigate noncovalently bound complexes and presents advantages compared to other techniques [8]. Using MS, the molecular weight of intact noncovalent complexes, their stoichiometry, and interactions between subunits can be established. Analysis of biomolecules bound by noncovalent interactions using MS requires only small amounts of sample (on the order of femto- to picomoles), directly provides stoichiometric information and can often be carried out more easily than analyses by other techniques.

Mass spectrometric analysis of noncovalently bound complexes requires a suitable soft ionization method. The two main methods in use are electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). MALDI and ESI allow, under appropriate conditions, the preservation of noncovalently

bound complexes in the gas phase (for reviews see [9–14]). ESI-MS is often considered to be the preferred method, although the results are affected by the nature of the intermolecular interaction, by the composition, ionic strength, and pH of the buffer in which the protein complexes are dissolved, and by the voltages and the pressures in the mass spectrometer. However, MALDI-MS is an attractive alternative to ESI-MS because it can overcome some of the difficulties of ESI-MS. Foremost, the mass spectra acquired by MALDI-MS are simpler to interpret than those acquired by ESI-MS due to the predominance of singly charged species. Furthermore, MALDI ionization is more tolerant to the presence of detergents and salts than ESI ionization. To date, ESI analyses are mainly carried out using volatile buffers such as ammonium acetate or ammonium carbonate.

In the current review we illustrate and selectively discuss the major advances in the application of MALDI-MS to study noncovalent complexes. First we describe the challenges in the analysis of noncovalent complexes (Sect. 2). Then we distinguish between two different approaches to analyze noncovalent assemblies: direct methods that allow a direct detection of the noncovalent assembly (Sect. 3) and indirect methods in which the presence of the complex is deduced indirectly, e.g., by the intensity decrease of an interacting partner (Sect. 4). The last section encompasses quantitative strategies to characterize noncovalent interactions (Sect. 5).

2 Challenges for the Detection of Noncovalent Interactions

Detection of noncovalent complexes by MALDI-MS generally involves the co-crystallization of the analyte with a dedicated matrix typically on a stainless steel plate, ionization and desorption of the crystals with a pulsed laser beam, acceleration of ions with an electrical field, their mass analysis, often using time-of-flight (TOF) instruments, and ion detection. This process requires properly adapted instrumental conditions and sample preparation protocols and, for a number of reasons, is not always straightforward.

The most obvious hurdle is that many MALDI matrices present an unnatural, acidic environment, and that they are often used in combination with organic solvents. This environment tends to denature proteins and is certainly not conducive to maintaining noncovalent interactions during sample preparation. Some remedies for preserving noncovalent interactions were proposed. For example, it is possible to carry out the co-crystallization of the sample with the matrix very rapidly, such that there may be insufficient time for complex dissociation. Aqueous solutions can be utilized. The most successful, however, is the use of special, nonacidic matrices for MALDI of weakly bound complexes. The crystalline environment of the matrix itself may cause disruption of complexes. Special sample preparation methods for MALDI-MS of noncovalent complexes will be discussed in detail in Sect. 3.2. It was observed that MALDI mass spectra obtained from a fresh area of the sample by a single laser shot are much more successful for generating intact complex peaks.

This is the so-called "first shot phenomenon" [15] and it is not general, but sample-, preparation-, and matrix-dependent. This will be discussed in detail in Sect. 3.1.3.

Another difficulty in the analysis of noncovalent assemblies is that large, multiprotein complexes generally require high laser pulse energies to be liberated from the condensed phase for the creation of detectable ions. At some point on the molecular weight scale, the required laser pulse energy may be so high that dissociation of the complex is favored over desorption/ablation. In order to stabilize high molecular weight noncovalent assemblies against dissociation, one often resorts to a "trick," chemical crosslinking [16]. Chemical crosslinking will be discussed more extensively in Sect. 3.3.

A third problem is that MALDI produces predominantly singly charged ions, and complexes of very high m/z (over 50,000–100,000 Da) are quite difficult to detect using TOF instruments. The most successful approach to address this problem has been the development of special high-mass detectors, notably cryogenic detectors [17–19] and ion-to-ion conversion dynode detectors [20, 21]. Using these technologies, it is now possible to measure in the hundreds of kDa to MDa range on the m/z scale. For details, see Sect. 3.1.4.

Generally, however, commercial TOF instruments are not equipped with these special detectors, chemical cross-linking is not widely used, and the temptation has been to work at high sample concentration to increase the chances to observe high-mass complexes by MALDI-MS. However, this leads to another problem, namely unwanted clustering in the expanding MALDI plume. The distinction between a specifically formed noncovalent complex and a nonspecific cluster can become difficult. For example, Chan et al. [22] observed clusters of up to 15-mers of chicken egg white lysozyme when working with a 126-pmol MALDI sample and a liquid matrix, 3-nitrobenzyl alcohol. There are several other examples of clustering in the literature, e.g., clusters between arginine-lysine/oligodeoxyribonucleotide or protein and oligothymidylic acid formed, for instance, by multiple ion-pair interactions [23, 24]. Usually it is impossible to distinguish specifically bound complexes from non-specific complexes in such spectra. Of course, there are also clusters between analyte and matrix, which, given the limited resolution at elevated mass, generally leads to some undesired peak broadening. Occasionally, however, one observes a specific complex peak "sticking out" of an otherwise exponentially decaying nonspecific cluster distribution. This was, for example, the case for avidin [25], where the specific tetramer peak at $\approx$63 kDa was clearly more intense than peaks of the dimer, trimer, pentamer, hexamer, etc. in the spectrum, which were all present due to nonspecific clustering in the plume.

An open question is whether the gas-phase structure of a biomolecule complex is identical to the structure in solution. This may depend on whether the complex is transferred into the gas phase by MALDI (which is somewhat less "soft") or by a very soft "native" ESI method [11, 26]. There are some indirect hints, for example, the charge state distribution in ESI mass spectra or the collision cross section as measured by ion mobility spectroscopy that the gas-phase conformation of such complexes is indeed very close to that in solution phase, but a rigorous proof is still lacking [27].

In conclusion, finding the appropriate strategies for preserving and detecting noncovalent interactions, the distinction between specific and nonspecific complexes, and the resemblance of gas-phase structure to the solution-phase structure are the main challenges during the analysis of noncovalent complexes by MALDI-MS.

3 Approaches for Direct Detection of Specific Complexes by MALDI-MS

The direct detection of noncovalent biomolecular complexes with MALDI-MS was accomplished in the early days after the development of this soft ionization technique. Intact molecular ions of streptavidin [28], glucose isomerase [29], and porin [30] forming quaternary structures with identical subunits could be observed using the water soluble matrix nicotinic acid and minimizing the amount of organic solvents. However, nicotinic acid suffers from several disadvantages, such as low salt tolerance, extensive matrix adduct formation, and lack of compatibility with widespread lasers such as nitrogen lasers or frequency-tripled Nd:YAG lasers with emissions at 337 nm or 355 nm, respectively. Experiments with more typical MALDI matrices (e.g., sinapinic acid) often yielded only signals of subunits rather than of the complex ions. Therefore, dedicated sample procedures were developed to preserve noncovalent biomolecular complexes, often largely dependent on the analyte itself. Thus, several experimental approaches should be tested for each analyte.

Below we describe the instrumental features that should be chosen to detect noncovalent complexes successfully (e.g., source pressure and laser energy). We also discuss how the sample should be prepared prior to the MALDI-MS analysis. Finally, we illustrate the achievements of chemical crosslinking.

3.1 Instrumental Parameters

3.1.1 Laser Pulse Energy and Wavelength

The laser pulse energy strongly influences the presence of the specific intact assembly [31]. In most cases, laser pulse energy values just above the detection threshold value are recommended to maintain reasonable relative intensities of the complex and to minimize the amount of nonspecific clusters formed in the gas-phase (Fig. 1) [32–34].

For the detection of intact double-stranded DNA, desorption with infrared (IR) and ultraviolet (UV) lasers was compared [33]. The IR laser in combination with glycerol/ammonium acetate as matrix was favored for higher numbers of base pairs

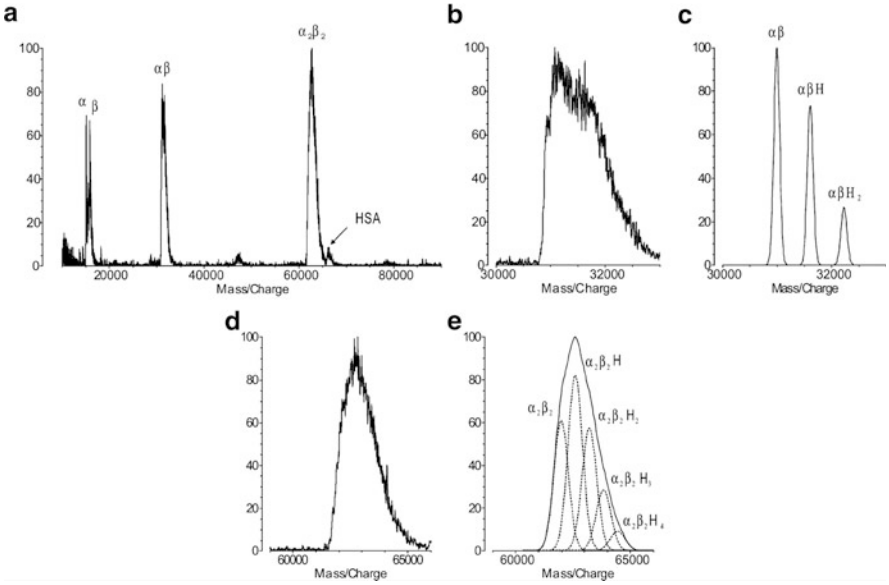


Fig. 1 MALDI-TOF mass spectrum of whole human blood diluted 1:500 in 20 mM ammonium acetate (**a**). In the spectrum the intact hemoglobin $\alpha_2\beta_2$, the hemoglobin α- and β-chains and human serum albumin (*HSA*) are present. The peaks of the heterodimeric subassembly (**b**) and the intact heterotetrameric assembly (**d**) are shown in detail. The corresponding calculated spectra are given (**c**, **e**). *α* indicates the hemoglobin α-chain, *β* the hemoglobin β-chain and *H* the heme b group. A saturated 2,6-dihydroxyacetophenone solution in acetonitrile/20 mM ammonium acetate 1:3 (v/v) was used as matrix. Laser energy values just above the threshold were applied. (Reproduced with permission from [34]. © (2004) John Wiley & Sons, Ltd)

and less fragmentation was observed compared to 6-aza-2-thiothymine (ATT) matrix in the UV range. However, with DNA being smaller than 70 mers, only the UV laser allowed the preservation of the intact assembly.

In contrast, Zehl and Allmaier observed a minor importance of the type of extraction (delayed or continuous), the choice of the acceleration voltage, and the use of reflectron mode concerning the relative intensity of the intact quaternary structure and nonspecific cluster ions, when studying tetrameric avidin [31].

3.1.2 Source Pressure

Few studies presented the utility of ionization at atmospheric pressure (AP). Due to collisions with ambient gases after ionization, AP-MALDI produces ions with lower internal energy than MALDI at ultrahigh vacuum. Thus, weak noncovalent interactions experience less fragmentation. AP-MALDI was successfully used for sugar–sugar [35] and sugar–peptide complexes [36] by IR-AP-MALDI or peptide–peptide complexes by UV-AP-MALDI [37].

3.1.3 The First Shot Phenomenon

In many cases the detection of intact noncovalent assemblies was only possible for the first or the first few laser shots on a non-irradiated sample spot [15, 38–44]. Successive laser shots at the same position mainly yielded monomeric ions (Fig. 2). This observation is generally described as "the first shot phenomenon" and was first investigated by Rosinke et al. for the homotrimeric OmpF porin protein [15]. Cohen et al. suggested segregation or precipitation of the quaternary complex at the crystal surface or dissociation of complexes around the ablation crater induced by laser irradiation [38]. Neither the macroscopic crystal structure, nor the type of substrate, nor the pH stability range of the protein samples had any influence. Detailed investigations of fluorescent or fluorescently labeled protein complexes with confocal laser scanning microscopy measurements confirmed the original assumption of Cohen et al. [38] that size segregation during crystal growth and dissociation of protein complexes in the crystal interior are responsible for the occurrence of the phenomenon [39, 40]. The first shot phenomenon is not generally observed during analysis of noncovalent interactions. Others report the lack of a clear-cut first shot phenomenon for certain analyte-matrix combinations [31, 45]. The extension to types of biomolecular interactions other than protein–protein assemblies has only been reported so far for adenylate kinase with the nucleotides adenosine-5′-monophosphate (AMP), adenosine-5′-diphosphate (ADP), and ATP [44].

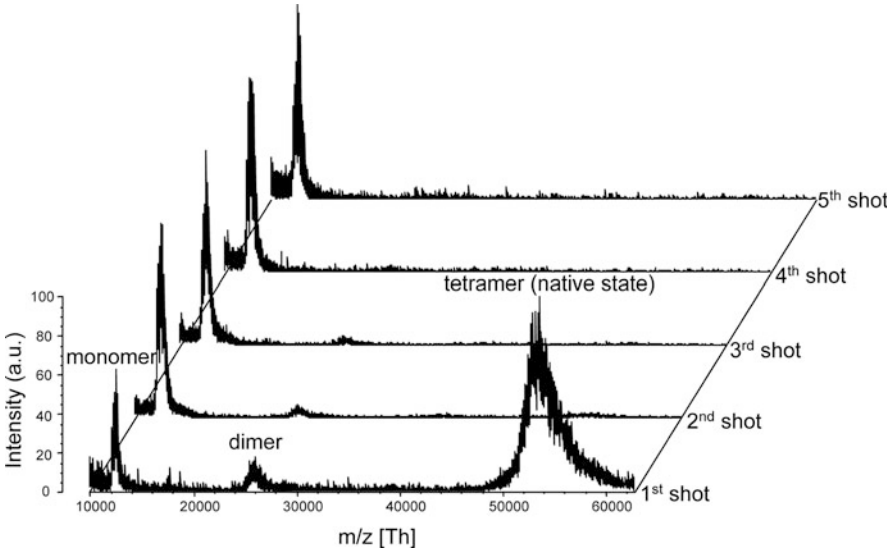


Fig. 2 MALDI mass spectra showing first and subsequent shot data of tetrameric streptavidin. Sinapinic acid was dissolved in triethylammonium bicarbonate at pH 8.5. Sample spots were prepared with a dried droplet preparation using a 1,000:1 matrix-to-analyte ratio. All spectra are normalized to the base peak. (Reproduced with permission from [40]. © (2007) The Royal Society of Chemistry)

3.1.4 High-Mass Detectors

The conventionally used microchannel plate detectors (MCP) suffer from low detection efficiency of high-mass ions due to their low velocity when impinging on the detector surface [46]. Additionally, detector saturation becomes a problem when low-mass ions are present. Ions of lower molecular weight arrive first at the detector. Since the recovery times of detector channels are typically on the order of microseconds [47], the detection sensitivity for high-mass ions with flight times on the order of several hundred microseconds is significantly lower.

Using special high mass detectors, it is possible to measure in the hundreds of kDa to MDa range on the *m*/*z* scale. For so-called superconducting tunnel junction cryodetectors, the impinging analyte ions cause a break-up of Cooper electron pairs under superconducting temperatures. The resulting free electrons can tunnel through a thin oxide layer and thus be detected as excess current [19]. Using this principle, the ion detection is mass-independent and detector saturation effects are avoided due to the fast refresh time of few microseconds. Cryogenic detectors allowed the detection of von Willebrand factor [17] and large molecules such as dendrimers [48]. In ion-to-ion conversion detectors, the analyte ions are converted into secondary ions on a conversion dynode, reaccelerated, and detected with a subsequent secondary electron multiplier [20]. Examples of high-mass MALDI mass spectra with an ion-to-ion conversion detector include the analysis of antibodies [49], PEGylated proteins and glycoproteins [50], and direct profiling and imaging of proteins on tissue [51]. The given examples, however, do not represent noncovalent complexes.

3.1.5 LILBID-MS

Promising results in the analysis of noncovalent interactions were obtained by an alternative ionization technique called laser induced liquid beam ionization/desorption (LILBID) [52]. Instead of a crystallized matrix-analyte mixture, Brutschy et al. used a free liquid beam consisting of a solvent as laser target. After seeding the beam with sample and injecting it into vacuum, partially solvated ions are desorbed by exciting a vibration of the solvent with pulsed infrared laser radiation and mass analyzed with a reflectron TOF tube. The use of an aqueous buffer as liquid beam solvent allowed the conservation of protein–protein interactions, such as hemoglobin [53], ribonuclease S, or calcium-dependent calmodulin/melittin complexes [54]. Shifting the temperature or pH value of the buffer solution to non-native solution-phase conditions induced dissociation into subunits, thus indicating specific complex detection. More recently, the miniaturization of the beam to micro droplets provided higher sensitivity and less sample consumption, extending the application range to more delicate biomolecules of low availability [55]. This variant of LILBID-MS is called laser-induced liquid *bead* ion desorption. Ranging from DNA duplexes to DNA/RNA-ligand complexes, LILBID-MS showed high

potential for the analysis of large macromolecular complexes of nucleic acids in the megadalton mass range [56]. In the case of the human immunodeficiency virus (HIV) Tat:TAR transactivation complex, a quantitative evaluation of the spectra allowed one to draw conclusions on binding specificity, on affinities of different ligands, and on efficiencies of potential inhibitors and the determination of apparent IC_{50} and K_d values [57].

Intact membrane protein complexes could be successfully ionized in the presence of detergent molecules, which partially remained bound to the macromolecule [58, 59]. By increasing the laser intensity, the detergent molecules were stripped off causing disintegration into different oligomeric association states. Being able to detect intact bacterial ATP synthase with a molecular weight of about 540 kDa (Fig. 3) [60], the complex subunit stoichiometries were determined and comparisons between bacterial and eukaryotic species made [61]. In comparison with ESI, LILBID-MS provides simpler spectra due to lower charge states and a higher salt tolerance.

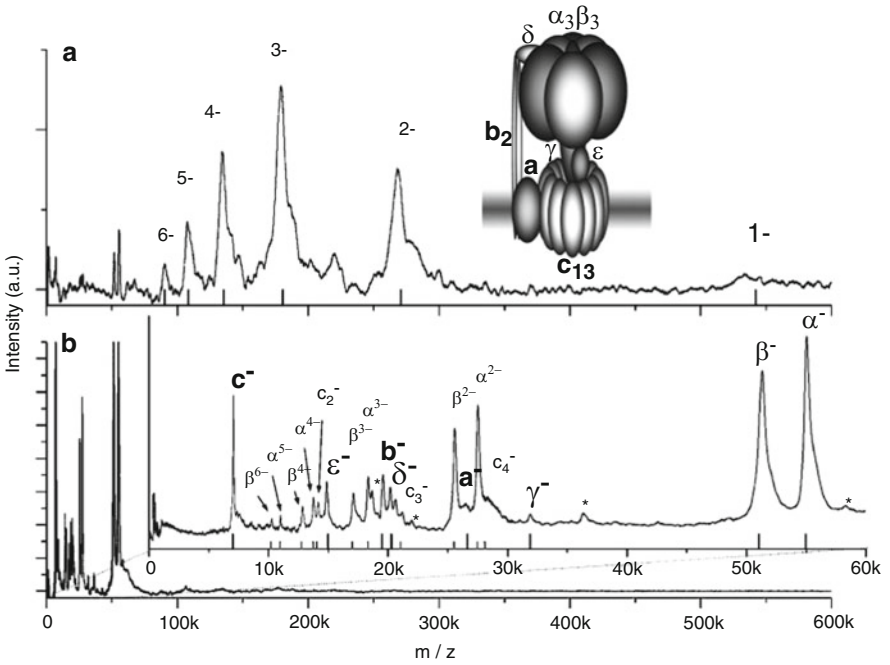


Fig. 3 LILBID anion spectrum of the intact ATP synthase measured under soft conditions (**a**). Using higher laser intensities, all eight subunits are visible (**b**). Additionally, the α and β subunits appear multiply charged and fragments consisting of oligomeric *c* subunits (c_n) are visible as well. Peaks indicated by *asterisks* could not yet be accounted for and probably represent minor impurities from the enzyme preparation. (Reproduced with permission from [60]. © (2008) Elsevier B.V.)

3.2 Dedicated Sample Preparation Procedures

The noncovalent complex has to survive all target preparation and laser desorption steps in order to be detected by MALDI-MS. Several influencing factors that prevent complex disruption could be pointed out. The choice of MALDI matrix [42, 45], the pH of the solution [62–67], the crystal morphology [15, 42], the sample spotting technique, the presence of organic solvent [15], the ionic strength [33, 68], the matrix/analyte ratio [32, 38, 69], and the speed of solvent evaporation [31, 42] are crucial for preserving the noncovalent interaction. Although numerous approaches turned out to be successful for certain complexes, no general protocol applicable for different kinds of biomolecular interactions has been obtained so far.

Note that the binding strength of a noncovalent interaction is not necessarily a good indicator for whether or not an intact complex will be successfully detected by MALDI-MS. For example, to the best of the authors' knowledge, no biotin-streptavidin complex, which represents one of the strongest known noncovalent interactions, has ever been measured in its assembled form so far by MALDI-MS.

3.2.1 Choice of MALDI Matrix

MALDI matrices enhance desorption and ionization of the analyte molecules by absorbing the laser light and are typically small organic molecules that have a high absorption at the corresponding laser wavelengths. The physicochemical nature of the MALDI matrix is one of the key factors for detecting intact noncovalent complexes. Depending on the solvent, the crystal morphology and the analyte incorporation change. For example, the intact porin trimer was only detectable in sufficient amounts of ferulic acid (FA) in tetrahydrofuran (THF), but not with the same matrix dissolved in other solvents, such as acetonitrile (ACN), ACN:0.1% trifluoroacetic acid (TFA) = 1:2 (v/v) or acetone [15]. Other tested matrices such as a 2,5-dihydroxybenzoic acid (DHB):2-hydroxy-5-methoxy benzoic acid = 9:1 (v/v) mixture dissolved in acetonitrile:0.1% TFA = 1:2 (v/v) did not provide any signal of intact trimers. For the ternary system porin/FA/THF, only the finely structured, microcrystalline areas of the sample spots yielded the trimer. In addition to FA [15, 38, 42], several other matrices allowed the detection of reasonable amounts of oligomeric species for certain protein quaternary structures. One of the more common ones are the less acidic hydroxyacetophenone derivatives such as isomers of dihydroxyacetophenone (DHAP) [31, 34, 38, 40, 42, 44] or trihydroxyacetophenones (THAP) [38], 3-hydroxypicolinic acid (HPA) [45], and ATT [39, 43, 69] dissolved in various solvent mixtures, preferably aqueous buffers containing ammonium salts. The latter also represents a favored matrix for biomolecular interactions of low-molecular weight complexes, such as those involving oligonucleotides [33, 41, 70–77], peptide–peptide [37, 67, 78–81], peptide-amino acid [82], or peptide–metal ion complexes [64, 66]. Additionally, other matrices have been used: *p*-nitroaniline (PNA) [64, 83–87] and 2-amino-4-methyl-5-

nitropyridine (AMNP) [65, 88] for the analysis of peptide–metal ion/ligand complexes, HPA for double-stranded DNA [89], and DHAP to preserve DNA-peptide interactions [68]. In a few cases, highly acidic matrices such as sinapinic acid (SA) [25, 32, 40, 62, 63, 90, 91] or α-cyano-4-hydroxycinnamic acid (CHCA) [92] showed promising results.

3.2.2 pH Value of the Matrix Solution

One would expect that acid-sensitive complexes tend to dissociate at low pH values when acidic matrices are used. Thus, raising the pH of the matrix solution to physiological values can facilitate the observation of oligomeric species [62, 69]. In the case of several enzyme-substrate complexes, Woods et al. tested SA as a matrix dissolved in ethanol:1 M ammonium citrate = 1:1 (v/v) or ethanol:water = 1:1 (v/v, pH < 2). Ions for the enzyme-substrate complex were only observed under the first conditions [62]. However, the stability gain of noncovalent complexes with increased pH values is often counterbalanced by a loss in ionization efficiency [93].

As demonstrated by Jespersen et al., the physicochemical properties of the matrix seem to have stronger influence than the acidity or basicity of the matrix itself [45]. Investigating glutathione-*S*-transferase and streptavidin with different acidic or basic matrices in a pH range of 2–7.1, species related to the protein quaternary structures were only detected using HPA (pH 3.8) dissolved in water, but not with DHB, CHCA, or basic matrices such as 2-aminonicotinic acid (ANA), PNA, and AMNP (Fig. 4). Matrices were dissolved in water except for CHCA, which was dissolved in ACN:water = 30:70.

In addition to ammonium salts, additives such as methylene blue, peptides, or spermine can stabilize a noncovalent assembly during crystal growth and/or desorption/ionization and thus facilitate its detection by MALDI-MS (Fig. 5) [76].

3.2.3 Sample Preparation

Different sample preparations have been used ranging from variations of the dried-droplet method [15, 25, 31–34, 38–40, 42–45, 62–70, 72–84, 87–89, 91] to layer techniques [37, 40–42, 92] such as thin layer or sandwich preparations. For the dried droplet method [94], the analyte is directly mixed with the matrix solution either on the MALDI plate or in an Eppendorf tube and dried under room temperature. The thin layer technique [95] utilizes a thin layer substrate of matrix crystals as a seeding ground for subsequent cocrystallization of the analyte solution. For the sandwich method [95], a second layer of matrix is added on top of the sample prepared by the thin layer technique. However, the dried-droplet method is most widely applied. The sandwich method is compatible with the matrixes CHCA and SA and has slightly higher tolerance to sample impurities. Fast sample spot drying under reduced pressure turned out to be more successful than slow drying [31, 42].

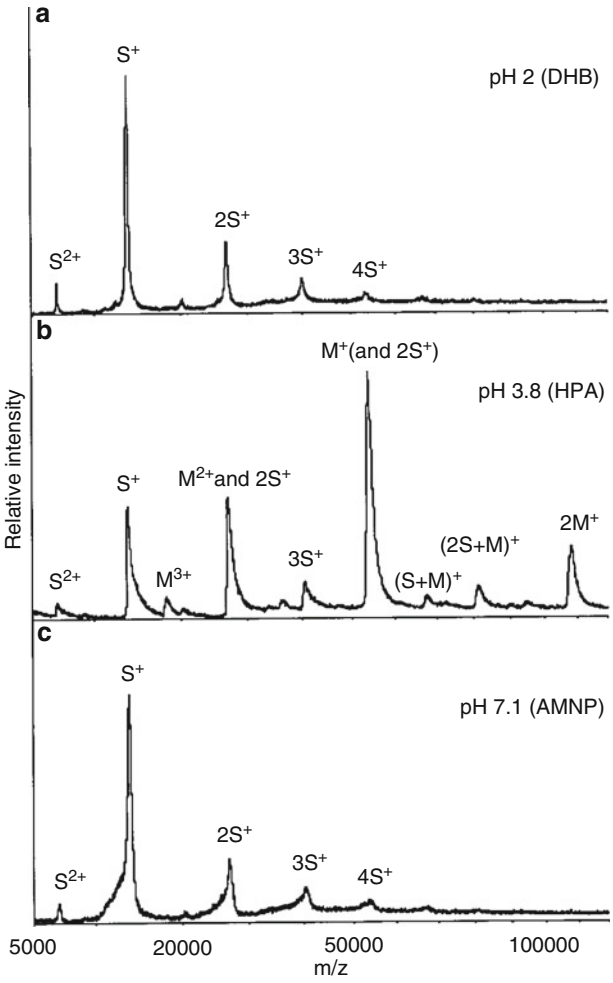


Fig. 4 Positive ion mode MALDI mass spectra of recombinant streptavidin obtained at different pH values with three different matrix compounds: (**a**) DHB in water (pH 2), (**b**) HPA in water (pH 3.8), and (**c**) AMNP in water (pH 7.1). Peaks corresponding to the dissociated and undissociated subunits are indicated as "S" and "M," respectively. (Reproduced with permission from [45]. © (1998) John Wiley & Sons, Ltd)

3.2.4 Sample Concentration

The detection of intact assemblies is often only possible for a limited concentration range. On one hand, a sufficient amount of matrix is required for isolating sample molecules from each other; on the other, the sample concentration should not be too low compared to the matrix concentration. Cohen et al. found for the homomeric protein complex streptavidin a valid range between 0.05 and 10 g/L (3.8–770 pmol/μL) and for alcohol dehydrogenase between 0.1 and 0.5 g/L

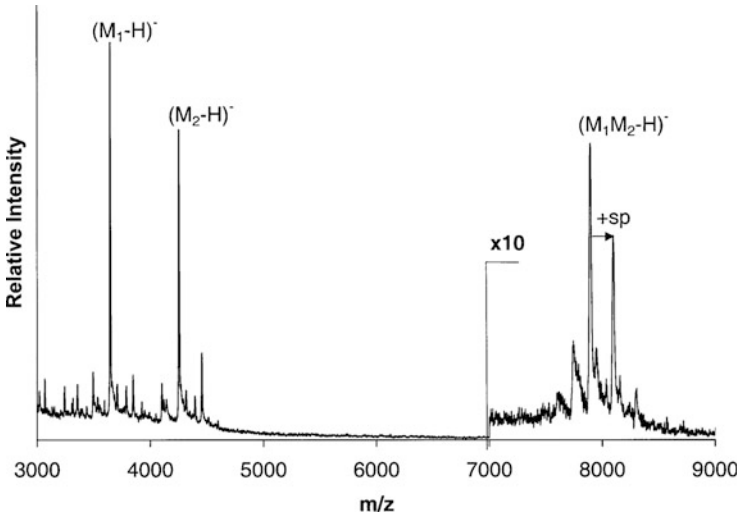


Fig. 5 Negative ion mode MALDI mass spectrum of a duplex oligonucleotide with the peptide β-melanocyte stimulating hormone as additive. The peptide:duplex ratio was 1:250. The mixture of 6-aza-2-thiothymine and spermine (*sp*) was used as matrix. (Reproduced with permission from [76]. © (2002) Elsevier B.V.)

(2.7–13.5 pmol/μL) [38]. For aerolysin, Moniatte et al. could measure the heptameric complex only in a concentration window of 0.5–1.5 pmol/μL [32]. In order to avoid suppression of analyte signals, matrix/analyte ratios as well as mixing ratios between different analytes in heteromeric complexes have to be carefully optimized. Testing several analyte/matrix ratios by preparing analyte solutions ranging from 0.1 to 10 pmol/μL was recommended.

3.2.5 Successful Examples

The direct detection of specific noncovalent complexes by MALDI-MS was accomplished for homomeric or heteromeric protein–protein [15, 25, 31, 32, 34, 38–45, 63, 69, 90, 91] (Fig. 1), protein–peptide [62, 69, 92], peptide–peptide [37, 67, 78–81], protein/peptide–ligand [44, 84–87], protein/peptide–metal ion complexes [62, 64–66, 83, 88], peptide-amino acid [82], oligonucleotide double strands [33, 68, 70, 74–76, 89] (Fig. 5), and protein/peptide/guanidinium derivatives-oligonucleotide interactions [41, 65, 68, 71–73, 77].

3.3 Chemical Crosslinking

Chemical crosslinking has been extensively used in the past to determine the stoichiometry of noncovalent complexes in combination with gel electrophoresis [96].

For this, biomolecule assemblies were covalently stabilized with homobifunctional reagents such as imidoesters. Starting in the early 1990s, chemical crosslinking has been successfully applied in order to prevent complex dissociation of protein quaternary structures during MALDI-MS analysis. In the first experiments, glutaraldehyde was selected as crosslinker [16]. Due to its polymerization reaction in solution, several lengths of bifunctional crosslinkers are generated that can react with ε-amino groups of lysines and α-amino groups of the N-termini. Thus, a variety of distances between the subunits can be bridged. Using this approach, Caprioli and Farmer detected dimeric and tetrameric complexes such as avidin and yeast alcohol dehydrogenase (Fig. 6) [16]. Only a few studies demonstrated the feasibility of this approach [97, 98]. One reason may be the high amount of generated polymerization products that deteriorates the spectra quality significantly.

Since the early days, numerous hetero- or homobifunctional linkers have been developed and applied to study biomolecular noncovalent interactions. In order to identify interaction partners and to determine the stoichiometry of complexes, nonselective photoreagents, as well as site-specific linkers reacting with a limited number of amino acid side chains, were examined. The underlying chemistry and applied functional groups are summarized in the literature [99, 100] and commercial product catalogs [101]. Among the photoreactive linkers, azido groups are particularly noteworthy. If higher selectivity is needed, *N*-hydroxy-succinimide (NHS) esters are often chosen as reactive end groups. The main targets of these

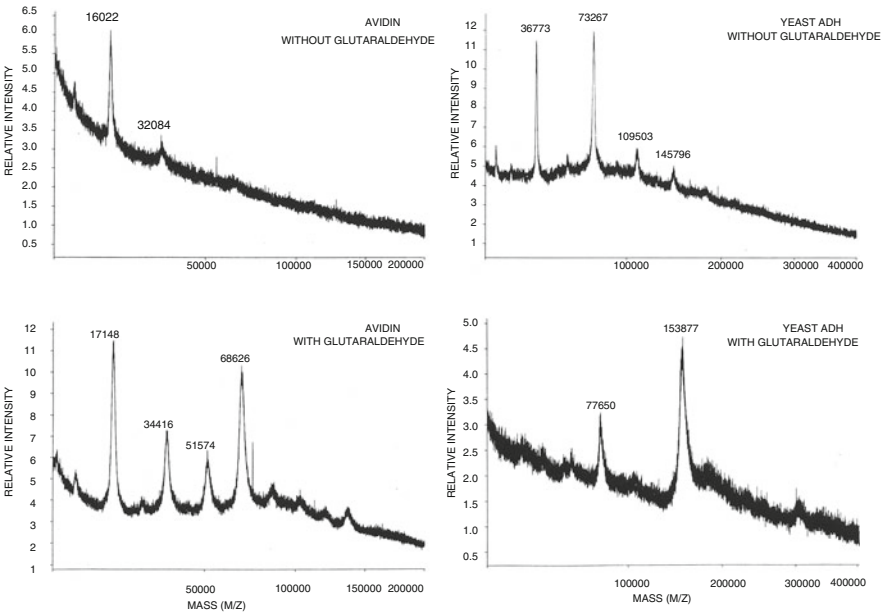


Fig. 6 MALDI mass spectra of avidin and yeast alcohol dehydrogenase before and after crosslinking with glutaraldehyde. (Reproduced with permission from [16]. © (1991) John Wiley & Sons)

esters are primary or secondary amines. However, side reactions with hydroxyl groups of serine, threonine and tyrosine have been reported as well [102, 103].

As an example, a heterobifunctional linker bearing both reactive groups, i.e., sulfosuccinimido-2-(7-azido-4-methylcoumarin-3-acetamido)ethyl-1,3′dithiopropionate, was used to confirm the 1:1 stoichiometry of the gp120 (HIV-1 virus) interaction with the CD4 receptor of T lymphocytes [104].

In combination with previously described high-mass detectors, the analysis of biomolecular complexes up to several MDa became possible. The application of high-mass MALDI and crosslinking for epitope mapping, kinetic studies, sandwich assays for immunocomplexes [49, 105], monitoring of ligand regulation mechanisms (Fig. 7) [106], screening of protease inhibitors [107], and determining association states [108] was demonstrated using NHS esters as crosslinkers.

In comparison with ESI analysis of noncovalent interactions under native conditions or direct MALDI-MS analysis, chemical crosslinking has a remarkable advantage. Since the stabilization of the complex is performed under solution conditions, gas-phase labile complexes (e.g., bound by the hydrophobic effect) are easily preserved during the analysis [109].

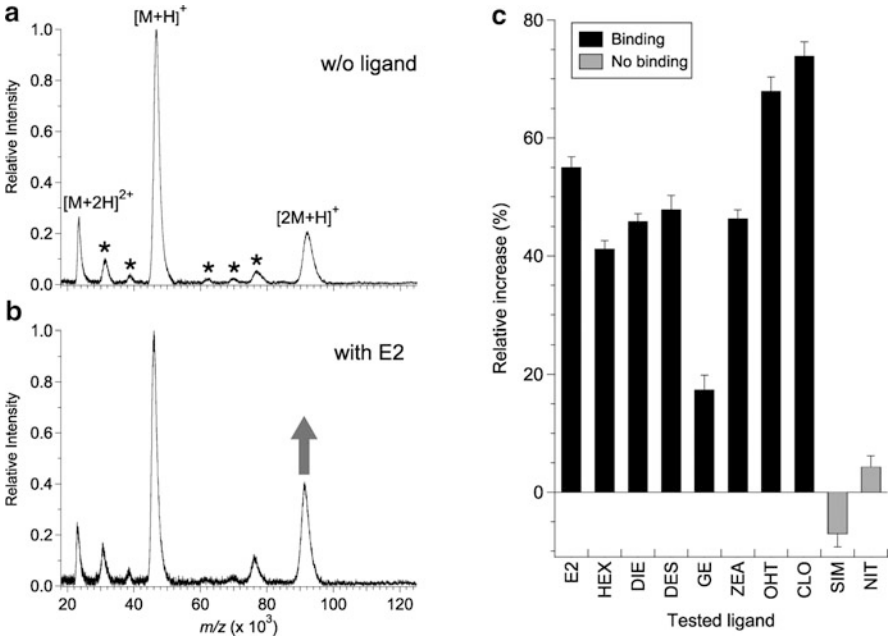


Fig. 7 High-mass MALDI-MS showing a ligand-dependent dimerization in solution for the mutant human estrogen receptor α ligand binding domain (hERα LBD) after chemical crosslinking with NHS esters. MALDI mass spectrum of crosslinked hERα LBD without ligand (**a**) and after adding the ligand E2 (**b**). After incubation with E2 and crosslinking, the MALDI mass spectrum clearly shows an increase of the homodimer, which is labeled $[2M + H]^+$. After incubation with different test compounds and crosslinking (**c**), only the ligands SIM and NIT did not increase the homodimer abundance relative to the aporeceptor significantly. The *asterisks* indicate sample impurities. (Reproduced with permission from [106]. © (2008) American Chemical Society)

Combining chemical crosslinking with electrophoretic separation, digestion and subsequent MALDI-MS of the excised bands to perform peptide mapping allowed the identification of several interacting partners [110, 111]. For example, experiments in living cells with paraformaldehyde revealed interactions of adhesin proteins with other membrane proteins [110]. Applying a heterobifunctional crosslinker comprising azido and an NHS ester functionalities allowed the identification of bacterial surface adhesins with a carbohydrate-containing crosslinking probe [111]. Potential candidates of triadic proteins interacting with RyR1 or TRPC3 in skeletal muscles were established using a maleimide crosslinker [112].

The crosslinker arm can also act as a ruler to map spatial proximities of amino acids in proteins and protein complexes [113]. Typical "bottom–up" protocols for the analysis of protein assemblies include the crosslinking reaction of the interacting proteins, purification of reaction products, and their digestion. The generated peptides are chromatographically separated and subjected to MS and tandem MS analysis. The data analysis using dedicated software [114, 115] allows the identification of intra-molecularly or inter-molecularly crosslinked peptides and modified amino acids. The derived distance constraints yield low-resolution tertiary structures of proteins. Using an NHS ester crosslinking agent and MALDI-post source decay analysis, the bovine basic fibroblast growth factor FGF-2 was identified as a member of the β-trefoil family [113]. The same methodology showed promising results for mapping binding interfaces of noncovalent protein–protein interactions [116–122]. Utilizing different crosslinkers, several subunits of the ATP synthase from *Saccharomyces cerevisiae* were investigated and their role in this yeast machinery was deduced [117–119]. With the same protocol, topology and spatial organization models of several multiprotein complexes were proposed [120]. Since the number of unmodified peptides or fragments by far exceeds the number of modified ones, data analysis is often the crucial step in this method. With the use of cleavable crosslinkers, affinity tags or isotopic labeling, improvements in the detection of intra-molecularly or inter-molecularly crosslinked species were obtained [123–135]. Recently, several crosslinkers have been optimized for MALDI-MS conditions. On one hand, this optimization was accomplished by inserting photolabile groups that give characteristic fragmentation patterns induced by the UV laser pulse in the mass spectrometer [136]. On the other hand, specific signal enhancement of peptides modified with crosslinker molecules was accomplished by incorporating a CHCA moiety in the linker [137].

In order to study protein-oligonucleotide binding interfaces, several crosslinking approaches were tested in combination with MALDI-MS. Most of them incorporate photochemical reactions, which are reviewed in [138]. A strategy avoids the use of any crosslinker molecules and relies on the natural UV reactivity of the nucleobases. Upon UV radiation of protein-oligonucleotide complexes, covalent linkages between the interaction partners are formed at the binding interface (Fig. 8). The resulting species can be analyzed directly by MALDI-MS to find interaction partners [139] or can be subjected to digestion procedures and MS [140–143] or tandem MS [144, 145] analysis to identify binding sites. Since direct UV-induced crosslinking often suffers from low yields of crosslinked products,

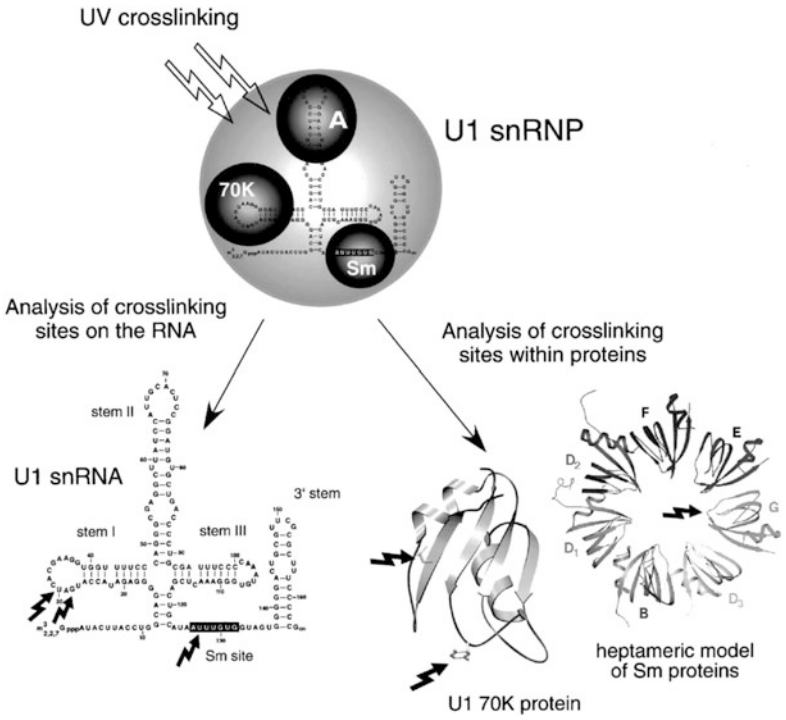


Fig. 8 Complementary strategy for identification of protein–RNA crosslinking sites in native ribonucleoprotein (RNP) particles as outlined for UV-irradiated U1 small nuclear (sn)RNPs. Arrows in the U1 snRNA secondary structure indicate the crosslinking sites on the RNA as identified by a immunoprecipitation/primer-extension method. Arrows at the 3D protein models of the U1 70K protein and of the heptameric Sm protein ring show the corresponding crosslinking sites within the protein as identified by Edman degradation combined with MALDI-MS. (Reproduced with permission from [143]. © (2002) Elsevier Science (USA))

dedicated enrichment protocols are preferable [146–148]. Alternatively, the oligonucleotides are chemically modified with reactive groups forming either specific [149, 150] or nonspecific photo-induced [139, 151] covalent linkages to the protein. Using this approach, amino acids in close proximity to these modifications in RNA/DNA-binding domains of proteins could be identified. The main challenges for the analysis of protein-oligonucleotide complexes are conflicting conditions for an optimum ionization of the interacting molecules. Thus, sample preparation conditions should be carefully chosen [140, 152].

In recent studies, MALDI-MS has been mainly used as a fast and sensitive tool to monitor the crosslinking step and to optimize its reaction conditions prior to ESI analysis [132, 153–155] or as a supportive tool to complement ESI data [156].

In order to avoid the detection of nonspecific assemblies, the concentration levels of biomolecules should be kept as low as possible. Control experiments using the same crosslinking conditions, but nonbinding biomolecules, are strongly

recommended. A study focusing on the application range of chemical crosslinking has been published recently. Mädler et al. pointed out that complexes with lower affinity than a K_d of 25 μM cannot be analyzed using standard protocols [157].

4 Indirect Methods to Detect Specific Complexes by MALDI-MS

In contrast to the described direct methodologies, alternative approaches have been developed to detect the presence of noncovalent complexes by the appearance or fading of certain binding partners in the mass spectra. In this review, these methods will be referred as "indirect methods."

4.1 Intensity Fading

In the so-called intensity fading approach, the formation of a complex between a target biomolecule and its ligands is monitored in the presence and absence of the biomolecule. When the target biomolecule is present, the relative intensity of a ligand decreases (i.e., fades), if compared to control mixtures where no target molecule is present.

In early immunoassay studies, proteolysis products of antigens after the reaction with a monoclonal antibody were analyzed by MALDI-MS and their relative intensities were compared to reaction mixtures where no antibody was present [158–160]. An intensity decrease for certain peptides revealed their participation in the epitope. This approach was also used for high-throughput screening [161, 162] and in combination with dedicated software [163].

More recently, the intensity fading of intact binding partners has been monitored. A non-binding molecule, with similar mass and ion intensity as the interacting partner, was used as internal control. Utilizing this strategy, complexes between proteases and their corresponding inhibitors, as well as protein-nucleic acid complexes, were successfully analyzed [164, 165]. In order to preserve the noncovalent interaction during crystallization, DHAP in ammonium buffers with low acetonitrile content was applied as matrix. Immobilizing the protein target on microbeads and incorporating prefractionation steps dramatically increased the efficiency of this approach [166]. For instance, 16 protein inhibitors of serine proteases among a complexity of nearly 2,000 molecular species were identified in the saliva of the leech *Hirudo medicinalis*. The validity of the intensity fading methodology was confirmed with high-mass MALDI measurements after chemical crosslinking [107]. However, the mechanism is still not fully understood. A detailed study on experimental conditions, which are necessary to observe intensity fading, revealed the necessity of sub-μM concentrations of the binding partners and the presence of several non-binding compounds [167].

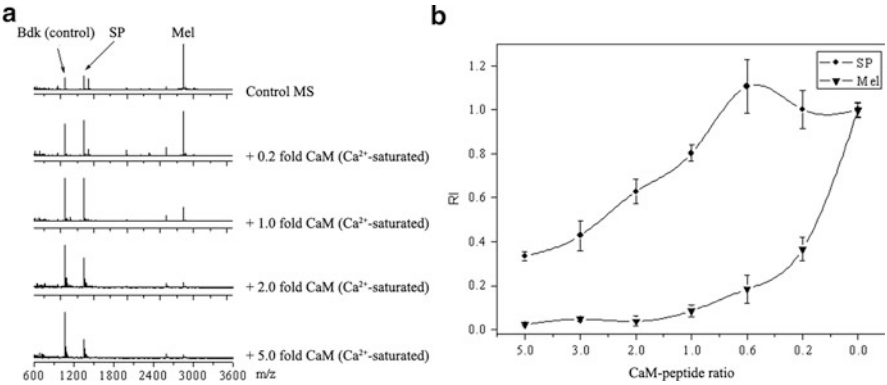


Fig. 9 MALDI mass spectra of mixtures of calmodulin-binding peptides melittin (*Mel*), substance P (*SP*), and a nonbinding control (bradykinin, *BDK*) after the addition of different concentrations of calcium-saturated calmodulin (**a**). Plot of the relative intensities (*RI*) of both melittin and substance P (corrected with the RI of the control) after the addition of calcium-saturated calmodulin with different concentrations (**b**). THAP was used as the matrix. (Reprinted with permission from [170]. © (2009) Elsevier Inc.)

Attempts to use the "intensity fading" strategy to gain quantitative information on a noncovalent complex were published for the interaction between the protease papain and its inhibitor cystatin [168]. In this study, the number of binding sites of cystatin for papain was determined by a modified Scatchard analysis. In order to draw the Scatchard plot, the relative intensity decrease of free cystatin was monitored. However, in the age of computer-based curve fitting, the linearization provided by the Scatchard plot is not really necessary, and it is somewhat problematic because the ordinate and abscissa are not independent. Today, nonlinear least squares fits can easily be used instead of linearization. More recent approaches demonstrated the feasibility of the "intensity fading" approach for studying metallopeptidase-inhibitor complexes [169], calcium-ion dependent calmodulin-peptide interactions (Fig. 9) [170], complexes between proteases and inhibitors from the plant *Capsicum annuum* [171], small molecules inhibiting the formation of a dimeric kinase [172], and RNA-polypeptide interactions [173].

4.2 Identification of Affinity-Separated Interaction Partners

Since its early days, MALDI-MS has served as a tool for the study of immunocomplexes. Instead of detecting the intact immunocomplexes, an indirect approach based on affinity-capture of the antigen on immobilized antibodies has often been applied [174]. In order to identify the antibody recognition site of the antigen, a proteolytic protection assay was used [175–178]. The antigen was captured by antibodies immobilized on beads, subsequently digested, and the

unbound peptides were washed off. Due to the high stability of the antibody to proteolysis, the epitope remained noncovalently attached to the antibody and was directly analyzed on the beads by MALDI-MS. This approach is generally termed as "epitope excision." If the antigen is proteolytically cleaved before binding to the antibody, "epitope extraction" is the term of choice. With the use of magnetic beads, the separation efficiency from unbound material is increased [179–182].

Competition assays allow the quantitation of binding efficiencies of different ligands [179, 180]. However, the immobilization step of one partner can alter the binding affinity due to covalent modifications. Binding the antibody to an immobilized protein, such as protein G, circumvents this problem [183, 184]. The same approach of noncovalent immobilization of an antibody was applied for immunoassays performed on self-assembled monolayers on a gold surface [185]. Using a porous gold layer, MALDI imaging techniques provided high-throughput and high sensitivity analysis [186]. A less costly alternative to antibody immobilization is the use of silica surfaces with covalently bound specific peptides [187].

MALDI-MS coupled with surface plasmon resonance (SPR) allows the simultaneous determination of binding kinetics during affinity separation, as reviewed in [188] (Fig. 10). In this method, a chip, which is directly analyzed by MS, replaces the beads [189]. An immobilized compound is used as a hook to fish unknown ligands from a complex biological sample [190]. MALDI-MS analysis again serves as a tool to identify the bound ligands and to map the recognition sites of the immobilized compound [191, 192]. High-throughput measurements can be easily carried out, when different target molecules are immobilized in distinct areas of the chip [193]. Moreover, the use of reflectometric interference spectroscopy was suggested as an alternative technique to SPR to investigate quantitative and qualitative binding of mixtures of ligands to target biomolecules [194].

MALDI-MS commonly serves as a tool to identify complex partners in diverse biological samples after electrophoretic separations [195] or tandem-affinity purification [196]. However, co-purified, nonspecific interactors are not easily distinguishable from specific ones. Additionally, changes in the complex composition can occur during sample preparation. In order to overcome these difficulties, a four-channel iTRAQ (isobaric tag for relative and absolute quantitation) approach provides in a single liquid chromatography-MALDI-TOF/TOF analysis the identification of genuine partners of the bait and the detection of variations in complex composition [197].

4.3 Hydrogen/Deuterium Exchange

MALDI-MS is used to study biomolecular interactions with hydrogen/deuterium (H/D) exchange strategies, as reviewed in [198]. In brief, deuterium atoms are integrated in interaction partners of interest by replacing backbone amide hydrogen atoms. As hydrogen is exchanged for deuterium, the increase in mass is

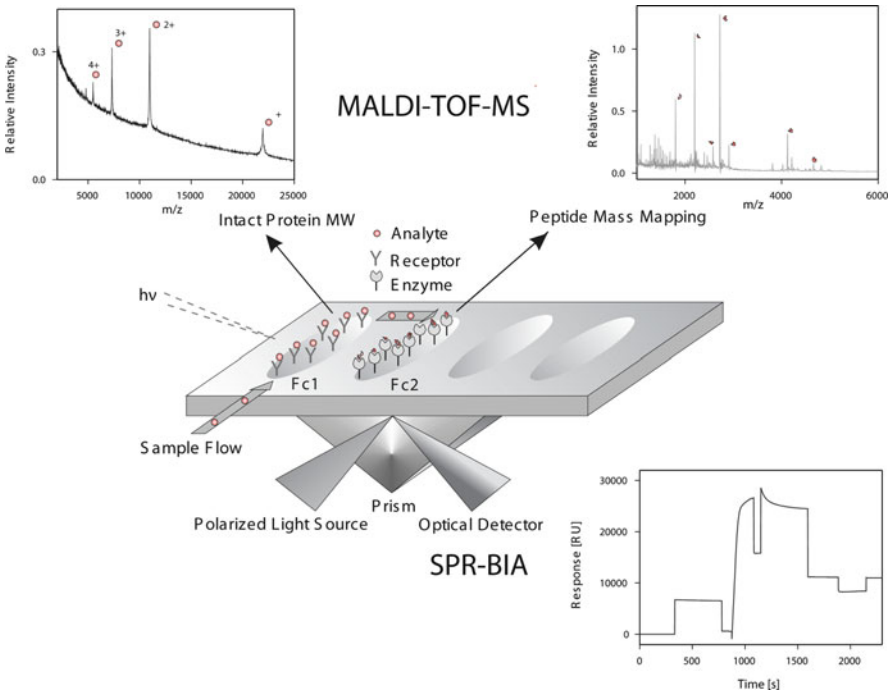


Fig. 10 Overview of MS coupled with SPR with on-chip incorporated proteolytic digestion. A receptor is covalently immobilized on the surface of the first flow cell (FC1). A second flow cell (FC2) is derivatized with a proteolytic enzyme. The analyte (ligand)-containing solution is routed through FC1 where the component of interest is affinity-captured. Following washing of non-specifically retained components, the ligand is eluted/routed from FC1 into FC2, where time for digestion is allowed. MALDI-MS analysis performed on the surface of FC2 yields accurate masses of the proteolytic peptide fragments that can be used for in-depth protein characterization. MALDI-MS performed on the surface of FC1 yields the mass of the intact protein. (Reprinted with permission from [188]. © (2000) WILEY-VCH Verlag GmbH)

monitored by MS. The exchange rates depend on the extent of inter- and intramolecular hydrogen bonding and solvent accessibility. Upon binding of an interaction partner, the H/D exchange rate at the binding site is altered. Komives and coworkers demonstrated the feasibility of MALDI-MS analysis to identify binding sites on a protein–protein interface. They studied the cyclic-AMP-dependent protein kinase complex with a kinase inhibitor and ATP [199] and the thrombin-thrombomodulin fragment complex [200]. Additionally, investigations on epitope mapping [201], assembly of viral capsids [202], conformational changes induced by aggregation [203–206], structural changes upon complex association (Fig. 11) [207–209], and the topology of supramolecular protein complexes [210] were conducted.

However, MALDI-MS analysis coupled with H/D exchange suffers from several disadvantages. One major problem is the back-exchange effect occurring during

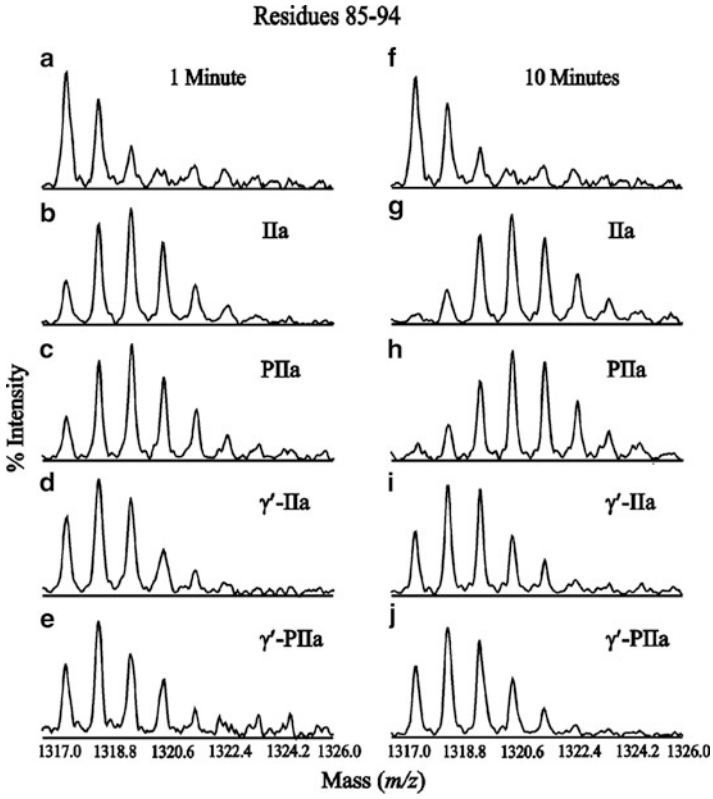


Fig. 11 Conformational analysis of γ′ peptide (410–427) interactions with thrombin anion binding exosite II (ABE-II). MALDI mass spectra representing residues 85–94 of ABE-II (*m/z* 1317.73) after 1 and 10 min of deuteration and subsequent proteolysis: (**a**, **f**) undeuterated peak cluster, (**b**, **g**) thrombin spectrum in the absence of ligands, (**c**, **h**) inhibited thrombin, (**d**, **i**) γ′ peptide bound to thrombin, (**e**, **j**) γ′ peptide bound to inhibited thrombin. This peak cluster contains the ABE-II residue R^{93} and experiences a significant degree of protection from deuterium in the presence of the γ′ peptide. Since this protection is maintained over 10 min, the HDX data present evidence that the γ′ peptide is interacting with R^{93}. (Reprinted with permission from [209]. © (2006) American Chemical Society)

analysis which is more prominent than in ESI experiments, although protocols have been developed to reduce [211] or to quantify its importance [212]. The application of collisional-induced dissociation to obtain site-specific information about the incorporation of deuterium into peptides and proteins is problematic due to "scrambling" of the deuterium position [213]. In contrast, in-source decay fragmentation induced less scrambling [214]. Although the higher salt tolerance and the simplified spectra are a strong advantage of MALDI and can often lead to abandonment of chromatographic purification, ESI is still the main ionization technique used for H/D exchange studies of biomolecular complexes.

5 Quantitative Characterization of Biomolecular Interactions

Noncovalently bound complexes are composed of interacting molecules (e.g., A and B) and it is possible to determine their binding affinities by MALDI-MS. The propensity of a noncovalently bound complex (e.g., AB) to dissociate into its components can be quantified by calculating an equilibrium constant named dissociation constant. The dissociation constant is usually indicated by K_d and is the inverse of the association constant K_a:

$$A + B \rightleftharpoons AB \qquad \frac{[A][B]}{[AB]} = K_d = \frac{1}{K_a}.$$

The method SUPREX (stability of unpurified proteins from rates of H/D exchange) was developed by Fitzgerald's group to investigate the strength of protein–ligand binding interactions in solution using H/D exchange and MALDI-MS [215]. The SUPREX protocol for the determination of K_d values includes an initial H/D exchange by adding a tenfold excess of deuterated exchange buffer to the target protein, usually at physiological pH. The exchange buffers contain varying concentrations of a chemical denaturant such as guanidinium chloride and urea. The denaturant leads to unfolding of the protein and to an enhancement of the H/D exchange rate, thus increasing the rate of deuterium incorporation into the protein. At a given exchange time, a small aliquot of the reaction mixture is added to the matrix (sinapinic acid) solution in a tenfold excess and the sample is subjected to MALDI analysis. The change in mass relative to the fully protonated sample (Δmass) in the spectra is plotted as a function of denaturant concentration, and the data is fitted to a sigmoidal function to obtain a transition mid-point. Solution-phase folding free energies (ΔG°f) in the presence and the absence of ligand can be calculated. The folding free energies ΔG°f of a protein and a protein–ligand complex are different, and this difference, ΔΔG°f, can be used to determine the protein–ligand binding constant. SUPREX works in a high-throughput automated fashion, requires only minute amounts of sample, and is applicable to purified as well as unpurified protein–ligand complexes. A prerequisite for SUPREX is that the protein must unfold in a two-state manner, i.e., only the fully folded and the fully unfolded states of the protein should be populated at equilibrium, thus posing a limit on its application.

As an example, the ternary protein–DNA complex formed of tryptophan repressor (TrpR), two molecules of L-tryptophan, and a 25-base pair duplex of DNA containing TrpR's cognate DNA sequence was analyzed by SUPREX [216]. The K_d values of the complexes, investigated in this study, were in agreement with previously established K_d values.

SUPREX was also used to measure quantitatively the stability of unpurified proteins in complex biological matrices [217]. Experiments of Fitzgerald's group in this field led to some excellent studies where SUPREX was applied to measure the thermodynamic stability of proteins both in vitro and in vivo with good accuracy and high precision [218, 219].

Related to SUPREX, Fitzgerald's group developed a technique called SPROX (stability of proteins from rates of oxidation) to determine the thermodynamic stability of proteins and protein–ligand complexes [220]. Proteins are oxidized with hydrogen peroxide in the presence of increasing concentrations of a denaturant (e.g., guanidine hydrochloride). Using MALDI-MS, the degree of oxidation is established at each oxidation time as a function of the denaturant concentration. By correlating denaturant concentration and oxidation rate, a folding free energy (ΔGf) and m value (δΔGf/δ[denaturant]) are measured during protein unfolding. If ΔGf and m values of the proteins are measured in the presence and absence of ligands, it is possible to evaluate protein–ligand affinities (e.g., ΔΔGf and K_d values). The main advantage of SPROX over SUPREX is the use of irreversible oxidation. The chemical stability of the oxidized proteins enables the manipulation of the modified proteins after oxidation. As an example, oxidized methionine amino acids were used to probe the solvent accessibility of these amino acidic residues as a function of temperature in order to construct thermal denaturation curves [221].

The Wanner group described a novel MALDI-based binding assay to determine affinity constants between small ligands and proteins in saturation and competition experiments, showing an excellent agreement between the K_d values determined by the two methods. A known ligand of the protein was used as an internal standard to generate a calibration function. They compared MS binding assays based on MALDI-MS/MS and those based on LC-ESI-MS/MS quantification. The instrument they used was a MALDI-triple quadrupole with high speed in the analysis of small molecules, commercialized as "FlashQuant system" [222].

6 Conclusions

In this review, we described the major applications of MALDI-MS to investigate noncovalent complexes of biomolecules. We highlighted the strengths and the limitations of the MALDI-based approaches to detect and quantify noncovalent complexes. We described not only the experimental approaches to detect complexes directly, but we also explained what the critical instrumental parameters are, and how intact complexes can be preserved by chemical cross-linking. Finally, we illustrated the possibility of detecting noncovalent complexes indirectly.

Overall, MALDI-MS represents an excellent method for studying noncovalent interactions. The substantial advantages of MALDI-MS over other techniques are sensitivity, speed, a higher salt tolerance, and the possibility to obtain precise molecular weights and stoichiometric information.

Acknowledgments We would like to thank Richard Caprioli and Dobrin Nedelkov for providing original versions of figures. Financial support for this work from the Swiss National Science Foundation (grant no. 200020_124663) is gratefully acknowledged.

References

1. Coux O, Tanaka K, Goldberg AL (1996) Structure and functions of the 20S and 26S proteasomes. Annu Rev Biochem 65:801–847
2. Davies DR, Padlan EA, Sheriff S (1990) Antibody-antigen complexes. Annu Rev Biochem 59:439–473
3. Scheeff ED, Bourne PE (2005) Structural evolution of the protein kinase-like superfamily. PLoS Comput Biol 1:e49
4. Terwilliger NB (1998) Functional adaptations of oxygen-transport proteins. J Exp Biol 201:1085–1098
5. McGhee JD, Felsenfeld G (1980) Nucleosome structure. Annu Rev Biochem 49:1115–1156
6. Thuong NT, Helene C (1993) Sequence-specific recognition and modification of double-helical DNA by oligonucleotides. Angew Chem Int Ed 32:666–690
7. Phizicky EM, Fields S (1995) Protein-protein interactions - methods for detection and analysis. Microbiol Rev 59:94–123
8. Wyttenbach T, Bowers MT (2007) Annu Rev Phys Chem 58:511–533
9. Bolbach G (2005) Matrix-assisted laser desorption/ionization analysis of non-covalent complexes: fundamentals and applications. Curr Pharm Des 11:2535–2557
10. Bich C, Zenobi R (2009) Mass spectrometry of large complexes. Curr Opin Struct Biol 19:632–639
11. Heck AJ (2008) Native mass spectrometry: a bridge between interactomics and structural biology. Nat Methods 5:927–933
12. Schalley CA (2001) Molecular recognition and supramolecular chemistry in the gas phase. Mass Spectrom Rev 20:253–309
13. Hardouin J, Hubert-Roux M, Delmas AF et al (2006) Identification of isoenzymes using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 20:725–732
14. Schermann SM, Simmons DA, Konermann L (2005) Mass spectrometry-based approaches to protein-ligand interactions. Expert Rev Proteomics 2:475–485
15. Rosinke B, Strupat K, Hillenkamp F et al (1995) Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) of membrane-proteins and non-covalent complexes. J Mass Spectrom 30:1462–1468
16. Farmer TB, Caprioli RM (1991) Assessing the multimeric states of proteins: studies using laser desorption mass spectrometry. Biol Mass Spectrom 20:796–800
17. Wenzel RJ, Matter U, Schultheis L et al (2005) Analysis of megadalton ions using cryodetection MALDI time-of-flight mass spectrometry. Anal Chem 77:4329–4337
18. Twerenbold D, Gerber D, Gritti D et al (2001) Single molecule detector for mass spectrometry with mass independent detection efficiency. Proteomics 1:66–69
19. Frank M, Labov SE, Westmacott G et al (1999) Energy-sensitive cryogenic detectors for high-mass biomolecule mass spectrometry. Mass Spectrom Rev 18:155–186
20. Hillenkamp F, Nazabal A, Roehling U, Wenzel R (2008) Method for analyzing ions of high mass in time-of-flight mass spectrometer, involves applying potential difference between front and back sides of electron multiplier to multiply number of electrons. WO2009086642-A1 WOCH000007
21. Spengler B, Kirsch D, Kaufmann R et al (1990) The detection of large molecules in matrix-assisted UV-laser desorption. Rapid Commun Mass Spectrom 4:301–305
22. Chan TWD, Colburn AW, Derrick PJ (1992) Matrix-assisted laser desorption/ionization using a liquid matrix: formation of high-mass cluster ions from proteins. Oms 27:53–56
23. Juhasz P, Biemann K (1994) Mass spectrometric molecular-weight determination of highly acidic compounds of biological significance via their complexes with basic polypeptides. Proc Natl Acad Sci USA 91:4333–4337

24. Tang X, Callahan JH, Zhou P et al (1995) Noncovalent protein-oligonucleotide interactions monitored by matrix-assisted laser desorption/ionization mass spectrometry. Anal Chem 67:4542–4548
25. Song F (2007) A study of noncovalent protein complexes by matrix-assisted laser desorption/ionization. J Am Soc Mass Spectrom 18:1286–1290
26. Benesch JLP, Ruotolo BT, Simmons DA et al (2007) Protein complexes in the gas phase: technology for structural genomics and proteomics. Chem Rev 107:3544–3567
27. Ruotolo BT, Giles K, Campuzano I et al (2005) Evidence for macromolecular protein rings in the absence of bulk water. Science 310:1658–1661
28. Karas M, Bahr U, Ingendoh A et al (1990) Principles and applications of matrix-assisted UV-laser desorption/ionization mass spectrometry. Anal Chim Acta 241:175–185
29. Karas M, Bahr U (1990) Laser desorption ionization mass spectrometry of large biomolecules. Trends Anal Chem 9:321–325
30. Hillenkamp F, Karas M, Beavis RC et al (1991) Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. Anal Chem 63:A1193–A1202
31. Zehl M, Allmaier G (2005) Instrumental parameters in the MALDI-TOF mass spectrometric analysis of quaternary protein structures. Anal Chem 77:103–110
32. Moniatte M, van der Goot FG, Buckley JT et al (1996) Characterisation of the heptameric pore-forming complex of the Aeromonas toxin aerolysin using MALDI-TOF mass spectrometry. FEBS Lett 384:269–272
33. Kirpekar F, Berkenkamp S, Hillenkamp F (1999) Detection of double-stranded DNA by IR- and UV-MALDI mass spectrometry. Anal Chem 71:2334–2339
34. Zehl M, Allmaier G (2004) Ultraviolet matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of intact hemoglobin complex from whole human blood. Rapid Commun Mass Spectrom 18:1932–1938
35. Von Seggern CE, Cotter RJ (2003) Fragmentation studies of noncovalent sugar-sugar complexes by infrared atmospheric pressure MALDI. J Am Soc Mass Spectrom 14:1158–1165
36. Von Seggern CE, Cotter RJ (2004) Study of peptide-sugar non-covalent complexes by infrared atmospheric pressure matrix-assisted laser desorption/ionization. J Mass Spectrom 39:736–742
37. Moyer SC, Marzilli LA, Woods AS et al (2003) Atmospheric pressure matrix-assisted laser desorption/ionization (AP MALDI) on a quadrupole ion trap mass spectrometer. Int J Mass Spectrom 226:133–150
38. Cohen LRH, Strupat K, Hillenkamp F (1997) Analysis of quaternary protein ensembles by matrix assisted laser desorption/ionization mass spectrometry. J Am Soc Mass Spectrom 8:1046–1052
39. Horneffer V, Strupat K, Hillenkamp F (2006) Localization of noncovalent complexes in MALDI-preparations by CLSM. J Am Soc Mass Spectrom 17:1599–1604
40. Wortmann A, Pimenova T, Alves S et al (2007) Investigation of the first shot phenomenon in MALDI mass spectrometry of protein complexes. Analyst 132:199–207
41. Gruic-Sovulj I, Lüdemann H-C, Hillenkamp F et al (1997) Detection of noncovalent tRNA-aminoacyl-tRNA synthetase complexes by matrix-assisted laser desorption/ionization mass spectrometry. J Biol Chem 272:32084–32091
42. Moniatte M, Lesieur C, Vécsey-Semjén B et al (1997) Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry in the subunit stoichiometry study of high-mass non-covalent complexes. Int J Mass Spectrom Ion Process 169(170):179–199
43. Vogl T, Roth J, Sorg C et al (1999) Calcium-induced noncovalently linked tetramers of MRP8 and MRP14 detected by ultraviolet matrix-assisted laser desorption/ionization mass spectrometry. J Am Soc Mass Spectrom 10:1124–1130
44. Strupat K, Sagi D, Bönisch H et al (2000) Oligomerization and substrate binding studies of the adenylate kinase from *Sulfolobus acidocaldarius* by matrix-assisted laser desorption/ionization mass spectrometry. Analyst 125:563–567

45. Jespersen S, Niessen WMA, Tjaden UR et al (1998) Basic matrices in the analysis of noncovalent complexes by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom 33:1088–1093
46. Brunelle A, Chaurand P, Della-Negra S et al (1993) Surface secondary electron and secondary ion emission induced by large molecular ion impacts. Int J Mass Spectrom Ion Process 126:65–73
47. Coeck S, Beck M, Delaure B et al (2006) Microchannel plate response to high-intensity ion bunches. Nucl Instrum Meth A 557:516–522
48. Clark CG Jr, Wenzel RJ, Andreitchenko EV et al (2007) Controlled MegaDalton assembly with locally stiff but globally flexible polyphenylene dendrimers. J Am Chem Soc 129:3292–3301
49. Nazabal A, Wenzel R, Zenobi R (2006) Immunoassays with direct mass spectrometric detection. Anal Chem 78:3562–3570
50. Seyfried BK, Siekmann J, Belgacem O et al (2010) MALDI linear TOF mass spectrometry of PEGylated (glyco)proteins. J Mass Spectrom 45:612–617
51. van Remoortere A, van Zeijl RJM, van den Oever N et al (2010) MALDI imaging and profiling MS of higher mass proteins from tissue. J Am Soc Mass Spectrom 21:1922–1929
52. Kleinekofort W, Avdiev J, Brutschy B (1996) A new method of laser desorption mass spectrometry for the study of biological macromolecules. Int J Mass Spectrom Ion Process 152:135–142
53. Wattenberg A, Sobott F, Brutschy B (2000) Detection of intact hemoglobin from aqueous solution with laser desorption mass spectrometry. Rapid Commun Mass Spectrom 14:859–861
54. Wattenberg A, Sobott F, Barth HD et al (2000) Studying noncovalent protein complexes in aqueous solution with laser desorption mass spectrometry. Int J Mass Spectrom 203:49–57
55. Morgner N, Barth HD, Brutschy B (2006) A new way to detect noncovalently bonded complexes of biomolecules from liquid micro-droplets by laser mass spectrometry. Aust J Chem 59:109–114
56. Hoffmann J, Schmidt TL, Heckel A et al (2009) Probing the limits of liquid droplet laser desorption mass spectrometry in the analysis of oligonucleotides and nucleic acids. Rapid Commun Mass Spectrom 23:2176–2180
57. Morgner N, Barth HD, Brutschy B et al (2008) Binding sites of the viral RNA element TAR and of TAR mutants for various peptide ligands, probed with LILBID: a new laser mass spectrometry. J Am Soc Mass Spectrom 19:1600–1611
58. Morgner N, Kleinschroth T, Barth HD et al (2007) A novel approach to analyze membrane proteins by laser mass spectrometry: from protein subunits to the integral complex. J Am Soc Mass Spectrom 18:1429–1438
59. Morgner N, Zickermann V, Kerscher S et al (2008) Subunit mass fingerprinting of mitochondrial complex I. Biochim Biophys Acta Bioenerg 1777:1384–1391
60. Morgner N, Hoffmann J, Barth HD et al (2008) LILBID-mass spectrometry applied to the mass analysis of RNA polymerase II and an F1Fo-ATP synthase. Int J Mass Spectrom 277:309–313
61. Hoffmann J, Sokolova L, Preiss L et al (2010) ATP synthases: cellular nanomotors characterized by LILBID mass spectrometry. Phys Chem Chem Phys 12:13375–13382
62. Woods AS, Buchsbaum JC, Worrall TA et al (1995) Matrix-assisted laser desorption/ionization of noncovalently bound compounds. Anal Chem 67:4462–4465
63. Reichenbecher W, Rüdiger A, Kroneck PMH et al (1996) One molecule of molybdopterin guanine dinucleotide is associated with each subunit of the heterodimeric Mo-Fe-S protein transhydroxylase of *Pelobacter acidigallici* as determined by SDS/PAGE and mass spectrometry. Eur J Biochem 237:406–413
64. Salih B, Masselon C, Zenobi R (1998) Matrix-assisted laser desorption/ionization mass spectrometry of noncovalent protein-transition metal ion complexes. J Mass Spectrom 33:994–1002

65. Lehmann E, Zenobi R (1998) Detection of specific noncovalent zinc finger peptide-oligodeoxynucleotide complexes by matrix-assisted laser desorption/ionization mass spectrometry. Angew Chem Int Ed 37:3430–3432
66. Lehmann E, Zenobi R, Vetter S (1999) Matrix-assisted laser desorption/ionization mass spectra reflect solution-phase zinc finger peptide complexation. J Am Soc Mass Spectrom 10:27–34
67. Woods AS, Koomen JM, Ruotolo BT et al (2002) A study of peptide-peptide interactions using MALDI ion mobility o-TOF and ESI mass spectrometry. J Am Soc Mass Spectrom 13:166–169
68. Terrier P, Tortajada J, Zin G et al (2007) Noncovalent complexes between DNA and basic polypeptides or polyamines by MALDI-TOF. J Am Soc Mass Spectrom 18:1977–1989
69. Glocker MO, Bauer SHJ, Kast J et al (1996) Characterization of specific noncovalent protein complexes by UV matrix-assisted laser desorption ionization mass spectrometry. J Mass Spectrom 31:1221–1227
70. Lecchi P, Pannell LK (1995) The detection of intact double-stranded DNA by MALDI. J Am Soc Mass Spectrom 6:972–975
71. Lin S, Cotter RJ, Woods AS (1998) Detection of non-covalent interaction of single and double stranded DNA with peptides by MALDI-TOF. Protein Struct Funct Genet Suppl 2:12–21
72. Ohara K, Smietana M, Vasseur JJ (2006) Characterization of specific noncovalent complexes between guanidinium derivatives and single-stranded DNA by MALDI. J Am Soc Mass Spectrom 17:283–291
73. Lin S, Long S, Ramirez SM et al (2000) Characterization of the "helix clamp" motif of HIV-1 reverse transcriptase using MALDI-TOF MS and surface plasmon resonance. Anal Chem 72:2635–2640
74. Bahr U, Aygün H, Karas M (2008) Detection and relative quantification of siRNA double strands by MALDI mass spectrometry. Anal Chem 80:6280–6285
75. Sudha R, Zenobi R (2002) The detection and stability of DNA duplexes probed by MALDI mass spectrometry. Helv Chim Acta 85:3136–3143
76. Distler AM, Allison J (2002) Additives for the stabilization of double-stranded DNA in UV-MALDI MS. J Am Soc Mass Spectrom 13:1129–1137
77. Luo SZ, Li YM, Qiang W et al (2004) Detection of specific noncovalent interaction of peptide with DNA by MALDI-TOF. J Am Soc Mass Spectrom 15:28–31
78. Woods AS, Huestis MA (2001) A study of peptide-peptide interaction by matrix-assisted laser desorption/ionization. J Am Soc Mass Spectrom 12:88–96
79. Zehl M, Allmaier G (2003) Investigation of sample preparation and instrumental parameters in the matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of noncovalent peptide/peptide complexes. Rapid Commun Mass Spectrom 17:1931–1940
80. Ciruela F, Burgueno J, Casado V et al (2004) Combining mass spectrometry and pull-down techniques for the study of receptor heteromerization. Direct epitope-epitope electrostatic interactions between adenosine A2A and dopamine D2 receptors. Anal Chem 76:5354–5363
81. Woods AS (2004) The mighty arginine, the stable quaternary amines, the powerful aromatics, and the aggressive phosphate: their role in the noncovalent minuet. J Proteome Res 3:478–484
82. Woods A, Zangen A (2001) A direct chemical interaction between dynorphin and excitatory amino acids. Neurochem Res 26:395–400
83. Masselon C, Salih B, Zenobi R (1999) Matrix-assisted laser desorption/ionization Fourier transform mass spectrometry of luteinizing hormone releasing hormone-metal ion complexes. J Am Soc Mass Spectrom 10:19–26
84. Friess SD, Zenobi R (2001) Protein structure information from mass spectrometry? Selective titration of arginine residues by sulfonates. J Am Soc Mass Spectrom 12:810–818
85. Friess SD, Daniel JM, Hartmann R et al (2002) Mass spectrometric noncovalent probing of amino acids in peptides and proteins. Int J Mass Spectrom 219:269–281

86. Friess SD, Daniel JM, Zenobi R (2004) Probing the surface accessibility of proteins with noncovalent receptors and MALDI mass spectrometry. Phys Chem Chem Phys 6:2664–2675
87. Salih B, Zenobi R (1998) MALDI mass spectrometry of dye-peptide and dye-protein complexes. Anal Chem 70:1536–1543
88. Greiner G, Seyfarth L, Poppitz W et al (2000) Complexation of metal ions by pseudotripeptides with different functionalized N-alkyl residues. Lett Pept Sci 7:133–141
89. Little DP, Jacob A, Becker T et al (1997) Direct detection of synthetic and biologically generated double-stranded DNA by MALDI-TOF MS. Int J Mass Spectrom Ion Process 169–170:323–330
90. Tissot B, Gonnet F, Iborra A et al (2005) Mass spectrometry analysis of the oligomeric C1q protein reveals the B chain as the target of trypsin cleavage and interaction with fucoidan. Biochemistry 44:2602–2609
91. Schlosser G, Pocsfalvi G, Malorni A et al (2003) Detection of immune complexes by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 17:2741–2747
92. Kiselar JG, Downard KM (2000) Preservation and detection of specific antibody-peptide complexes by matrix-assisted laser desorption ionization mass spectrometry. J Am Soc Mass Spectrom 11:746–750
93. Cohen SL, Chait BT (1996) Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. Anal Chem 68:31–37
94. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10000 Daltons. Anal Chem 60:2299–2301
95. Kussmann M, Nordhoff E, Rahbek-Nielsen H et al (1997) Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. J Mass Spectrom 32:593–601
96. Davies GE, Stark GR (1970) Use of dimethyl suberimidate, a cross-linking reagent, in studying subunit structure of oligomeric proteins. Proc Natl Acad Sci USA 66:651–656
97. Evans JT, Rohrmann GF (1997) The baculovirus single-stranded DNA binding protein, LEF-3, forms a homotrimer in solution. J Virol 71:3574–3579
98. Helin J, Caldentey J, Kalkkinen N et al (1999) Analysis of the multimeric state of proteins by matrix assisted laser desorption/ionization mass spectrometry after cross-linking with glutaraldehyde. Rapid Commun Mass Spectrom 13:185–190
99. Wong SS (1991) Chemistry of protein conjugation and cross-linking. CRC, Boca Raton
100. Hermanson GT (2008) Bioconjugate techniques. Academic, San Diego
101. Pierce TS (2009) Crosslinking technical handbook. Thermo Fisher Scientific, Rockford, IL, USA
102. Kalkhof S, Sinz A (2008) Chances and pitfalls of chemical cross-linking with amine-reactive N-hydroxysuccinimide esters. Anal Bioanal Chem 392:305–312
103. Mädler S, Bich C, Touboul D et al (2009) Chemical cross-linking with NHS esters: a systematic study on amino acid reactivities. J Mass Spectrom 44:694–706
104. Borchers C, Tomer KB (1999) Characterization of the noncovalent complex of human immunodeficiency virus glycoprotein 120 with its cellular receptor CD4 by matrix-assisted laser desorption/ionization mass spectrometry. Biochemistry 38:11734–11740
105. Bich C, Scott M, Panagiotidis A et al (2008) Characterization of antibody-antigen interactions: comparison between surface plasmon resonance measurements and high-mass matrix-assisted laser desorption/ionization mass spectrometry. Anal Biochem 375:35–45
106. Bovet C, Ruff M, Eiler S et al (2008) Monitoring ligand modulation of protein-protein interactions by mass spectrometry: estrogen receptor α-SRC1. Anal Chem 80:7833–7839
107. Yanes O, Nazabal A, Wenzel R et al (2006) Detection of noncovalent complexes in biological samples by intensity fading and high-mass detection MALDI-TOF mass spectrometry. J Proteome Res 5:2711–2719

108. Pimenova T, Pereira CP, Schaer DJ et al (2009) Characterization of high molecular weight multimeric states of human haptoglobin and hemoglobin-based oxygen carriers by high-mass MALDI MS. J Sep Sci 32:1224–1230
109. Bich C, Baer S, Jecklin MC et al (2010) Probing the hydrophobic effect of noncovalent complexes by mass spectrometry. J Am Soc Mass Spectrom 21:286–289
110. Layh-Schmitt G, Podtelejnikov A, Mann M (2000) Proteins complexed to the P1 adhesin of *Mycoplasma pneumoniae*. Microbiology 146:741–747
111. Larsson T, Bergström J, Nilsson C et al (2000) Use of an affinity proteomics approach for the identification of low-abundant bacterial adhesins as applied on the Lewisb-binding adhesin of *Helicobacter pylori*. FEBS Lett 469:155–158
112. Woo JS, Kim DH, Allen PD et al (2008) TRPC3-interacting triadic proteins in skeletal muscle. Biochem J 411:399–405
113. Young MM, Tang N, Hempel JC et al (2000) High throughput protein fold identification by using experimental constraints derived from intramolecular cross-links and mass spectrometry. Proc Natl Acad Sci USA 97:5802–5806
114. Schilling B, Row RH, Gibson BW et al (2003) MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically crosslinked peptides. J Am Soc Mass Spectrom 14:834–850
115. de Koning LJ, Kasper PT, Back JW et al (2006) Computer-assisted mass spectrometric analysis of naturally occurring and artificially introduced cross-links in proteins and protein complexes. FEBS J 273:281–291
116. Yang T, Horejsh DR, Mahan KJ et al (1996) Mapping cross-linking sites in modified proteins with mass spectrometry: an application to cross-linked hemoglobins. Anal Biochem 242:55–63
117. Soubannier V, Rusconi F, Vaillier J et al (1999) The second stalk of the yeast ATP synthase complex: identification of subunits showing cross-links with known positions of subunit 4 (subunit b). Biochemistry 38:15017–15024
118. Fronzes R, Chaignepain S, Bathany K et al (2003) Topological and functional study of subunit h of the F1Fo ATP synthase complex in yeast *Saccharomyces cerevisiae*. Biochemistry 42:12038–12049
119. Schäfer I, Rössle M, Biuković G et al (2006) Structural and functional analysis of the coupling subunit F in solution and topological arrangement of the stalk domains of the methanogenic A1A0 ATP synthase. J Bioenerg Biomembr 38:83–92
120. Rappsilber J, Siniossoglou S, Hurt EC et al (2000) A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. Anal Chem 72:267–275
121. Chang Z, Kuchar J, Hausinger RP (2004) Chemical cross-linking and mass spectrometric identification of sites of interaction for UreD, UreF, and Urease. J Biol Chem 279:15305–15313
122. Kitatsuji C, Kurogochi M, Nishimura S-I et al (2007) Molecular basis of guanine nucleotide dissociation inhibitor activity of human neuroglobin by chemical cross-linking and mass spectrometry. J Mol Biol 368:150–160
123. Bennett KL, Kussmann M, Bjork P et al (2000) Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping - a novel approach to assess intermolecular protein contacts. Protein Sci 9:1503–1518
124. Müller DR, Schindler P, Towbin H et al (2001) Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis. Anal Chem 73:1927–1934
125. Back JW, Notenboom V, de Koning LJ et al (2002) Identification of cross-linked peptides for protein interaction studies using mass spectrometry and 18O labeling. Anal Chem 74:4417–4422
126. Petrotchenko EV, Olkhovik VK, Borchers CH (2005) Isotopically coded cleavable cross-linker for studying protein-protein interaction and protein complexes. Mol Cell Proteomics 4:1167–1179

127. Seebacher J, Mallick P, Zhang N et al (2006) Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing. J Proteome Res 5:2270–2282
128. Ihling C, Schmidt A, Kalkhof S et al (2006) Isotope-labeled cross-linkers and Fourier transform ion cyclotron resonance mass spectrometry for structural analysis of a protein/peptide complex. J Am Soc Mass Spectrom 17:1100–1113
129. Sinz A (2007) Isotope-labeled photoaffinity reagents and mass spectrometry to identify protein-ligand interactions. Angew Chem Int Ed 46:660–662
130. Hurst GB, Lankford TK, Kennel SJ (2004) Mass spectrometric detection of affinity purified crosslinked peptides. J Am Soc Mass Spectrom 15:832–839
131. Pearson KM, Pannell LK, Fales HM (2002) Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method. Rapid Commun Mass Spectrom 16:149–159
132. Sinz A, Kalkhof S, Ihling C (2005) Mapping protein interfaces by a trifunctional cross-linker combined with MALDI-TOF and ESI-FTICR mass spectrometry. J Am Soc Mass Spectrom 16:1921–1931
133. Ahrends R, Kosinski J, Kirsch D et al (2006) Identifying an interaction site between MutH and the C-terminal domain of MutL by crosslinking, affinity purification, chemical coding and mass spectrometry. Nucleic Acids Res 34:3169–3180
134. King GJ, Jones A, Kobe B et al (2008) Identification of disulfide-containing chemical cross-links in proteins using MALDI-TOF/TOF-mass spectrometry. Anal Chem 80:5036–5043
135. Krauth F, Ihling CH, Rüttinger HH et al (2009) Heterobifunctional isotope-labeled amine-reactive photo-cross-linker for structural investigation of proteins by matrix-assisted laser desorption/ionization tandem time-of-flight and electrospray ionization LTQ-Orbitrap mass spectrometry. Rapid Commun Mass Spectrom 23:2811–2818
136. Petrotchenko EV, Xiao KH, Cable J et al (2009) BiPS, a photocleavable, isotopically coded, fluorescent cross-linker for structural proteomics. Mol Cell Proteomics 8:273–286
137. Paramelle D, Cantel S, Enjalbal C et al (2009) A new generation of cross-linkers for selective detection by MALDI MS. Proteomics 9:5384–5388
138. Steen H, Jensen ON (2002) Analysis of protein-nucleic acid interactions by photochemical cross-linking and mass spectrometry. Mass Spectrom Rev 21:163–182
139. Jensen ON, Barofsky DF, Young MC et al (1993) Direct observation of UV-crosslinked protein-nucleic acid complexes by matrix-assisted laser desorption ionization mass spectrometry. Rapid Commun Mass Spectrom 7:496–501
140. Bennett SE, Jensen ON, Barofsky DF et al (1994) UV-catalyzed cross-linking of *Escherichia coli* uracil-DNA glycosylase to DNA. Identification of amino acid residues in the single-stranded DNA binding site. J Biol Chem 269:21870–21879
141. Connor DA, Falick AM, Young MC et al (1998) Probing the binding region of the single-stranded DNA-binding domain of rat DNA polymerase beta using nanosecond-pulse laser-induced cross-linking and mass spectrometry. Photochem Photobiol 68:299–308
142. Wang Q, Shoeman R, Traub P (2000) Identification of the amino acid residues of the amino terminus of vimentin responsible for DNA binding by enzymatic and chemical sequencing and analysis by MALDI-TOF. Biochemistry 39:6645–6651
143. Urlaub H, Hartmuth K, Lührmann R (2002) A two-tracked approach to analyze RNA-protein crosslinking sites in native, nonlabeled small nuclear ribonucleoprotein particles. Methods 26:170–181
144. Kühn-Hölsken E, Lenz C, Sander B et al (2005) Complete MALDI-ToF MS analysis of cross-linked peptide-RNA oligonucleotides derived from nonlabeled UV-irradiated ribonucleoprotein particles. RNA 11:1915–1930
145. Pingoud V, Geyer H, Geyer R et al (2005) Identification of base-specific contacts in protein-DNA complexes by photocrosslinking and mass spectrometry: a case study using the restriction endonuclease SsoII. Mol Biosyst 1:135–141

146. Kühn-Hölsken E, Dybkov O, Sander B et al (2007) Improved identification of enriched peptide–RNA cross-links from ribonucleoprotein particles (RNPs) by mass spectrometry. Nucleic Acids Res 35:e95
147. Kühn-Hölsken E, Lenz C, Dickmanns A et al (2010) Mapping the binding site of snurportin 1 on native U1 snRNP by cross-linking and mass spectrometry. Nucleic Acids Res 38:5581–5593
148. Geyer H, Geyer R, Pingoud V (2004) A novel strategy for the identification of protein-DNA contacts by photocrosslinking and mass spectrometry. Nucleic Acids Res 32:e132
149. Naryshkin NA, Farrow MA, Ivanovskaya MG et al (1997) Chemical cross-linking of the human immunodeficiency virus type 1 Tat protein to synthetic models of the RNA recognition sequence TAR containing site-specific trisubstituted pyrophosphate analogues. Biochemistry 36:3496–3505
150. Farrow MA, Aboul-ela F, Owen D et al (1998) Site-specific cross-linking of amino acids in the basic region of Human Immunodeficiency Virus type 1 Tat peptide to chemically modified TAR RNA duplexes. Biochemistry 37:3096–3108
151. Qin J, Chait BT (1997) Identification and characterization of posttranslational modifications of proteins by MALDI ion trap mass spectrometry. Anal Chem 69:4002–4009
152. Jensen ON, Kulkarni S, Aldrich JV et al (1996) Characterization of peptide-oligonucleotide heteroconjugates by mass spectrometry. Nucleic Acids Res 24:3866–3872
153. Pimenova T, Nazabal A, Roschitzki B et al (2008) Epitope mapping on bovine prion protein using chemical cross-linking and mass spectrometry. J Mass Spectrom 43:185–195
154. Dihazi GH, Sinz A (2003) Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry. Rapid Commun Mass Spectrom 17:2005–2014
155. Gomes AF, Gozzo FC (2010) Chemical cross-linking with a diazirine photoactivatable cross-linker investigated by MALDI- and ESI-MS/MS. J Mass Spectrom 45:892–899
156. Onisko B, Fernández EG, Freire ML et al (2005) Probing PrPSc structure using chemical cross-linking and mass spectrometry: evidence of the proximity of Gly90 amino termini in the PrP 27–30 aggregate. Biochemistry 44:10100–10109
157. Mädler S, Seitz M, Robinson J et al (2010) Does chemical cross-linking with NHS esters reflect the chemical equilibrium of protein-protein noncovalent interactions in solution? J Am Soc Mass Spectrom 21:1775–1783
158. Kiselar JG, Downard KM (1999) Antigenic surveillance of the influenza virus by mass spectrometry. Biochemistry 38:14185–14191
159. Kiselar JG, Downard KM (1999) Direct identification of protein epitopes by mass spectrometry without immobilization of antibody and isolation of antibody-peptide complexes. Anal Chem 71:1792–1801
160. Morrissey B, Downard KM (2008) Kinetics of antigen-antibody interactions employing a MALDI mass spectrometry immunoassay. Anal Chem 80:7720–7726
161. Morrissey B, Downard KM (2006) A proteomics approach to survey the antigenicity of the influenza virus by mass spectrometry. Proteomics 6:2034–2041
162. Morrissey B, Streamer M, Downard KM (2007) Antigenic characterisation of H3N2 subtypes of the influenza virus by mass spectrometry. J Virol Methods 145:106–114
163. Ho JWK, Morrissey B, Downard KM (2007) A computer algorithm for the identification of protein interactions from the spectra of masses (PRISM). J Am Soc Mass Spectrom 18:563–566
164. Villanueva J, Yanes O, Querol E et al (2003) Identification of protein ligands in complex biological samples using intensity-fading MALDI-TOF mass spectrometry. Anal Chem 75:3385–3395
165. Yanes O, Villanueva J, Querol E et al (2007) Detection of non-covalent protein interactions by 'intensity fading' MALDI-TOF mass spectrometry: applications to proteases and protease inhibitors. Nat Protocol 2:119–130

166. Yanes O, Villanueva J, Querol E et al (2005) Functional screening of serine protease inhibitors in the medical leech *Hirudo medicinalis* monitored by intensity fading MALDI-TOF MS. Mol Cell Proteomics 4:1602–1613
167. Yanes O, Aviles FX, Roepstorff P et al (2007) Exploring the "intensity fading" phenomenon in the study of noncovalent interactions by MALDI-TOF mass spectrometry. J Am Soc Mass Spectrom 18:359–367
168. Shabab M, Kulkarni MJ, Khan MI (2008) Study of papain-cystatin interaction by intensity fading MALDI-TOF-MS. Protein J 27:7–12
169. Sanglas L, Aviles FX, Huber R et al (2009) Mammalian metallopeptidase inhibition at the defense barrier of *Ascaris parasite*. Proc Natl Acad Sci USA 106:1743–1747
170. Wang Z, Yu X, Cui M et al (2009) Investigation of calmodulin-peptide interactions using matrix-assisted laser desorption/ionization mass spectrometry. J Am Soc Mass Spectrom 20:576–583
171. Mishra M, Tamhane VA, Khandelwal N et al (2010) Interaction of recombinant CanPIs with Helicoverpa armigera gut proteases reveals their processing patterns, stability and efficiency. Proteomics 10:2845–2857
172. Gimenez-Oya V, Villacanas O, Fernandez-Busquets X et al (2009) Mimicking direct protein-protein and solvent-mediated interactions in the CDP-methylerythritol kinase homodimer: a pharmacophore-directed virtual screening approach. J Mol Model 15:997–1007
173. Sugaya M, Saito R, Matsumura Y et al (2008) Facile detection of specific RNA-polypeptide interactions by MALDI-TOF mass spectrometry. J Pept Sci 14:978–983
174. Papac DI, Hoyes J, Tomer KB (1994) Direct analysis of affinity-bound analytes by MALDI/TOF MS. Anal Chem 66:2609–2613
175. Papac DI, Hoyes J, Tomer KB (1994) Epitope mapping of the gastrin-releasing peptide/anti-bombesin monoclonal antibody complex by proteolysis followed by matrix-assisted laser desorption ionization mass-spectrometry. Protein Sci 3:1485–1492
176. Zhao YM, Muir TW, Kent SBH et al (1996) Mapping protein-protein interactions by affinity-directed mass spectrometry. Proc Natl Acad Sci USA 93:4020–4024
177. Parker CE, Tomer KB (2002) MALDI/MS-based epitope mapping of antigens bound to immobilized antibodies. Mol Biotechnol 20:49–62
178. Raska CS, Parker CE, Sunnarborg SW et al (2003) Rapid and sensitive identification of epitope-containing peptides by direct matrix-assisted laser desorption/ionization tandem mass spectrometry of peptides affinity-bound to antibody beads. J Am Soc Mass Spectrom 14:1076–1085
179. Schlosser G, Vékey K, Malorni A et al (2005) Combination of solid-phase affinity capture on magnetic beads and mass spectrometry to study non-covalent interactions: example of minor groove binding drugs. Rapid Commun Mass Spectrom 19:3307–3314
180. Rüdiger AH, Rüdiger M, Carl UD et al (1999) Affinity mass spectrometry-based approaches for the analysis of protein-protein interaction and complex mixtures of peptide-ligands. Anal Biochem 275:162–170
181. Legros C, Guette C, Martin-Eauclaire M-F et al (2009) Affinity capture using chimeric membrane proteins bound to magnetic beads for rapid ligand screening by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Rapid Commun Mass Spectrom 23:745–755
182. Kull S, Pauly D, Störmann B et al (2010) Multiplex detection of microbial and plant toxins by immunoaffinity enrichment and matrix-assisted laser desorption/ionization mass spectrometry. Anal Chem 82:2916–2924
183. Peter J, Unverzagt C, Lenz H et al (1999) Purification of prostate-specific antigen from human serum by indirect immunosorption and elution with a hapten. Anal Biochem 273:98–104
184. Neubert H, Jacoby ES, Bansal SS et al (2002) Enhanced affinity capture MALDI-TOF MS: orientation of an immunoglobulin G using recombinant protein G. Anal Chem 74:3677–3683

185. Patrie SM, Mrksich M (2007) Self-assembled monolayers for MALDI-TOF mass spectrometry for immunoassays of human protein antigens. Anal Chem 79:5878–5887
186. Evans-Nguyen KM, Tao S-C, Zhu H et al (2008) Protein arrays on patterned porous gold substrates interrogated with mass spectrometry: detection of peptides in plasma. Anal Chem 80:1448–1458
187. Johnson EM, Ellis WR, Powers LS et al (2009) Affinity capture mass spectrometry of biomarker proteins using peptide ligands from biopanning. Anal Chem 81:5999–6005
188. Nelson RW, Nedelkov D, Tubbs KA (2000) Biosensor chip mass spectrometry: a chip-based proteomics approach. Electrophoresis 21:1155–1163
189. Nelson RW, Krone JR, Jansson O (1997) Surface plasmon resonance biomolecular interaction analysis mass spectrometry. 1. Chip-based analysis. Anal Chem 69:4363–4368
190. Krone JR, Nelson RW, Dogruel D et al (1997) BIA/MS: interfacing biomolecular interaction analysis with mass spectrometry. Anal Biochem 244:124–132
191. Nelson RW, Jarvik JW, Taillon BE et al (1999) BIA/MS of epitope-tagged peptides directly from *E. coli* lysate: multiplex detection and protein identification at low-femtomole to subfemtomole levels. Anal Chem 71:2858–2865
192. Bellon S, Buchmann W, Gonnet F et al (2009) Hyphenation of surface plasmon resonance imaging to matrix-assisted laser desorption ionization mass spectrometry by on-chip mass spectrometry and tandem mass spectrometry analysis. Anal Chem 81:7695–7702
193. Nedelkov D, Nelson RW (2003) Design and use of multi-affinity surfaces in biomolecular interaction analysis-mass spectrometry (BIA/MS): a step toward the design of SPR/MS arrays. J Mol Recognit 16:15–19
194. Mehlmann M, Garvin AM, Steinwand M et al (2005) Reflectometric interference spectroscopy combined with MALDI−TOF mass spectrometry to determine quantitative and qualitative binding of mixtures of vancomycin derivatives. Anal Bioanal Chem 382:1942–1948
195. Wang X, Chen G, Liu H et al (2010) Four-dimensional orthogonal electrophoresis system for screening protein complexes and protein-protein interactions combined with mass spectrometry. J Proteome Res 9:5325–5334
196. Shevchenko A, Schaft D, Roguev A et al (2002) Deciphering protein complexes and protein interaction networks by tandem affinity purification and mass spectrometry. Mol Cell Proteomics 1:204–212
197. Pflieger D, Junger MA, Müller M et al (2008) Quantitative proteomic analysis of protein complexes. Mol Cell Proteomics 7:326–346
198. Wales TE, Engen JR (2006) Hydrogen exchange mass spectrometry for the analysis of protein dynamics. Mass Spectrom Rev 25:158–170
199. Mandell JG, Falick AM, Komives EA (1998) Identification of protein-protein interfaces by decreased amide proton solvent accessibility. Proc Natl Acad Sci USA 95:14705–14710
200. Mandell JG, Baerga-Ortiz A, Akashi S et al (2001) Solvent accessibility of the thrombin-thrombomodulin interface. J Mol Biol 306:575–589
201. Baerga-Ortiz A, Hughes CA, Mandell JG et al (2002) Epitope mapping of a monoclonal antibody against human thrombin by H/D-exchange mass spectrometry reveals selection of a diverse sequence in a highly conserved protein. Protein Sci 11:1300–1308
202. Tuma R, Coward LU, Kirk MC et al (2001) Hydrogen-deuterium exchange as a probe of folding and assembly in viral capsids. J Mol Biol 306:389–396
203. Hosia W, Johansson J, Griffiths WJ (2002) Hydrogen/deuterium exchange and aggregation of a polyvaline and a polyleucine alpha-helix investigated by matrix-assisted laser desorption ionization mass spectrometry. Mol Cell Proteomics 1:592–597
204. Nazabal A, Dos RS, Bonneu M et al (2003) Conformational transition occurring upon amyloid aggregation of the HET-s prion protein of *Podospora anserina* analyzed by hydrogen/deuterium exchange and mass spectrometry. Biochemistry 42:8852–8861
205. Kraus M, Bienert M, Krause E (2003) Hydrogen exchange studies on Alzheimer's amyloid-β peptides by mass spectrometry using matrix-assisted laser desorption/ionization and electrospray ionization. Rapid Commun Mass Spectrom 17:222–228

206. Nazabal A, Hornemann S, Aguzzi A et al (2009) Hydrogen/deuterium exchange mass spectrometry identifies two highly protected regions in recombinant full-length prion protein amyloid fibrils. J Mass Spectrom 44:965–977
207. Turner BT Jr, Maurer MC (2002) Evaluating the roles of thrombin and calcium in the activation of coagulation factor XIII using H/D exchange and MALDI-TOF MS. Biochemistry 41:7947–7954
208. Catalina MI, Fischer MJE, Dekker FJ et al (2005) Binding of a diphosphorylated-ITAM peptide to spleen tyrosine kinase (Syk) induces distal conformational changes: a hydrogen exchange mass spectrometry study. J Am Soc Mass Spectrom 16:1039–1051
209. Sabo TM, Farrell DH, Maurer MC (2006) Conformational analysis of gamma' peptide (410-427) interactions with thrombin anion binding exosite II. Biochemistry 45:7434–7445
210. Nazabal A, Laguerre M, Schmitter J-M et al (2003) Hydrogen/deuterium exchange on yeast ATPase supramolecular protein complex analyzed at high sensitivity by MALDI mass spectrometry. J Am Soc Mass Spectrom 14:471–481
211. Kipping M, Schierhorn A (2003) Improving hydrogen/deuterium exchange mass spectrometry by reduction of the back-exchange effect. J Mass Spectrom 38:271–276
212. Hotchko M, Anand GS, Komives EA et al (2006) Automated extraction of backbone deuteration levels from amide H/2 H mass spectrometry experiments. Protein Sci 15:583–601
213. Jørgensen TJD, Bache N, Roepstorff P et al (2005) Collisional activation by MALDI tandem time-of-flight mass spectrometry induces intramolecular migration of amide hydrogens in protonated peptides. Mol Cell Proteomics 4:1910–1919
214. Bache N, Rand KD, Roepstorff P et al (2008) Gas-phase fragmentation of peptides by MALDI in-source decay with limited amide hydrogen (1H/2H) scrambling. Anal Chem 80:6431–6435
215. Powell KD, Wales TE, Fitzgerald MC (2002) Thermodynamic stability measurements on multimeric proteins using a new H/D exchange- and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry-based method. Protein Sci 11:841–851
216. Powell KD, Ghaemmaghami S, Wang MZ et al (2002) A general mass spectrometry-based assay for the quantitation of protein-ligand binding interactions in solution. J Am Chem Soc 124:10256–10257
217. Ghaemmaghami S, Fitzgerald MC, Oas TG (2000) A quantitative, high-throughput screen for protein stability. Proc Natl Acad Sci USA 97:8296–8301
218. Ghaemmaghami S, Oas TG (2001) Quantitative protein stability measurement in vivo. Nat Struct Biol 8:879–882
219. Powell KD, Fitzgerald MC (2003) Accuracy and precision of a new H/D exchange- and mass spectrometry-based technique for measuring the thermodynamic properties of protein-peptide complexes. Biochemistry 42:4962–4970
220. West GM, Tang L, Fitzgerald MC (2008) Thermodynamic analysis of protein stability and ligand binding using a chemical modification- and mass spectrometry-based strategy. Anal Chem 80:4175–4185
221. West GM, Thompson JW, Soderblom EJ et al (2010) Mass spectrometry-based thermal shift assay for protein-ligand binding analysis. Anal Chem 82:5573–5581
222. Hofner G, Merkel D, Wanner KT (2009) MS binding assays-with MALDI toward high throughput. ChemMedChem 4:1523–1528

Top Curr Chem (2013) 331: 37–54
DOI: 10.1007/128_2012_321
© Springer-Verlag Berlin Heidelberg 2012
Published online: 1 May 2012

Application of MALDI-TOF-Mass Spectrometry to Proteome Analysis Using Stain-Free Gel Electrophoresis

Iuliana Susnea, Bogdan Bernevic, Michael Wicke, Li Ma, Shuying Liu, Karl Schellander, and Michael Przybylski

Abstract The combination of MALDI-TOF-mass spectrometry with gel electrophoretic separation using protein visualization by staining procedures involving such as Coomassie Brilliant Blue has been established as a widely used approach in proteomics. Although this approach has been shown to present high detection sensitivity, drawbacks and limitations frequently arise from the significant background in the mass spectrometric analysis. In this chapter we describe an approach for the application of MALDI-MS to the mass spectrometric identification of proteins from one-dimensional (1D) and two-dimensional (2D) gel electrophoretic separation, using stain-free detection and visualization based on native protein fluorescence. Using the native fluorescence of aromatic protein amino acids with UV transmission at 343 nm as a fast gel imaging system, unstained protein spots are localized and, upon excision from gels, can be proteolytically digested and analyzed by MALDI-MS. Following the initial development and testing with

I. Susnea, B. Bernevic and M. Przybylski (✉)
Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, 78457 Konstanz, Germany
e-mail: Michael.Przybylski@uni-konstanz.de

M. Wicke
Institute of Animal Breeding and Genetics, University of Göttingen, Göttingen, Germany

L. Ma
Laboratory of Analytical Chemistry and Biopolymer Structure Analysis, Department of Chemistry, University of Konstanz, 78457 Konstanz, Germany

Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, People's Republic of China

S. Liu
Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun, People's Republic of China

K. Schellander
Department of Animal Physiology and Veterinary Medicine, University of Bonn, Bonn, Germany

standard proteins, applications of the stain-free gel electrophoretic detection approach to mass spectrometric identification of biological proteins from 2D-gel separations clearly show the feasibility and efficiency of this combination, as illustrated by a proteomics study of porcine skeleton muscle proteins. Major advantages of the stain-free gel detection approach with MALDI-MS analysis are (1) rapid analysis of proteins from 1D- and 2D-gel separation without destaining required prior to proteolytic digestion, (2) the low detection limits of proteins attained, and (3) low background in the MALDI-MS analysis.

Keywords Gel electrophoresis · MALDI-TOF-mass spectrometry · Native fluorescence · Protein identification · Skeleton muscle proteomics

Contents

Abbreviations

1D	One-dimensional gel electrophoresis
2D	Two-dimensional gel electrophoresis
MALDI-TOF	Matrix assisted laser desorption/ionization–time-of-flight
MS	Mass spectrometry
PMF	Peptide mass fingerprinting
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis

1 Introduction

Matrix assisted laser desorption/ionization–mass spectrometry (MALDI-MS), introduced by Karas and Hillenkamp in 1988 [1, 2], is now widely used in proteomics studies. Initially developed for the ionization of large polypeptides

and proteins [3], applications of MALDI-MS have significantly broadened and incorporated glycoproteins, oligonucleotides, carbohydrates, and small biomolecules [4]. MALDI is referred to as a "soft" ionization technique, because it causes minimal or no fragmentation and allows the molecular ions of analytes to be identified, even in complex mixtures of biopolymers. The ionization–desorption principle of MALDI-MS is based on the co-crystallization of analytes with an organic, light-absorbing matrix (e.g., α-cyano-4-hydroxy cinnamic acid or sinapinic acid) which, when activated by a laser, ionizes the analyte as it enters the gas phase. The ions once formed are accelerated in an electric field and separated according to their mass-to-charge ratio (m/z) in the mass spectrometer analyzer. Typically, MALDI is coupled with time-of-flight (TOF) analyzers that determine the mass of intact biopolymers.

The most common lasers used in MALDI-MS are ultraviolet (UV) lasers. Most of the commercially available MALDI mass spectrometers are equipped with nitrogen lasers ($\lambda = 337$ nm) which are used as the standard device, although Nd:YAG lasers ($\lambda = 266$ or 355 nm) are also employed. MALDI-MS can also use infrared (IR) lasers such as Er:YAG lasers ($\lambda = 2.94$ μm) or CO_2 lasers ($\lambda = 10.6$ μm), and thus can be employed in applications to proteome analysis [5, 6].

MALDI-TOF-MS is a well established method in peptide and protein analysis because of its robust, simple operation and high sensitivity, and the coupling of MALDI-TOF as well as high resolution analyzers, such as FTICR with gel electrophoretic separation has enabled successful protein identifications in recent years [7–14]. The sequence of steps in a typical proteomics experiment is schematically outlined in Fig. 1: (1) first, proteins of interest from a biological mixture are separated by one-dimensional (1D) or two-dimensional (2D) gel electrophoresis; (2) following the gel electrophoretic separation, proteins are visualized using a staining procedure; (3) the protein bands (spots) of interest are excised from the gel and digested by a protease of high specificity (e.g., trypsin); (4) the resulting mixture of proteolytic peptides is analyzed by MALDI-MS yielding a peptide mass map; and (5) identification of proteins is obtained by searching for the best match between the experimentally determined masses of the peptide map and peptide masses calculated from theoretical cleavage of proteins in an appropriate sequence database [15].

In order to visualize proteins separated by gel electrophoresis a number of techniques have been developed in recent years. Most mass spectrometric proteomics studies employ staining procedures with such as Coomassie Brilliant Blue or silver salts, but fluorescent dyes of high detection sensitivity have also been used (Flamingo, SYPRO® Ruby) [16–18]. Although several of these approaches provide high sensitivity and are easy to use, major problems are frequently encountered with the compatibility of staining procedures with the mass spectrometric analysis [19]. A more recently explored alternative to omit the use of dyes in the visualization procedure has been the development of methods based on the fluorescent properties of proteins [20, 21]. During fluorescence labeling studies of glycoproteins, Zhao and co-workers observed a fluorescent signal for non-glycosylated proteins such as hen eggwhite lysozyme, which was attributed to

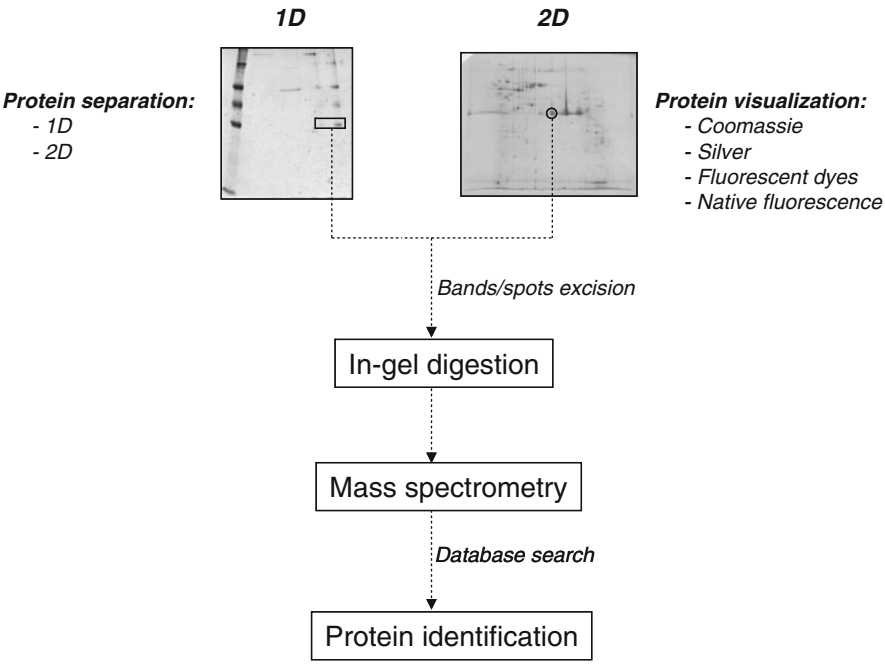


Fig. 1 Scheme of the steps involved in a proteomics experiment. After gel separation (1D or 2D) proteins are visualized by staining methods such as Coomassie, silver, and fluorescent dyes, or by "stain-free" native fluorescence. Protein spots are excised from gels, digested with trypsin, and digestion mixtures are analyzed by MALDI-MS

intrinsic (native) protein fluorescence [22]. A first detection method for unstained proteins based on UV fluorescence was developed by Roegener et al. who used laser excitation with 280 nm UV light and demonstrated protein visualization in both 1D- and 2D-gel separations with detection limits in the low nanogram range (1–5 ng) [23]. Recently, a commercial gel-analyzer based on native fluorescence has been developed (LaVision-BioTec; Bielefeld, Germany) and employed in the present study.

In this work we have developed and applied native fluorescence detection of proteins in stain-free one- and two-dimensional gel electrophoretic separations as a sensitive and efficient approach for mass spectrometric identifications in proteome analysis. In initial testing experiments 1D-gels of model proteins were analyzed to investigate (1) the relation between fluorescence intensity observed and the relative amounts of aromatic amino acids in proteins, (2) detection sensitivity of the native fluorescence in comparison with Coomassie and silver staining sensitivities, and (3) the applicability of native fluorescence detection to mass spectrometric protein identification. In a second part, the stain-free gel bioanalyzer was successfully employed in applications to porcine skeleton muscle proteomics, providing

identifications of proteins at high detection sensitivity, without the need for staining and destaining isolated protein bands.

2 Methods

2.1 Protein Separation by Gel Electrophoresis

Model proteins used for evaluation in the stain-free gel bioanalyzer were separated by 15% sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on 1-mm gels using the standard Laemmli method with a Mini-PROTEAN®3 cell gel system (Bio-Rad, München, Germany). Myoglobin, ubiquitin, bovine serum albumin (BSA), carbonic anhydrase, lysozyme, and α-casein were purchased from Sigma-Aldrich Chemie GmbH (Taufkirchen, Germany). Pepsin was from Fluka Chemie GmbH (Buchs, Switzerland), and human γ-globulin was from Merck4Biosciences (Darmstadt, Germany).

Porcine skeleton muscle samples for 2D-gel separations were prepared as previously described [24] (the samples were isolated from *Longissimus dorsi* muscle and were provided by the Department of Animal Breeding, University of Bonn, Germany). Samples of 800 μg total protein were applied for 12 h on 17-cm IPG strips (pH 3–10) using a passive in-gel rehydration method. Isoelectric focusing (IEF) was carried out using a Multiphor horizontal electrophoresis system (Amersham Biosciences, München, Germany). For the second separation step, the Bio-Rad Protean-II-xi vertical electrophoresis system was used, and 12.5% SDS-PAGE gels of 1.5 mm thickness were prepared. Electrophoresis was performed in two steps: (1) 25 mA/gel for approximately 30 min, and (2) 40 mA/gel until the dye front reached the anodic end of the gels. All buffers and solutions used for 2D-gel electrophoresis have been described elsewhere [7].

2.2 Gel Bioanalyzer for Protein Detection and Visualization

Proteins separated by 1D- or 2D-gel electrophoresis were visualized with the gel bioanalyzer (LaVision-Biotec, Bielefeld, Germany; http://www.lavisionbiotec.com/en/microscopy-products/gelreader/). The experimental setup of the gel bioanalyzer is based on a UV excitation source and a detection system within the UV range. The UV excitation light was generated by a 300-W xenon lamp (265–680 nm). The irradiation area was set to 1 cm^2 at 35 mW/cm^2 and imaged by three lenses onto a photomultiplier detector. A UV bandpass filter (280–400 nm) is incorporated to block the excitation light from the detection system. From four filter positions (one for UV excitation, three for visible fluorescence), the UV filter transmitting light at $\lambda = 343 \pm 65/2$ nm was employed. The large reading area

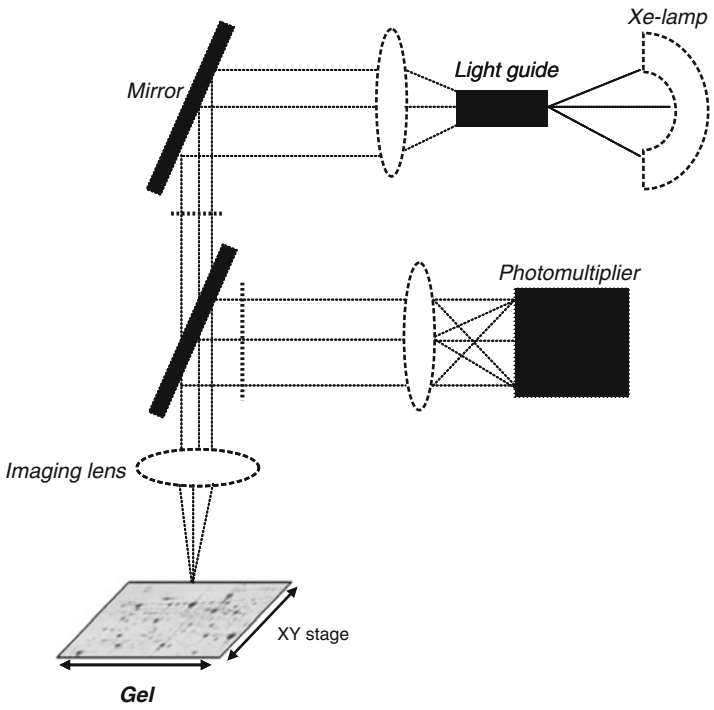


Fig. 2 Scheme of the gel bioanalyzer (LaVision-Biotec, Bielefeld, Germany), modified after http://www.lavisionbiotec.com/en/microscopy-products/gelreader/

(30×35 cm^2) provided scanning of both 1D- and 2D-gels. The instrument has a removable gel tray and is equipped to read unstained as well as stained protein gels (Fig. 2). In the present study only scanning of unstained gels was performed. High precision polycarbonate tools for localization and isolation of protein spots were prepared in our Laboratory [7]. Following fixation in position on the gel tray, localization and excision of gel spots was carried out by moving the gel tray, with positioning and scanning of the gel controlled by the LaVision-Biotec scanning software. Using polycarbonate tools, different small holes were made in the scanned gel in order to isolate the protein bands (four fine holes were necessary for the localization of each protein band), and upon gel ejection from the gel bioanalyzer bands were excised with a scalpel.

2.3 *In-Gel Proteolytic Digestion*

Following detection, visualization, and localization, spots were manually excised from the gels and subjected to in-gel trypsin digestion according to Mortz et al. [25].

No destaining steps were required, since no visible staining procedure was used. The complete protocol has been previously described [7]. The resulting supernatant and elution fractions were combined and lyophilized to dryness for mass spectrometric investigation.

2.4 MALDI-TOF-Mass Spectrometry

MALDI-TOF-MS was carried out with a Waters VG-Micromass TOFSpec-2DE mass spectrometer (Waters Micromass, Manchester, UK) equipped with a nitrogen UV laser (337 nm), channel plate detector, and MASSLynx 4.0 data system for spectra acquisition and instrument control. A saturated solution of α-cyano-4-hydroxy-cinnamic acid (HCCA) in acetonitrile/0.1% trifluoroacetic acid in water (2:1 v/v) was used as the matrix. Aliquots of 0.8 μL of the sample solution and saturated matrix solution were mixed on the stainless steel MALDI target and allowed to dry. Acquisition of spectra was carried out at an acceleration voltage of 20 kV.

2.5 Database Search

Digestion mixtures determined by MALDI-MS were directly used for a database search employing the MASCOT peptide mass fingerprinting (PMF) search engine (http://www.matrixscience.com), employing search and acceptance criteria for protein identification as follows: 0.5–1.2 Da mass error tolerance; two missed cleavage sites permitted; methionine oxidation as variable modification; carbamidomethyl (cysteine) as fixed modification. The database employed was NCBInr 20060712 (3,783,042 sequence entries, 1,304,471,729 residues), a compilation of several databases including SWISS-PROT, PIR, PRF, PDB, and GenBank CDS translations.

3 Results and Discussion

3.1 Evaluation of Stain-Free Native Fluorescence for Protein Detection and Visualization

Conventional staining procedures used to visualize proteins within gel electrophoretic separations present a number of problems, such as high background and compatibility problems with mass spectrometry procedures (e.g., solvents) [19], high costs of fluorescent dyes, and extensive analysis times required for staining

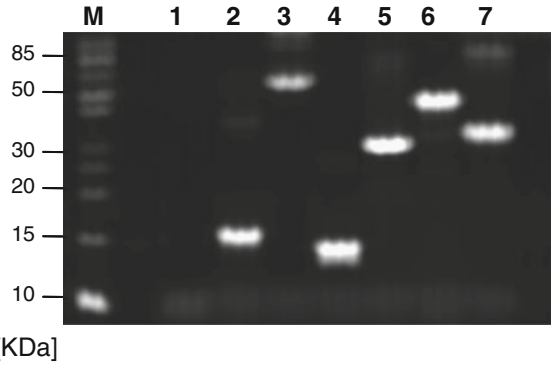


Fig. 3 Native fluorescence visualization and detection for a 15% SDS-PAGE separation. Fluorescence intensity depends on the amount of aromatic amino acids in proteins (see Table 1 for aromatic amino acid and tryptophan amounts, given in pmol). *M* – molecular weight marker (Fermentas; 10–200 kDa; 5 μL); (*1*) ubiquitin (5 μg); (*2*) myoglobin (5 μg); (*3*) bovine serum albumin (BSA; 5 μg); (*4*) lysozyme (5 μg); (*5*) carbonic anhydrase (5 μg); (*6*) pepsin (5 μg); (*7*) α-casein (5 μg)

Table 1 Tryptophan and aromatic amino acid amounts (pmol) in seven different proteins. Proteins were separated in 15% SDS-PAGE gels (see Fig. 3). Fluorescence intensity values exhibited by these proteins and protein amounts (pmol) applied on each band are also listed

Lane	Protein	MW (Da)	Protein amount/ band (pmol)	Amount of aromatic aa/ band (pmol)	Amount of tryptophan/ band (pmol)	Fluorescence intensity (no. of counts)
1	Ubiquitin	8,565	583.8	22.8	0	2,816
2	Myoglobin	17,070	292.9	20.8	3.8	16,820
3	BSA	69,239	72.2	6.4	0.4	10,381
4	Lysozyme	14,309	349.4	32.5	16.4	19,028
5	Carbonic anhydrase	29,100	171.8	17.2	4.6	21,086
6	Pepsin	41,300	121.1	12.4	3.3	19,132
7	α-Casein	26,019	192.2	20	1.7	14,081

aa Amino acids

and destaining of gels [25]. The native fluorescence of a protein is a composite of the fluorescence from individual aromatic residues. Most of the native fluorescence emission of a protein is due to tryptophan residues, with minor contribution from tyrosine and phenylalanine residues. In order to characterize the native fluorescence contributions, several model proteins with different contents of aromatic amino acids (percentage of Trp) were separated by gel electrophoresis using 15% SDS-PAGE, and gels were scanned with the gel-bioanalyzer (Fig. 3, Table 1). From each protein 5 μg were applied on the gel. In the gel shown in Fig. 3, ubiquitin (lane 1) which has only three aromatic amino acids (22.8 pmol) and no tryptophan gave only a weak fluorescence signal (see Table 1). When comparing fluorescence intensities for carbonic anhydrase (lane 5) and α-casein (lane 7), both having similar molecular weights and similar contents of aromatic amino acids (carbonic

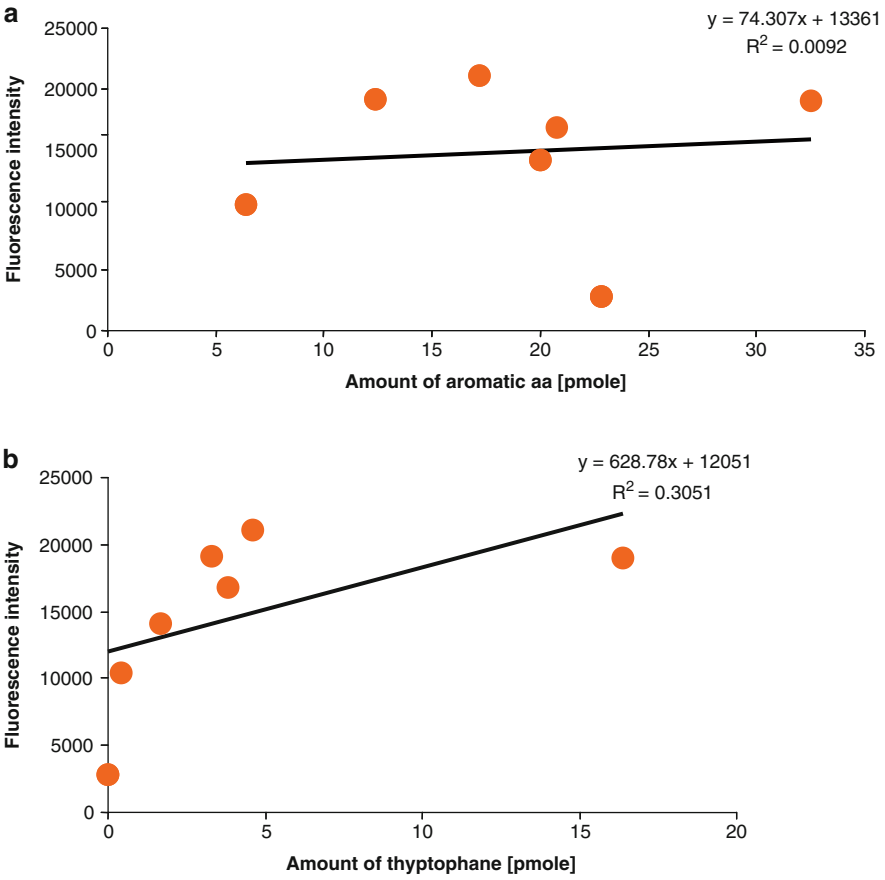


Fig. 4 (**a**) Linear regression for the investigation of the contribution of aromatic amino acid amounts (pmol) to the fluorescence signal intensity. (**b**) Linear regression for the investigation of the contribution of tryptophan amounts (pmol) to the fluorescence signal intensity. Fluorescence intensity values, aromatic amino acid amounts, and tryptophan amounts (pmol) correspond to the proteins separated in the gel presented in Fig. 3 (see Table 1)

anhydrase – 17.2 pmol; α-casein – 20 pmol) but different contributions of tryptophan to the aromatic protein content (4.6 pmol for carbonic anhydrase which represents 26.7% tryptophan contribution to the total aromatic protein amount; 1.7 pmol, 8.5%, respectively, for α-casein), a higher fluorescence intensity signal was observed for carbonic anhydrase (see Table 1 for fluorescence intensity values). Another comparison was made between myoglobin (lane 2) and lysozyme (lane 4). Lysozyme with 50.5% tryptophan contribution to the fluorescence signal (32.5 pmol aromatic amino acids and 16.4 pmol tryptophan) showed a higher fluorescence than myoglobin (18.3% tryptophan contribution to the fluorescence signal; 20.8 pmol aromatic amino acids and 3.8 pmol tryptophan) (Fig. 3,

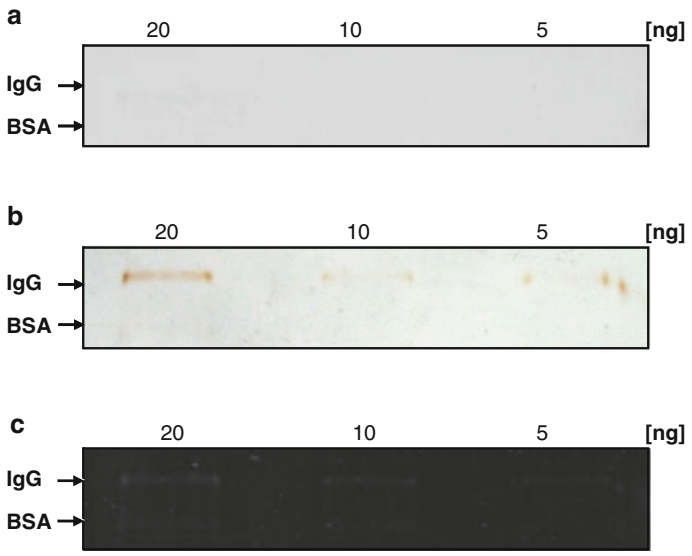


Fig. 5 Sensitivity of stain-free fluorescence detection and visualization in comparison with Coomassie and silver visualizations. Protein samples, IgG (150 kDa heavy and light chain dimer) and BSA (67 kDa) were separated in 3 lanes at 20–5 ng. Gel areas presented are zoomed regions from 12% SDS-PAGE separations. (**a**) Coomassie stained gel; (**b**) silver stained gel; (**c**) native fluorescence gel

Table 1). These results clearly illustrate the dependence of fluorescence intensity on the amounts of tryptophan and other aromatic amino acid residues in proteins.

Furthermore, linear regression was used in order to find the relationship between the amount of aromatic amino acids and tryptophan (given in pmol) towards the detection signal intensity of the proteins separated in the gel from Fig. 3 (Fig. 4a, b).

A further step in the evaluation of the stain-free native fluorescence detection method was to test its sensitivity. Sensitivity tests were performed with 1D-gel separations of mixtures of two model proteins, immunoglobulin-G and BSA, which were scanned with the gel bioanalyzer (Fig. 5c) at concentrations of 20–5 ng/band and compared with gels prepared at identical conditions but visualized using standard staining procedures (Coomassie – Fig. 5a and silver – Fig. 5b). These results showed comparable sensitivities for the UV fluorescence detection and silver staining, with detection limits of approximately 1–5 ng [7]. The detection limit in the low nanogram range is in good agreement with sensitivity data reported by Roegener et al. [23].

In another set of experiments it was shown that the fluorescence intensity of the protein bands in 1D-gels increases linearly as the protein amount, the amount of aromatic amino acids, and the mass of tryptophan increases, as shown in Fig. 6a–d.

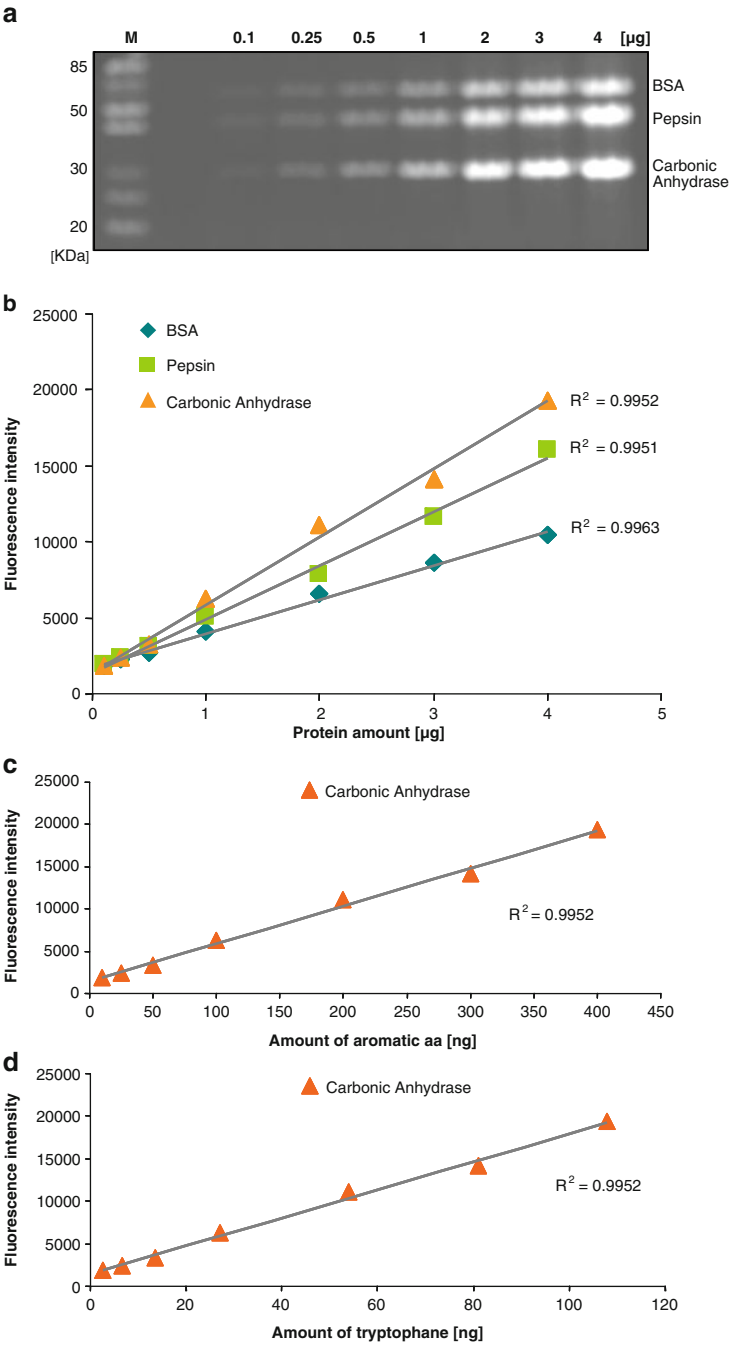


Fig. 6 (**a**) 15% SDS-PAGE separation of a protein mixture (bovine serum albumin – *BSA*; pepsin; carbonic anhydrase) at different concentrations (0.1–4 µg). Proteins were visualized by native fluorescence using the gel bioanalyzer instrument. Fluorescence intensity increases linearly with protein concentration (**b**), with the amount of aromatic amino acids (**c**), and with the content of tryptophan in a protein band (**d**). *M* – molecular weight marker (Fermentas; 5 µL)

Fifteen percent SDS-PAGE was used to separate a protein mixture (BSA; pepsin; carbonic anhydrase) at different concentrations (0.1–4 μg) (Fig. 6a). Proteins were visualized by native fluorescence using the gel bioanalyzer instrument. All the proteins in the gel were detectable at 0.1 μg. UV-fluorescence detection offers a linear dynamic range from 0.1 to 4 μg with a correlation coefficient of 0.99 (Fig. 6b–d).

3.2 *Application of Stain-Free Native Fluorescence Detection to MALDI-TOF-MS Identification of 1D-Gel Separated Proteins*

Following development and optimization, the stain-free detection method in gel electrophoresis was subjected to mass spectrometric identifications of 1D-gel separated proteins. The identification of horse heart myoglobin (5 μg/gel) from an unstained gel (see Fig. 3, lane 2) was successfully achieved. After gel bioanalyzer scanning the protein band was excised, subjected to in-gel digestion with trypsin, and the digestion mixture analyzed by MALDI-TOF-MS. The resulting masses were used for a database search with the MASCOT PMF search engine, and provided unequivocal identification of horse heart myoglobin with 18 identified peptides (data not shown). Since no destaining step was required for the in-gel digestion, high sensitivity and considerably lower sample preparation time were needed compared to conventional Coomassie staining. Identifications were obtained with significantly lower protein amounts used for 1D-gel separation, with a score of 86 (64% sequence coverage) and 320 ng protein band (Fig. 7a). For comparison reasons the same amount of myoglobin (320 ng) but from a gel stained with Coomassie was used and led to protein identification with a score of 117% and 82% sequence coverage (Fig. 7b). In summary, these model studies suggested that gel separation of proteins with native fluorescence detection represents an efficient and sensitive approach for MALDI-MS identification in proteomics.

3.3 *Application of Stain-Free Native Fluorescence to Mass Spectrometric Proteome Analysis of Porcine Muscle Tissue*

In subsequent proteomics studies, the stain-free detection approach was successfully applied to protein identifications from 2D gels by MALDI-TOF-MS (Fig. 8). Examples of proteome analyses of porcine muscle skeleton proteins isolated post-mortem using the stain-free gel bioanalyzer are summarized in Figs. 9, 10, and 11. The rate and extent of post-mortem metabolic processes of skeleton muscle proteins have recently found increasing interest, and it is generally believed that structural

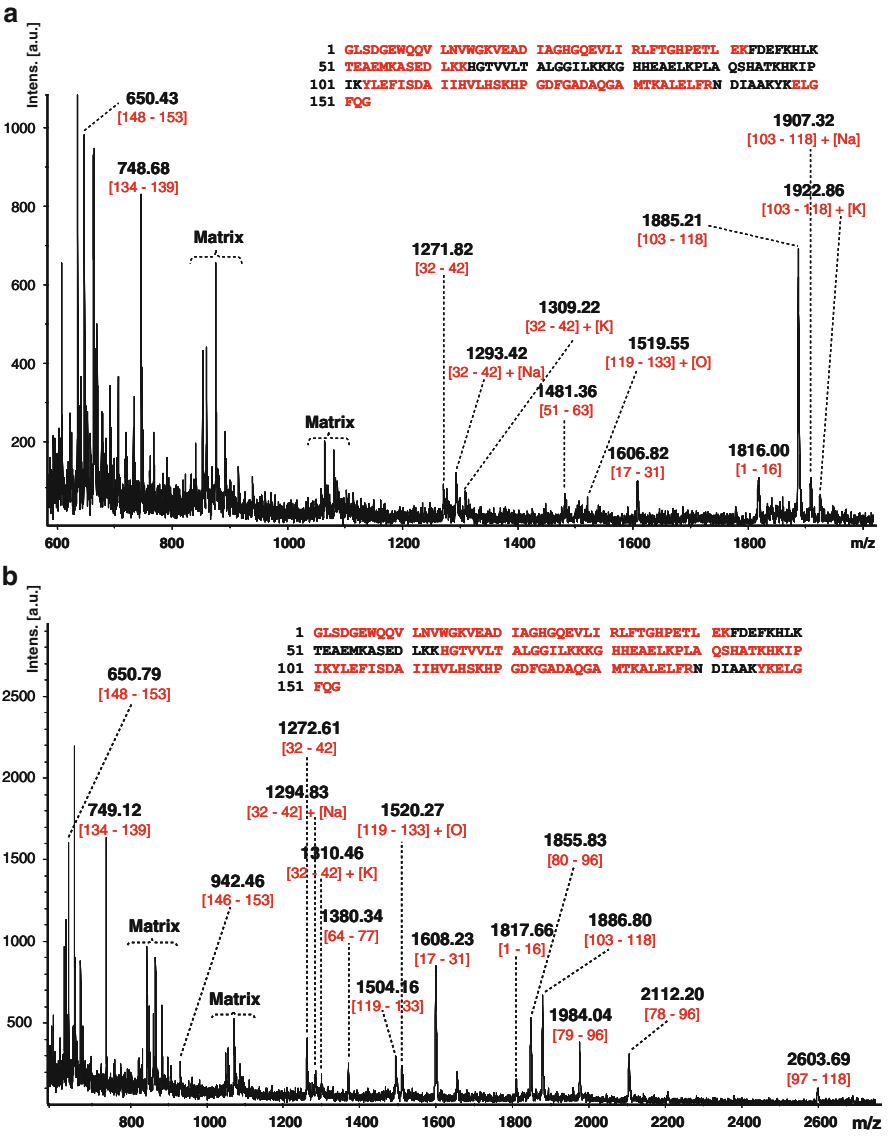


Fig. 7 MALDI-TOF-mass spectrometric identification of horse heart myoglobin (320 ng) from a stain-free gel (**a**) and from a Coomassie stained gel (**b**)

changes such as degradation and oxidation forms post-mortem may be indicative as biomarkers and affect meat properties [26]. Thus, tenderization processes have been associated with calpains and calpain inhibitors, calpastatins, that potentially influence proteolytic changes, and with proteins involved in carbonylation that may be potential oxidation biomarkers [27–29].

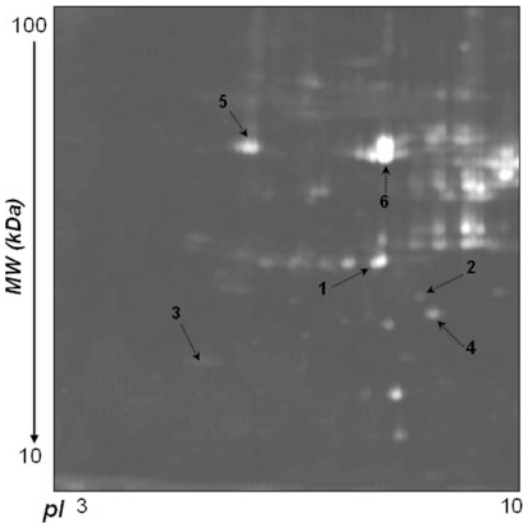


Fig. 8 The 2D-gel of a post-mortem porcine muscle sample (12.5% SDS-PAGE; 800 µg total protein per gel) was visualized by native fluorescence. Spots 1–6 were excised from the gel, digested with trypsin, and used for protein identification

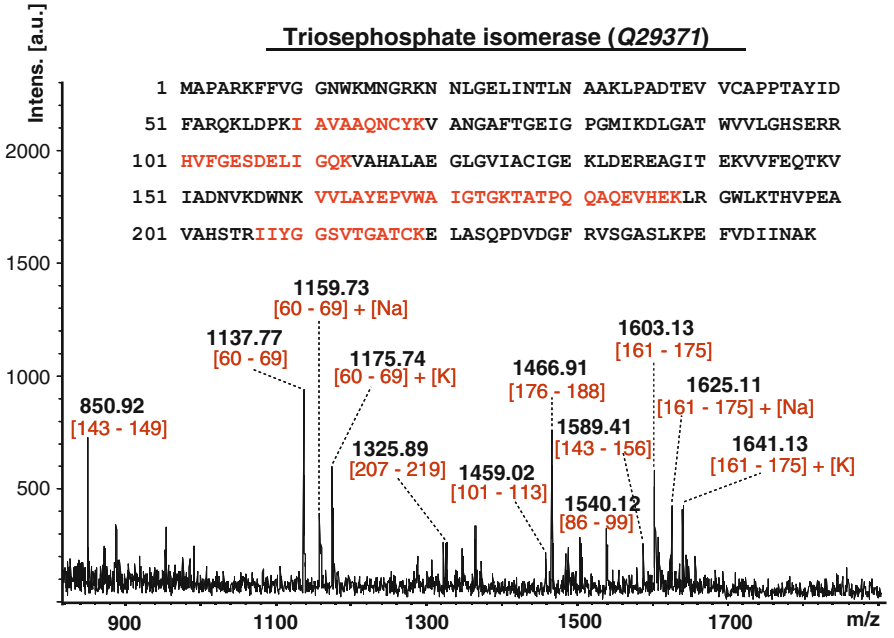


Fig. 9 MALDI-TOF-mass spectrum of the digestion mixture of spot number 1 (see Fig. 8). Labeled peaks correspond to the identified peptides from porcine skeletal triosephosphate isomerase (identified peptides are shown in *red* in the amino acid sequence of the protein)

A total amount of 800 µg was used for the 2D-gel electrophoretic separation of porcine muscle proteins (see Fig. 8). The gel was scanned with the gel bioanalyzer and proteins to be analyzed by MALDI-TOF-MS were excised using high-precision

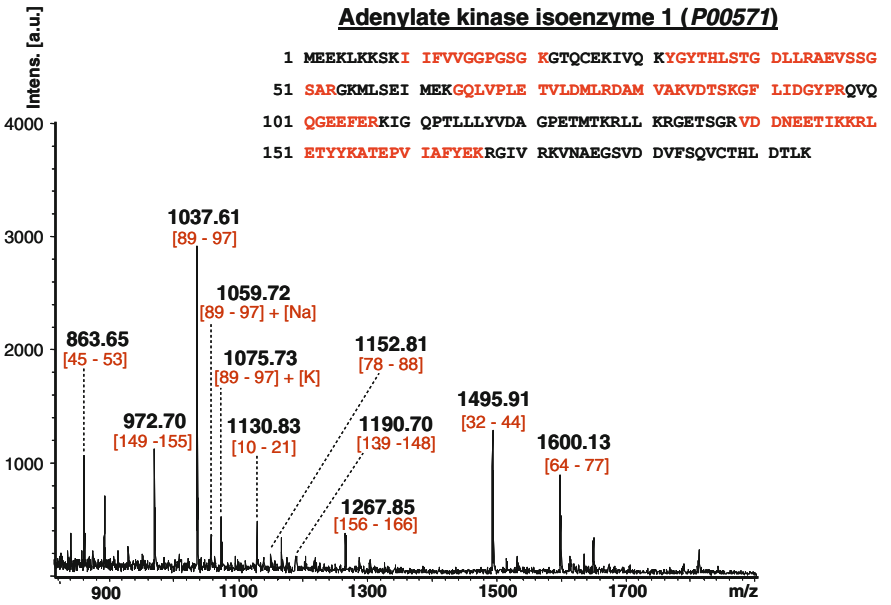


Fig. 10 MALDI-TOF-mass spectrometric identification of porcine adenylate kinase isoenzyme 1. Following gel reader visualization, spot 2 (from Fig. 8) was excised, in-gel digested with trypsin and analyzed by MALDI-MS. Labeled peaks denote the identified peptides (identified peptides are also shown in the amino acid sequence of the protein)

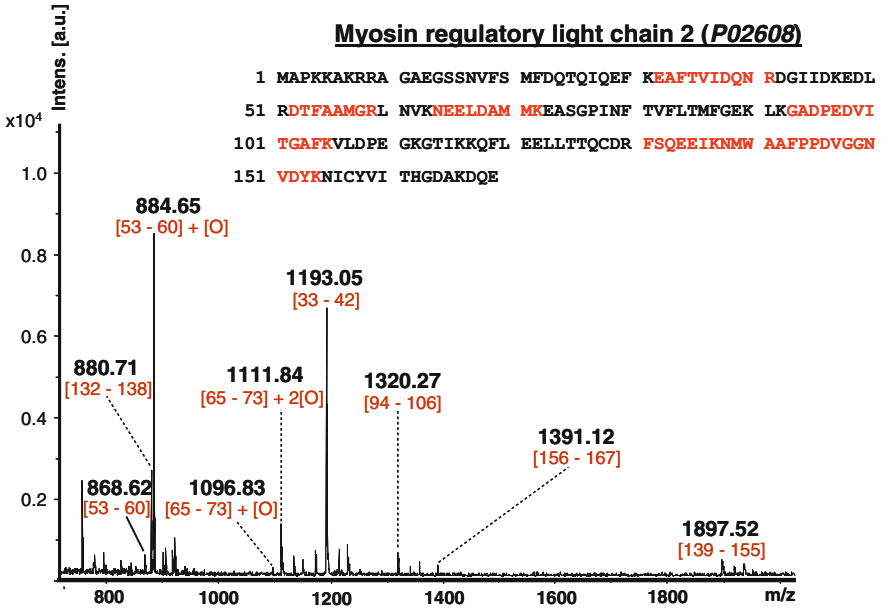


Fig. 11 Mass spectrometric identification of spot 3 (from Fig. 8) as porcine myosin regulatory light chain 2 (with labels for the identified peptides). Identified peptides are denoted in the amino acid sequence

Table 2 Protein identifications in proteome application to post-mortem porcine muscle sample. After native fluorescence visualization and localization, spots 1–6 were excised from 2D-gel (see Fig. 8), in-gel digested with trypsin, and measured by MALDI-TOF-mass spectrometry. Upon database search, proteins were successfully identified

Spot no.[a]	Protein	Score	No. of identified peptides	Sequence coverage (%)	Accession no.[b]
1	Triosephosphate isomerase	78	12	28	Q29371
2	Adenylate kinase isoenzyme 1	98	9	49	P00571
3	Myosin regulatory light chain 2	76	7	44	P02608
4	Alpha-crystallin	70	6	29	P02470
5	Skeletal alpha actin	92	13	70	P68137
6	Creatine kinase M chain	78	20	90	Q5XLD3

[a]Spot numbers correspond to the 2D-gel shown in Fig. 8
[b]Accession numbers are from SWISS-PROT or TrEMBL database

spot-picking tools [7]. Following tryptic digestion of isolated gel spots, the MALDI-MS analysis provided unequivocal identifications of several proteins, as summarized in Table 2. Figure 9 shows the identification of triosephosphate isomerase from spot 1 (see Fig. 8). Nine peptides (labeled in Fig. 10) provided unambiguous identification of adenylate kinase isoenzyme 1 (spot 2 in Fig. 8), while myosin regulatory light chain 2 was identified from spot 3 based on seven peptides (Figs. 8 and 11). From the proteins identified, alpha-actin (spot 5), creatine kinase M (spot 6), and myosin regulatory light chain 2 (spot 3) showed modifications by oxidation (Table 2, Fig. 11) [24].

4 Concluding Remarks

In this study we show stain-free detection and visualization of proteins in gels using native protein fluorescence as an efficient and sensitive approach for MALDI-mass spectrometric proteome analysis. The stain-free gel bioanalyzer enabled the detection and MALDI-MS identification of proteins from gel spots at detection limits in the low nanogram range, comparable to silver staining. Moreover, this approach does not require any post-electrophoretic manipulation by destaining, thus enabling direct MALDI-MS analysis with reduced background and time needed for sample preparation. The use of fluorescence detection with two-dimensional gel electrophoresis should be feasible for the development of automated, high-throughput technologies in proteome analysis. Thus, the stain-free fluorescence visualization approach should prove useful as both a complement and an alternative to staining techniques for mass spectrometric proteome analysis.

Acknowledgments We thank Martin Schütte and Bernd Müller-Zülow, LaVision-BioTec for technical support regarding the gel bioanalyzer. This work has been partially supported by the Deutsche Forschungsgemeinschaft, Bonn, Germany (PR-175-14/1), and the University of Konstanz (Proteostasis Research Center).

References

1. Karas M, Hillenkamp F (1988) Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. Anal Chem 60:2299–2301
2. Hillenkamp F, Karas M (1990) Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization. Methods Enzymol 193:280–295
3. Spengler B, Cotter RJ (1990) Ultraviolet laser desorption/ionization mass spectrometry of proteins above 100,000 daltons by pulsed ion extraction time-of-flight analysis. Anal Chem 62:793–796
4. Cohen LH, Gusev AI (2002) Small molecule analysis by MALDI mass spectrometry. Anal Bioanal Chem 373:571–586
5. Schleuder D, Hillenkamp F, Strupat K (1999) IR-MALDI-mass analysis of electroblotted proteins directly from the membrane: comparison of different membranes, application to on-membrane digestion, and protein identification by database searching. Anal Chem 71:3238–3247
6. 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 11:513–518
7. Susnea I, Bernevic B, Svobodova E, Simeonova DD, Wicke M, Werner C, Schink B, Przybylski M (2011) Mass spectrometric protein identification from two-dimensional gel separation with stain-free detection and visualization using native fluorescence. Int J Mass Spectrom 301:22–28
8. Aebersold R, Goodlett DR (2001) Mass spectrometry in proteomics. Chem Rev 101:269–295
9. Jungblut P, Thiede B (1997) Protein identification from 2-DE gels by MALDI mass spectrometry. Mass Spectrom Rev 16:145–162
10. Krutchinsky AN, Kalkum M, Chait BT (2001) Automatic identification of proteins with a MALDI-quadrupole ion trap mass spectrometer. Anal Chem 73:5066–5077
11. Bai Y, Galetskiy D, Damoc E, Ripper J, Woischnik M, Griese M, Liu Z, Liu S, Przybylski M (2007) Lung alveolar proteomics of bronchoalveolar lavage from a pulmonary alveolar proteinosis patient using high-resolution FTICR mass spectrometry. Anal Bioanal Chem 389:1075–1085
12. Damoc E, Youhnovski N, Crettaz D, Tissot JD, Przybylski M (2003) High resolution proteome analysis of cryoglobulins using Fourier transform-ion cyclotron resonance mass spectrometry. Proteomics 3(8):1425–1433
13. Sun JF, Shi ZX, Guo HC, Li S, Tu CC (2011) Proteomic analysis of swine serum following highly virulent classical swine fever virus infection. Virol J 8:107
14. Takagi T, Naito Y, Okada H, Okayama T, Mizushima K, Yamada S, Fukumoto K, Inoue K, Takaoka M, Oya-Ito T, Uchiyama K, Ishikawa T, Handa O, Kokura S, Yagi N, Ichikawa H, Kato Y, Osawa T, Yoshikawa T (2011) Identification of dihalogenated proteins in rat intestinal mucosa injured by indomethacin. J Clin Biochem Nutr 48:178–182
15. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20:3551–3567
16. Neuhoff V, Arold N, Taube D, Ehrhardt W (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie Brilliant Blue G-250 and R-250. Electrophoresis 9:255–262
17. Heukeshoven J, Dernick R (1988) Improved silver staining procedure for fast staining in PhastSystem Development Unit. I. Staining of sodium dodecyl sulfate gels. Electrophoresis 9:28–32
18. Nock CM, Ball MS, White IR, Skehel JM, Bill L, Karuso P (2008) Mass spectrometric compatibility of Deep Purple and SYPRO Ruby total protein stains for high-throughput proteomics using large-format two-dimensional gel electrophoresis. Rapid Commun Mass Spectrom 22:881–886

19. Lin JF, Chen QX, Tian HY, Gao X, Yu ML, Xu GJ, Zhao FK (2008) Stain efficiency and MALDI-TOF MS compatibility of seven visible staining procedures. Anal Bioanal Chem 390:1765–1773
20. Ladner CL, Yang J, Turner RJ, Edwards RA (2004) Visible fluorescent detection of proteins in polyacrylamide gels without staining. Anal Biochem 326:13–20
21. Sluszny C, Yeung ES (2004) One- and two-dimensional miniaturized electrophoresis of proteins with native fluorescence detection. Anal Chem 76:1359–1365
22. Zhao Z, Aliwarga Y, Willcox MD (2007) Intrinsic protein fluorescence interferes with detection of tear glycoproteins in SDS-polyacrylamide gels using extrinsic fluorescent dyes. J Biomol Tech 18:331–335
23. Roegener J, Lutter P, Reinhardt R, Bluggel M, Meyer HE, Anselmetti D (2003) Ultrasensitive detection of unstained proteins in acrylamide gels by native UV fluorescence. Anal Chem 75:157–159
24. Bernevic B, Petre BA, Galetskiy D, Werner C, Wicke M, Schellander K, Przybylski M (2010) Degradation and oxidation postmortem of myofibrillar proteins in porcine skeleton muscle revealed by high resolution mass spectrometric proteome analysis. Int J Mass Spectrom 305:217–227
25. Mortz E, Vorm O, Mann M, Roepstorff P (1994) Identification of proteins in polyacrylamide gels by mass spectrometric peptide mapping combined with database search. Biol Mass Spectrom 23:249–261
26. Koohmaraie M (1996) Biochemical factors regulating the toughening and tenderization processes of meat. Meat Sci 43:193–201
27. Huang J, Forsberg NE (1998) Role of calpain in skeletal-muscle protein degradation. Proc Natl Acad Sci USA 95:12100–12105
28. Doumit ME, Koohmaraie M (1999) Immunoblot analysis of calpastatin degradation: evidence for cleavage by calpain in postmortem muscle. J Anim Sci 77:1467–1473
29. Lametsch R, Roepstorff P, Bendixen E (2002) Identification of protein degradation during post-mortem storage of pig meat. J Agric Food Chem 50:5508–5512

Top Curr Chem (2013) 331: 55–78
DOI: 10.1007/128_2012_366
© Springer-Verlag Berlin Heidelberg 2012
Published online: 28 September 2012

MALDI Mass Spectrometry for Nucleic Acid Analysis

Xiang Gao, Boon-Huan Tan, Richard J. Sugrue, and Kai Tang

Abstract With the discovery of several matrices which enable the ionization of DNA and RNA, matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) has become a powerful platform for the study of nucleic acid sequence changes (e.g., mutations, single nucleotide polymorphisms (SNPs), insertion/deletion, alternative splicing, etc.), amount changes (e.g., copy number variation, gene expression, allele expression, etc.), as well as modifications (e.g., methylation of genomic DNA, post transcriptional modification of tRNAs and rRNAs). Two major strategies have been employed to characterize these changes. Primer extension reactions are designed for genotyping of known polymorphic sites and determining the levels of gene or allele expressions. Base-specific cleavage reactions are used for discovery of unknown polymorphisms and characterization of modifications. These two assays usually generate nucleic acid fragments less than 30 bases in length, which is the ideal mass range for MALDI-MS. Here we review the basic concepts of these assays, sample analysis techniques, and their applications published in recent years.

Keywords Base-specific cleavage · Comparative sequencing · MALDI-MS · Nucleic acids · Primer extension · SNP

X. Gao and K. Tang (✉)
Division of Chemical Biology and Biotechnology, School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551, Singapore
e-mail: ktang@pmail.ntu.edu.sg

B.-H. Tan
Detection and Diagnostic Laboratory, DSO National Laboratories, Singapore, Singapore

R.J. Sugrue
Division of Molecular and Cellular Biology, School of Biological Sciences, Nanyang Technological University, Singapore, Singapore

Contents

Abbreviations

3-HPA	3-Hydroxypicolinic acid
ATT	6-Aza-2-thiothymine
ddNTP	Dideoxynucleotide triphosphate
FFPE	Formalin-fixed paraffin-embedded
MALDI-MS	Matrix-assisted laser desorption/ionization mass spectrometry
MLEE	Multilocus enzyme electrophoresis
MLST	Multilocus sequence typing
m/*z*	Mass to charge ratio
SNP	Single nucleotide polymorphism
tRNA	Transfer RNA

1 Introduction

1.1 Nucleic Acid Sequences and Genes

Nucleic acids are biological molecules which are polymers of nucleotides. A nucleotide consists of a base, a pentose sugar, and a phosphate group. Nucleic acids are divided into two groups – DNA (deoxyribonucleic acids) and RNA (ribonucleic acid), according to the type of the pentose sugar that builds up the nucleotide – a D-ribose or a D-deoxyribose.

A single-stranded DNA consists of a poly-nucleotide chain, with a backbone made of pentose sugars and phosphate groups. There are four types of bases that can be attached to each pentose sugar – cytosine (C), guanine (G), adenine (A), and thymine (T). A double-stranded DNA consists of two single-stranded DNA molecules running anti-parallel via specific pairing of bases: C is paired with G and A is paired with T. The main role of DNA molecules is the long-term storage of genetic information which is encoded by the sequence of the bases along the DNA backbone.

RNA is also made up of a chain of nucleotides, but the type of bases in RNA is slightly different from that in DNA – thymine is replaced by uracil (U). Some RNA can carry genetic information in the sequence of nucleotides like DNA [1], while some other RNA can have enzymatic activity just like proteins.

The sequence of nucleotides makes one DNA or RNA molecule different from another. It carries genetic information and thus is critical for all living organisms. The unit of heredity that resides on a stretch of nucleic acid sequence is called a "gene." Genes carry the basic instructions which encode for the synthesis and assemblies of all bio-molecules, thus deciding the behavior of one cell or organism [2–4].

1.2 SNP

A number of variations could be found when DNA sequences from different individuals are compared. A variation found in more than 1% of the population is defined as a polymorphism [5]. In 1978 the discovery of "restriction fragment length polymorphisms" (RFLP) was reported [6] and the advancement of the RFLP technique opened the door for the study of polymorphisms. Homologous DNA molecules are digested by restriction enzymes and the resulting restriction fragments appear different in length after gel electrophoresis separation. This difference in the length of restriction fragments could happen as a result of a single nucleotide polymorphism (SNP), which could either introduce a new restriction enzyme recognition site or abolish an existing one [5, 7]. These SNPs are of the highest abundance in all types of variations and a lot of related databases are available via the internet [8–22]. The frequency of SNPs in the human genome is thought to be at least 1 per 1,000 base-pairs [23–27]. These SNPs are unevenly distributed across the genome, with only a small portion located in the coding region [28–30]. SNPs in coding regions can alter the structure and the function of the protein, which might result in the occurrence of a certain disease or a different response to a certain drug, while the remaining abundant SNPs in non-coding regions are also widely used as important biomarkers in genetic and genomic studies [31].

1.3 Genotype and Genotyping

Genotype is the genetic constitution of a cell, an individual, or an organism. Genotyping is the process of determining the genes or genotypes of an individual by examining its DNA sequence with a molecular biological assay. Sanger sequencing, also well-known as "dideoxy sequencing," was first devised in the 1970s by Frederick Sanger and has become the standard technique for the determination of the sequence of DNA molecules [32–35]. This sequencing method uses dideoxynucleotide triphosphates (ddNTPs) as chain terminators in an enzymatic DNA synthesis procedure. The Sanger sequencing method has been the "gold standard" in the identification of both known and unknown sequence specific nucleotide variations for many years, against which new emerging technologies need to be judged. However, as the Sanger sequencing requires radioactive or fluorescent reporters, and uses separation methods like gel or capillary electrophoresis which generates massive quantity of reads to be analyzed later, the whole process is always rather time-consuming, laborious, and costly. For this reason, researchers are constantly searching for new genotyping technologies [7, 36, 37] which can meet as many as possible the following requirements:

1. The new assay should be cost efficient.
2. The assay should be of sufficient sensitivity and reproducibility so that even tiny amounts of DNA samples can still be detected.
3. The assay should be able to detect the relative frequency of individual alleles mixed in pooled samples.
4. The assay should be able to achieve a high level of throughput and automation.
5. The analysis of the result should be simple and automated.

It is not easy to find a single perfect genotyping method which can meet all those different requirements. When a new method is being evaluated, the level of throughput and the efficiency in time and cost are normally of most importance [35].

2 Analysis of Nucleic Acid with MALDI-MS

2.1 Matrix for Nucleic Acid Analysis

The analysis of oligonucleotides using MALDI-TOF mass spectrometry (MS) was first reported in 1990 [38–40]. In the beginning, only combinations of matrix and laser wavelength that worked for protein and peptide analyses were tested. However, researchers soon noticed the problems of extensive ion fragmentation and adduction of metal cations to analyte molecules, which resulted in reduced quality of oligonucleotide mass spectra compared to those from proteins and peptides. It was obvious that nucleic acids behave differently from proteins and peptides under MALDI and thus new MALDI conditions, especially new matrices, were required for better ionization.

In 1992 Tang et al. reported the use of a mixture of 3-methylsalicylic acid and 3-hydroxy-4-methoxybenzaldehyde as matrix resulted in the detection of oligonucleotides up to a 34-mer [41], and Nordhoff et al. reported the use of succinic acid, urea, and nicotinic acid as matrices in combination with an infrared (IR) laser at wavelength 2.94 μm which led to the detection of tRNAs and 5S rRNA from *Escherichia coli* [42]. In 1994, Wu et al. reported the use of 3-hydroxypicolinic acid (3-HPA) as matrix to analyze single-stranded DNA oligomers up to 67 nucleotides in length [43]. Although other compounds, such as 2,4,6-trihydroxyacetophenone (2,4,6-THAP) [44], 2,3,4-trihydroxyacetophenone (2,3,4-THAP) [45], picolinic acid (PA) [46, 47], 3-aminopicolinic acid (3-APA) [48], 6-aza-2-thiothymine (ATT) [49], 5-methoxysalicylic acid (5-MSA) [50], quinaldic acid (QA) [51], pyrazinecarboxylic acid (PCA) [52], 3-hydroxycoumarin (3-HC) [53], and 3,4-diaminobenzophenone (DABP) [54], were found to be effective matrices later, 3-HPA turned out to be the most popular matrix for nucleic acids. Glycerol was found to be a superior matrix for IR-MALDI at 2.94 μm, and was demonstrated to be able to produce molecular ions for RNA transcripts as large as 2,180 nucleotides [55]. However, applications for this technique have been lacking so far.

Matrix additives have been proven to be important in MALDI of nucleic acids. Ammonium salts such as ammonium acetate [56], ammonium citrate, ammonium tartrate [44], and ammonium fluoride [57] were found to be useful additives which significantly suppress the adduction of metal cations. However, molar concentration of ammonium salt in the matrix solution should be limited to about 10% of that of the matrix, as excess amounts of ammonium salt, when crystallized with matrix, reduce the absorption of laser energy and quench the ionization of nucleic acids. Spermine as a cation exchanger has also been used as an additive to several matrices [50, 58], but not to 3-HPA [58]. Other compounds, such as PA, various sugars [59], and PCA [52], has been added to 3-HPA to enhance its performance. However, the mechanism of such improvement remains elusive.

2.2 Sample Preparation

Small-scale sample preparation is usually performed manually, whereas automatic nanoliter dispensing systems are required for large scale high-throughput preparation. Several manual preparation methods are available for different matrices. The most common method is the traditional dried-droplet method, which is performed by spotting a droplet of the analyte-matrix mixture on the MS sample stage and drying the spot at room temperature. This method is modified for matrices that are soluble only in aqueous solvent, such as 3-HPA. The first step is to deposit a droplet of matrix in aqueous solution onto the sample plate and allow the sample to crystallize at room temperature. The second step is to add the aqueous analyte solution onto the crystallized matrix. This partially re-dissolves the matrix crystals and re-crystallizes them together with analyte molecules. In general, the two-step dried-droplet method gives more reproducible results than the traditional one.

Another method called thin-layer matrix preparation is reported to use ATT matrix dissolved in organic solvent to form a seed layer of thin homogeneous film, and then saturated ATT solution containing analyte is added to form the second layer of thin film. The thin layer preparation method can generate highly reproducible MS result [60].

For high-throughput MALDI-MS, sample spots have to be miniaturized to improve homogeneity of crystal distribution. The surface of the MALDI sample plate (metal or silicon chip) is usually coated with highly hydrophobic material [61, 62], with precisely located and miniaturized sample pads (e.g., 200 μm × 200 μm) that are hydrophilic or less hydrophobic than their surroundings. Matrix solution can be transferred to the MALDI sample plate by high speed serial dispensing using a nano-dispensing device such as a piezoelectric pipette or solenoid dispensing valve. The aqueous matrix solution initially occupies an area larger than the sample pad. While the solvent evaporates, the droplet size shrinks, and eventually crystallization of the matrix occurs only inside the sample pad. The concentration of the matrix is optimized so that the crystallized matrix covers the entire sample pad. The aqueous analyte solution, usually from a microtiter plate, can be transferred to the MALDI sample plate preloaded with matrix by parallel dispensing using a pintool device with an array of precisely mounted pins [63]. The analyte solution partially re-dissolves the 3-HPA matrix. The droplet shrinks again toward the hydrophilic pad while solvent evaporates, and then re-crystallizes to cover the pad. This miniaturized sample spot can provide more homogeneous sample distribution and thus yield more reproducible MALDI results.

Another important factor of sample preparation is to desalt nucleic acid samples before mixing them with matrix, as sodium and potassium cations form multiple adducts with the negatively charged phosphate backbone easily and degrade the quality of mass spectra. High concentration of salt can also affect matrix crystallization and prevent successful ionization of analyte molecules. In biological assays the use of sodium or potassium salt should be avoided as much as possible, or replaced with ammonium salt as it is volatile and usually does not form adducts to MALDI ions. As sodium and potassium are ubiquitous, desalting before MALDI-MS is always recommended. This can be achieved by using chromatographic pipette tips [64], solid-phase extraction [65], or by adding cation exchange resins to the sample solution [42].

2.3 Mass Spectrometry Instrumentation

Nucleic acids are less stable compared to proteins and peptides. The N-glycosidic bond linking the base and the pentose sugar is relatively weak, especially in DNA. Therefore, in MALDI, base loss, and subsequent backbone fragmentation, could happen [66]. Usually MALDI ion source is coupled with a time-of-flight (TOF) mass spectrometer. The instability of DNA ions could cause significant amounts of post-source decay (PSD) and result in reduced sensitivity and resolution if a

reflectron TOF is used. Therefore, a linear TOF is commonly used to minimize this effect. However, for small oligonucleotides, such as those generated by base-specific cleavage, PSD does not seem to be extensive and reflectron can be used to improve mass resolution [60, 67]. It is also expected that high performance can be achieved in a quadrupole-TOF if collisional cooling is employed. Of course the quadrupole must be designed to transmit high *m*/*z* ions generated by MALDI. When collisional cooling is effectively used, MALDI-generated DNA ions can be stable for a significantly long period (milliseconds), which allows for high resolution detection in Fourier transform ion cyclotron resonance (FTICR)-MS [68], or in the newly developed Orbitrap with a MALDI ion source.

3 Primer Extension Assays

3.1 SNP Genotyping

To genotype known polymorphic sites, the primer extension assay combined with MALDI-MS analysis is widely agreed to be a powerful platform. The assay is usually performed in three major steps. The first step is PCR amplification of the region which contains the SNP of interest. Second, phosphatase is added to inactivate the remaining dNTPs. Third, a primer extension reaction is performed using a specific genotyping primer positioned with its 3′-end adjacent to the polymorphic site. This reaction generates allele-specific primer extension products with different masses which can be interrogated by MALDI-MS for genotype determination. There are two different versions of primer extension assay depending on the type of dNTP and ddNTP used in the primer extension reaction:

1. Primer extension with variable number of nucleotides (Fig. 1a):
   In this primer extension reaction, a mixture of dNTP and ddNTP is used. The content of the mixture is selected so that the primer extension stops at the polymorphic site for one allele and at another site downstream for the other allele. Therefore the SNP resulted in allele-specific primer extension products with different lengths – one extended by a single base, the other by two or more bases. The large mass difference between the products makes identification by MALDI-MS a trivial task, even for heterozygous samples. Braun et al. first used this strategy to detect mutations in the CFTR gene [69]. It was later dubbed as homogenous MassEXTEND (hME) by Sequenom for multiplexed genotyping of different gene loci. The main drawback of this strategy is that in multiplexing, where multiple primers are used in a single reaction to target many different SNPs, careful planning is required to avoid peak overlapping. Even so, the level of multiplexing is moderate (up to 12-plex), and only those using the same type of dNTP and ddNTPs can be multiplexed.

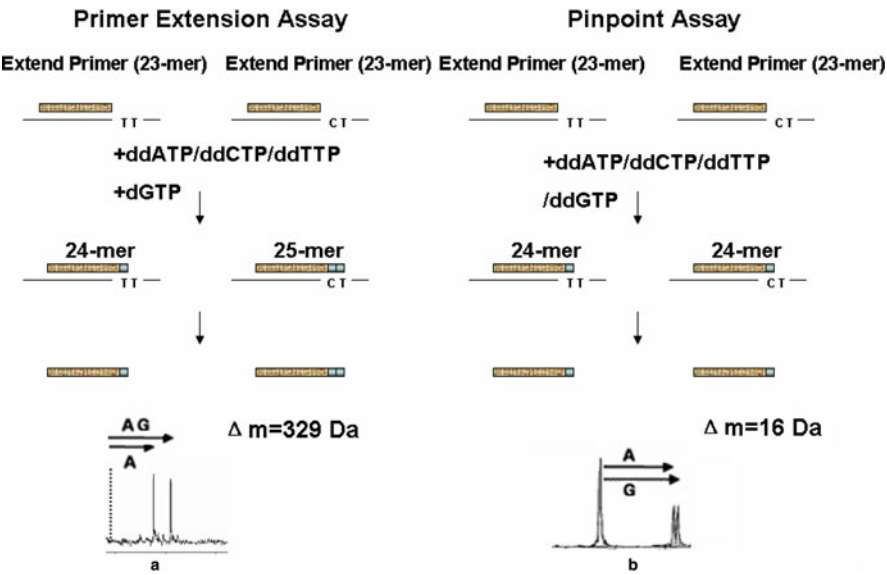


Fig. 1 Comparison of the primer extension assays. (**a**) In the primer extension assay with variable number of nucleotides, at least one dNTP is used. The extension products differ by at least one base, a larger mass shift which can be easily resolved. (**b**) In single-base primer extension assay, only ddNTPs are used and the primers will always extend by one base no matter what the sequence is. The natural mass differences of the extended bases are used for SNP genotyping

2. Single-base primer extension (Fig. 1b):
   This type of primer extension uses only ddNTPs. Therefore, the genotyping primer is only extended by a single base onto the polymorphic site. The genotype is differentiated by the mass of the ddNTP incorporated [70]. This strategy is relatively simple as a common ddNTP mixture is used for all different types of SNPs. When single-base extension is combined with different primers with various 5′-overhangs (addition of non-templated bases), a high level of multiplexing can be achieved. Ross et al. first demonstrated this potential in 1998 [71]. However, as most primer extension reactions are measured in linear TOF mass spectrometers to minimize the effect of PSD of MALDI ions, mass resolution is usually low and it is not always possible to resolve an A/T polymorphism where the mass difference between the two alleles differs by only 9 Da. This limitation can be overcome by using a modified ddATP or ddTTP. This modified ddATP or ddTTP must have mass value larger than ddGTP by at least 16 Da, and can be incorporated efficiently by DNA polymerase. Recently, Sequenom used a similar strategy in iPLEX genotyping design although detailed modification has not been published [72].

The MS-based primer extension approach was widely applied in a variety of fields that require the study of SNPs [73, 74]. Jaremko et al. reported successful SNP genotyping of DNA samples isolated from formalin-fixed paraffin-embedded tissues (FFPET) [75]. FFPE materials are usually not suitable for molecular disease

approaches, as they tend to yield degraded DNA, but they perfectly meet the need for long-term clinical research. With MALDI-TOF MS and Taqman assisted SNP genotyping approaches, the FFPET can be genotyped in a relatively high throughput format. Horn et al. further proved the reliability of the approach [76]. Five DNA samples were extracted from FFPE tumor specimens using both phenol–chloroform extraction and a commercial extraction kit. Thirty one SNPs from 25 genes were analyzed using the multiplex primer extension assay and 64 FFPE tumor specimens were compared with matched germline DNA samples, and the genotyping result appeared to be reliable.

The MALDI-TOF MS-based approach is also used for oncogenic human papillomavirus (HPV) genotyping, providing an economical and high-throughput platform for the cervical screening programs and HPV vaccination programs. Soderlund-Strand et al. applied the primer extension assay to the detection of 14 oncogenic HPV genotypes [77]. The MS-based genotyping result was compared with the reverse dot blot hybridization (RDBH) result for 532 cervical cell samples. After comparison it is found that the MS-based primer extension assay is sensitive enough to cover not only all samples appeared positive under RDBH, but five more cases of cervical disease that RDBH had missed as well. All 20 discrepancies were finally proved to be positive by MS-based primer extension assay.

MALDI-TOF MS-based primer extension assay also had good performance in the analysis of mitochondrial DNA coding region SNP (mtSNP) variation. Cerezo et al. compared MS-base primer extension assay with other techniques for the mtSNP genotyping and found that to genotype large amount of SNPs in large collections of samples, MS-based techniques appeared to be more suitable [78]. Thompson et al. also employed the MS-based primer extension assay for SNP genotyping of 635 trees from 15 populations and the result was used as indirect evidence of realized F1 fertility for *populus* hybrids [79].

Thomas et al. employed the primer extension method for high-throughput genotyping to analyze oncogenic mutation in human cancer [80] and this method is also useful in other mutation profiling [81, 82].

Huebner et al. optimized the MS-based SNP genotyping method to prevent the genotyping errors caused by tri-allelic SNPs [83]. Sasayama's group reported an alternative MS-based genotyping approach using Lu(III) ions as molecular scissors [84]. Compared with primer extension assay, their approach is independent of the sequence investigated and thus is able to genotype not only SNPs but indels as well.

3.2 Genome-Wide Association Study

For large-scale population SNP analysis, DNA samples are commonly pooled as a mixture for cost considerations, and the allele frequency of a certain SNP is usually calculated from the ratio of the alleles. The SNP frequency analysis can be used in the genome-wide association study, aiming to identify genes involved in certain diseases. Association is a statistical statement about the co-occurrence of alleles and/or

phenotypes. The genome-wide association studies usually examine most, if not all, of the genes from different individuals of one particular population. Allele frequencies of all genes from the patients group (with a certain disease) are compared with the control (without the disease) group to identify the association between SNPs in different alleles and the disease. A large number of SNPs, most of which are in non-coding regions, are found to be associated with certain diseases in this way [85, 86].

Kammerer et al. performed a large-scale association study in order to identify genes that influence the risk of developing breast and prostate cancer [87]. More than 25,000 SNPs from 16,000 genes were screened and the *ICAM* gene region was finally found to have the most significant association with breast and prostate cancer risk.

Suttner et al. reported an association study between *TBX21* gene and asthma [88]. In their work, the whole *TBX21* gene was resequenced and screened for polymorphisms. Forty three polymorphisms were identified using MALDI-TOF MS and three of them were found to be associated strongly with asthma. These polymorphic alleles are combined with SNPs within the *HLX1* genes, which had been shown to have association with asthma in previous study. The combination resulted in the increased risk to develop asthma.

Abel and coworkers analyzed pooled DNA with genome-wide SNPs using the primer extension assay and proved its usefulness for discovering candidate susceptibility genes for common diseases [89]. It is unlikely to get the unambiguous identification of the related gene or genetic variations for those common diseases using this genetic analysis only, but it is possible to locate the region of most potential for further experiments to identify the genes and variations responsible for the disease susceptibility. Osteoarthritis, for example, is a common disease with which neither traditional family-based linkage studies nor candidate gene studies can identify genes of significant association. Genome-wide association approach is performed and only a small amount of SNPs were selected for genotyping from the thousands of SNPs that exist. Finally, an intronic SNP in a gene, which is on chromosome 13 and encodes an unknown new protein with a calponin homology domain, was considered as the marker of the strongest association with the risk of knee osteoarthritis [90].

Similar methodologies are also adopted for pathogenic organisms research. Volkman et al. genotyped about 47,000 SNPs across the genome of the malaria parasite *Plasmodium falciparum* [91]. Their study generated a large data set which is of great value for genome-wide diversity analysis. Herring et al. performed comparative genome sequencing to *E. coli* for bacterial evolution study and the MS-based primer extension assay was applied for the mutation validation [92].

3.3 Gene Expression Analysis

The primer extension assay works well not only in SNP genotyping research but also in quantitative gene expression studies. For quantitative gene expression analysis, primer extension assay is combined with competitive real-time PCR,

followed by MALDI-MS analysis. Extracted RNA is used for reverse transcription and a synthetic DNA of about 90 nucleotides in length, referred as the "competitor," is spiked into the cDNA product. This competitor differs from the cDNA by only one base in the center region and is used as an internal standard. The competitor internal standard serves as an artificial allele so that the method used for allele frequency analysis can be used here for gene expression analysis. The competitor and cDNA is co-amplified by PCR, after which shrimp alkaline phosphatase (SAP) is added to inactivate the unused nucleotides. Primer extension assay is then performed using the PCR products from the competitor and cDNA as templates. The primer extension products are submitted to MALDI-TOF analysis and the related mass signals are resolved to calculate the peak area ratios [93].

Rode et al. compared this MS-based primer extension assay with the microarray assay, which is a commonly used method for gene expression analysis, and tested the results from both MassARRAY and microarray against quantitative real-time PCR (qRT-PCR), which is widely accepted as a standard method for single gene expression quantification [94]. Their study showed a high rate of concordance among the MS-base primer extension assay, microarray, and qRT-PCR in most of the data. The MS-based primer extension assay has consistently proved suitable for prokaryotes gene expression research and became a better choice than microarray when analyzing a limited set of genes instead of investigating global gene expression study. The MS-based primer extension assay is also advantageous over microarray in the low requirement of sample amount.

Duffield et al. combined the MS-based primer extension and the molecular affinity isolation for multiple mRNA levels determination [95]. The mRNA levels of four genes with significant variation were determined simultaneously in their assay. Biotinylated ddNTPs were used in the single base extension which not only enlarged the mass difference between the extension products of the samples and the competitors but also facilitated a purification step before MS analysis. Unextended primers, which had the potential to overlap with other extended products, were all removed, releasing plenty of space in the spectrum to enable a higher level of multiplexing.

3.4 Non-invasive Prenatal Diagnosis

The primer extension assay combined with the MS technique is of great potential in risk-free, non-invasive prenatal diagnosis. Ding et al. introduced the use of single allele base extension reaction (SABER) and MS for prenatal diagnosis in 2004 [96]. Circulating fetal DNA can be detected from the maternal plasma. However, if the standard primer extension assay is performed, the maternal allele can always form an overwhelming background. The SABER protocol is able to extend the paternal-specific fetal allele, avoiding interference caused by the background of maternal allele.

Daelemans et al. analyzed candidate imprinted genes in human placenta and the allele-specific gene expression [97]. Sequenom and illumina assays, two typical allele-specific high-throughput technologies, are employed and both of them are of sufficient sensitivity for imprinting and allelic bias studies.

Lo et al. applied the MS-based primer extension assay to analyze the ratio between alleles of an SNP in *PLAC4* mRNA in order to perform non-invasive diagnosis of fetal trisomy 21 [98]. The same technique can also be used in non-invasive prenatal diagnosis of trisomy 13 and 18. Further development of noninvasive prenatal diagnosis depending on MS-based primer extension technology for gene expression analysis has been reviewed elsewhere [99].

4 Base-Specific Cleavage Assays

4.1 SNP Discovery

In the past two decades a number of different methods have been developed to achieve base-specific cleavage of both large DNA and RNA products using an enzymatic approach as well as chemical means [100–104] in order to make it possible for MALDI-TOF MS measurement. Within all these methods, the most prominent approach introduces an intermediate RNA transcription step after PCR amplification of the target sequence region. Compared with direct cleavage of DNA product, this assay has several advantages. First of all, the cleavage of double-stranded DNA will certainly result in an overlapping cleavage patterns from each of the forward and reverse strands. After the in vitro transcription step, the single-stranded RNA product is ready for cleavage, avoiding the elaborate strands separation steps. Second, the transcription step could further amplify the analyte, thus making it possible to add ion exchange treatment and dilution, which are important conditioning steps before MALDI-TOF analysis. Third, RNA is more stable than DNA under standard UV-MALDI condition as RNA is less prone to depurination.

The concept of base-specific cleavage has also been combined with MALDI-TOF MS [103, 105] to discover new SNPs and detect mutations. In this process there are two major enzymatic approaches to achieve base-specific cleavage of RNA transcripts. The first uses base-specific RNase; for example, the use of RNase T1 could yield G-specific cleavage in normal working condition of the enzyme. This G-specific cleavage reaction could also be used in the reverse strand of the target region. The second method uses a mutant RNA polymerase – T7 or SP6 R&DNA polymerase, which can incorporate some of the dNTPs in transcription, combined with pyrimidine-specific RNase, for example RNase A. When mixed with normal transcript, RNase A would cleave after every rC and rU residue, resulting in more frequent fragmentation of RNA and loss of sequence information. However, as the mutant transcriptase could incorporate certain kinds of dNTPs

(dCTP, dTTP, dUTP), base-specific cleavage can be achieved when either rC or rU is fully replaced by dC or dT/dU, respectively. In one reaction, dCTP is used instead of rCTP during transcription, which results in the U-specific cleavage of the transcript by RNase A, and in the other reaction, rUTP is fully replaced by dTTP or dUTP, and the transcript will finally be C-specifically cleaved. Considering that T7 promoter tag can be added to either the forward primer or the reverse primer in the PCR amplification step, one dsDNA sequence can actually be cleaved in both forward and reverse strands in four separate reactions. The cleavage products will be conditioned for mass spectrometric analysis by addition of ion exchange resin and the supernatant will be transferred onto a silicon chip containing a 16×24 array of matrix spots, which can be automatically scanned in a MALDI mass spectrometer.

The combination of base-specific cleavage and MALDI-TOF MS offers extremely high sample throughput for large-scale comparative sequence analysis. It takes only a few seconds to measure each sample spot. If there is a sequence change in the target region, no matter what it is, up to ten "observations" can be found in the mass spectra from the four base-specific cleavage reactions, two reactions will have mass shift, one reaction will introduce a cleavage site, and the last one will lose a cleavage site. If the sequence change does not alter a cleavage site, a mass peak shift will appear in one spectrum. The mass shift is actually two observations: the disappearance of a theoretical peak and the appearance of a new peak. If the sequence change alters a cleavage site, three changes will appear in one spectrum: if a new cleavage site is introduced by the sequence change, a peak with high mass will be replaced by two peaks with lower masses; if a cleavage site is abolished by the sequence change, two peaks with lower masses will disappear and a peak with higher mass will appear. For example, there is a sequence change of A to G in the forward direction transcript; it is still an A to G change and in both rC and rU specific cleavage reaction there is only a mass shift of 16 Da. While in the reverse direction transcript, it becomes a U to C change. In the rU specific cleavage reaction, one cleave site will disappear and in the rC specific cleavage reaction, a new cleavage site will be introduced. This comparative sequencing with multiple "observations" shows great advantage for diagnostic applications (Fig. 2).

Similar to the identification of an unknown protein by the peptide mass fingerprinting (PMF) approach [106–108], theoretical RNA sequences of a certain size are in silico digested with base specific enzyme like RNase A or RNase T1, and the molecular weight patterns for the predicted cleavage products are collected in the database. The mass spectra for the cleavage products from an unknown RNA sequence can be searched against the existing whole prokaryotic genomes or RNA FASTA sequence databases to find the best match or even identification using certain software such as "RNA mass mapping" (RMM) developed by Matthiesen and Kirpekar [109].

For SNP screening from a known sequence, a software package called "RNaseCut" was developed by Krebs et al. [104]. It can be used to design primers for those regions of high SNP discovery score and validation of individual SNPs.

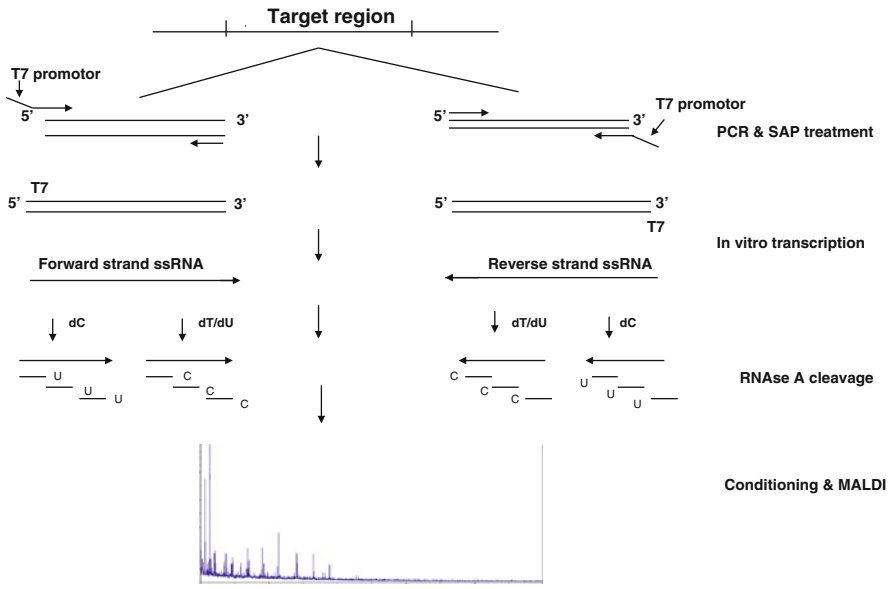


Fig. 2 Schematic diagram of the base-specific cleavage procedures using mutant transcriptase and RNase A. T7 promoter tag is added to the 5′ end of either forward primer or reverse primer to amplify the target region. After PCR, SAP is added in order to degrade the unused dNTP. After SAP treatment, the PCR product is directly used for in vitro transcription to produce single strand transcript using a mutant T7 transcriptase. The transcription is performed in two groups, one of which uses dCTP instead of rCTP, while the other uses dUTP or dTTP instead of rUTP. As a result, the transcript is made up of three rNTPs and one dNTP. The transcript is then digested with RNase A, which cleaves after each rC and rU, and the cleavage will be driven to completion. Introduction of either dC or dT/dU during the transcription will render the RNase A cleavage base-specific, as these dNTPs are insensitive under RNase treatment. With the combination of the transcripts from forward and reverse direction and either rC or rU specific cleavage, one sequence could be cleaved in four independent ways. It must be noted that, due to charge competition in MALDI, some fragments may not appear in the spectrum for a single reaction. Sequence coverage from one cleavage reaction is usually not high enough to identify certain mutations. Two base-specific cleavage reactions from the same transcript (e.g., forward transcript) or all four reactions from both forward and reverse transcripts are necessary to achieve high sequence coverage

When a target sequence is given, the software first determines the best position for PCR amplification, in which more SNPs are theoretically able to be detected and validated. Then the transcribed RNA product is in silico digested to generate reference mass peak patterns. When the experimental observed spectra have differences compared to the reference, the software will calculate the possible solutions for an SNP which can cause the observed variations [110]. A more sophisticated algorithm for SNP discovery has also been developed [102] but the software itself remains proprietary.

4.2 Identification of tRNA

It is difficult to get high resolution separation of transfer RNAs (tRNAs) using traditional methods because of the high level of similarity in their secondary and tertiary structure. Using base-specific cleavage and MALDI-MS assay, tRNAs are digested with RNase T1 and the generated fragments are found to contain unique masses for each individual tRNA. In this way, molecular mass peak patterns can be used as fingerprints for detection and identification of individual tRNAs from a total tRNA pool. This protocol does not require a separation or purification step for the tRNAs sample and is thus of relatively high efficiency [111].

4.3 RNA Modifications Analysis

MS nowadays serves as a powerful technology for the analysis of posttranscriptional modifications of different types of RNA [112–114], especially for those modifications that are not detectable by the RT-primer extension assay and the chromatographic assay (e.g., m^5U, m^5C, m^6A, etc.) [115]. Kirpekar et al. was the first to apply the base-specific cleavage and MALDI-MS assay to screen the 5S rRNA, in order to identify posttranscriptional modifications [116]. The whole approach was of high speed and sensitivity while the procedure was relatively simple with only one purification step using ion-exchange resin. Suppression of the mass signal for some fragment might be observed but the total sequence coverage remained satisfactory. However, this approach lacks the ability to identify mass-neutral modifications. For example, the replacement of uridine by its structural isomer, pseudouridine will not lead to any mass difference and thus is not detectable by MS. Several methods were reported in order to identify the pseudouridine, as its presence accounted for a large part of the RNA modification in Eukarya. Mengel-Jorgensen and Kirpekar generated a 53-Da mass increase to pseudouridine through cyanoethylation [117]. Durairaj et al. utilized *N*-cyclohexyl-*N*′-β-(4-methylmorpholinium) ethylcarbodiimide (CMC) derivatization and alkaline treatment which introduced a 252-Da mass shift to the pseudouridine and also make it detectable [118].

Auxilien et al. used MALDI-MS to identify the target region of m^5U methyltransferases by pinpointing the base-specific cleaved methylated RNA transcript. Target specificities of several m^5U methyltransferase were analyzed and some shifts of target specificity were observed in archaeal RNA m^5U methyltransferase [119]. Using MALDI-MS, Goll et al. found that the target of human DNA methyltransferase-2 (DNMT2), which should be cytosine in the anticodon loop, was actually an RNA instead of DNA [120]. This RNA was identified to be aspartic acid transfer RNA ($tRNA^{ASP}$). On the basis of this observation, they suggested that eukaryotic DNA cytosine methyltransferase might be derived from an RNA methyltransferase instead of a prokaryotic DNA methyltransferase.

4.4 DNA Methylation Analysis

Base-specific nucleic acid cleavage combined with MALDI-MS techniques has also been used for DNA methylation analysis. DNA methylation occurs by adding a methyl group to the number 5 carbon of the cytosine pyrimidine ring or the number 6 nitrogen of the adenine purine ring. DNA methylation in somatic tissues is typically observed at the cytosine base in a CpG dinucleotide context [121, 122], while non-CpG methylation occurs mostly in stem cells [122–124].

The target DNA is first treated with bisulfite, which can mediate a deamination of unmethylated cytosines to uracils, while the methylated cytosines stay unchanged. After the bisulfite treatment, the target DNA is PCR amplified and used for in vitro transcription. The generated single-stranded RNA product will then undergo RNase A mediated base specific cleavage procedure by replacing either rCTP or rUTP with their dNTP counterparts. The CpG methylation could be detected in both the forward and the reverse strand. For forward strand as an example, in the C-specific cleavage reaction, methylated cytosines remain unchanged and will be cleaved by the RNase while the unmethylated cytosines will finally be transcribed using dUTPs and these cleavage sites are abolished. For the reverse strand, the methylation site will appear as G/A variation and the fragments containing these sites will have a mass shift of 16 Da. By analyzing these mass shifts, methylation could easily be detected, and by comparing with the signal intensity of mass peaks from fully methylated and unmethylated templates, quantitative analysis of methylation is possible [121, 125–128].

4.5 MLST by Base-Specific Cleavage for Bacterial Studies

At the end of the last century, in order to discriminate the strains of infectious agents which caused disease, there were two different approaches. The first approach identified individual loci or the uncharacterized genomic regions which were variable among bacterial populations. Several methods were developed based on this approach for the determination of the relatedness of pathogens. These included pulsed-field gel electrophoresis (PFGE), ribotyping, and PCR-based fingerprinting. PFGE was widely considered as “gold standard” even though this technique was unable to type many strains as a result of DNA degradation during the electrophoresis process. The second approach was based on the characterization of variations in the sequence of many loci, which were slowly accumulating in the microbial population. Multilocus enzyme electrophoresis (MLEE) was developed based on this approach. Because it can analyze several loci at the same time, MLEE was able to achieve high levels of discrimination. However, like many other typing methods, the result of the MLEE was not portable between different laboratories and thus presented comparison difficulties [129].

Multilocus sequence typing (MLST) was a comparative sequence typing approach on the principle of MLEE for the typing of multiple loci. It was first introduced by Martin C. Maiden and the coworkers in 1998. In their research, MLST was used to characterize the population structure of bacterial isolates with the analysis of DNA sequences of internal fragments of multiple housekeeping genes [129–131].

The proposal of MLST has improved knowledge of bacterial evolution and population biology [132–135] while at the same time making the high-throughput determination of nucleotide sequence much more available and cost-efficient [136, 137]. The most significant advantage of MLST over other molecular typing approaches mentioned above was that it could be applied to almost all bacterial species, even those that are very difficult to cultivate, and the sequence data were highly unambiguous and portable between laboratories – all the materials for MLST could be exchanged between labs and the primer sequences and the protocol could be shared via internet, thereby, for the first time, making it possible to establish a global database for any individual species on the website. The MLST database covers 27 species today, and will continue to expand in the future.

As the typical sequence-based typing methodology [138], MLST could be combined with base-specific cleavage assay and MALDI-TOF MS for microbial typing [139]. This nucleic acid-based typing approach did not need any prior target information, which was critical for traditional microbial typing methods that were derived from in situ hybridization [140]. The 16S small ribosomal RNA gene, which was commonly used for bacterial identification [141], was PCR amplified for base-specific-MS analysis. The data acquired were searched against the database containing the in silico digest result of existing 16S rRNA sequences. The whole process took only a few hours and was compatible with both bacterial isolates and uncultured clones [142–146].

4.6 Pathogenic Viruses Study

The base-specific cleavage and MALDI-MS approach has already been used for hepatitis B virus (HBV) genotyping [147]. Ganova-Raeva's group compared the MS-based approach with traditional sequencing for HBV genotyping and proved that the new approach is much more cost-effective but of equal reliability. For large scale analysis which requires high throughput, the cost for base-specific cleavage and MALDI-MS assay can be much lower than that for sequencing.

This new approach also has great potential in the rapid identification and mutation surveillance of other pathogenic viruses like influenza. Commonly used methods for rapid diagnosis of influenza are usually based on antigen-specific antibody probes [148], or real-time reverse transcription PCR (RT-PCR) of the M gene to identify viral species [149], and then use the second round of RT-PCR assay for HA and NA genes to determine the subtypes [150, 151]. Compared with these methods, the approach based on base specific cleavage and MALDI-MS not only

has advantages in the throughput and speed, but also in the ability for the identification of new emerging genetic variants with mutations in the drug-resistant mutation site [152, 153].

5 Conclusion and Perspectives

The MALDI-TOF MS-based technique has provided a platform to analyze nucleic acids with high degrees of throughput, sensitivity, accuracy, and automation. The primer extension assay is highly flexible and can be employed for multiple purposes. It can be applied to genotype individual SNP for allele determination, or to analyze the allele frequency in population studies. The MS-based primer extension strategy has good performance in quantitative gene expression analysis and is a good substitute of traditional microarray techniques in cases where only a small set of genes of interest are analyzed. The modified primer extension assay has been applied to identify several fetal diseases and will play an increasingly important role in non-invasive prenatal diagnosis.

The comparative sequence analysis using base-specific cleavage and MALDI-MS is nowadays more widely used for routine SNP genotyping, mutation surveillance, and pathogen identification. It acts as a relatively low cost tool for a single analyte and at the same time a powerful platform for high-throughput analysis for large scale samples from multiple populations, thus being suitable for epidemiologic surveillance. The most striking advantage over other genotyping schemes is the collateral security. Once there is a sequence change, it can results in as much as ten "observations" in the mass spectra for the cleavage products. These observations could be a mass peak shift, appearance of new mass peaks, or disappearance of a theoretical mass peak. The identification of the sequence change is always based on the combination of multiple observations from several spectra, which is extremely important for diagnostic applications.

Acknowledgements KT is supported by Ministry of Health of Singapore, BHT by the Ministry of Defense, and RJS by Ministry of Health of Singapore, National Research Foundation, and Defense Science & Technology Agency.

References

1. Rassoulzadegan M, Grandjean V, Gounon P, Vincent S, Gillot I, Cuzin F (2006) Nature 441:469
2. Pennisi E (2007) Science 316:1556
3. Pearson H (2006) Nature 441:398
4. Noble D (2008) Philos Trans A Math Phys Eng Sci 366:3001
5. Brookes AJ (1999) Gene 234:177
6. Waikan Y, Dozy AM (1978) Proc Natl Acad Sci USA 75:5631

7. Syvanen AC (2001) Nat Rev Genet 2:930
8. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K (2001) Nucleic Acids Res 29:308
9. Thorisson GA, Smith AV, Krishnan L, Stein LD (2005) Genome Res 15:1592
10. Fredman D, Siegfried M, Yuan YP, Bork P, Lehvaslaiho H, Brookes AJ (2002) Nucleic Acids Res 30:387
11. Hinds DA, Stuve LL, Nilsen GB, Halperin E, Eskin E, Ballinger DG, Frazer KA, Cox DR (2005) Science 307:1072
12. Hirakawa M, Tanaka T, Hashimoto Y, Kuroda M, Takagi T, Nakamura Y (2002) Nucleic Acids Res 30:158
13. Dooley EE (2003) Environ Health Perspect 111:A575
14. Hernandez-Boussard T, Whirl-Carrillo M, Hebert JM, Gong L, Owen R, Gong M, Gor W, Liu F, Truong C, Whaley R, Woon M, Zhou T, Altman RB, Klein TE (2008) Nucleic Acids Res 36:D913
15. Hewett M, Oliver DE, Rubin DL, Easton KL, Stuart JM, Altman RB, Klein TE (2002) Nucleic Acids Res 30:163
16. Packer BR, Yeager M, Staats B, Welch R, Crenshaw A, Kiley M, Eckert A, Beerman M, Miller E, Bergen A, Rothman N, Strausberg R, Chanock SJ (2004) Nucleic Acids Res 32: D528
17. Livingston RJ, von Niederhausern A, Jegga AG, Crawford DC, Carlson CS, Rieder MJ, Gowrisankar S, Aronow BJ, Weiss RB, Nickerson DA (2004) Genome Res 14:1821
18. Ingelman-Sundberg M, Oscarson M, Daly AK, Garte S, Nebert DW (2001) Cancer Epidemiol Biomarkers Prev 10:1307
19. Hollegaard MV, Bidwell JL (2006) Genes Immun 7:269
20. Haukim N, Bidwell JL, Smith AJP, Keen LJ, Gallagher G, Kimberly R, Huizinga T, McDermott MF, Oksenberg J, McNicholl J, Pociot F, Hardt C, D'Alfonso S (2002) Genes Immun 3:313
21. Bidwell J, Keen L, Gallagher G, Kimberly R, Huizinga T, McDermott MF, Oksenberg J, McNicholl J, Pociot F, Hardt C, D'Alfonso S (2001) Genes Immun 2:61
22. Bidwell J, Keen L, Gallagher G, Kimberly R, Huizinga T, McDermott MF, Oksenberg J, McNicholl J, Pociot F, Hardt C, D'Alfonso S (1999) Genes Immun 1:3
23. Schuler GD, Boguski MS, Stewart EA, Stein LD, Gyapay G, Rice K, White RE, RodriguezTome P, Aggarwal A, Bajorek E, Bentolila S, Birren BB, Butler A, Castle AB, Chiannilkulchai N, Chu A, Clee C, Cowles S, Day PJR, Dibling T, Drouot N, Dunham I, Duprat S, East C, Edwards C, Fan JB, Fang N, Fizames C, Garrett C, Green L, Hadley D, Harris M, Harrison P, Brady S, Hicks A, Holloway E, Hui L, Hussain S, LouisDitSully C, Ma J, MacGilvery A, Mader C, Maratukulam A, Matise TC, McKusick KB, Morissette J, Mungall A, Muselet D, Nusbaum HC, Page DC, Peck A, Perkins S, Piercy M, Qin F, Quackenbush J, Ranby S, Reif T, Rozen S, Sanders C, She X, Silva J, Slonim DK, Soderlund C, Sun WL, Tabar P, Thangarajah T, VegaCzamy N, Vollrath D, Voyticky S, Wilmer T, Wu X, Adams MD, Auffray C, Walter NAR, Brandon R, Dehejia A, Goodfellow PN, Houlgatte R, Hudson JR, Ide SE, Iorio KR, Lee WY, Seki N, Nagase T, Ishikawa K, Nomura N, Phillips C, Polymeropoulos MH, Sandusky M, Schmitt K, Berry R, Swanson K, Torres R, Venter JC, Sikela JM, Beckmann JS, Weissenbach J, Myers RM, Cox DR, James MR et al (1996) Science 274:540
24. Collins FS, Lander ES, Rogers J, Waterston RH, Conso IHGS (2004) Nature 431:931
25. Sachidanandam R, Weissman D, Schmidt SC, Kakol JM, Stein LD, Marth G, Sherry S, Mullikin JC, Mortimore BJ, Willey DL, Hunt SE, Cole CG, Coggill PC, Rice CM, Ning ZM, Rogers J, Bentley DR, Kwok PY, Mardis ER, Yeh RT, Schultz B, Cook L, Davenport R, Dante M, Fulton L, Hillier L, Waterston RH, McPherson JD, Gilman B, Schaffner S, Van Etten WJ, Reich D, Higgins J, Daly MJ, Blumenstiel B, Baldwin J, Stange-Thomann NS, Zody MC, Linton L, Lander ES, Altshuler D, Grp ISMW (2001) Nature 409:928

26. Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen NP, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998) Science 280:1077
27. Schmid KJ, Sorensen TR, Stracke R, Torjek O, Altmann T, Mitchell-Olds T, Weisshaar B (2003) Genome Res 13:1250
28. Li WH, Sadler LA (1991) Genetics 129:513
29. Nickerson DA, Taylor SL, Weiss KM, Clark AG, Hutchinson RG, Stengard J, Salomaa V, Vartiainen E, Boerwinkle E, Sing CF (1998) Nat Genet 19:233
30. Holden AL (2002) Forensic DNA Typing: Biology, Technology, and Genetics of Str Markers. Biotechniques, Suppl 22–26
31. Kim S, Misra A (2007) Annu Rev Biomed Eng 9:289
32. Sanger F, Nicklen S, Coulson AR (1977) Proc Natl Acad Sci USA 74:5463
33. Dianzani I, Camaschella C, Ponzone A, Cotton RG (1993) Trends Genet 9:403
34. Cotton RG (1997) Trends Genet 13:43
35. Oberacher H (2008) Anal Bioanal Chem 391:135
36. Kristensen VN, Kelefiotis D, Kristensen T, Borresen-Dale AL (2001) Biotechniques 30:318
37. Sobrino B, Carracedo A (2005) Methods Mol Biol 297:107
38. Spengler B, Pan Y, Cotter RJ, Kan LS (1990) Rapid Commun Mass Spectrom 4:99
39. Bornsen KO, Schar M, Widmer HM (1990) Chimia 44:412
40. Karas M, Bahr U (1990) Trac-Trends Anal Chem 9:321
41. Tang K, Allman SL, Chen CH (1992) Rapid Commun Mass Spectrom 6:365
42. Nordhoff E, Ingendoh A, Cramer R, Overberg A, Stahl B, Karas M, Hillenkamp F, Crain PF (1992) Rapid Commun Mass Spectrom 6:771
43. Wu KJ, Steding A, Becker CH (1993) Rapid Commun Mass Spectrom 7:142
44. Pieles U, Zurcher W, Schar M, Moser HE (1993) Nucleic Acids Res 21:3191
45. Zhu YF, Chung CN, Taranenko NI, Allman SL, Martin SA, Haff L, Chen CH (1996) Rapid Commun Mass Spectrom 10:383
46. Tang K, Taranenko NI, Allman SL, Chen CH, Chang LY, Jacobson KB (1994) Rapid Commun Mass Spectrom 8:673
47. Tang K, Taranenko NI, Allman SL, Chang LY, Chen CH (1994) Rapid Commun Mass Spectrom 8:727
48. Taranenko NI, Tang K, Allman SL, Ch'ang LY, Chen CH (1994) Rapid Commun Mass Spectrom 8:1001
49. Lecchi P, Pannell LK (1995) J Am Soc Mass Spectrom 6:972
50. Distler AM, Allison J (2001) J Am Soc Mass Spectrom 12:456
51. Song FH (2003) Rapid Commun Mass Spectrom 17:1802
52. Zhou LH, Deng HM, Deng QY, Zhao SK (2004) Rapid Commun Mass Spectrom 18:787
53. Zhang ZY, Zhou LH, Zhao SK, Deng HM, Deng QY (2006) J Am Soc Mass Spectrom 17:1665
54. Yu F, Xu SY, Pan CS, Ye ML, Zou HF, Guo BC (2006) Nucleic Acids Res 34:e94
55. Berkenkamp S, Kirpekar F, Hillenkamp F (1998) Science 281:260
56. Currie GJ, Yates JR (1993) J Am Soc Mass Spectrom 4:955
57. Li YCL, Cheng SW, Chan TWD (1998) Rapid Commun Mass Spectrom 12:993
58. Asara JM, Allison J (1999) Anal Chem 71:2866
59. Shahgholi M, Garcia BA, Chiu NHL, Heaney PJ, Tang K (2001) Nucleic Acids Res 29:e91
60. Garcia BA, Heaney PJ, Tang K (2002) Anal Chem 74:2083
61. Schuerenbeg M, Luebbert C, Eickhoff H, Kalkum M, Lehrach H, Nordhoff E (2000) Anal Chem 72:3436
62. Nordhoff E, Schurenberg M, Thiele G, Lubbert C, Kloeppel KD, Theiss D, Lehrach H, Gobom J (2003) Int J Mass Spectrom 226:163
63. Tang K, Opalsky D, Abel K, van den Boom D, Yip P, Del Mistro G, Braun A, Cantor CR (2003) Int J Mass Spectrom 226:37

64. Kirpekar F, Nordhoff E, Larsen LK, Kristiansen K, Roepstorff P, Hillenkamp F (1998) Nucleic Acids Res 26:2554
65. Gilar M, Belenky A, Wang BH (2001) J Chromatogr A 921:3
66. Nordhoff E, Kirpekar F, Roepstorff P (1996) Mass Spectrom Rev 15:67
67. Tang K, Shahgholi M, Garcia BA, Heaney PJ, Cantor CR, Scott LG, Williamson JR (2002) Anal Chem 74:226
68. Li YZ, Tang K, Little DP, Koster H, Hunter RL, McIver RT (1996) Anal Chem 68:2090
69. Braun A, Little DP, Koster H (1997) Clin Chem 43:1151
70. Haff LA, Smirnov IP (1997) Genome Res 7:378
71. Ross P, Hall L, Smirnov I, Haff L (1998) Nat Biotechnol 16:1347
72. Sequenom (2009) iPLEX® gold application guide. San Diego
73. Wang L, Luhm R, Lei M (2007) Microarray Technology and Cancer Gene Profiling 593:105
74. Vogel N, Schiebel K, Humeny A (2009) Transfus Med Hemother 36:253
75. Jaremko M, Justenhoven C, Abraham BK, Schroth W, Fritz P, Brod S, Vollmert C, Illig T, Brauch H (2005) Hum Mutat 25:232
76. Horn H, Pott C, Kalla J, Dreyling M, Rosenwald A, Ott G, Schwab M, Schaeffeler E (2010) Pharmacogenet Genomics 20:598
77. Soderlund-Strand A, Dillner J, Carlson J (2008) Clin Chem 54:86
78. Cerezo M, Cerny V, Carracedo A, Salas A (2009) Electrophoresis 30:3665
79. Thompson SL, Lamothe M, Meirmans PG, Perinet P, Isabel N (2010) Mol Ecol 19:132
80. Thomas RK, Baker AC, DeBiasi RM, Winckler W, LaFramboise T, Lin WM, Wang M, Feng W, Zander T, MacConnaill LE, Lee JC, Nicoletti R, Hatton C, Goyette M, Girard L, Majmudar K, Ziaugra L, Wong KK, Gabriel S, Beroukhim R, Peyton M, Barretina J, Dutt A, Emery C, Greulich H, Shah K, Sasaki H, Gazdar A, Minna J, Armstrong SA, Mellinghoff IK, Hodi FS, Dranoff G, Mischel PS, Cloughesy TF, Nelson SF, Liau LM, Mertz K, Rubin MA, Moch H, Loda M, Catalona W, Fletcher J, Signoretti S, Kaye F, Anderson KC, Demetri GD, Dummer R, Wagner S, Herlyn M, Sellers WR, Meyerson M, Garraway LA (2007) Nat Genet 39:567
81. Lubezky N, Ben-Haim M, Marmor S, Brazowsky E, Rechavi G, Klausner JM, Cohen Y (2011) J Gastrointest Surg 15:503
82. Kiessling MK, Oberholzer PA, Mondal C, Karpova MB, Zipser MC, Lin WM, Girardi M, MacConaill LE, Kehoe SM, Hatton C, French LE, Garraway LA, Polier G, Suss D, Klemke CD, Krammer PH, Gulow K, Dummer R (2011) Blood 117:2433
83. Huebner C, Petermann I, Browning BL, Shelling AN, Ferguson LR (2007) Cancer Epidemiol Biomarkers Prev 16:1185
84. Sasayama T, Kato M, Aburatani H, Kuzuya A, Komiyama M (2006) J Am Soc Mass Spectrom 17:3
85. Risch NJ (2000) Nature 405:847
86. Goldstein DB, Ahmadi KR, Weale ME, Wood NW (2003) Trends Genet 19:615
87. Kammerer S, Roth RB, Reneland R, Marnellos G, Hoyal CR, Markward NJ, Ebner F, Kiechle M, Schwarz-Boeger U, Griffiths LR, Ulbrich C, Chrobok K, Forster G, Praetorius GM, Meyer P, Rehbock J, Cantor CR, Nelson MR, Braun A (2004) Cancer Res 64:8906
88. Suttner K, Rosenstiel P, Depner M, Schedel M, Pinto LA, Ruether A, Adamski J, Klopp N, Illig T, Vogelberg C, Schreiber S, von Mutius E, Kabesch M (2009) J Allergy Clin Immunol 123:1062
89. Abel K, Kammerer S, Hoyal C, Reneland R, Marnellos G, Nelson MR, Braun A (2005) Identification of Susceptibility Genes and Genetic Modifiers of Human Diseases. Imaging, Manipulation, and Analysis of Biomolecules and Cells: Fundamentals and Applications III, Proc. SPIE Vol. 5699. doi:10.1117/12.604620
90. Abel K, Reneland R, Kammerer S, Mah S, Hoyal C, Cantor CR, Nelson MR, Braun A (2006) Autoimmun Rev 5:258
91. Volkman SK, Sabeti PC, DeCaprio D, Neafsey DE, Schaffner SF, Milner DA, Daily JP, Sarr O, Ndiaye D, Ndir O, Mboup S, Duraisingh MT, Lukens A, Derr A, Stange-Thomann N,

Waggoner S, Onofrio R, Ziaugra L, Mauceli E, Gnerre S, Jaffe DB, Zainoun J, Wiegand RC, Birren BW, Hartl DL, Galagan JE, Lander ES, Wirth DF (2007) Nat Genet 39:113
92. Herring CD, Raghunathan A, Honisch C, Patel T, Applebee MK, Joyce AR, Albert TJ, Blattner FR, van den Boom D, Cantor CR, Palsson BO (2006) Nat Genet 38:1406
93. Ding CM, Cantor CR (2003) Proc Natl Acad Sci USA 100:3059
94. Rode TM, Berget I, Langsrud S, Moretro T, Holck A (2009) J Microbiol Methods 78:86
95. Duffield DS, Cai L, Kim S (2010) RNA 16:1285
96. Ding CM, Chiu RWK, Lau TK, Leung TN, Chan LC, Chan AYY, Charoenkwan P, Ng ISL, Law HY, Ma ESK, Xu XM, Wanapirak C, Sanguansermsri T, Liao C, Ai MATJ, Chui DHK, Cantor CR, Lo YMD (2004) Proc Natl Acad Sci USA 101:10762
97. Daelemans C, Ritchie ME, Smits G, Abu-Amero S, Sudbery IM, Forrest MS, Campino S, Clark TG, Stanier P, Kwiatkowski D, Deloukas P, Dermcitzakis ET, Tavare S, Moore GE, Dunham I (2010) BMC Genet Apr 19;11:25
98. Lo YMD, Tsui NBY, Chiu RWK, Lau TK, Leung TN, Heung MMS, Gerovassili A, Jin YJ, Nicolaides KH, Cantor CR, Ding C (2007) Nat Med 13:218
99. Zhong XY, Holzgreve W (2009) Transfus Med Hemother 36:263
100. Ehrich M, Hillenkamp F, van den Boom D (2007) Perspectives in bioanalysis, vol 2. Elsevier, Amsterdam
101. Fan AXC, Garritsen HSP, Tarhouny SEL, Morris M, Hahn S, Holzgreve W, Zhong XY (2008) Clin Chem Lab Med 46:299
102. Bocker S (2003) Bioinformatics 19:i44
103. Tang K, Oeth P, Kammerer S, Denissenko MF, Ekblom J, Jurinke C, van den Boom D, Braun A, Cantor CR (2004) J Proteome Res 3:218
104. Krebs S, Medugorac I, Seichter D, Forster M (2003) Nucleic Acids Res 31:e37
105. Stanssens P, Zabeau M, Meersseman G, Remes G, Gansemans Y, Storm N, Hartmer R, Honisch C, Rodi CP, Bocker S, van den Boom D (2004) Genome Res 14:126
106. Henzel WJ, Billeci TM, Stults JT, Wong SC, Grimley C, Watanabe C (1993) Proc Natl Acad Sci USA 90:5011
107. Mann M, Hojrup P, Roepstorff P (1993) Biol Mass Spectrom 22:338
108. Pappin DJC, Hojrup P, Bleasby AJ (1993) Curr Biol 3:327
109. Matthiesen R, Kirpekar F (2009) Nucleic Acids Res 37:e48
110. Bocker S (2007) Bioinformatics 23:E5
111. Hossain M, Limbach PA (2007) RNA 13:295
112. Benitez-Paez A, Villarroya M, Douthwaite S, Gabaldon T, Armengod ME (2010) RNA 16:2131
113. Purta E, Kaminska KH, Kasprzak JM, Bujnicki JM, Douthwaite S (2008) RNA 14:2234
114. Purta E, O'Connor M, Bujnicki JM, Douthwaite S (2009) Mol Microbiol 72:1147
115. Douthwaite S, Kirpekar F (2007) Methods Enzymol 425:3
116. Kirpekar F, Douthwaite S, Roepstorff P (2000) RNA 6:296
117. Mengel-Jorgensen J, Kirpekar F (2002) Nucleic Acids Res 30:e135
118. Durairaj A, Limbach PA (2008) Anal Chim Acta 612:173
119. Auxilien S, Rasmussen A, Rose S, Brochier-Armanet C, Husson C, Fourmy D, Grosjean H, Douthwaite S (2011) RNA 17:45
120. Goll MG, Kirpekar F, Maggert KA, Yoder JA, Hsieh CL, Zhang XY, Golic KG, Jacobsen SE, Bestor TH (2006) Science 311:395
121. Schatz P, Dietrich D, Schuster M (2004) Nucleic Acids Res 32:e167
122. Dodge JE, Ramsahoye BH, Wo ZG, Okano M, Li E (2002) Gene 289:41
123. Haines TR, Rodenhiser DI, Ainsworth PJ (2001) Dev Biol 240:585
124. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM, Edsall L, Antosiewicz-Bourget J, Stewart R, Ruotti V, Millar AH, Thomson JA, Ren B, Ecker JR (2009) Nature 462:315
125. Radpour R, Haghighi MM, Fan AXC, Torbati PM, Hahn S, Holzgreve W, Zhong XY (2008) Mol Cancer Res 6:1702

126. Ehrich M, Nelson MR, Stanssens P, Zabeau M, Liloglou T, Xinarianos G, Cantor CR, Field JK, van den Boom D (2005) Proc Natl Acad Sci USA 102:15785
127. Ehrich M, Turner J, Gibbs P, Lipton L, Giovanneti M, Cantor C, van den Boom D (2008) Proc Natl Acad Sci USA 105:4844
128. Coolen MW, Statham AL, Gardiner-Garden M, Clark SJ (2007) Nucleic Acids Res 35:e119
129. Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou JJ, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG (1998) Proc Natl Acad Sci USA 95:3140
130. Urwin R, Maiden MCJ (2003) Trends Microbiol 11:479
131. Maiden MCJ (2006) Annu Rev Microbiol 60:561
132. Levin BR, Lipsitch M, Bonhoeffer S (1999) Science 283:806
133. Musser JM (1996) Emerg Infect Dis 2:1
134. Spratt BG, Maiden MCJ (1999) Philos Trans R Soc Lond B Biol Sci 354:701
135. Enright MC, Spratt BG (1999) Trends Microbiol 7:482
136. Maiden MCJ (2000) Int J Med Microbiol 290:183
137. Jefferies J, Clarke SC, Diggle MA, Smith A, Dowson C, Mitchell T (2003) Mol Biotechnol 24:303
138. Underwood A, Green J (2011) J Clin Microbiol 49:23
139. Honisch C, Chen Y, Mortimer C, Arnold C, Schmidt O, van den Boom D, Cantor CR, Shah HN, Gharbia SE (2007) Proc Natl Acad Sci USA 104:10649
140. Amann R, Fuchs BM, Behrens S (2001) Curr Opin Biotechnol 12:231
141. Clarridge JE (2004) Clin Microbiol Rev 17:840
142. Jackson GW, McNichols RJ, Fox GE, Willson RC (2007) Int J Mass Spectrom 261:218
143. von Wintzingerode F, Bocker S, Schlotelburg C, Chiu NHL, Storm N, Jurinke C, Cantor CR, Gobel UB, van den Boom D (2002) Proc Natl Acad Sci USA 99:7039
144. Zhang ZD, Jackson GW, Fox GE, Willson RC (2006) BMC Bioinformatics 7:117
145. Hartmer R, Storm N, Boecker S, Rodi CP, Hillenkamp F, Jurinke C, van den Boom D (2003) Nucleic Acids Res 31:e47
146. Lefmann M, Honisch C, Bocker S, Storm N, von Wintzingerode F, Schlotelburg C, Moter A, van den Boom D, Gobel UB (2004) J Clin Microbiol 42:339
147. Ganova-Raeva L, Ramachandran S, Honisch C, Forbi JC, Zhai X, Khudyakov Y (2010) J Clin Microbiol 48:4161
148. Ruest A, Michaud S, Deslandes S, Frost EH (2003) J Clin Microbiol 41:3487
149. van Elden LJ, Nijhuis M, Schipper P, Schuurman R, van Loon AM (2001) J Clin Microbiol 39:196
150. Wright KE, Wilson GA, Novosad D, Dimock C, Tan D, Weber JM (1995) J Clin Microbiol 33:1180
151. Schweiger B, Zadow I, Heckler R, Timm H, Pauli G (2000) J Clin Microbiol 38:1552
152. Barr IG, Komadina N, Hurt AC, Iannello P, Tomasov C, Shaw R, Durrant C, Sjogren H, Hampson AW (2005) J Med Virol 76:391
153. Dharan NJ, Gubareva LV, Meyer JJ, Okomo-Adhiambo M, McClinton RC, Marshall SA, George KS, Epperson S, Brammer L, Klimov AI, Bresee JS, Fry AM, Grp O-RW (2009) JAMA 301:1034

Top Curr Chem (2013) 331: 79–116
DOI: 10.1007/128_2012_384
© Springer-Verlag Berlin Heidelberg 2012
Published online: 30 October 2012

Determination of Peptide and Protein Disulfide Linkages by MALDI Mass Spectrometry

Hongmei Yang, Ning Liu, and Shuying Liu

Abstract The disulfide bond is one of the most common post-translational modifications in proteins, of which determination is essential to the comprehensive understanding of protein structures. Disulfide bond analysis has undergone great improvement due to the development of matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS), especially in terms of speed and sensitivity. In general, the characterization of disulfide-containing peptides is achieved by the reduction of disulfide bonds followed by alkylation. In this review we focus on the analysis of disulfide-containing proteins/peptides by some unique methods in MALDI MS. The MALDI in-source decay (ISD) of disulfide bonds and adducts of the matrix and sulfhydryl-containing peptide are discussed in detail. The mechanism of each method is discussed so as to help the reader gain greater insight into it, and examples of its application are also presented. The goal of this review is to provide an understanding of these techniques for analysis of disulfide-linked proteins/peptides in MALDI MS.

Keywords Application · Disulfide bond analysis · Disulfide-containing proteins/peptides · MALDI MS · Mechanism

H. Yang
Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, 5625 Renmin Street, Changchun 130022, China

N. Liu
Central Laboratory, The Second Hospital of Jilin University, Changchun 130041, China

S. Liu (✉)
Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, 5625 Renmin Street, Changchun 130022, China

Changchun University of Chinese Medicine, Changchun 130117, China
e-mail: syliu19@yahoo.com.cn; syliu@ciac.jl.cn

Contents

Abbreviations

3-HCCA	α-Cyano-3-hydroxycinnamic acid
APTA	(3-Acrylamidopropyl)trimethylammonium chloride
ASA	5-Aminosalicylic acid
BSA	Bovine serum albumin
CA	Caffeic acid
CAD	Collision-activated dissociation
CDAP	1-Cyano-4-(dimethylamino)pyridinium tetrafluoroborate
CHCA	α-Cyano-4-hydroxycinnamic acid
CID	Collision-induced dissociation
Cys	Cysteine
Da	Dalton
DAN	1,5-Diaminonaphthalene
DHB	2,5-Dihydroxybenzoic acid
DTT	Dithiothreitol
ECD	Electron capture dissociation
ESI	Electrospray ionization
FA	Ferulic acid
FAB	Fast atom bombardment
FT-ICR	Fourier transform ion cyclotron resonance
IAM	Indoacetamide
ISD	In-source decay
MALDI MS	Matrix-assisted laser desorption/ionization mass spectrometry
ME	β-Mercaptoethanol
NEM	*N*-Ethylmaleimide

NMR	Nuclear magnetic resonance
NSMA	Negative signature mass algorithm
NTCB	2-Nitro-5-thiocyanobenzoic acid
PSD	Post-source decay
RNAseA	Ribonuclease A
S/N	Signal-to-noise ratio
SA	Sinapinic acid
SDS	Sodium dodecyl sulfate
THAP	2, 4, 6-Trihydroxyacetophenone
TOF	Time-of-flight

1 Introduction

Disulfide bonds in proteins are formed between the thiol groups of cysteine residues. The formation of disulfide bonds is one of the most notable post-translational modifications in proteins, which is critical for stabilizing the native structures of some proteins. The disulfide bridges in a protein cannot be predicted from its amino acid sequence; therefore, the determination of disulfide bonds in a protein will provide useful information on its 3D structure and stability, and contribute to the understanding of its structural-functional relationship [1–3]. The study of disulfide bonds is gaining in importance in many areas of biochemistry. Unfortunately, the determination of the correct disulfide linkages between multiple cysteine residues has always been challenging in structure elucidation of natural proteins/peptides.

There has been extensive work done in the development of analytical techniques which can identify disulfide bonds. Among this, the nuclear magnetic resonance (NMR)-based approach is efficient for analyzing disulfide-rich peptides [4–6]. Unfortunately, NMR is less sensitive than MS methods and certainly cannot meet the requirements of proteomic analysis. During the last two decades, great improvements in speed and sensitivity for disulfide bond analysis have been made by MS. The first investigation of reduction reactions of the disulfide bonds in bovine insulin on a mass spectrometry probe appeared in 1986 [7], in which the traditional reduction method was employed. Disulfide bridges have been extensively studied by methods in either MS or MS/MS mode, such as fast atom bombardment (FAB) [8–11] or electrospray ionization (ESI) [12–18]. Although FAB-MS does not require substantial sample preparation or derivatization, it has limitations such as a high chemical background due to the matrix employed and low sensitivity, which has hampered its usefulness for biologically-derived samples. A strategy for neutral loss of H_2S_2 and H_2S by cleavage of the S–C bond is widely employed via low-energy collision-induced dissociation (CID) in ESI, and alkali

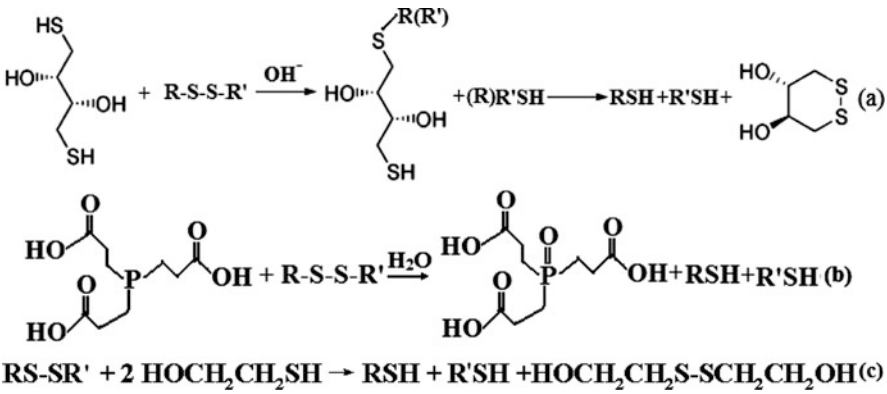
and alkaline earth metal enolate complexes were used to enhance the selective cleavages [15–18]. In addition, fragments due to the breakdown of S–S bonds were prominent in electron capture dissociation (ECD) mass spectra of multiply charged proteins, which allowed an easy determination of the presence and position of each S–S bond [19, 20]. The cleavage appeared to be due to the high S–S affinity for H· atoms. Recently, some novel strategies for the assignment of proteins/peptides containing disulfide linkages were presented by use of a new algorithm called "D-Bond" [21], a computational program named RADAR [22], a software system based on the Fenyö disulfide bond assignment algorithm [23], and the MassMatrix MS/MS Search Engine [24]. These approaches dramatically facilitated rapid and automatic disulfide bond assignment.

The advent of matrix-assisted laser desorption ionization (MALDI), which is most commonly coupled with a time-of-flight (TOF) analyzer, in the late 1980s [25, 26] marked the beginning of a new era for the study of biological macromolecules, especially polypeptides and proteins, as well as synthetic polymers [27–30]. The MALDI-TOF MS has its favorable properties including broad mass range, high sensitivity, tolerance to salts and buffers, simplicity, and fast data acquisition [29, 31–33]. It is therefore an increasingly popular technique for analysis of proteins/peptides. The common approach for elucidation of the disulfide bridge pattern of a protein is usually a combination of enzymatic degradation, Edman degradation, and MS [34–38]. However, this strategy requires further purification of the peptides after proteolytic digestion and relatively large amounts of samples due to the complex procedure. Compared to this, some methods are very attractive. For example, protease-mediated incorporation of ^{18}O into the terminal carboxylates of peptides, the only strategy used for analysis of disulfides in the reported review, was demonstrated to analyze disulfide-containing proteins/peptides [39, 40]. This procedure can be performed at an acid pH and succeeds in preventing disulfide scrambling. However, the isotope distribution may complicate the interpretation of the mass spectra.

The goal of this review is for biochemists to gain a better understanding of characteristics and analytical methods for disulfide. We describe in this chapter not only the partial reduction and alkylation strategy but also some novel strategies for identifying disulfide bonds and/or cysteine residues in proteins by MALDI MS. The phenomena regarding prompt disulfide fragmentation and adducts of matrix and sulfhydryl-containing peptide are discussed in detail.

2 Reduction Followed by Alkylation of Cysteines

In practice, most of the applications of MALDI MS to characterizing proteins that contain disulfides and free cysteines have usually been conducted with a variety of reagents for reducing the disulfides, as well as alkylating the hydrosulfide groups. One of the most widely used reducing reagents in protein chemistry is dithiothreitol (DTT) [41–44], which is often used under denaturing conditions, such as at elevated



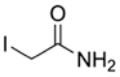
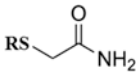
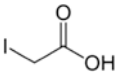
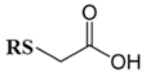
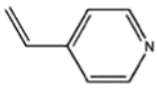
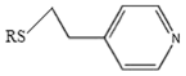
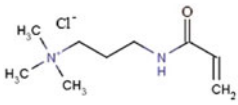
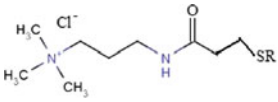
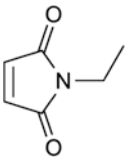
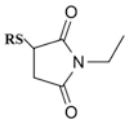
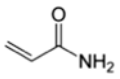
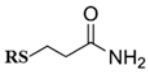
Scheme 1 Reactions between disulfide-containing proteins/peptides and (*a*) DTT, (*b*) TCEP, (*c*) ME

temperatures or in 6 M guanidinium hydrochloride, 8 M urea, or 1% sodium dodecyl sulfate (SDS). Besides DTT, tris(2-carboxyethyl)phosphine (TCEP) is another reducing agent commonly used to reduce disulfide bonds in proteins [45–50]. Unlike DTT, which becomes less potent as with lower pH and should be used in a buffer at pH 7.5–8.5, TCEP is more stable and can work even at low pH. β-Mercaptoethanol (ME) used to be used as a reagent to reduce disulfide bonds, having been traditionally used as a reducing agent in sample loading buffer for SDS polyacrylamide gel electrophoresis. However, because of its toxicity and smell, as well as its tendency to form adducts with free cysteines, nowadays it is not used as widely as DTT or TCEP. While less potent, ME is only used in the first step of sample pretreatment during the extraction of whole proteins [51]. The reactions between disulfide-containing proteins/peptides and DTT, TCEP, and ME are as shown in Scheme 1.

For alkylating the free cysteines in proteins, there are a variety of reagents that are being utilized in the field of protein chemistry and proteomics, including indoacetamide (IAM) or indoacetic acid [51–57], 4-vinylpyridine [58], (3-acrylamidopropyl)trimethylammonium chloride (APTA) [59], *N*-ethylmaleimide (NEM) [60, 61], and acrylamide [62], among which, IAM is the most widely used, and NEM can be used under acidic conditions to avoid disulfide scrambling. For example, Horn et al. [61] used a method for partial reduction and alkylation at acidic pH, coupled with peptide mapping with MALDI MS and NMR, to determine the disulfide bonds contained in the N-terminal somatomedin B domain to be Cys5:Cys9, Cys19:Cys31, Cys21:Cys32, and Cys25:Cys39. The information about the alkylating reagent is summarized in Table 1. In addition, methyl methanethiosulfonate is commonly used in isobaric tags for relative and absolute protein quantification [63], and isotope-coded affinity tag reagents are used for protein quantification through alkylation reaction [64].

Some researchers used an alternative strategy to determine the number of cysteine residues contained in proteins. In the study by Juárez et al. [58] the purified

Table 1 Structures of commonly used alkylating reagents and cysteine alkylation products

Alkylating reagent	Structure	Alkylation product
Indoacetamide (IAM)		
Indoacetic acid		
4-Vinylpyridine		
(3-Acrylamidopropyl) trimethylammonium chloride (APTA)		
N-Ethylmaleimide (NEM)		
Acrylamide		

protein was denatured and then treated with either 4-vinylpyridine, or with DTT, followed by further alkylation by excess 4-vinylpyridine. MALDI TOF MS was used to measure the exact masses of the non-reduced but alkylated protein (M_{VP}), the reduced and alkylated protein (M_{PE}), and the purified protein without any treatment (M_{NAT}). The quantitation of free cysteine residues and disulfide bonds can be achieved by simple calculations with the obtained mass data of the three kinds of the same protein as follows: $N_{SH} = (M_{VP} - M_{NAT})/105.3$, where N_{SH} represents the number of free cysteine residues and 105.3 is the mass increment due to the pyridylethylation of one thiol group (molecular weight of ethylpyridine is 107.1); $N_{Cys} = [(M_{PE} - M_{VP})/106.3] + N_{SH}$, where N_{Cys} is the number of total cysteine residues and 106.3 is the mass increment due to the pyridylethylation of a cysteine residue; $N_{S-S} = (N_{Cys} - N_{SH})/2$, where N_{S-S} is the number of disulfide bonds.

Another study used a similar strategy but the alkylation reagent was APTA [59]. After reduction, lysozyme was treated with APTA for 1 h and then dialyzed. MALDI-MS spectra of both untreated and derivatized lysozyme were compared, indicating that all the eight cysteine residues in lysozyme were derivatized by APTA. NEM has also been used to alkylate cysteines in proteins and peptides, being derived from maleic acid. It is an alkene that is reactive toward thiols. Mandal et al. [60] used a strategy of differential cysteine labeling to determine the disulfide linkage pattern in T-superfamily conotoxins from *Conus virgo*. The conditions of partial reduction and alkylation were accomplished with several dedicated combinations of DTT, IAM, TCEP, and NEM. The differentially reduced and alkylated peptide was analyzed by MALDI-MS.

Acrylamide has long been used as an alkylation reagent for cysteine, the reaction first being thought as an undesired reaction that frequently occurred during polyacrylamide gel electrophoresis. Recently, acrylamide has been widely used as an isotopic tagging reagent [62, 65] for quantitative studies rather than for cysteine/disulfide mapping studies. The applications of acrylamide in quantitative studies were summarized and reviewed by Turko and Sechi [62]. The proteins isolated by one- and two-dimensional polyacrylamide gel electrophoresis were successfully quantified by the use of acrylamide and deuterated acrylamide as a cysteine alkylating reagent, which was simple and inexpensive.

Recently, adding DTT to the sample followed by online microwave heating and online LC-MS analysis was described [55]. Disulfide bonds were reduced at the same time as the hydrolysis at aspartic acid. Hence, the need for alkylation to prevent reformation of disulfide bonds was eliminated.

3 Fundamental Studies of Prompt Disulfide Fragmentation

Early in the development of MALDI, it was considered a soft ionization technique that produced almost exclusively intact protonated species [25, 26]. With the gradual application of MALDI techniques, a significant degree of metastable decay behavior called post-source decay (PSD), which corresponds to the fragmentation of the selected precursor ion, can occur in the TOF analyzer [66–68], by which peptide or protein sequencing can be achieved [69, 70]. However, it is not commonly used due to the drawbacks, such as broad precursor selection width and relatively poor quality spectra. Another kind of metastable fragmentation, called in-source decay (ISD), occurring within several hundred nanoseconds after laser irradiation in the source, was first observed by Brown and Lennon in 1995 [71]. The minimum pulsed ion extraction delay time was 320 ns with their instrument when the dissociation product ions were observed for peptides and proteins, this fast dissociation process being influenced by the laser fluence and the matrix used [72–74]. This metastable ion decay process was consistent with a unimolecular ion activation process, possibly associated with the proton transfer step in MALDI, while the PSD process was most probably dominated by bimolecular collisional

processes. Their experimental results showed that the metastable fragmentation in the source was also quite different from the prompt fragmentation [75] which we will discuss in the following section. Prompt fragmentation also occurs in the source but within several nanoseconds after laser irradiation.

3.1 Prompt Fragmentation of Disulfide-Linked Proteins/Peptides

One has to be aware that prompt fragmentation can be observed in linear MALDI-TOF MS equipped with not only delayed extraction but also continuous acceleration [75, 76]. In early 1989, Beavis and Chait observed prompt disulfide fragmentation (Fig. 1) when they investigated factors affecting the ultraviolet laser desorption of proteins using a simple linear instrument equipped with an HY-400 Nd-YAG laser [78]. No distinct fragments were detected for any of the proteins studied in the positive mode. Fragment ions were observed for the intact A and B chains due to symmetric cleavage at the two interchain disulfide linkages in the negative-ion-mode spectrum of insulin. Zhou and co-workers detected A and B chain fragments with various numbers of sulfurs attached in the negative ion mode, corresponding to cleavage ions of the interchain disulfide bridges at different points [79] (Table 2). The mirror voltage was optimized at −9.7 kV for chain B, while it was −6.7 kV for chain A. A and B represent A and B chain fragments due to symmetric cleavage of the two interchain disulfides in this table. They found two characteristic quadruplet of peaks with separation of 32 u, which correspond to the fragment with retention of zero (B_0, A_0), one (B_1, A_1), two (B_2, A_2), and three sulfurs (B_3, A_3) at the two interchain disulfide linkages, respectively. This type of decomposition is very useful for identification of the presence of disulfide bonds. Fragments with much lower intensities were also observed in the positive spectrum, and no A-chain fragments were detected.

In other articles, prominent ions from prompt disulfide fragmentation were reported in the positive MALDI spectra [75–77]. The parameters such as matrix, accelerating voltage, detection mode, and laser fluence were investigated [75, 76]. Prompt fragments were not always observed in the spectra using sinapinic acid (SA) as matrix, and the spectra had poor signal-to-noise ratio (S/N). Laser fluence played an essential role in obtaining prompt fragmentation of disulfide-linked peptides because ISD fragments appeared only after increasing the fluence of the laser beam to above the threshold. By contrast, other parameters had little effect on the extent of prompt fragmentation. To confirm whether the fragmentation was a result of chemical reduction by the matrix, ESI-MS was employed for analysis before and after mixing, incubating, and drying with matrix [75]. The results indicated that the matrix did not cause any reduction. Analysis of the same peptides or tryptic peptide mixtures by ESI-MS indicated that the disulfide bonds were not cleaved during the sample treatment [76, 77]. The specific cleavage at the S–S with increased laser fluence did occur in the MALDI MS analysis though its mechanism was not clear.

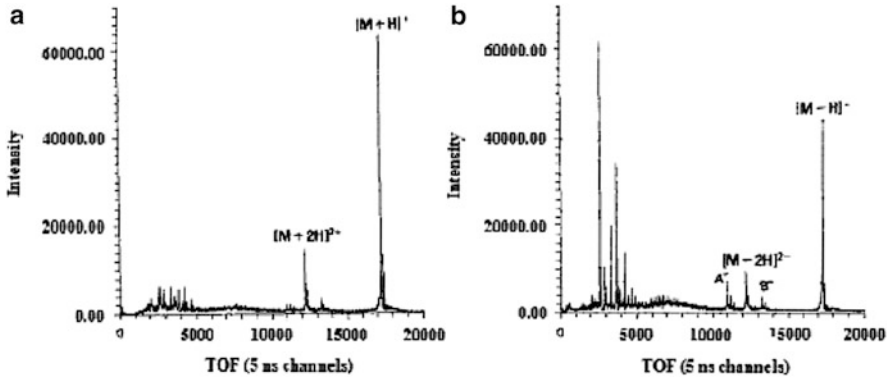


Fig. 1 MALDI-TOF mass spectra of bovine insulin using 2-pyrazinecarboxylic acid as matrix in (**a**) positive ion linear mode and (**b**) negative ion linear mode. The mass spectra were generated from the sum of 100 laser shots. A, B, and M denote A-chain, B-chain, and intact insulin molecules, respectively (adapted with permission from [77]. © 1989 Heyden & Son Limited)

Table 2 Fragment ions of bovine insulin observed in negative-ion MALDI MS (adapted with permission from [80]. © 1993 Elsevier Science Publishers B.V.)

	Peak			
	B_3	B_2	B_1	B_0
Assignment	$[B - H + S]^-$	$[B - H]^-$	$[B - H - S]^-$	$[B - H - 2S]^-$
Calculated mass (u) (chemical)	3,429.0	3,396.9	3,364.8	3,332.8
Experimental mass (u)	3,429.6	3,397.1	3,365.3	3,333.2
Δm (u)	0.6	0.2	0.5	0.4
Experimental spacing (u)	$\Delta B_{3.2} = 32.5$	$\Delta B_{2.1} = 31.8$	$\Delta B_{1.0} = 32.1$	
	Peak			
	A_3	A_2	A_1	A_0
Assignment	$[A - H + S]^-$	$[A - H]^-$	$[A - H - S]^-$	$[A - H - 2S]^-$
Calculated mass (u) (chemical)	2,366.7	2,334.6	2,302.5	2,270.5
Experimental mass(u)	2,367.4	2,336.0	2,303.0	2,272.3
Δm (u)	0.7	1.4	0.5	1.8
Experimental spacing (u)	$\Delta A_{3.2} = 31.4$	$\Delta A_{2.1} = 33.0$	$\Delta A_{1.0} = 30.7$	

The authors' group found the characteristic "triplet" ions with a mass difference of a single sulfur atom due to MALDI ISD in positive ion reflectron mode for the first time [77] (Fig. 2). According to the argument applied to determine single interchain bridges by collisional dissociation in positive ion mode [80, 81], the symmetric cleavage should be strongly favored, i.e., P1 > P0 and P1 > P2, where P0, P1, and P2 are probabilities of bearing zero, one, and two sulfurs for the charged fragment, respectively. Our observation was in agreement with the above deduction. The resolution in previous reports [75, 76, 78, 79] was insufficient to determine whether capture or loss of hydrogens occurred in prompt disulfide fragments. Judging from Table 3, we can conclude that the fragments from

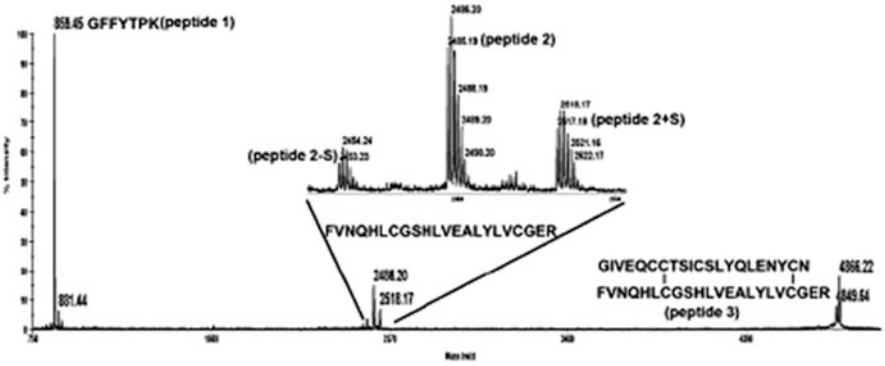


Fig. 2 MALDI-TOF mass spectrum of human insulin tryptic digests using CHCA as matrix. The *inset* shows enhanced signals of the triplet ions. The ion at *m/z* 2,485.19 is produced by symmetrical fragmentation of the two inter-chain disulfide bonds (adapted with permission from [81]. © 2007 Wiley InterScience)

Table 3 The fragmentation pattern of disulfide bonds with a mass separation of a single sulfur atom obtained by MALDI-TOF-MS (adapted with permission from [81]. © 2007 Wiley InterScience)

	Monoisotopic $[M + H]^{+a}$	
Peptide sequence	Calculated (Da)	Observed (Da)
FVNQHLCGSHLVEALYLVCGER	2,485.21	2,485.19
	2,485.21 + S (31.97) = 2,517.18	2,517.18
	2,485.21 − S (31.97) = 2,453.24	2,453.23

[a]Calculated based on all Cys residues being involved in disulfide bonds

human insulin tryptic digests are not protonated at the sulfur atom. To investigate the matter further, β_2-microglobulin was selected to be another model protein. The sequences of prompt fragments from β_2-microglobulin tryptic digests are DEYACR and SNFLNCYVSGFHPSDIEVDLLK, respectively. Calculated based on the two fragments protonated at the sulfur atom, the *m/z* values are 756.30 and 2,497.21, respectively. The observed values from Fig. 3 are consistent with the calculated ones, which indicate cleavage of the disulfide bond and addition of a proton to the sulfur atom. This discrepancy was rationalized as described in Sect. 3.2 in the following section.

When UV-MALDI MS was used to analyze disulfide-containing proteins/peptides, disulfides can readily undergo ISD with fragmentation efficiency being related to the matrix used [82]. This phenomenon could be of much analytical value for automated screening for the presence of disulfide bonds in heteropeptides or complex peptide mixtures [76, 83]. In order to prove its feasibility for identifying peptides with a disulfide linkage in complex mixtures, bovine ribonuclease A (RNase A) and bovine serum albumin (BSA) were used as model systems containing 8 and 35 cysteine residues, respectively. All 4 disulfide bonds in

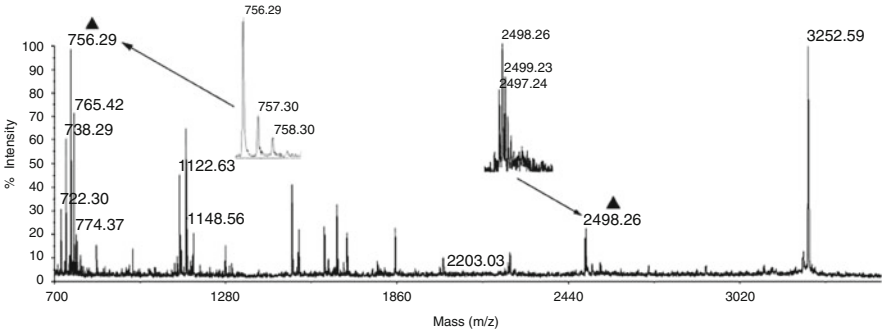


Fig. 3 MALDI-TOF mass spectrum of β_2-microglobulin tryptic digests using CHCA as matrix in positive ion reflectron mode. Signals marked with *triangles* are assigned to prompt fragments (adapted with permission from [81]. © 2007 Wiley InterScience)

RNase A and 8 two-disulfide clusters comprising 16 out of the 17 disulfide bonds in BSA were successfully identified by MALDI ISD and LIFT-TOF/TOF-MS [83].

3.2 Mechanism for Disulfide Scrambling in the Gas Phase

Disulfide scrambling means disulfide interchange which can occur at neutral and alkaline pH, resulting in incorrect formation of disulfide bonds. Several probable mechanisms of the prompt fragmentation of disulfide-containing proteins/peptides were proposed, which could help us answer some questions. Possible reactions for producing the negative fragment ions are shown in Fig. 4 [79]. Several experiments suggest that disulfide bond cleavages observed in MALDI are due to photodegradation in the solid prior to desorption [72, 75, 76, 84]. At higher laser fluences it was possible that more matrix molecules going into the gas phase resulted in a plume with higher density [75]. The accelerating analyte ions were in conditions similar to high-energy CID and underwent prompt fragmentation. Figure 5 represents laser-induced reduction mass spectra of native oxytocin [84]. The ^{13}C isotopic distribution of the quasimolecular ions of oxytocin matched the calculated isotopic relative natural abundances within 5% experimental error for the single laser shot spectrum. However, after about 100 subsequent laser shots on the same laser spot, the relative abundances of the ion $[M + 2 + Na]^+$ and the ion $[M + 2 + H]^+$ (both denoted by asterisks) increase. As evidenced by increasing abundance at 2 Da above the monoisotopic mass, disulfide bond reduction was induced by laser. In contrast, obvious disulfide bond reduction of isotocin was not observed. Therefore, the extent of disulfide bond cleavage also appears to depend on the analyte.

In a recent article, disulfide bond scrambling in the gas phase during MALDI-MS analysis was reported [86]. The fragmentation of disulfide bonds during

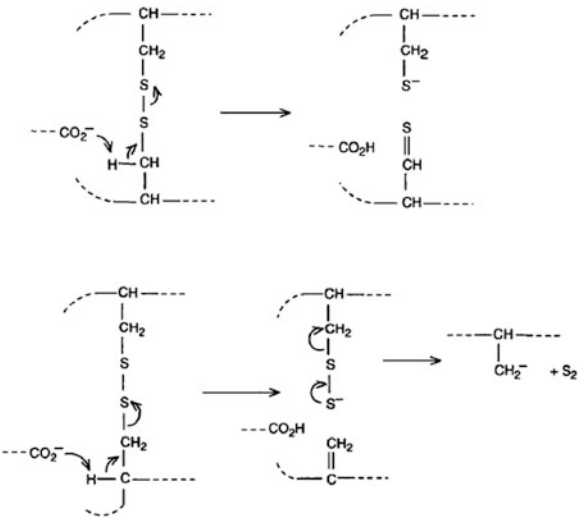


Fig. 4 The two possible reactions for cleavage of disulfide bonds in negative-ion MALDI MS. The *top* symmetrical cleavage is promoted by a carboxylate ion of either chain. The *down* asymmetrical cleavages might be promoted first by a carboxylate ion and then the negative charge developed at the first cystine residue (adapted with permission from [80]. © 1993 Elsevier Science Publishers B.V.)

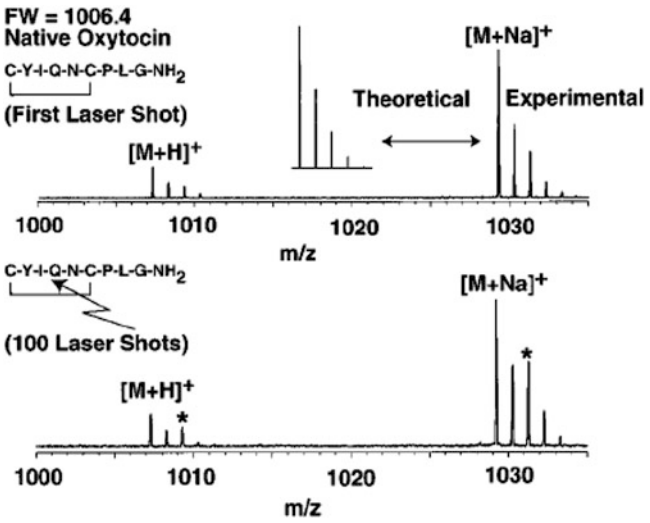


Fig. 5 MALDI FTICR mass spectra of native oxytocin from a single laser shot (*top*) and from 100 laser shots averaged (*bottom*). Reduced oxytocin in situ during the MALDI process was denoted by *asterisks* at 2 Da above the monoisotopic mass (adapted with permission from [85]. © 1997 American Chemical Society)

MALDI proceeded with high probability through radical intermediates. For example, the commonly used matrix 2,5-dihydroxybenzoic acid (DHB) with OH-groups in *para*-position relative to each other has the potential to undergo redox reaction just as the hydroquinone is oxidized to benzoquinone in the solution phase. This process furnishes two protons and two electrons. Either α-cyano-4-hydroxycinnamic acid (CHCA) or DHB matrix can reduce the

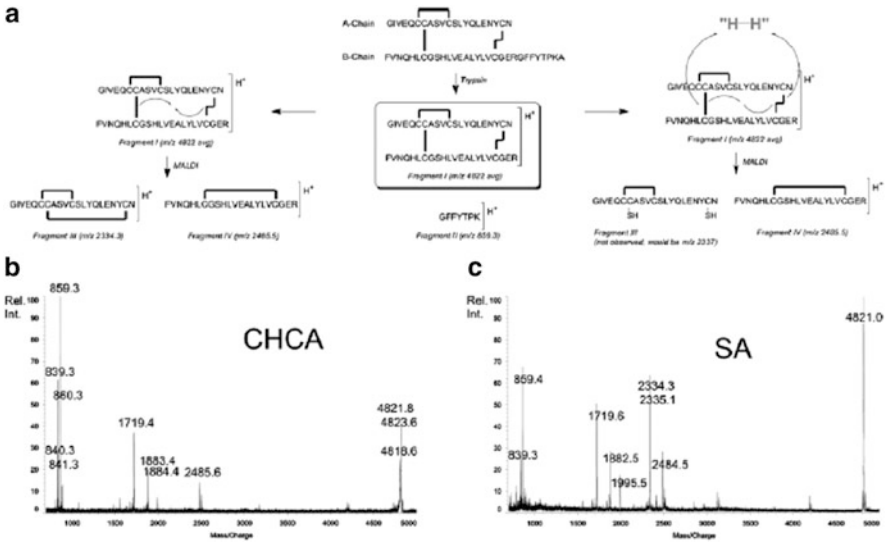


Fig. 6 A possible mechanism for this gas-phase disulfide bond scrambling is based on photo-induced disulfide scrambling (*left* of **a**) or reduction by the matrix during desorption/ionization (*right* of **a**). MALDI MS of bovine insulin tryptic digests using (**b**) CHCA and (**c**) SA as matrices. Fragment III in (**a**) observed in SA provides the evidence for photo-induced mechanism (adapted with permission from [87]. © 2009 American Society for Mass Spectrometry. Published by Elsevier Inc.)

oxidation product (Cys–Cys) of cystine to cystine during desorption/ionization. However, this was not so with 2, 4, 6-trihydroxyacetophenone (THAP). It did not undergo redox reactions probably because THAP has three OH-groups in *meta*-position relative to each other. Evidence for a radical disulfide reduction reaction initiated by the matrix in the gas phase was found.

Still, a point that remains puzzling is whether prompt disulfide fragments are protonated at the sulfur atom. This question will be answered in the two following paragraphs. The presence of interchain disulfide bonds in the case of bovine insulin is an excellent example for solving the problem. A possible mechanism for the gas-phase disulfide bond scrambling and the mass spectra of bovine insulin tryptic digests are as indicated in Fig. 6. The less intense peak at *m/z* 2,485.6 in Fig. 6b corresponds to a new peptide (IV) with an intrachain disulfide bond. The observed fragment ion IV at *m/z* 2,485.6 results from photoinitiated formation of sulfhydryl radicals and rapid intramolecular recombination and is a non-reductive process consistent with previously published results from collisionally activated dissociation (CAD) [85] and MALDI [77] experiments. An A-chain fragment ion III with a theoretical *m/z* value of 2,336.9 was not observed in the mass spectrum using CHCA as matrix. However, the oxidized A-chain fragment ion III is observed at *m/z* 2,335.1 in prominent abundance in Fig. 6c. This phenomenon may be due to the differences in the efficiency of light-absorption in the matrix and the subsequent physical desorption/ionization process. A reasonable explanation is that the

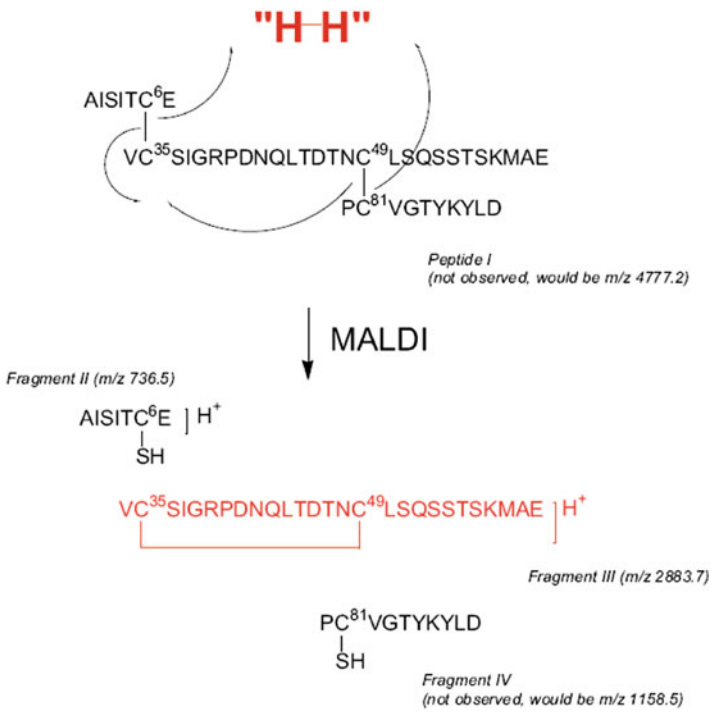


Fig. 7 Possible reaction mechanism for the photo-induced scrambling of disulfide bonds of a peptide from the Glu-C digestion of SEL24K (adapted with permission from [87]. © 2009 American Society for Mass Spectrometry. Published by Elsevier Inc.)

A-chain fragment ion III lacking basic amino acids is ionized more easily in the SA matrix than in the CHCA matrix. Evidence for photo-induced disulfide scrambling was observed from fragment III in SA.

Next, a peptide from the Glu-C digestion of the rhamnose-binding lectin SEL24K was used as a simple model peptide [86] (Fig. 7). Hydrogen radicals from the matrix attack sulfur atoms under laser irradiation, which induce disulfide bond fission. The formation of the scrambled intrachain fragment ion at *m/z* 2,883.7 was possibly due to spatial disposition of the two sulfhydryl radicals. The reductive ISD fragment ion at *m/z* 736.5 was the result of reduction by the matrix. All other peptides studied with an odd number of cysteines became subject to reduction, and those with an even number of cysteines (fully-paired) displayed mainly scrambled disulfide bonds for thermodynamic reasons. The results of β_2-microglobulin tryptic digests in [80] complied with the rule described here. In short, radical recombination scrambling may possibly be predominant if peptides with fully-paired intrachain disulfide bridges can be formed, and matrix/ISD-initiated radical redox scrambling may possibly be predominant as long as chains possess an odd number of cysteines.

3.3 Comparison of ECD and ISD

The ECD technique, developed by McLafferty's group in 1998 [87], is based on fragmentation of multiply charged cations that interact with electrons. Thus, typically, the ESI technique capable of generating multiply charged cations was employed in electron capture mass spectrometry experiments. Up to now, efficient ECD is performed only with Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) for the following two reasons. First, at least several milliseconds are required to ensure electron capture by the majority of the multiply charged precursor ions [88]. Second, precursor ions are irradiated with low-energy electrons (<0.2 eV) to obtain efficient ion-electron interactions [89]. Hence ECD and ESI FT-ICR are an ideal combination. It is difficult to meet these prerequisites in many other types of mass spectrometers, such as TOF and quadrupole instruments. ECD is believed to be nonergodic [87, 90, 91], i.e., the cleavage happens rapidly prior to intramolecular vibrational energy redistribution. Owing to this nature, c and z• (or c• and z) ions resulting from the cleavage of the backbone (N–Cα) of a peptide or protein are mainly formed during the ECD process [87, 90–92]. ECD therefore provides information complementary to that obtained with other dissociation techniques. In the case of labile post-translational modifications such as glycosylation and phosphorylation, they are preserved in ECD, which are often not in CID [93–95].

Interestingly, it is a distinctive feature of ECD that preferential cleavage of S–S bonds in disulfide-linked peptides occurred, leading to the formation of a neutral sulfhydryl (R-SH) and a thiyl radical species (·S-R) [20, 88–90, 96, 97]. This feature allows for an easy determination of the S–S bond presence and position as shown in Fig. 8 [20]. The most abundant fragment corresponds to $(RSH + 2H)^{2+}$ ion in the ECD spectrum. When porcine insulin, which contains two S–S bonds connecting its A and B chains, was analyzed by ECD, the major products of the more highly charged B-chain were the even-electron ions $(HSBSH + 3H)^{3+}$ (most abundant) and $(SBS + 3H)^{3+}$ (possibly cyclic) and the corresponding 2^{+} ions; those of the much less abundant A-chain ions were mainly the odd-electron ions $[S(A - H)S\cdot + H]^{1+}$ and $(HSAS\cdot + H)^{1+}$. It was speculated that the formation of the most abundant $(HSBSH + 3H)^{3+}$ ion involved transfer of an extra H· to the most highly charged B-chain, as indicated in Scheme 2. The preference for cleavage of disulfide bonds is true for other disulfide-linked proteins in ECD.

Protonation at S–S is improbable theoretically because the proton affinity of S–S is lower than many other protein functional groups; for example, CH_3SSCH_3 is 23.6 kcal/mol less than that of $CH_3CONHCH_3$ [20]. For the explanation of protonation at S–S, the hot hydrogen-atom mechanism was proposed [20], which states that disulfide bond cleavage was likely due to the high affinity of disulfide for the H··atom resulting from neutralization of a proton by electron capture. Then a hypervalent intermediate formed, leading to rapid dissociation of disulfide bond. Subsequently, the hot hydrogen-atom mechanism was supported by another

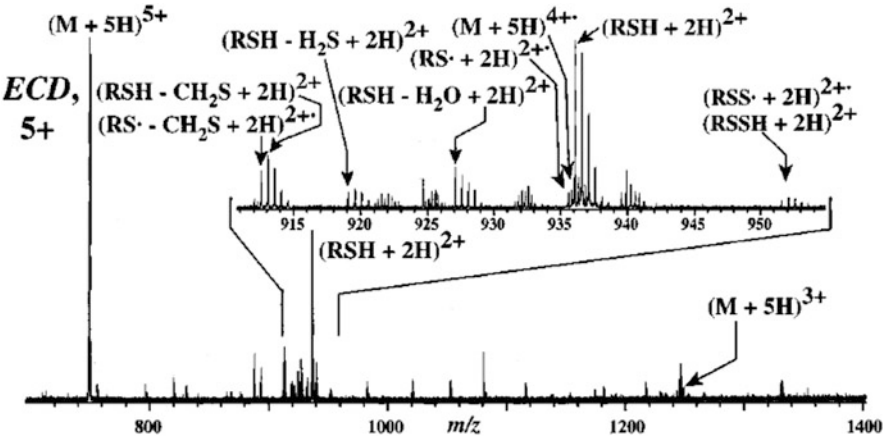
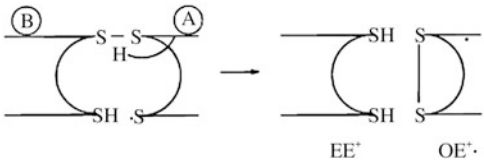


Fig. 8 ECD spectrum of this symmetrical R-S-S-R peptide (R is -CLKMAGNGRQLREILLG). Numerical values refer to product charge states (adapted with permission from [20]. © 1999 American Chemical Society)

Scheme 2 The mechanism of the formation of HSBSH (Adapted with permission from [20]. © 1999 American Chemical Society)



group's results [98]. Recently, direct dissociative electron attachment has been suggested as an alternative to the hot hydrogen mechanism [99, 100]. Results showed that electron attachment mainly occurred at the positive sites (e.g., protonated side chains). Electron transfer can thus take place from a positively charged site to a disulfide sigma* orbital provided that this orbital experiences sufficient Coulomb stabilization from proximal positively charged groups, through which disulfide cleavage occurs. In fact, multiple mechanisms may play a role in the fragmentation of ECD.

Consistent with the preferential cleavage of disulfide by ECD, fragmentation of disulfide bonds, which occurs within a very fast time scale in the MALDI ion source, is often observed as described above. In comparison with ECD, no odd electron species (·S-R) was detected after the cleavage of the S–S bond by MALDI-ISD. Probably ·S-R is quickly converted into a neutral sulfhydryl (R-SH) by picking up a proton from the dense MALDI plume [83]. Multiply-charged molecular ions are present in most reported ISD spectra, and there are plenty of electrons, hot hydrogen atoms, and matrix radicals in the MALDI plume [101–103]. Therefore, it is reasonable to assume that ISD is induced by a similar mechanism as proposed for ECD. Speculation was confirmed by Zubarev's group

showing that that ISD was a radical reaction, following a fragmentation mechanism analogous to the hot hydrogen atom model in ECD [104]. The hot hydrogen atom in ISD is generated by the matrix, presumably resulting from a photochemical reaction.

4 Application of the Adducts of Matrix and Sulfhydryl-Containing Peptide to Study Disulfides/Cysteines

It is well known among mass spectrometrists that matrix association is commonly observed to some extent in the gas phase. Most of the matrix adduct signals reported in the literatures [78, 105–108] are formed as non-covalent species in the gas phase. However, the peptide-matrix adducts observed in our study are based on the interaction between the -SH group of Cys and an α, β-unsaturated matrix in the condensed phase.

4.1 Observation of the Peptide-Matrix Adducts and the Factors of the Adduct Formation

When on-target reduction was performed with TCEP in the presence of the CHCA matrix, peptide-matrix adduct ions were observed [109, 110]. Figure 9 shows MALDI mass spectra of a disulfide-containing peptide from α-amylase tryptic digests before and after on-target reduction. In Fig. 9a, the disulfide-linked peptide [101–104]-S–S-[517–534] ion at *m/z* 2,544 as the base peak is shown together with small signals for its prompt disulfide fragment at *m/z* 2,070 and another peptide from α-amylase tryptic digests at *m/z* 1,350. The ion signal at *m/z* 2,544 disappears almost completely, and two new abundant ion signals for the peptide [517–534] at *m/z* 2,070 and its HCCA matrix adduct at *m/z* 2,259 are shown in Fig. 9b. The presence of TCEP on the target did not interfere with analyte ions detection. However, only peptide-HCCA 1:1 adducts were detected, and the adducts of other matrices and cysteine-containing peptides were not observed.

Recently the authors' group investigated this matrix addition phenomenon in detail [111]. Figure 10a, b compare MALDI mass spectra of oxytocin (simplified as M_1) with CHCA matrix before and after reduction by DTT, and Fig. 10c, d correspond to those of octreotide (M_2), respectively. Interestingly, oxytocin produces sodium-coordinated species as the major ion. Comparing Fig. 10a with Fig. 10b as well as Fig. 10c with Fig. 10d, it can be seen that the new ions are 2 u higher than their native peptides, resulting from reduction of the disulfide bond. The mass spectrum in Fig. 10b contains three dominant new signals at *m/z* 1,220.50, 1,409.53, and 1,387.54, which are attributed to $[M_1 + 2 + CHCA + Na]^+$, $[M_1 + 2 + 2\ CHCA + Na]^+$, and $[M_1 + 2 + 2\ CHCA + H]^+$, respectively. Similarly, two distinct adduct ions $[M_2 + 2 + CHCA + H]^+$ and $[M_2 + 2 + 2\ CHCA + H]^+$ at *m/z* 1,210.47 and

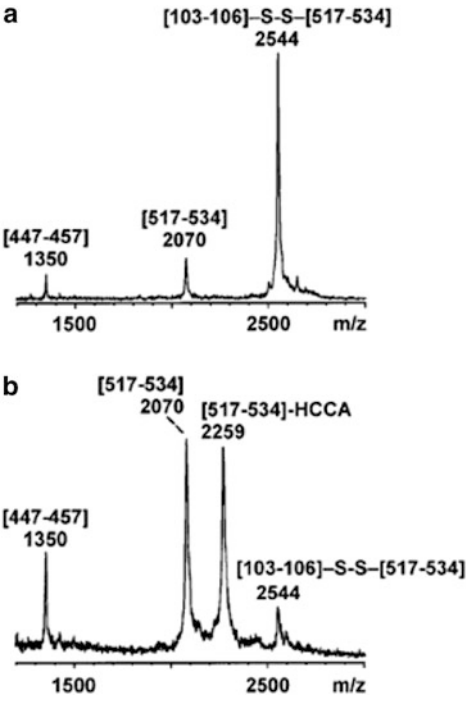


Fig. 9 MALDI mass spectra of an HPLC-fractionated tryptic digest of a disulfide-containing peptide from α-amylase (**a**) before and (**b**) after on-target reduction with TCEP in the presence of CHCA matrix (adapted with permission from [111]. © 1997 The American Society for Biochemistry and Molecular Biology, Inc.)

1,399.50 are observed in Fig. 10d. In order to gain a better insight into this phenomenon, several other commonly used matrices were tested. SA, ferulic acid (FA), α-cyano-3-hydroxycinnamic acid (3-HCCA), and caffeic acid (CA) were also found able to form matrix-adducts of reduced forms of disulfide-containing peptides at pH 7.5.

To clarify the question as to whether peptide-matrix adducts were formed in solution or in gas phase, octreotide solution after treatment with DTT and incubation with CHCA was analyzed by LC/ESI-MS [111]. The reduced octreotide-adducts of one CHCA and two CHCA were easily detected by this method. LC/ESI-MS analysis provided evidence for the generation of peptide-matrix adducts in the solution phase. A systematic study of the observed fragmentations was presented by the authors' laboratory [111]. The pH of the solution is a very important factor in controlling the formation of the adducts as evidenced by the mass spectra in Fig. 11. Through comparison of the mass spectra (Fig. 11a, e, Fig. 11b, f, Fig. 11c, g, and Fig. 11d, h), it can be readily found that the higher the pH, the higher the intensity of peptide-CHCA adducts. CHCA was less sensitive to pH than the other matrices. Distinct peptide-matrix adducts for SA, FA, and CA were only formed under alkaline conditions, and no peptide-matrix adducts were observed for other matrices even under strong alkaline conditions. Therefore, CHCA is the best candidate for producing intense adducts.

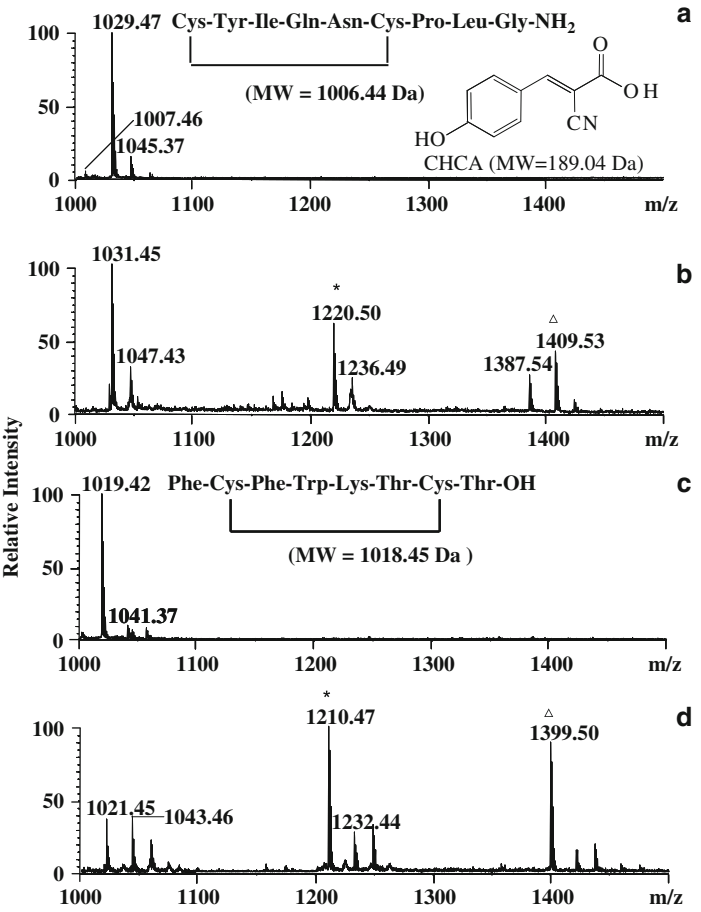


Fig. 10 MALDI mass spectra in positive ion mode of (**a**) oxytocin, (**b**) oxytocin after on-target reduction by DTT, (**c**) octreotide, and (**d**) octreotide after on-target reduction by DTT using CHCA as the matrix under pH 7.5. Signals marked with *asterisks* are assigned to peptide-matrix 1:1 adducts. Signals corresponding to peptide-matrix 1:2 adducts are marked with *triangles* (adapted with permission from [112]. © 2009 American Society for Mass Spectrometry. Published by Elsevier Inc.)

Several factors including molar ratio of matrix to peptide, incubation time, and reductant were also investigated. An increase of the molar ratio of matrix to peptide could enhance the intensities of adducts in some conditions. Incubation times (0.5–4 h) did not significantly affect the addition reaction. Among the three commonly used reductants, ME was not found to be favorable because formation of mixed disulfides between the cysteinyl thiol-containing peptide and ME was dominant at higher molar ratios of ME to disulfide-linked peptide (>150:1). The data supported the report of Spiess et al. in which peptide-CHCA adduct ions were not observed at the molar ratios of ME/peptide >200:1 [109]. Peptide-matrix

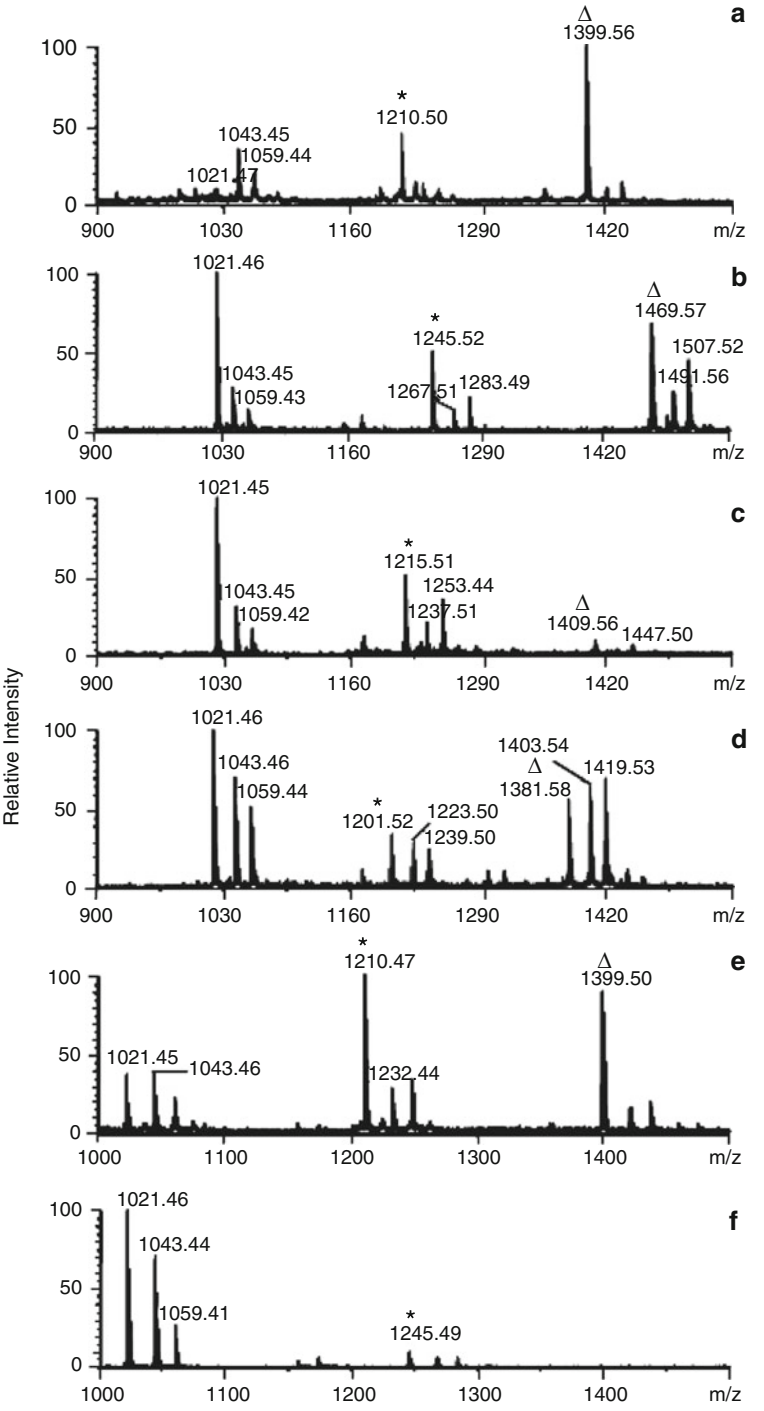


Fig. 11 (continued)

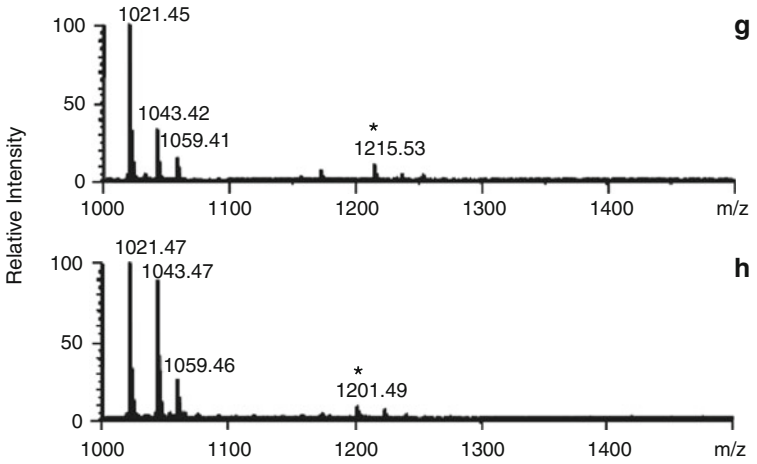


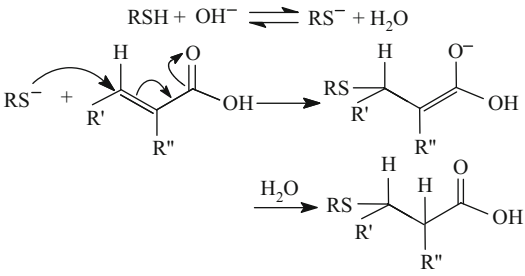
Fig. 11 MALDI mass spectra in positive ion mode of octreotide after on-target reduction by DTT using (**a**, **e**) CHCA, (**b**, **f**) SA, (**c**, **g**) FA, and (**d**, **h**) CA as matrices. The mass spectra (**a**–**d**) and (**e**–**h**) are acquired under pH 8.5 and 7.5, respectively. Annotation is performed similar to Fig. 10 (adapted with permission from [112]. © 2009 American Society for Mass Spectrometry. Published by Elsevier Inc.)

adducts were observed at the molar ratios of ME/oxytocin <100:1, but here great amounts of disulfide-linked peptides were not reduced. In contrast, DTT and TCEP were recommended for analysis of disulfide-linked proteins/peptides.

Adducts were observed when cysteine-containing protein biomarkers from lysates of *E. coli* were analyzed by MALDI MS equipped with a 355-nm pulsed solid-state YAG laser using SA as matrix [113]. The additional peaks at *m*/*z* ~208 greater than the *m*/*z* of a cysteine-containing protein ion and their doubly charged ions at △ *m*/*z* ~105, corresponding to the protein-SA adducts, were observed when using SA. However, they were absent when using CHCA. Minor adducts of cysteine-containing protein and FA may also be detected although perhaps not as obviously as the protein-SA adducts. It was not clear whether pH or solvent would affect the formation of the cysteine residue-SA adducts. This MALDI analysis of cysteine-containing proteins was at low pH (2–3) conditions, very different from previous reports [109–111].

4.2 Mechanism of Forming the Adducts of Matrix and Sulfhydryl-Containing Protein/Peptide

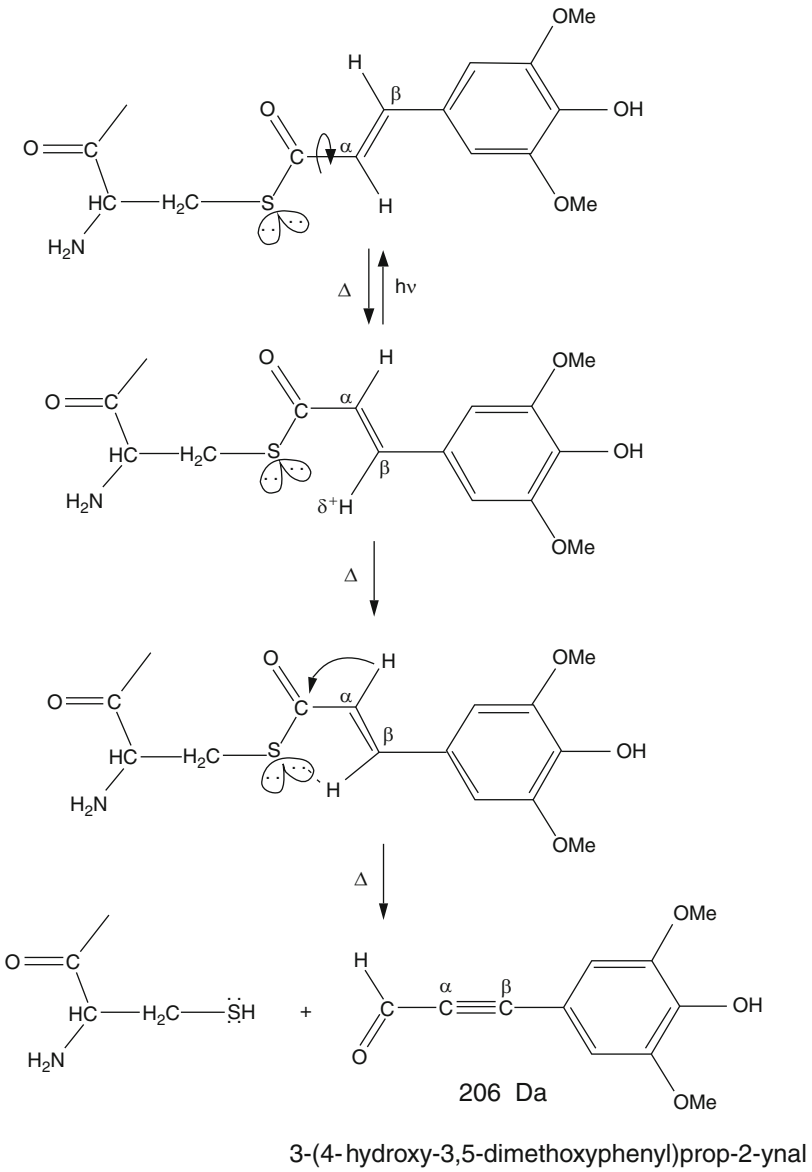
It is well known that cysteines in proteins can easily be alkylated by α, β-unsaturated compounds (notably acrylonitrile, acrylamide, and methyl acrylate) [112, 114–117]. It was assumed that these alkylations were susceptible to a lot of



Scheme 3 The proposed reaction mechanism of a free -SH group of proteins and an α, β-unsaturated carboxylic matrix (adapted with permission from [112]. © 2009 American Society for Mass Spectrometry. Published by Elsevier Inc.)

experimental parameters including reaction time, the composition of the reaction medium, and, in particular, its pH, the molarity of the monomers, intrinsic characteristics, and the amino acid composition of the protein. The Cys-acrylamide adduct had been identified by NMR, MS, and elemental analysis of the purified compound [118]. Similarly, only α, β-unsaturated carboxylic matrices were able to form matrix-adducts of reduced forms of disulfide-containing peptides under mildly alkaline conditions [111]. The phenomenon of peptide-matrix appeared to be consistent with that of protein alkylation by α, β-unsaturated compounds. Based on the existing data together with those reported by other groups, a reaction mechanism of a free -SH group in proteins and an α, β-unsaturated matrix is proposed as shown in Scheme 3. First, a free -SH group of proteins releases a proton under alkaline conditions, and the product is RS^-. Cysteine with the proton affinity at 216.1 kcal mol^{-1} is rather basic [119], which rationalized the key reaction factor, i.e., alkaline conditions. Then RS^- attacks at the double bond of the α, β-unsaturated matrix, and the addition reaction occurs. The nucleophilic addition reaction can occur easily when -R″ is an electrophilic group (cyano radical -CN, for example). CHCA and 3-HCCA are considered as derivatives of acrylonitrile, which are favorable to the Michael addition reaction; thus CHCA or 3-HCCA are more effective for forming distinct peptide-matrix adducts.

Soon after the above phenomenon was reported, the protein biomarkers containing two cysteine residues were identified as having reactivity to SA as described in Sect. 3.1 [113]. As noted above, the mass differences between ions with and without SA attachment was ~208 Da; therefore SA may be covalently bound to one cysteine residue by a thiol ester bond according to the previous reports [120–123]. In an attempt to investigate it further, the cysteine-containing proteins and the protein-SA adducts in the spectra of MALDI MS were then analyzed by MALDI-TOF-TOF MS. Through comparison of the two spectra, the fragment ions at *m/z* approximately 206 larger than their adjacent ions appeared in the MALDI-TOF-TOF mass spectra of the protein-SA adducts. These fragment ion comparisons unambiguously confirmed that the protein-SA adducts may result from covalent attachment of SA via a thiol ester linkage to the side-chain of cysteine residues. The mass differences (~206 Da) between fragment ions with and without SA attachment revealed dissociative loss of SA via a molecular ion rearrangement followed by cleavage of thiol ester linkage from the cysteine residue as illustrated in Scheme 4.



Scheme 4 The proposed mechanism of dissociative loss of the SA-adduct of a free -SH group of protein (adapted with permission from [114]. © 2010 American Society for Mass Spectrometry. Published by Elsevier Inc.)

Thiol ester-linked SA underwent *cis–trans* photoisomerization of the *trans* vinyl bond in the solution phase [121, 123]. Such an isomerization may occur for thiol ester-linked SA in the gas phase during the 5-ns desorption/ionization laser pulse.

As shown in Scheme 4, the *trans* vinyl bond in thiol ester-linked SA may be isomerized to the *cis* configuration by laser. Then the stable five-membered transition-state consisting of the sulfur, carbonyl carbon of the thiol ester bond, the α-carbon, the β-carbon, and β-hydrogen forms at temperatures as high as 700–900 K in a few nanoseconds after sample irradiation [103]. In this pentacyclic transition-state, proton abstraction of the labile β-hydrogen by the sulfur atom and transfer of the α-hydrogen to the carbonyl carbon resulted in the formation of a free sulfhydryl group and a propargyl aldehyde.

The molecular weight of 3-(4-hydroxy-3,5-dimethoxyphenyl)prop-2-ynal is 206 Da, which is in accordance with the observed mass of the leaving group. CHCA did not form a thiol ester linkage at the cysteine residues under acidic conditions in consequence of the absence of an α-hydrogen on the basis of the proposed mechanism. This appeared to be an acid catalyzed mechanism and different from the previous one [111]. Still, there was a discrepancy between this observation and a previous report [124], where no thioester bond was formed by the direct reaction of the carboxyl group of *p*-coumaric acid and the thiol group of cysteine under the experimental conditions used. It appears that the question of thiol ester linkage is more complex than expected.

4.3 Application for the Analysis of Disulfide-Rich Proteins

Disulfide linkages in tryptic digests of insulin and lysozyme had been successfully examined using the specific modification of protein sulfhydryl groups with CHCA [111]. Figure 12 represents the MALDI mass spectra of insulin and lysozyme under different conditions. Comparing the mass spectrum of protein tryptic digests and that of reduced protein tryptic digests, we can obtain the total number of disulfide bonds in the protein by 2 Da mass increase after reduction of a disulfide bond. For example, the sum of *m*/*z* of the two new ions at *m*/*z* 2,383.00 and 2,487.23 in Fig. 12b is 4,870.23, and the *m*/*z* of the disulfide-linked tryptic peptide is 4,863.23 in Fig. 12a. Then 4,870.23 minus 4,863.23 equals 7, and there should be three disulfide bonds in human insulin. Unfortunately, the number of cysteines in each new peptide could not be determined. This problem can be resolved by using the peptide-CHCA adduct method as shown in Fig. 12c. Molecular mass shifts in increments of 189 (molecular mass of CHCA) × n Da corresponded to n CHCA molecule additions to n cysteine sites ($n = 1, 2, \ldots, n$). The measured monoisotopic mass peak at *m*/*z* 2,383.01 migrated to 3,139.23, and the resulting mass shift of 756 Da indicates four cysteine residues. A further monoisotopic mass peak at *m*/*z* 2,487.23 is shifted by 378 Da, corresponding to two cysteine residues. The results are consistent with the known structure.

Another typical example is represented in Fig. 12d–f. The spectra shown in Fig. 12e, f both include an ion at *m*/*z* 2,337, which is shifted to 2,904.28 in Fig. 12f, indicating three cysteine residues in this fragment. Compared to Fig. 12e, the new peaks at *m*/*z* 439.17, 524.20, 1,025.47, 1,125.43, and 1,457.70 in Fig. 12f indicate

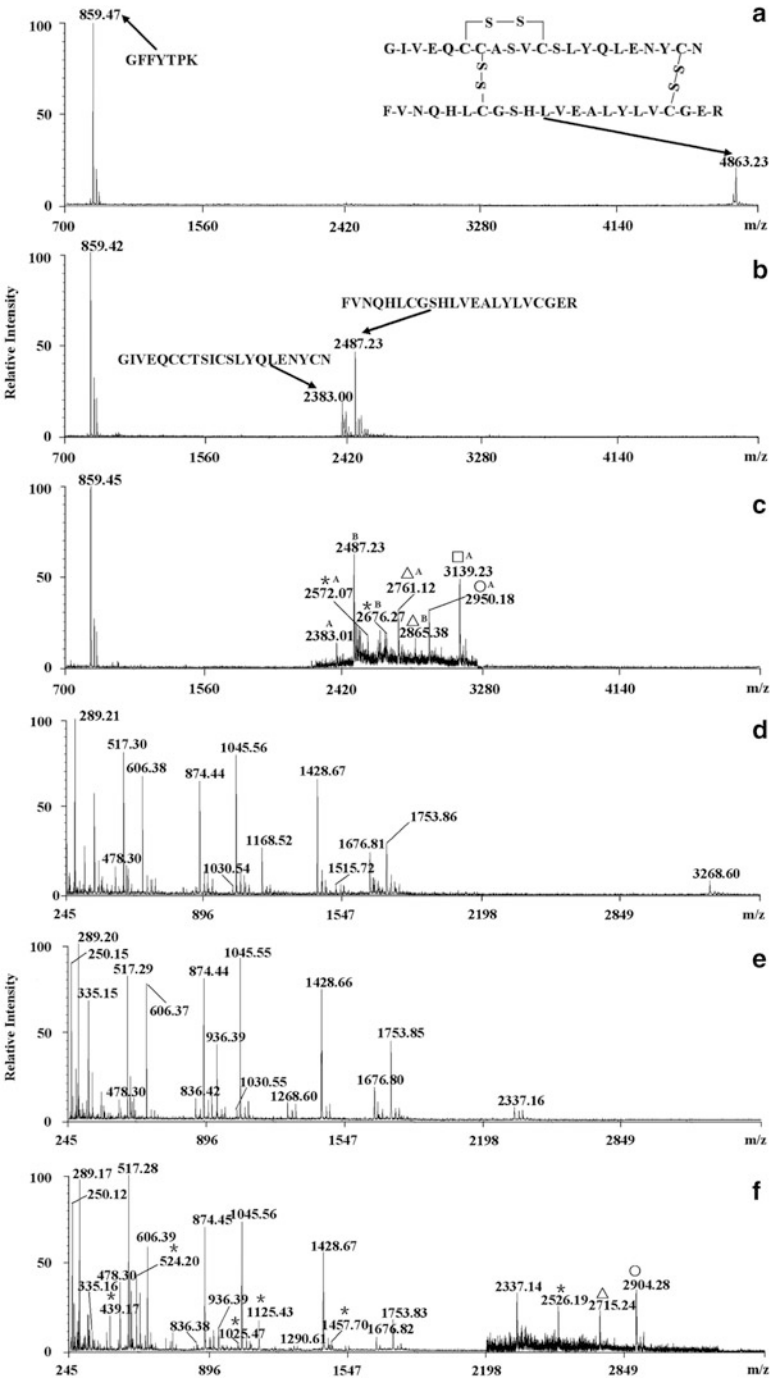


Fig. 12 MALDI mass spectra of insulin (**a**–**c**) and lysozyme (**d**–**e**) tryptic digests using (**a**, **d**) CHCA as matrix, (**b**, **e**) obtained after on-target reduction by DTT using DHB as matrix, and (**c**, **f**) obtained after on-target reduction by DTT using CHCA as matrix in alkaline conditions. The *open*

that there is only one cysteine residue in each fragment. The observed values are in agreement with those calculated and summarized in Table 4. This method is expected to be applied to unknown species. As for another method mentioned in this section, it has been exploited for analysis of two protein biomarkers from bacterial cell lysates of *E. coli* [113].

5 Other Approaches

5.1 Cyanylation-Based Methods

The cyanylation of sulfhydryl groups was first investigated by Jacobson et al. [125] using 2-nitro-5-thiocyanobenzoic acid (NTCB) under mildly alkaline conditions. Although this approach had been used with success, it suffered from serious drawbacks. For example, a main side reaction, β-elimination of HSCN from the S-cyanocysteine group, had been reported [126]. An alternative method for determining disulfide linkages is partial reduction of a peptide/protein followed by cyanylation of nascent sulfhydryls with the reagent 1-cyano-4-(dimethylamino) pyridinium tetrafluoroborate (CDAP) at acidic pH [127], essentially eliminating the potential for disulfide scrambling and generating few side reactions. The cyanylated analyte was subjected to specific chemical cleavage at the modified cysteinyl residues in aqueous ammonia. Especially for a protein containing multiple disulfide bonds, the cleaved peptides, which may be linked by remaining disulfide bonds, were then completely reduced and analyzed by MALDI-MS. Subsequent cleavage at the modified cysteinyl residues followed by complete reduction of the remaining disulfide bridges and MS analysis allowed the recognition of the initially reduced disulfide bonds. An overview of the cyanylation and cleavage reactions is illustrated in Scheme 5 [129, 130]. From Scheme 5 it is clearly seen that CDAP selectively cyanylates free cysteines and reduces cysteine residues under acidic conditions, resulting in molecular mass shifts in increments of 25 Da, corresponding to a single cyanylated species; 50 Da corresponding to a doubly cyanylated species, etc. Subsequent basic cleavage of the peptide bonds at the N-terminal side of the modified cysteinyl residues involves nucleophilic attack to the carbonyl carbon of the amide followed by a concerted cyclization. Then, an amino-terminal peptide and 2-iminothiazolidine-4-carboxyl peptides (itz-peptides) are produced, providing direct information about the location of sulfhydryl groups.

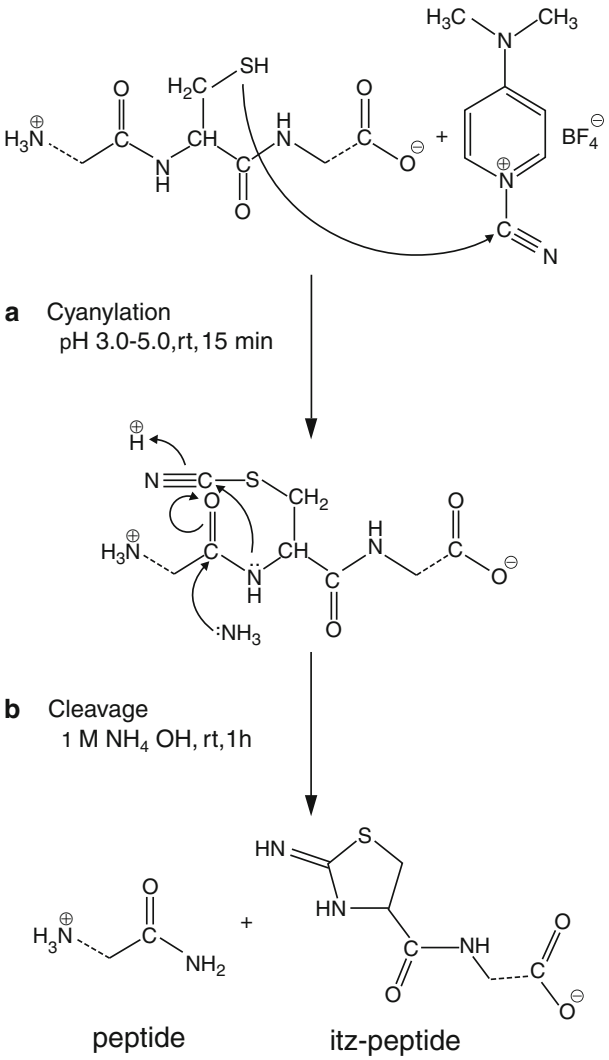
Fig. 12 (continued) circle indicates the ion peaks of peptide-CHCA 1:3 adducts. The ion peak of peptide-CHCA 1:4 adduct is marked with an *open square*. Other annotation is performed similar to Fig. 11. To make a clear annotation, signals in the mass range 2,250–3,250 in panel **c** and 2,220–3,280 in panel **f** are enhanced by three times and six times, respectively (adapted with permission from [112]. © 2009 American Society for Mass Spectrometry. Published by Elsevier Inc.)

Table 4 Disulfide-linked tryptic peptides generated from lysozyme before and after chemical reduction and CHCA adducts of cysteine-containing peptide (adapted with permission from [112]. © 2009 American Society for Mass Spectrometry. Published by Elsevier Inc.)

Peptide position	Peptide sequence	Measured mass (MH$^+$)[a]	Calculated mass (MH$^+$)[a]
Cys 6-cys 127	GCR CELAAAMK[b]	1,168.52	1,168.53
Cys 30-cys 115	CK GYSLGNWVCAAK[b]	1,515.72	1,515.71
Cys 64-cys 80	WWCNDGR NLCNIPCSALLSSDITASVNCAK[b]	3268.60	3268.56
115–116	CK[c]	250.15	250.12
126–128	GCR[c]	335.15	335.15
6–13	CELAAAMK[c]	836.42	836.40
62–68	WWCNDGR[c]	936.39	936.38
22–33	GYSLGNWVCAAK[c]	1,268.60	1,268.61
74–96	NLCNIPCSALLSSDITASVNCAK[c]	2,337.16	2,337.12
115–116	CK + 1 CHCA[d]	439.17	439.16
126–128	GCR + 1 CHCA[d]	524.20	524.19
6–13	CELAAAMK + 1 CHCA[d]	1,025.47	1,025.44
62–68	WWCNDGR + 1 CHCA[d]	1,125.43	1,125.42
22–33	GYSLGNWVCAAK + 1 CHCA[d]	1,457.70	1,457.65
74–96	NLCNIPCSALLSSDITASVNCAK + 1 CHCA[d]	2,526.19	2,526.16
74–96	NLCNIPCSALLSSDITASVNCAK + 2 CHCA[d]	2,715.24	2,715.20
74–96	NLCNIPCSALLSSDITASVNCAK + 3 CHCA[d]	2,904.28	2,904.24

[a]Monoisotopic mass
[b]Disulfide-linked tryptic peptides
[c]The peptides containing cysteine residues derived from reducing disulfide-linked tryptic peptides
[d]Peptide-matrix adducts

Systematic investigations have been carried out to optimize cleavage of cyanylated cystinyl proteins and improve the detection of the resulting cleavage products [127–129, 131]. It was found that CDAP had a maximum labeling efficiency at pH 5.0. In addition, its reactivity was not affected by excipients, salts, or denaturants [129]. In contrast with the commonly used conditions, low temperature (2 °C instead of room temperature), short reaction time (10 min instead of 60 min), and high concentration of ammonia (5 M instead of 1 M) greatly increased the cleavage product yields and minimized side reactions [128]. Further, with methylamine instead of ammonia as the nucleophile, the cleavage reaction was



Scheme 5 Reaction mechanism between cysteine residue and CDAP. (*A*) Cyanylation of sulfhydryl group by CDAP and (*B*) specific peptide bond cleavage catalyzed by ammonia (adapted with permission from [128]. © 2008 Humana Press, Totowa, NJ)

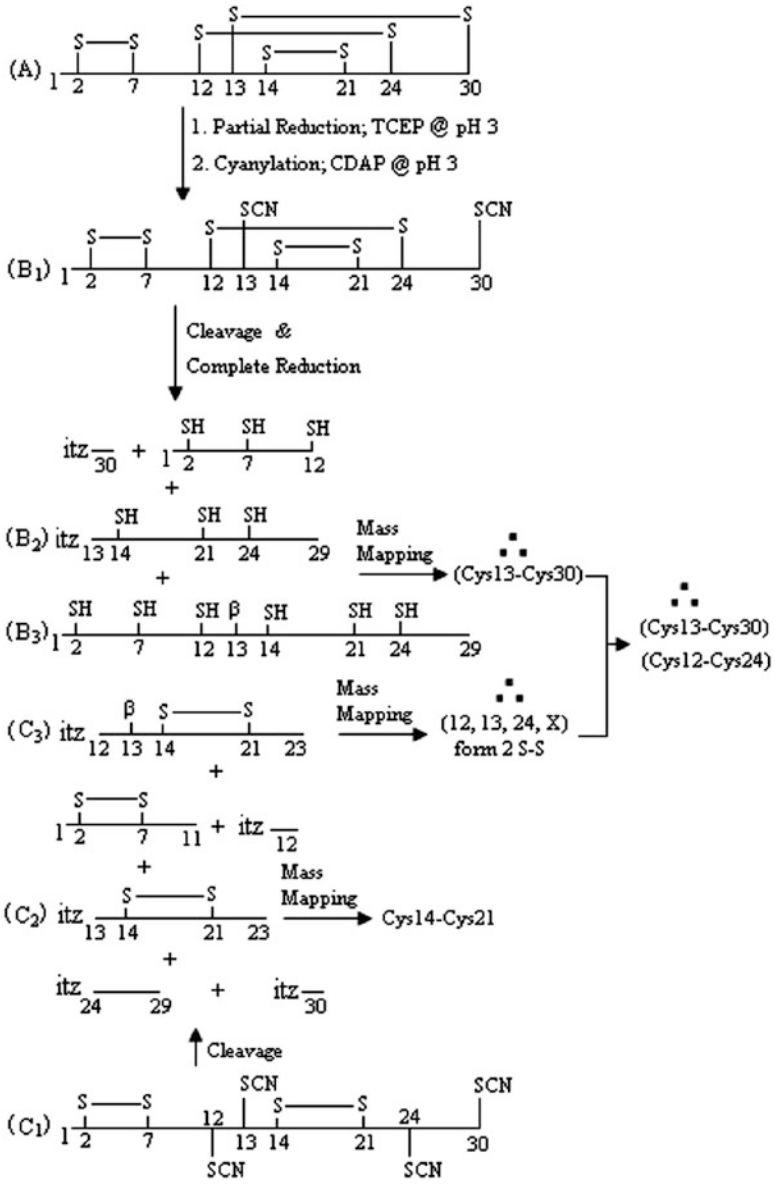
much faster. Interestingly, cleavage products by pairs of mass spectral peaks separated by 14 Da were obtained by the concurrent use of homologous nucleophiles such as ammonia and methylamine. This strategy facilitated recognition and identification of structurally diagnostic species [128].

With the cyanylation-based methodology, the identification of both free cysteines and disulfide bond linkages in proteins by MALDI MS was reported [132–134]. It was demonstrated that this approach was simple and effective.

Furthermore, the proteins contained closely spaced or adjacent cysteines, whereas conventional approaches often failed to profile these [135, 136]. Sillucin, for example, is a highly knotted antimicrobial peptide, in which four disulfide bonds involving three adjacent cysteines are present. In fact, four major components, consisting of the intact peptide, singly, doubly, and completely reduced and cyanylated isoforms of sillucin, were detected by HPLC separation after partial reduction and cyanylation of sillucin (data not shown). The singly, doubly reduced, and cyanylated species of sillucin are useful to deduce the disulfide linkages in sillucin, the analytical procedures of which are described in Scheme 6 [136]. As shown in Scheme 6, the cleavage followed by complete reduction of B_1 will produce four compounds. However, only two fragments (B_2 and B_3) were observed in the MALDI MS spectrum (data not shown). The disulfide bond between Cys13 and Cys30 was deduced from the mass-to-charge ratios of B_2 and B_3. Likewise, two of the cleavage products (before complete reduction) of the doubly reduced and cyanylated sillucin, i.e., C_2 and C_3 in Scheme 6, were detected by MALDI MS (data not shown). The disulfide bond between Cys14 and Cys 21 was directly indicated by the presence of C_2. According to the *m/z* of C_3 (data not shown), Cys12, Cys13, Cys24, and CysX must have formed two disulfide bonds. Considering the existence of Cys13–Cys30 deduced above, X must be 30. Hence the third disulfide bond between Cys12 and Cys24 was deduced. The Cys2–Cys7 disulfide bridge was deduced by default. In sum, the linkages of four disulfide bonds in sillucin were successfully identified as Cys2–Cys7, Cys12–Cys24, Cys13–Cys30, and Cys14–Cys21 by the cyanylation-based method.

With the aim of facilitating data processing and interpretation for a protein containing more than four cysteines by cyanylation-based methodology, the concept of "negative signature mass algorithm" (NSMA) [138] was introduced. The NSMA did not directly determine disulfide linkages but rather eliminated the existence of theoretically possible linkages by inputting an amino acid sequence and the obtained mass spectral data. Finally, only one disulfide structure can be constructed. In addition to convenience and automation, the greatest advantage of the NSMA is that it does not require physical separation nor determination of the state of reduction of partially reduced and cyanylated protein isoforms. The capability of the NSMA approach to data interpretation has been demonstrated by analyzing Ribonuclease A containing 8 cysteines in the form of 4 disulfide bonds [138] and the extracellular domain of the transforming growth factor β type II receptor containing 12 cysteines in the form of 6 disulfide bonds [137].

Compared to the conventional methods, the cyanylation-based method offers several advantages such as the elimination of the possibility for disulfide scrambling and the determination of disulfide bonds in proteins containing adjacent or closely spaced cysteines. However, the CN-based method has the potential drawback that occasionally some CN-induced cleavage fragments may not be detected. The "signature set" concept [139] was introduced to show that the correct disulfide structure can be deduced as long as a small, critical fraction of CN-induced cleavage fragments was obtained and matched with one of its unique "signature sets". The signature set model was analyzed in three different ways. First, some mathematical proofs about signature sets and their properties were provided.



Scheme 6 Overview of chemical reactions involved in the disulfide mass mapping of a singly reduced and cyanylated isoform (B_1) of sillucin (A) as well as a doubly reduced and cyanylated isoform (C_1). β is designated to the β-elimination of HSCN from the S-cyanocysteine group. X could be 2, 7, or 30 (adapted with permission from [137]. © 2001 American Chemical Society)

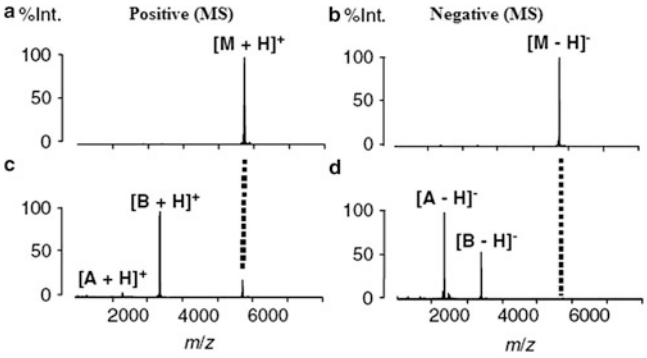
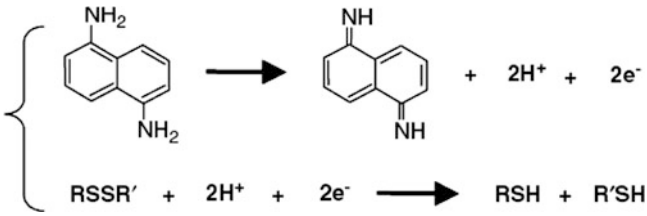


Fig. 13 Mass spectra of bovine insulin by using DHB (**a**, **b**) and 1,5-DAN (**c**, **d**) as matrices in positive and negative ion modes, respectively. Annotation is performed similar to Fig. 1 (adapted with permission from [148]. © 2005 John Wiley & Sons, Ltd.)



Scheme 7 Reduction reaction of disulfide bonds by 1,5-DAN (adapted with permission from [148]. © 2005 John Wiley & Sons, Ltd.)

Second, a computational analysis of all possible signature sets for 4-disulfide proteins was performed in order to identify patterns in the composition of signature sets. Finally, an experimental verification of the signature set model was provided. It was shown that the disulfide structure could be unequivocally determined, even when several CN-induced cleavage fragments were not detected [139]. Thus, the concept of signature sets made the cyanylation-based disulfide mass mapping methodology more robust.

5.2 Novel Reduction or Oxidation Methods

It is well known that with methods for analyzing disulfide-containing proteins/peptides, DTT, TCEP, or ME can be used to reduce disulfide bonds [140–143]. The total number of disulfide bonds in a protein/peptide can be deduced by mass increase after reduction of disulfide bonds. Recently, 1,5-diaminonaphthalene (1,5-DAN), a basic matrix previously used for gangliosides in negative ion mode [144, 145], was found to have the capability of reducing disulfide bonds in the laser plume [146, 147]. The mass spectra of bovine insulin in Fig. 13 [146] indicates the reductive property of 1,5-DAN as a matrix. As shown in Fig. 13a, b, the intact

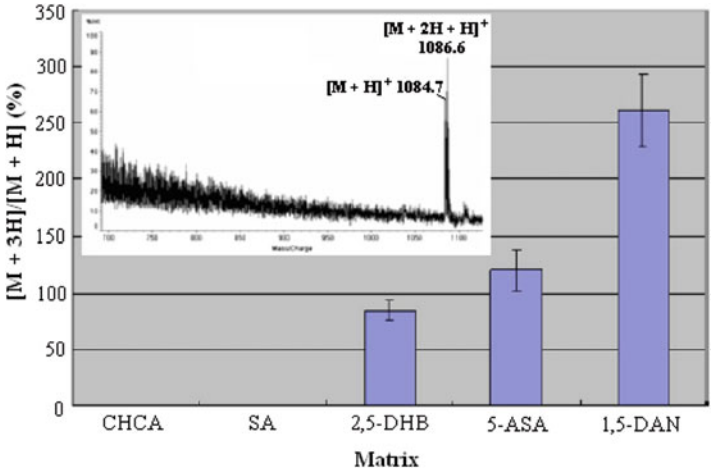
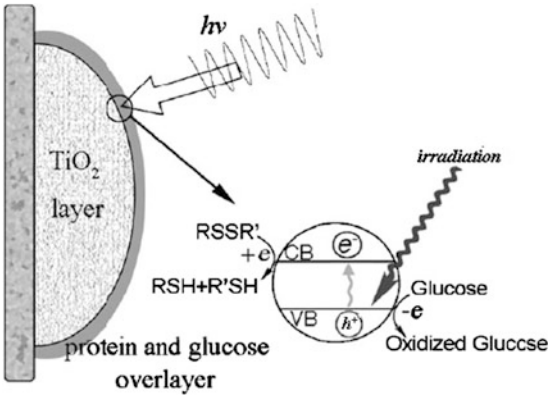


Fig. 14 Diagram of reductive ability of [Arg^8]-vasopressin (*m/z* 1,084.2) containing a disulfide bond for several MALDI matrices. The *inset* indicates the MALDI mass spectrum of vasopressin using 5-ASA as matrix (adapted with permission from [150]. © 2010 American Society for Mass Spectrometry)

insulin molecule is detected using DHB as a matrix in both positive and negative ion modes. Interestingly, under identical experiment conditions, two fragments derived from insulin A- and B-chains are mainly detected by using 1,5-DAN as a matrix (Fig. 13c, d). In addition, both the A- and B-chain cleavages also occurred when two isomeric forms of 1,5-DAN – 1,8-DAN and 2,3-DAN – were used as matrices (data not shown). Thus, the reductive property of 1,5-DAN matrix was further supported by those results. A proposed reaction mechanism between a disulfide-containing protein/peptide and 1,5-DAN is given in Scheme 7 [146]. From Scheme 7 it is clear that 1,5-DAN functions as a proton donor.

Furthermore, in a detailed study of the reducing matrix, it was found that 1,5-DAN dramatically enhanced ISD fragmentation [74]. Thus, one must exercise extreme caution during the interpretation of the obtained data. Using a combination of the two characteristics of 1,5-DAN, toxins from crude venoms were identified successfully [147]. The number of disulfide bridges in toxins was counted owing to their partial in-plume reduction by this particular matrix. In the meantime, novel conotoxin sequences were found by ISD using 1,5-DAN. This approach had also proven to be useful for an easy de novo sequencing of higher molecular weight (>6,500 Da) peptides from 1 pmol of material, where high sequence coverage was obtained from the c-ion series. However, 1,5-DAN is a suspected carcinogen and endocrine toxicant [148]. Subsequently, Takayama et al. [149] reported that 5-aminosalicylic acid (5-ASA) with less toxicity was suitable as a novel matrix for the MALDI ISD of peptides including phosphorylated peptides. The hydrogen-donating ability of 5-ASA together with several other matrices was evaluated by using the peak abundance of a reduction product $[M + 2H + H]^+$ compared to that of non-reduced protonated molecule $[M + H]^+$ of vasopressin containing a disulfide bond. The inset spectrum in Fig. 14 exhibits an abundant peak at *m/z* 1,086.6



Scheme 8 Schematic representation of the in-source photocatalytic reduction of disulfide bond with the assistance of glucose on a mesoporous TiO_2-modified target plate (adapted with permission from [152]. © 2008 The Royal Society of Chemistry)

corresponding to reduced [Arg^8]-vasopressin. Thus, the chemical 5-ASA can also be classified as a reductive matrix although with a lower reductive ability compared to 1,5-DAN. CHCA and SA did not show any reductive or ISD reactions for [Arg^8]-vasopressin. The order of hydrogen-donating ability is found to be: 1,5-DAN > 5-ASA > 2,5-DHB > SA ≈ CHCA, as depicted in Fig. 14. Interestingly, fragment ions originating from ISD were not observed in the spectrum of [Arg^8]-vasopressin. Conversely, the MALDI-ISD spectra of other tested peptides with 5-ASA showed useful fragment ions for sequence analysis without peak broadening or any interference peak.

As an alternative, photo-catalytic reaction using a TiO_2-modified plate was developed to carry out the in-source reduction of disulfide bonds during laser desorption ionization with the assistance of glucose as a sacrificial hole scavenger and proton donor [151]. As depicted in Scheme 8, under laser irradiation the TiO_2 nanoparticles absorb photons and electron-hole pairs are generated [150]. In this case, glucose is oxidized by the photo-induced valence band holes giving out electrons and protons, while the disulfide bonds are reduced by capturing the excited electrons from the conduction band and protons. The case of human insulin was studied as an example of the efficiency of present photo-reductive protocol (data not shown). To detect the role of glucose, citric acid instead of glucose was selected as a sacrificial hole scavenger, where a minor peak from B-chain was observed. Therefore, glucose played an important role in enhancing the photocatalytic reduction of disulfide bonds. This strategy offered both benefits and disadvantages when used for analysis of protein/peptide containing disulfide bonds. On the one hand the high redox capacity provided allowed effective determination of the reduction products of any given molecule by mass spectrometry. However, on the other hand the resolution was poor.

It is well known that performic oxidation is extensively used to cleave disulfide bonds [152–154]. Thiol groups of cysteines were transformed into sulfonic acids, which resulted in a 48-Da mass shift per cysteine residue. In comparison with traditional reduction/alkylation, this method can significantly enhance the sequence coverage. For example, this advantage was demonstrated by applying it to BSA

[152]. Due to its efficiency and many advantages, performic oxidation deserves more attention and research on degradation products formed during performic oxidation of peptides and proteins has already been carried out [155]. However, it suffered from some drawbacks such as long reaction times ranging from 2 to 4 h and the requirement of several steps to remove the performic acid reagent prior to MS analysis [152, 154]. An on-target oxidation method for analysis of disulfide-containing peptides was also reported [156]. However, S/N ratios of the oxidized peptide ions and the overall amino acid sequence coverage decreased because of ion suppression effects.

Recently, on-target oxidation using performic acid vapor was successfully applied to the disulfide analysis of intact bovine insulin and bovine ribonuclease A proteolytic digests [157], performed by placing 40 μL of a 1:1 (v/v) mixture of acetone:performic acid directly on the MALDI target in an area that does not contain sample. The target was placed in a −20 °C freezer for 10–60 min followed by room temperature (24 °C) oxidation for an additional 60 min. This strategy was preferred over solution phase oxidation methods because of less sample handling, increased oxidation throughput, and reduced ion suppression effects. In addition, the high efficiency of this method was demonstrated by simultaneous oxidation of multiple disulfide-linked peptides.

6 Conclusion and Outlook

From the above description it can be seen that the use of MALDI, particularly MALDI-TOF, plays an important role in assignment of disulfide bonds in proteins. The MALDI ISD and the adducts between cysteine sulfhydryl group and matrix are interesting tools, as we have shown in this review, providing rapidity, usefulness, and robustness. Several probable mechanisms of the prompt fragmentation of disulfide-containing proteins/peptides, such as laser-induced reduction, a mechanism similar to the hot hydrogen atom model in ECD, were proposed. In the case of forming the adducts of matrix and sulfhydryl-containing protein/peptide, there are two distinct mechanisms which need to be further investigated. All the investigations of peptides or proteins described here have established the potential analytical utility of the various methods mentioned above. Although MALDI is now well established for determining disulfide-containing peptides, there is still room for development, particularly in the analysis of very small amounts of material. Over the next few years, work is expected to see considerable improvements in this area, together with many more applications of MALDI MS to more complex molecules.

Acknowledgments This work was supported by China Postdoctoral Science Foundation funded project (No. 2012M511355), the National Natural Science Foundation of China (No. 21175127, 21175055), and Jilin Province Science and Technology Department (YYZX201131). We are very much indebted to all scientists quoted in the references.

References

1. Piatek R, Bruździak P, Wojciechowski M, Zalewska-Piatek B, Kur J (2010) Biochemistry 49:1460
2. Wedemeyer WJ, Welker E, Narayan M, Scheraga HA (2000) Biochemistry 39:4207
3. Reinders J, Sickmann A (2007) Biomol Eng 24:169
4. Watkins M, Han TS, Yoshikami D, Olivera BM, Bulaj G (2008) J Am Chem Soc 130:14280
5. von Ossowski L, Tossavainen H, von Ossowski I, Cai C, Aitio O, Fredriksson K, Permi P, Annila A, Keinänen K (2006) Biochemistry 45:5567
6. Collins ES, Wirmer J, Hirai K, Tachibana H, Segawa S, Dobson CM, Schwalbe H (2005) Chembiochem 6:1619
7. Chait BT, Field FH (1986) Biochem Biophys Res Commun 134:420
8. el Bouyoussfi M, Laus G, Verheyden P, Wyns L, Tourwe D, Van Binst G (1997) J Pept Res 49:336
9. Das PR, Pramanik BN (1998) Mol Biotechnol 9:141
10. Vaccaro AM, Salvioli R, Barca A, Tatti M, Ciaffoni F, Maras B, Siciliano R, Zappacosta F, Amoresano A, Pucci P (1995) J Biol Chem 270:9953
11. Poerio E, Caporale C, Carrano L, Pucci P, Buonocore V (1991) Eur J Biochem 199:595
12. Zhang M, Kaltashov IA (2006) Anal Chem 78:4820
13. Thakur SS, Balaram P (2008) J Am Soc Mass Spectrom 19:358
14. Clauser KR, Baker P, Burlingame AL (1999) Anal Chem 71:2871
15. Kim H, Beauchamp JL (2008) J Am Chem Soc 130:1245
16. Mihalca R, van der Burgt YE, Heck AJ, Heeren RM (2007) J Mass Spectrom 42:450
17. Kim HI, Beauchamp JL (2009) J Am Soc Mass Spectrom 20:157
18. Gunawardena HP, O'Hair RA, McLuckey SA (2006) J Proteome Res 5:2087
19. Zubarev RA (2003) Mass Spectrom Rev 22:57
20. Zubarev RA, Kruger NA, Fridriksson EK, Lewis MA, Horn DM, Carpenter BK, McLafferty FW (1999) J Am Chem Soc 121:2857
21. Choi S, Jeong J, Na S, Lee HS, Kim HY, Lee KJ, Paek E (2010) J Proteome Res 9:626
22. Huang SY, Wen CW, Li DT, Hsu JL, Chen C, Shi FK, Lin YY (2008) Anal Chem 80:9135
23. Craig R, Krokhin O, Wilkins J, Beavis RC (2003) J Proteome Res 2:657
24. Xu H, Zhang L, Freitas MA (2008) J Proteome Res 7:138
25. Karas M, Hillenkamp F (1988) Anal Chem 60:2299
26. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T (1988) Rapid Commun Mass Spectrom 2:151
27. Schoneich C (2005) Mass Spectrom Rev 24:701
28. Akashi S (2006) Med Res Rev 26:339
29. Evans-Nguyen KM, Tao SC, Zhu H, Cotter RJ (2008) Anal Chem 80:1448
30. Reyzer ML, Chaurand P, Angel PM, Caprioli RM (2010) Methods Mol Biol 656:285
31. Amini A, Dormady SJ, Riggs L, Regnier FE (2000) J Chromatogr A 894:345
32. Gatlin-Bunai CL, Cazares LH, Cooke WE, Semmes OJ, Malyarenko DI (2007) J Proteome Res 6:4517
33. van Kampen JJ, Burgers PC, de Groot R, Gruters RA, Luider TM (2011) Mass Spectrom Rev 30:101
34. Li SS, Claeson P (2003) Phytochemistry 63:249
35. John H, Forssmann WG (2001) Rapid Commun Mass Spectrom 15:1222
36. Trachsel C, Kämpfer U, Bechtold R, Schaller J, Schürch S (2009) Anal Biochem 390:103
37. Zhang W, Marzilli LA, Rouse JC, Czupryn MJ (2002) Anal Biochem 311:1
38. Matsushita N, Miyashita M, Sakai A, Nakagawa Y, Miyagawa H (2007) Toxicon 50:861
39. Gorman JJ, Wallis TP, Pitt JJ (2002) Mass Spectrom Rev 21:183
40. Wallis TP, Huang CY, Nimkar SB, Young PR, Gorman JJ (2004) J Biol Chem 279:20729
41. Ezure T, Suzuki T, Shikata M, Ito M, Ando E, Nishimura O, Tsunasawa S (2007) Proteomics 7:4424

42. Connors LH, Jiang Y, Budnik M, Théberge R, Prokaeva T, Bodi KL, Seldin DC, Costello CE, Skinner M (2007) Biochemistry 46:14259
43. Matsunaga H, Sadakane Y, Haginaka J (2004) Anal Biochem 331:358
44. Kalkhof S, Haehn S, Ihling C, Paulsson M, Smyth N, Sinz A (2008) Rapid Commun Mass Spectrom 22:1933
45. López-Ferrer D, Heibeck TH, Petritis K, Hixson KK, Qian W, Monroe ME, Mayampurath A, Moore RJ, Belov ME, Camp DG 2nd, Smith RD (2008) J Proteome Res 7:3860
46. Zhang HM, McLoughlin SM, Frausto SD, Tang H, Emmett MR, Marshall AG (2010) Anal Chem 82:1450
47. Korfali N, Wilkie GS, Swanson SK, Srsen V, Batrakou DG, Fairley EA, Malik P, Zuleger N, Goncharevich A, de Las Heras J, Kelly DA, Kerr AR, Florens L, Schirmer EC (2010) Mol Cell Proteomics 9:2571
48. Huang HZ, Nichols A, Liu D (2009) Anal Chem 81:1686
49. Koehn H, Clerens S, Deb-Choudhury S, Morton JD, Dyer JM, Plowman JE (2010) J Proteome Res 9:2920
50. Zhang X, Chien EY, Chalmers MJ, Pascal BD, Gatchalian J, Stevens RC, Griffin PR (2010) Anal Chem 82:1100
51. Ritter A, Ubertini M, Romac S, Gaillard F, Delage L, Mann A, Cock JM, Tonon T, Correa JA, Potin P (2010) Proteomics 10:2074
52. Liu M, Hou J, Huang L, Huang X, Heibeck TH, Zhao R, Pasa-Tolic L, Smith RD, Li Y, Fu K, Zhang Z, Hinrichs SH, Ding SJ (2010) Anal Chem 82:7160
53. Welton JL, Khanna S, Giles PJ, Brennan P, Brewis IA, Staffurth J, Mason MD, Clayton A (2010) Mol Cell Proteomics 9:1324
54. Chi F, Yang P, Han F, Jing Y, Shen S (2010) Proteomics 10:1861
55. Hauser NJ, Basile F (2008) J Proteome Res 7:1012
56. Siddiqui KS, Poljak A, Guilhaus M, Feller G, D'Amico S, Gerday C, Cavicchioli R (2005) J Bacteriol 187:6206
57. Chong JM, Speicher DW (2001) J Biol Chem 276:5804
58. Juárez P, Sanz L, Calvete JJ (2004) Proteomics 4:327
59. Ren D, Julka S, Inerowicz HD, Regnier FE (2004) Anal Chem 76:4522
60. Mandal AK, Ramasamy MR, Sabareesh V, Openshaw ME, Krishnan KS, Balaram P (2007) J Am Soc Mass Spectrom 18:1396
61. Horn NA, Hurst GB, Mayasundari A, Whittemore NA, Serpersu EH, Peterson CB (2004) J Biol Chem 279:35867
62. Turko IV, Sechi S (2007) Methods Mol Biol 359:1
63. Patterson J, Ford K, Cassin A, Natera S, Bacic A (2007) Plant Physiol 144:1612
64. Gygi SP, Rist B, Gerber SA, Turecek F, Gelb MH, Aebersold R (1999) Nat Biotechnol 17:994
65. Cahill MA, Wozny W, Schwall G, Schroer K, Hölzer K, Poznanovic S, Hunzinger C, Vogt JA, Stegmann W, Matthies H, Schrattenholz A (2003) Rapid Commun Mass Spectrom 17:1283
66. Spengler B, Kirsch D, Kaufmann R (1991) Rapid Commun Mass Spectrom 5:198
67. Spengler B, Kirsch D, Kaufmann R, Jaeger E (1992) Rapid Commun Mass Spectrom 6:105
68. Kaufmann R, Kirsch D, Spengler B (1994) Int J Mass Spectrom Ion Process 131:355
69. Chaurand P, Luetzenkirchen F, Spengler B (1999) J Am Soc Mass Spectrom 10:91
70. Sedo O, Novotná K, Havel J (2004) Rapid Commun Mass Spectrom 18:339
71. Brown RS, Lennon JJ (1995) Anal Chem 67:3990
72. Brown RS, Carr BL, John JJ (1996) J Am Soc Mass Spectrom 7:225
73. Yang H, Yu Y, Song F, Liu S (2011) J Am Soc Mass Spectrom 22:845
74. Demeure K, Quinton L, Gabelica V, De Pauw E (2007) Anal Chem 79:8678
75. Patterson SD, Katta V (1994) Anal Chem 66:3727
76. Crimmins DL, Saylor M, Rush J, Thoma RS (1995) Anal Biochem 226:355
77. Qiu X, Cui M, Li H, Liu Z, Liu S (2007) Rapid Commun Mass Spectrom 21:3520

78. Beavis RC, Chaitt BT (1989) Rapid Commun Mass Spectrom 3:233
79. Zhou J, Ensa W, Poppe-Schriemer N, Standing KG, Westmore JB (1993) Int J Mass Spectrom Ion Process 126:115
80. Bean MF, Carr SA (1992) Anal Biochem 201:216
81. Stults JT, Bourell JH, Canova-Davis E, Ling VT, Laramee GR, Winslow JW, Griffin PR, Rinderknecht E, Vandlen RL (1990) Biomed Environ Mass Spectrom 19:655
82. Huwiler KG, Mosher DF, Vestling MM (2003) J Biomol Tech 14:289
83. Schnaible V, Wefing S, Resemann A, Suckau D, Bücker A, Wolf-Kümmeth S, Hoffmann D (2002) Anal Chem 74:4980
84. Solouki T, Emmett MR, Guan S, Marshall AG (1997) Anal Chem 69:1163
85. Fagerquist CK (2004) Rapid Commun Mass Spectrom 18:685
86. Zhao L, Almaraz RT, Xiang F, Hedrick JL, Franza AH (2009) J Am Soc Mass Spectrom 20:1603
87. Zubarev RA, Kelleher NL, McLafferty FW (1998) J Am Chem Soc 120:3265
88. Zubarev RA (2004) Curr Opin Biotechnol 15:12
89. Cooper HJ, Håkansson K, Marshall AG (2005) Mass Spectrom Rev 24:201
90. Zubarev RA, Horn DM, Fridriksson EK, Kelleher NL, Kruger NA, Lewis MA, Carpenter BK, McLafferty FW (2000) Anal Chem 72:563
91. Zubarev RA, Haselmann KF, Budnik B, Kjeldsen F, Jensen F (2002) Eur J Mass Spectrom 8:337
92. Tsybin YO, Håkansson P, Budnik BA, Haselmann KF, Kjeldsen F, Gorshkov M, Zubarev RA (2001) Rapid Commun Mass Spectrom 15:1849
93. Hakansson K, Cooper HJ, Emmett MR, Costello CE, Marshall AG, Nilsson CL (2001) Anal Chem 73:4530
94. Shi SD, Hemling ME, Carr SA, Horn DM, Lindh I, McLafferty FW (2001) Anal Chem 73:19
95. Kjeldsen F, Haselmann KF, Budnik BA, Sorensen ES, Zubarev RA (2003) Anal Chem 75:2355
96. Kleinnijenhuis AJ, Mihalca R, Heeren RMA, Heck AJR (2006) Int J Mass Spectrom 253:217
97. Mirgorodskaya OA, Haselmann KF, Kjeldsen F, Zubarev RA, Roepstorff P (2003) Eur J Mass Spectrom 9:139
98. al-Khalili A, Thomas R, Ehlerding A, Hellberg F, Geppert WD, Zhaunerchyk V, af Ugglas M, Larsson M, Uggerud E, Vedde J, Adlhart C, Semaniak J, Kamińska M, Zubarev RA, Kjeldsen F, Andersson PU, Österdahl M, Bednarska VA, Paál A (2004) J Chem Phys 121:5700
99. Anusiewicz I, Berdys-Kochanska J, Simons J (2005) J Phys Chem A 109:5801
100. Neff D, Simons J (2010) J Phys Chem A 114:1309
101. Knochenmuss R, Zenobi R (2003) Chem Rev 103:441
102. Karas M, Krüger R (2003) Chem Rev 103:427
103. Zenobi R, Knochenmuss R (1998) Mass Spectrom Rev 17:337
104. Köcher T, Engström Å, Zubarev RA (2005) Anal Chem 77:172
105. Beavis RC, Chait BT (1989) Rapid Commun Mass Spectrom 3:432
106. Hutchens TW, Nelson RW, Li CM, Yip TT (1992) J Chromatogr A 604:125
107. Huth-Fehre T, Gosine JN, Wu KJ, Becker CH (1992) Rapid Commun Mass Spectrom 6:209
108. Fitzgerald MC, Parr GR, Smith LM (1993) Anal Chem 65:3204
109. Spiess C, Happersberger HP, Glocker MO, Spiess E, Rippe K, Ehrmann M (1997) J Biol Chem 272:22125
110. Happersberger HP, Bantscheff M, Barbirz S, Glocker MO (2000) Methods Mol Biol 146:167
111. Yang H, Liu N, Qiu X, Liu S (2009) J Am Soc Mass Spectrom 20:2284
112. Hamdan M, Bordini E, Galvani M, Righetti PG (2001) Electrophoresis 22:633
113. Fagerquist CK, Garbus BR, Williams KE, Bates AH, Harden LA (2010) J Am Soc Mass Spectrom 21:819
114. Bordini E, Hamdan M, Righetti PG (2000) Rapid Commun Mass Spectrom 14:840
115. Friedman M, Cavins JF, Wall JS (1965) J Am Chem Soc 187:3672

116. Chiari M, Righetti PG, Negri A, Ceciliani F, Ronchi S (1992) Electrophoresis 13:882
117. Bordini E, Hamdan M, Righetti PG (1999) Rapid Commun Mass Spectrom 13:1818
118. Chiari M, Manzocchi A, Righetti PG (1990) J Chromatogr A 500:697
119. Hunter EPL, Lias SG (1998) J Phys Chem Ref Data 27:413
120. van der Horst MA, Arents JC, Kort R, Hellingwerf KJ (2007) Photochem Photobiol Sci 6:571
121. Hoff WD, Devreese B, Fokkens R, Nugteren-Roodzant IM, Van Beeumen J, Nibbering N, Hellingwerf KJ (1996) Biochemistry 35:1274
122. Baca M, Borgstahl GEO, Boissinot M, Burke PM, Williams DR, Slater KA, Getzoff ED (1994) Biochemistry 33:14369
123. Hoff WD, Düx P, Hård K, Devreese B, Nugteren-Roodzant IM, Crielaard W, Boelens R, Kaptein R, Van Beeumen J, Nibbering N, Hellingwerf KJ (1994) Biochemistry 33:13959
124. Imamoto Y, Ito T, Kataoka M, Tokunaga F (1995) FEBS Lett 374:157
125. Jacobson GR, Schaffer MH, Stark GR, Vanaman TC (1973) J Biol Chem 248:6583
126. Wu J, Gage DA, Watson JT (1996) Anal Biochem 235:161
127. Wu J, Watson JT (1997) Protein Sci 6:391
128. Gallegos-Perez JL, Rangel-Ordonez L, Bowman SR, Ngowe CO, Watson JT (2005) Anal Biochem 346:311
129. Pipes GD, Kosky AA, Abel J, Zhang Y, Treuheit MJ, Kleemann GR (2005) Pharm Res 22:1059
130. Wu J (2008) Methods Mol Biol 446:1
131. Wu J, Watson JT (1998) Anal Biochem 258:268
132. Schutte CG, Lemm T, Glombitza GJ, Sandhoff K (1998) Protein Sci 7:1039
133. Watson JT, Yang Y, Wu J (2001) J Mol Graph Model 19:119
134. Li X, Chou YT, Husain R, Watson JT (2004) Anal Biochem 331:130
135. Yang Y, Wu J, Watson JT (1998) J Am Chem Soc 120:5834
136. Qi J, Wu J, Somkuti GA, Watson JT (2001) Biochemistry 40:4531
137. Borges CR, Qi J, Wu W, Torng E, Hinck AP, Watson JT (2004) Anal Biochem 329:91
138. Qi J, Wu W, Borges CR, Hang D, Rupp M, Torng E, Watson JT (2003) J Am Soc Mass Spectrom 14:1032
139. Wu W, Huang W, Qi J, Chou YT, Torng E, Watson JT (2004) J Proteome Res 3:770
140. McAlpin CR, Cox CR, Matyi SA, Voorhees KJ (2010) Rapid Commun Mass Spectrom 24:11
141. Manabe T, Jin Y (2005) Electrophoresis 26:257
142. van Montfort BA, Schuurman-Wolters GK, Duurkens RH, Mensen R, Poolman B, Robillard GT (2001) J Biol Chem 276:12756
143. Pitt JJ, Da Silva E, Gorman JJ (2000) J Biol Chem 275:6469
144. Juhasz P, Costello CE (1992) J Am Soc Mass Spectrom 3:785
145. Stemmler EA, Hettich RL, Hurst GB, Buchanan MV (1993) Proceeding of 41st American Society of Mass Spectrometry (ASMS) Conference, San Francisco, 249a
146. Fukuyama Y, Iwamoto S, Tanaka K (2006) J Mass Spectrom 41:191
147. Quinton L, Demeure K, Dobson R, Gilles N, Gabelica V, De Pauw E (2007) J Proteome Res 6:3216
148. See the pollution information site. http://www.scorecard.org/chemical-profiles/
149. Sakakura M, Takayama M (2010) J Am Soc Mass Spectrom 21:979
150. Hennig H (1999) Coord Chem Rev 182:101
151. Qiao L, Bi H, Busnel JM, Liu B, Girault HH (2008) Chem Commun 47:6357
152. Dai J, Wang J, Zhang Y, Lu Z, Yang B, Li X, Cai Y, Qian X (2005) Anal Chem 77:7594
153. Samgina TY, Artemenko KA, Gorshkov VA, Poljakov NB, Lebedev AT (2008) J Am Soc Mass Spectrom 19:479
154. Matthiesen R, Bauw G, Welinder KG (2004) Anal Chem 76:6848
155. Dai J, Zhang Y, Wang J, Li X, Lu Z, Cai Y, Qian X (2005) Rapid Commun Mass Spectrom 19:1130
156. Lucas JE (2003) Masters Thesis, Texas A&M University, College Station, TX
157. Williams BJ, Russell WK, Russell DH (2010) J Mass Spectrom 45:157

Top Curr Chem (2013) 331: 117–142
DOI: 10.1007/128_2012_363
© Springer-Verlag Berlin Heidelberg 2012
Published online: 14 September 2012

MALDI In-Source Decay, from Sequencing to Imaging

Delphine Debois, Nicolas Smargiasso, Kevin Demeure, Daiki Asakawa, Tyler A. Zimmerman, Loïc Quinton, and Edwin De Pauw

Abstract Matrix-assisted laser desorption/ionization (MALDI) is now a mature method allowing the identification and, more challenging, the quantification of biopolymers (proteins, nucleic acids, glycans, etc). MALDI spectra show mostly intact singly charged ions. To obtain fragments, the activation of singly charged precursors is necessary, but not efficient above 3.5 kDa, thus making MALDI MS/MS difficult for large species. In-source decay (ISD) is a prompt fragmentation reaction that can be induced thermally or by radicals. As fragments are formed in the source, precursor ions cannot be selected; however, the technique is not limited by the mass of the analyzed compounds and pseudo MS3 can be performed on intense fragments. The discovery of new matrices that enhance the ISD yield, combined with the high sensitivity of MALDI mass spectrometers, and software development, opens new perspectives. We first review the mechanisms involved in the ISD processes, then discuss ISD applications like top-down sequencing and post-translational modifications (PTMs) studies, and finally review MALDI-ISD tissue imaging applications.

Keywords Imaging • In-source decay • MALDI • Post-translational modifications • Sequencing

Contents

D.Debois, N.Smargiasso, K.Demeure, D.Asakawa, T.A.Zimmerman, L.Quinton and E.De Pauw (✉)
Mass Spectrometry Laboratory, GIGA-R, Department of Chemistry
University of Liège, Allée de la Chimie 3, 4000 Liège, Belgium
e-mail: e.depauw@ulg.ac.be

Abbreviations

2-AA	Anthranilic acid
2-AB	2-Aminobenzamide
5-ASA	5-Aminosalicylic acid
ATT	6-Aza-2-thiothymine
BLAST	Basic Local Alignment Search Tool
CHCA	α-Cyanohydroxycinnamic acid
CID	Collision induced dissociation
1,5-DAN	1,5-Diaminonaphthalene
2,5-DHB	2,5-Dihydroxybenzoic acid
ECD	Electron capture dissociation
ED	Edman degradation
ETD	Electron transfer dissociation
(HP)LC	High performance liquid chromatography
ISD	In-source decay
MALDI	Matrix-assisted laser desorption/ionization
MS/MS	Tandem mass spectrometry
PMF	Peptide-mass fingerprinting
PSD	Post-source decay
PTM	Post-translational modification
TDS	Top-down sequencing

1 Introduction

The field of biopolymer analysis has evolved considerably over the past 10 years [1]. A major example is the explosion of “omics” techniques. Mass spectrometry has had a major impact resulting from the continuous development of MS-based methodologies, including instrumentation and comprehensive strategies for the identification and quantification of a large number of molecules during a single experiment [2–4]. The next challenge is their localization in biological samples with their simultaneous identification and quantification.

Matrix-assisted laser desorption/ionization (MALDI) played a large part in these developments [5]. In UV-MALDI, the energy of a laser pulse is dissipated by a UV-absorbing matrix containing the analytes and a dense gas phase is produced (plume) from which intact ions survive. The primary mechanisms are not yet fully understood but probably consist of mixed processes involving photo-excitation, collision-induced desolvation of clusters, and gas phase acid–base chemistry leading to charge transfer. The role of the MALDI matrix is crucial to enhance the ion signal, and besides "general purpose" matrices, specific matrices are adapted for various families of analytes [6]. Matrices have been classified as hot or cold according to the internal energy transferred to the analytes, and many papers describe methods to study the energetics and kinetics of the MALDI process [7]. Among important processes, gas phase reactions can increase the internal energy of ions by their exothermicity. In contrast, multiple gas phase collisions have a cooling effect [8]. The balance between these processes will govern the final energy state of the ion population, and therefore the fragmentation rate constants and the resulting mass spectra. If fragmentation occurs in the source, the ions will be detected at their expected *m*/*z* values. The decay of this population of ions is called in-source decay (ISD). When the fragmentation reaction is too slow to occur in the source but fast enough to occur before detection, the ions are called metastable and the decay of their population is called post-source decay (PSD) [9]. Dedicated methods of detection are required [10].

Two major strategies rely on mass spectrometry in biopolymer analysis – the classical bottom-up approach and the top-down strategy. In the "bottom-up" approach, biopolymers are first digested into oligomers. In the case of proteins, the sample preparation consists of the reduction of the disulfide bridges, followed by alkylation of the sulfhydryl groups and an enzymatic digestion. The resulting mix of peptides is then analyzed by MALDI. If the sample is complex, a 2D gel separation is performed. This method is known as PMF for "peptide-mass fingerprinting". It can be followed by MS/MS for sequencing the peptides using PSD to increase the identification score. Another mass spectrometric approach commonly employed to sequence proteins is "top-down" sequencing (TDS) [11]. This strategy generally relies on the direct fragmentation of the intact protein in the gas-phase, i.e., without any enzymatic digestion before mass spectrometry. The fact that the whole sequence is accessible to fragmentation is one of its major advantages. TDS is usually performed on multiply charged ions generated from electrospray ion sources. Fragmentations occur by collision induced dissociation (CID), by electrons as in electron capture dissociation (ECD) [6, 12, 13], or by electron transfer dissociation (ETD) [14]. Until recently, TDS was not intensively used with MALDI due to the fact that the produced ions are singly charged and MS/MS techniques available on MALDI mass spectrometers are efficient only for peptides up to *m*/*z* 3,500 [12]. Indeed, PSD and CID fragmentation processes rely on a slow increase of the ion internal energy by multiple collisions with gas molecules until the lower energy bonds are dissociated (mainly labile or peptide bonds). In addition, for higher masses, the energy deposited is not sufficient to induce dissociation. However, as the ion formation in ISD occurs before the extraction, fragments even from large proteins receive the full kinetic energy,

fly at their own velocity, and are detected directly at the true *m*/*z* value. The abundance of these fragments is usually low and ISD is generally seen as an unwanted side process and is therefore minimized. The ISD fragment nature depends on the MALDI matrix chosen and the chemistry leading to the fragmentation.

In this chapter we will show that ISD can be a powerful structural analysis tool when promoted by specific MALDI matrices presenting high ISD yield. The following review is divided into three parts:

1. The description of the mechanisms leading to ISD fragmentation during the MALDI process. We will describe the two main mechanisms of ISD fragmentation and the influence of the matrix used.
2. The description of typical MALDI-ISD spectra obtained with peptides and proteins. We will demonstrate the interest in using MALDI-ISD to sequence proteins as well as the limitations of this approach and the possible solutions. We will also review the application of ISD to biopolymers.
3. The application of MALDI-ISD to tissue imaging. We describe a new application of ISD to mass spectrometry imaging (MSI). We demonstrate that ISD can be performed directly on a tissue slice and that the correlation between a fragment and its localization leads to identification of different protein isoforms.

2 ISD Mechanisms

Ions produced by soft ionization methods like MALDI are generally closed valence shell ions, which mostly undergo rearrangement type fragmentations. These reactions are often characterized by a low activation enthalpy and a negative activation entropy. Under low energy activation conditions, they give access to low energy fragmentation channels. Fast thermal activation can be achieved when sufficient energy is transferred from the matrix (hard matrices) but fast reactions have also been shown to occur through radical chemistry. Recently, new matrices able to release hydrogen radicals upon laser irradiation have been developed showing an increase in the yield of ISD fragments [15].

Mechanistically, ISD is not yet fully understood. In MALDI, the prompt in-source fragmentation can have two origins – thermal or chemical activation. Once the laser energy is dissipated into the matrix, a "plume" containing intact ions is produced [6]. Several phenomena can then occur, such as the fragmentation of matrix clusters and gas phase acid–base chemistry (proton transfer). Different matrices have been classified as "hot" or "cold" according to the internal energy transferred to the analytes. Chemical activation by hydrogen radicals is another mechanism inducing a prompt "in source" fragmentation reaction [16, 17]. The matrix releases hydrogen radicals. Radicals can react with the analyte to produce radical species that will undergo specific fragmentation reactions. Using radical scavengers, it has been shown that the ability of matrices to release hydrogen radicals is directly linked to the efficiency of ISD fragmentation. The mass accuracy and resolution of reflectron

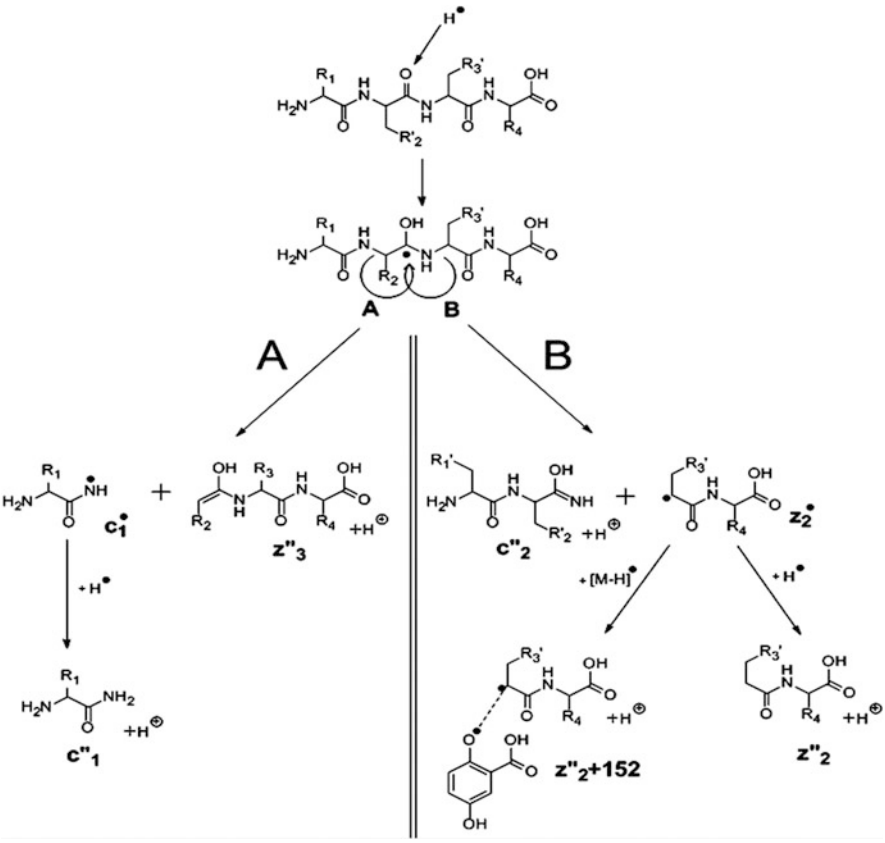
TOF instruments allow the unambiguous observation of the isotopic distribution of these fragments and clearly shows reduction reactions to occur.

The balance between thermal and radical activation will govern the chemical state (closed shell or radical) and the internal energy of the ion population, and therefore the fragmentation channels and their respective rate constants. Interestingly, these radical species are generated and induce fragmentations independently from the mass of the analytes. In the case of proteins, large *c*-ion and *z*-ion series detected up to *m*/*z* 10,000 allow for fast and efficient sequencing of the N- and C-termini extremities. The range $0 < m/z < 800$ is not accessible as intense matrix signals hide both C-terminal and N-terminal ISD fragments. This limitation can be overcome by the use of pseudo-MS^3 strategies.

A study on the internal energy build-up of benzylpyridinium ions formed in MALDI revealed that fragmentation can be due to thermal activation, and the effective temperatures were estimated [17]. It has been proposed that the exothermicity of acid–base reactions can lead to an internal energy increase [6]. The ion's initial velocity is directly proportional to the rate of dissipation of the plume [13, 17]. Thus, matrices characterized by a low exothermicity of acid–base reaction and by a high initial velocity are usually colder because the analytes have less probability of interacting with the rest of the plume.

ISD-suitable MALDI matrices revived interest in using the ISD for the de novo sequencing of peptides and proteins [18, 19]. Indeed, ISD experiments can be realized on all MALDI-TOF mass spectrometers, widely available in mass spectrometry laboratories and in proteomics facilities, just by using a suitable matrix. Sakakura et al. recently used the 5-aminosalicylic acid matrix (5-ASA) and demonstrated that this matrix is more efficient than the classical 2,5-DHB to produce *c*- and *z*-fragments [20]. However, this matrix remains less efficient than the 1,5-DAN matrix and rather difficult to use (low solubility) but possesses the advantage of being non-carcinogenic. Very recently, two additional molecules were shown to induce radical-induced fragmentation of peptides and proteins – 2-aminobenzoic acid (2-AA, also known as anthranilic acid) and 2-aminobenzamide (2-AB) [21].

The 1,5-DAN matrix mainly leads to (1) a radical mediated fragmentation pathway, leading to the formation of *c*- and *z*-fragments that can be fully annotated (C′ and Z′) and (2) a thermal activation induced by the laser irradiation, and mediated by the matrix which leads mainly to CID-like *y*-, *a*-, and *b*-fragments. In some cases, the very useful *d*- and *w*-type fragments are also observed [15]. These latter imply a side-chain loss whose mass depends on the amino acid residue and can therefore be useful to confirm a residue determined by two consecutive *c*- or *z*-fragments or even to distinguish the isobaric amino acid residues leucine and isoleucine [22]. Bache et al. have demonstrated that ISD fragmentation occurs with a low level of hydrogen scrambling, conserving the solution deuteration pattern in the gas phase fragments formed. This fragmentation technique could therefore be applied to hydrogen/deuterium exchange studies of peptides and small proteins [23, 24].



Scheme 1 Mechanism of the formation of radical-induced *c*- and *z*-fragments

2.1 Radical-Induced Pathway

Hydrogen radicals formed by the matrix are at the origin of the formation of the *c*- and *z*-fragments [25–27], as depicted in a mechanism proposed by Takayama [25]. In the matrix crystals, hydrogen bonds are formed between the matrix and the analytes. During UV irradiation, a hydrogen radical transfer occurs from the matrix to the peptide, leading to an unstable radical that is rapidly cleaved to give *c*- and *z*-fragments (Scheme 1). *z*-fragments can also form adducts with a DHB molecule to give $[z+2] + 152$ fragments using the 2,5-dihydroxybenzoic acid matrix. The authors also conclude that the photoelectrons are involved in the ISD fragmentation mechanism [27].

The radical-induced ISD efficiency of a matrix can be correlated to its tendency to transfer hydrogen radicals, and this ability can be tested with hydrogen scavengers [28, 29] such as spirooxazines [18]. The radical-induced ISD efficiency of a matrix can also be correlated to its tendency to reduce the disulfide bond of

peptide [15] and this strategy was used by Sakakura et al. to classify 5-ASA amongst the other ISD matrices [20].

Based on these results, a ranking of the different matrices, from the least to the most able to transfer hydrogen radicals, was deduced: CHCA $\ll$ 2,5-DHB $<$ 2-AA $<$ 5-ASA $\approx$ 2-AB $<$ 1,5-DAN. This ranking is in good agreement with the increasing order of their radical-induced ISD abilities. Another characteristic of radical-induced ISD is its independence from the ionization process [30]. Indeed, Takayama et al. suggest that basic amino acids near the N-terminus favor production of *c*-type fragments and basic amino acids near the C-terminus favor production of *z*- and *y*-type fragments in positive ion mode [31, 32]. As ISD is independent of the ionization process, radical-induced ISD fragments are also observed in the negative ion mode according to the ability of the matrix to produce negatively charged molecular ions. In negative mode, the 1,5-DAN matrix is far more efficient than the 2,5-DHB matrix [15, 18]. The radical-induced fragmentation mechanism is exactly the same in both ion modes [15].

2.2 *Collisionally Activated Pathway*

For the thermal pathway, the activation of the fragmentation process of peptides occurs by a direct energy transfer from the matrix to the peptide and by collisions with matrix molecules in the expanding MALDI plume formed following laser irradiation [16, 17]. This thermal pathway leads to the CID-like *a*-, *b*- and *y*-ions corresponding to their PSD analogs. PSD is the metastable fragmentation occurring outside the MALDI source. This fragmentation process is activated by bimolecular collisions with the plume components during the ion extraction. The PSD fragmentation leads to the formation of the CID-like *a*-, *b*- and *y*-fragments.

2.3 *Influence of the Matrix*

The two kinds of ISD pathways are differently favored according to the matrix used. The in-source formation of CID-like fragments by the thermal pathway has not been widely studied [33]. Because this thermal process is caused by vibrational activation of closed valence shell ions, the fragmentation mechanism can be explained by the "mobile proton" model [34]. These fragments are the result of the cleavage of the CO–NH bond that is weakened by the protonation of the nitrogen of the amide group [35, 36]. The CID-like fragments (*a*-, *b*-, and *y*-type) can then be favored by the acidity of the matrix. The 2,5-DHB matrix is then more efficient than the 1,5-DAN matrix for production of these ions. However, it was shown by Sachon et al. with labile peptides that the formation of these fragment types cannot be explained only by the proton affinities of the matrices [37]. The low energy bimolecular collisions between the peptide and the matrix can activate

the fragmentation and are mediated by the initial axial velocity of the ions of the MALDI plume. The lower abundance fragments issued from this pathway using the 1,5-DAN matrix can make ISD spectra easier to interpret. Nevertheless, use of the 1,5-DAN and 2,5-DHB matrices can be complementary to facilitate the interpretation of ISD spectra of unknown peptides or proteins. It is noted that the formation of CID-like fragments tends to disappear with the increasing molecular weight of the studied compound.

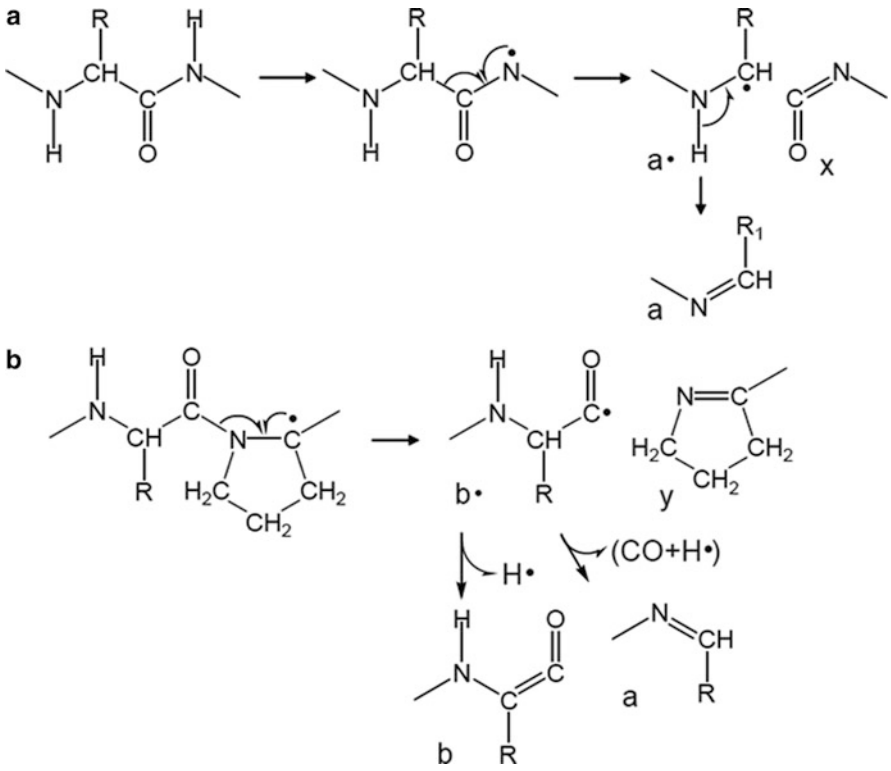
2.4 Radical-Induced Pathway via Hydrogen Abstraction

Recently it was found that the oxidizing matrices 5-formylsalicylic acid and 5-nitrosalicylic acid for MALDI-ISD resulted in the generation of *a* and *x* ions by cleavage of the C_α–C bond. MALDI-ISD with an oxidizing matrix is initiated by hydrogen abstraction from amide portion of the peptide backbone onto the matrix, as depicted in a mechanism proposed by Asakawa and Takayama [38, 39]. Hydrogen abstraction from peptides resulted in the formation of oxidized peptide molecules $[M–H]^{\bullet}$ containing a radical site on the amide nitrogen. Subsequently, the $[M–H]^{\bullet}$ radical principally generated the *a* and *x* ions, indicating cleavage of the C_α–C bonds on the peptide backbone. However, the *x* ions originating from the cleavage of the C_α–C bonds at Xxx–Pro sequences were absent, because the Pro residue cannot have a nitrogen-centered radical site. The C_α–C bond cleavage at Xxx–Pro and Pro–Xxx bonds would lead to *a*/*x*· and *a*/*x* fragment pairs, respectively. The absence of *x* ions originating from the cleavage of the C_α–C bonds at Xxx–Pro indicates that the fragmentation leading to an *a*/*x*·ion pair does not occur. However, the radical fragments *a*·ions were not observed in MALDI-ISD spectra with oxidizing matrix, and instead the *a* ions were detected. It is likely that there are sufficient amounts of exited matrix molecules and matrix radicals in the MALDI plume to form the *a* ions via further hydrogen abstraction after the C_α–C bond cleavage. The fragmentation mechanism in MALDI-ISD with oxidizing matrix is shown in Scheme 2a. In contrast, the cleavage of the C_α–C bond at Xxx–Pro did not occur, and instead CO–N bond cleavage at Xxx–Pro was observed via hydrogen abstraction from the C_α–H bond in the Pro residue. The CO–N bond cleavage leads to the formation of the *b*/*y* fragment pair and the *b*·ions undergo further degradation to form the *b* and *a* ions after the CO–N bond cleavage, as shown in Scheme 2b [39].

3 Proteomic Applications of MALDI-ISD

3.1 MALDI-ISD and Bottom-Up Strategies

As MALDI-ISD occurs in the source, this fragmentation method does not allow any precursor selection. In consequence, when working on enzymatic digests ISD



Scheme 2 Mechanism of MALDI-ISD with oxidizing matrix. (**a**) C_α–C bond cleavage and (**b**) CO–N bond cleavage at Xxx–Pro bond

signals from peptides will co-occur and the resulting overlapping spectra will be difficult to untangle. A purification step and separation of the peptides by LC-MALDI is thus necessary to overcome this limitation. The presence of matrix clusters in the low *m*/*z* range (<700 Da) may also burden the observation of low-mass fragments. Information cannot be easily obtained from this mass range, reducing the experimental observation window. However, Reiber and Brown have shown the possibility to sequence de novo purified small peptides in spite of matrix adducts in the spectral low mass range [40]. The authors combined the information coming from N-terminus (*c*-ions) and C-terminus (*y*-ions) fragments to identify six unknown peptides. Even if the authors underlined several ambiguities in their sequences, this approach highlights that two types of fragments can be used to bypass the loss of information caused by the low mass matrix clusters. Another methodology for de novo sequencing of peptides using MALDI-ISD has been recently demonstrated by Quinton et al. [19]. The authors have used PSD and ISD spectra in combination to determine sequences which were indecipherable by a single fragmentation method.

3.2 MALDI-ISD and Top-Down Strategies

MALDI-ISD appears to be a good answer to perform TDS because, as explained above, the mechanism partially relies on a transfer of radical species and the fragmentation process is consequently not limited by the mass. Thus, different groups tried to sequence purified proteins spotted in sinapinic acid (SA) or 2,5-DHB. In 1995, MALDI-ISD was performed on different proteins such as bovine cytochrome *c* (12.2 kDa), bovine superoxide dismutase (15.6 kDa), and equine apomyoglobin (16.9 kDa) [41, 42]. Numerous intense *c*-, *z*-, and *y*- fragments were observed between 850 *m*/*z* and 11,100 *m*/*z*. Figure 1 displays two MALDI-ISD spectra acquired for horse heart cytochrome *c* (12.2 kDa) and myoglobin (16.9 kDa). Clear *c*-, *z*-, and *y*-ion series are detected and lead to the characterization of 78 amino acids out of 104 for cytochrome *c* (75%) and 93 out of 153 for myoglobin (61%).

These signals enable the building of long sequence tags for each protein, which can be submitted in BLAST in order to compare them to the sequences present in the databases and to identify the protein of interest. However, the full sequence of the protein is not always accessible by ISD. For example, as the ISD process leads to breaking the N–C_α bond, a proline residue cannot produce ISD fragments due to its cyclic nature. However, this observation is predictable, and a gap in *c*-ions of an ISD-sequencing experiment can be interpreted as a proline residue on the N-side followed by another amino acid, Pro–X, the mass and position of which in the sequence can easily be determined.

3.3 MALDI-ISD and Pseudo-MS^3 Strategies

Protein ISD generally leads to an intense *c*-ion series and allows a precise characterization of the N-terminus part of the protein(s). However, the full N- and C-termini cannot be determined due to matrix background at low *m*/*z* ratios that prevents the observation of the very first fragments. Different "pseudo-MS^3" (pMS^3) strategies have therefore been imagined to describe the sequences extremities [43, 44]. These strategies exploit the fact that an ISD fragment can be selected and fragmented by CID like a classical ion generated by a MALDI source. Simply, an ISD fragment ($m/z < 3{,}500$) is isolated before subjection to PSD fragmentation. If the selected ion is a *c*-type, then the resulting MS/MS spectrum will characterize the N-terminus. Similarly, if the precursor ion is a *z*- or a *y*-ion, then the C-terminus part will be characterized. Figure 2 shows an example of a pMS^3-experiment. Aldolase protein (39.2 kDa) has been fragmented by MALDI-ISD using 2,5-DHB. A tag of 56 consecutive amino acids has been obtained from this single spectrum. However, as already discussed above, the N- and the C-termini are missing. To characterize the C-terminus part, a *y*-ion (y_{12}) was then selected and fragmented. The MS/MS spectrum obtained (Fig. 2,

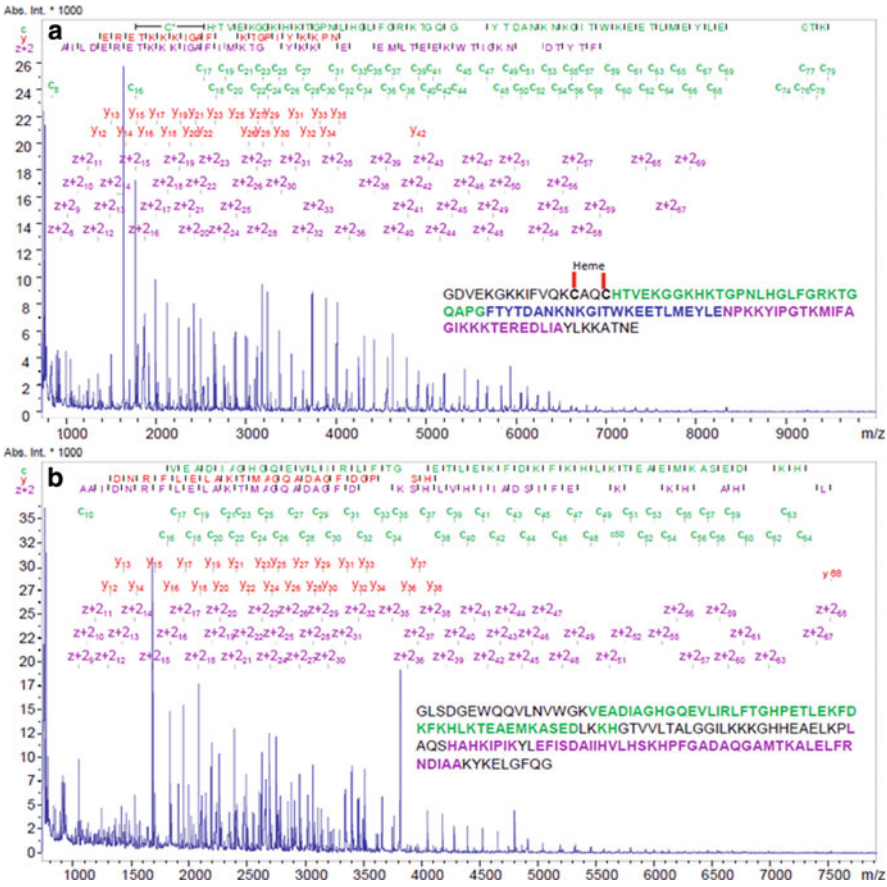


Fig. 1 MALDI-ISD spectra of (**a**) horse heart cytochrome *c* (12.2 kDa) and (**b**) horse heart myoglobin (16.9 kDa). One picomole of each protein was spotted onto the MALDI plate and mixed together with 1,5-DAN saturated in FA (0.2%)/ACN. Using intense *c*- and *z*-ion series, 77% and 61% of sequence coverage were characterized for cytochrome *c* and myoglobin, respectively. C- and N-terminus extremities are missing because fragments are hidden by intense matrix signals

bottom panel) clearly displays intense *y*- and *b*-ion types which enable an easy characterization of the missing C-terminus part.

As the use of pMS^3-sequencing allows full characterization of the protein extremities, it constitutes a competitive methodology to replace Edman degradation (ED). ED was the most used protein sequencer during the past few decades. First of all, the time scale of each experiment is in favor of ISD/pMS^3. Whereas ED takes more than 30 min per amino acid, MALDI-ISD takes around 10 min from the sample deposition to determination of a tag of several tens of amino acids. Moreover, from tens of picomoles of purified proteins, ED allows the sequencing of the first 50 amino acids except when the N-terminal amino acid is modified. With

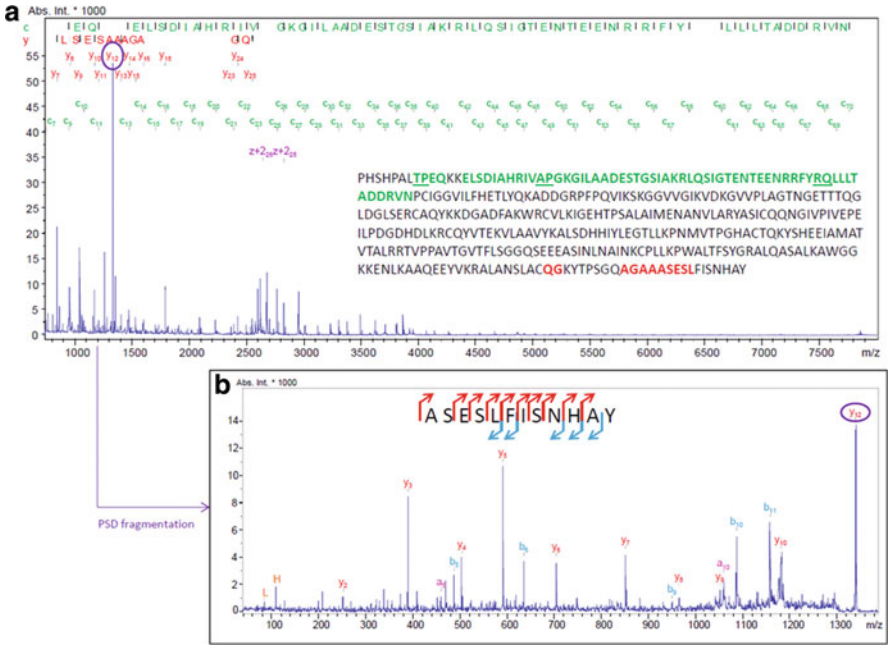


Fig. 2 Example of a pMS^3 experiment on Aldolase (39.2 kDa). (**a**) MALDI-ISD spectra of the protein displaying a long series of *c*-ions but also several *y*- and *z*-type ions. (**b**) y_{12} ion is fragmented by post-source decay, unambiguously revealing the whole missing N-terminus extremity

the same amount of compound or even less, MALDI-ISD combined with pMS^3-sequencing gives access in the more favorable cases to the first 80 amino acids from the N-terminus but also to the last 40 amino acids from the C-terminus.

It is important to highlight that pMS^3-sequencing also works if the extremities are modified and, thus, can reveal post-translational modifications (PTMs). The approach of combining ISD and pMS^3 has recently been highlighted by Resemann et al. [45]. In this work, the authors demonstrated the ability of MALDI-TDS to sequence fully a Camelid single heavy chain antibody of 13.6 kDa. The authors indicate that larger proteins may also be sequenced by MALDI-TDS but the use of reference sequence information in databases or complementary results from bottom-up approaches becomes unavoidable.

3.4 MALDI-ISD and PTMs

Several teams tried to apply MALDI-ISD to the detection and localization of PTMs. In 1997, Lennon and Walsh demonstrated that sequence determination by MALDI-ISD was stopped near the cysteins involved in disulfide bonds or covalent bonds to a heme molecule. These results demonstrated that the peptide

modifications were able to interfere with the ISD promoted by DHB [42, 44]. Indeed, a disulfide bond tends to be reduced during the ISD process [46] and therefore acts as a quencher of the fragmentation of neighboring amino acids. This observation was exploited by Schnaible et al. to localize disulfide bonds on HPLC-separated peptides [47, 48]. Moreover, PTMs that do not interfere with the ISD fragmentation of the peptide backbone can be studied by ISD as well. For example, Lennon et al. demonstrated in 1999 that labile modifications such as serine phosphorylations were not affected by the ISD process [49]. These PTMs can therefore be characterized either by a direct MALDI-ISD experiment or after a pMS^3-sequencing methodology [50].

Concerning glycosylation, MALDI-ISD was applied both on glycoproteins/glycopeptides and on isolated glycans. Collisionally-induced MALDI-ISD of glycopeptides led to the observation of both glycan and peptide fragmentations [51]. On the other hand, radical-induced ISD allowed Hanisch sequencing glycoproteins and localizing O-glycosylation sites, confirming again the ability of this technique to preserve labile PTMs [52]. However, Chaurand et al. observed that N-glycosylation was able to stop the protein sequencing, allowing localization of this PTM, but not further sequencing [53]. These contradictory results suggest a potential effect of the carbohydrate moiety on the sequencing of the glycoproteins.

As glycans can be easily isolated and purified, the ISD and pMS^3-sequencing approaches were also employed to characterize thoroughly the structures of both native [43, 54–60] and permethylated glycans [61–66], allowing unambiguous discrimination of isobaric species. In ref. [64] it was also mentioned that 6-aza-2-thiothymine (ATT) was able to promote ISD fragmentation of permethylated glycans. The mechanism involves in this case the protonation of the permethylated glycan, leading to its fragmentation. In addition, 2,5-DHB could act as a dual matrix here, allowing ISD of permethylated glycans when laser shots are made on crystals (where protons are available) but impeding ISD when laser shots are made in the central amorphous zone of the spots. Indeed, this part is sodium-rich and protonation is therefore unfavored compared to cationization [67].

Other classical PTMs such as pyroglutamic acid [53], formylation [68], deamidation [44], and acetylation [2] have also been described and studied by MALDI-ISD. Finally, Yoo and collaborators showed recently that MALDI-ISD was also an appropriate tool to localize PEGylation of protein [69], which is a chemical modification widely employed in the pharmaceutical field [70].

In conclusion, MALDI-ISD can be employed in a wide panel of applications such as the characterization of recombinant proteins, the full de novo sequencing of unknown proteins, and the localization of PTMs.

3.5 MALDI-ISD of Oligonucleotides

Besides peptides and proteins, substantial work was also achieved in order to use ISD to sequence oligonucleotides. In 1995, Nordhoff et al. reported the prompt fragmentation of oligonucleotides in an IR-MALDI source [71] while Zhu et al. used 2,5-DHB to induce in-source fragmentation in a UV-MALDI source [71, 72].

Following these observations, delayed extraction was applied and sequencing of an 11-mer DNA was achieved using a 266 nm laser with picolinic acid as matrix [73]. According to the McLuckey nomenclature [74], mainly *w*-type ions (resulting from the cleavage of a phosphodiester bond and containing a 5′-terminal phosphate group) were observed. A base loss is considered to be a critical step of this "thermal" fragmentation pathway [75, 76]. Alternatively, a pathway of electronic energy transfer was also proposed to occur in positive ion mode with 2-aminobenzoic acid and leads to the formation of supplementary ions (*d*-, *b*-, and *y*-ions) [77]. Concerning applications of ISD in the nucleic acids field, ISD was successfully applied by Wang et al. to sequence covalently modified (phosphorothioate instead of phosphodiester bonds) oligonucleotides [78] but failed to allow locating covalent rhodium adducts on DNA [79].

4 MALDI Mass Spectrometry Imaging

MALDI-MSI is a powerful technique that allows mapping of many classes of compounds directly from a tissue slice of several micrometers thickness. The principle of MALDI-MSI relies on a pixelization of the surface to be analyzed. A mass spectrum is recorded at each pixel, providing a triplet of information (spatial coordinates, mass, and intensity of the peak) for each signal present in the mass spectrum. Accumulating and integrating these data allows building a 2D cartography for any detected molecule, without the need for previous knowledge of the molecular content of the sample. On the image, the intensity of the signal is depicted by a color scale.

MALDI-MSI is suitable for high molecular weight biomolecules (proteins) but also for small molecules such as lipids, drugs, and their metabolites, etc. In most published applications, proteins are targeted as potential biomarkers of pathology, as in many studies, for ovarian tumors [80, 81], renal carcinoma [82], or Parkinson's disease [83, 84]. For small molecules analysis, two main classes may be distinguished – exogenous drugs and endogenous metabolites. Exogenous drugs have been studied in MALDI-MSI far later than proteins (since 2004) for different targets such as olanzapine (antipsychotic) [85–87], imatinib [87], vinblastine [88], and banoxantrone [89] (cancer drugs). Among endogenous metabolites, lipids such as glycerophospholipids and glycosphingolipids or simple lipids like cholesterol, diacylglycerols, or triacylglycerols have been extensively investigated [90]. Also studied are primary metabolites (ADP, ATP, . . .) as targets for MALDI-MSI [91].

Concerning MALDI-MSI, a lot of work has been done to improve the experimental workflow, such as optimization of sample preparation and matrix deposition [92], and much focus has been on the identification of the detected molecules. When small molecules are analyzed, the identification is rather straightforward where tandem mass spectrometry can be easily used, leading to the fragmentation of the detected compounds, allowing their unambiguous

characterization. The identification is even more reliable when high resolution exact mass measurements can be performed. If proteins are analyzed, it is rather difficult to perform MS/MS analysis directly from proteins that are more difficult to fragment and the unique information of the *m*/*z* value is not sufficient as it may correspond to several primary sequences. To achieve protein identification, in situ digestion has been used since 2007 [93]. This method consists of the deposition (more often by micro-spotting or nano-spotting) of a solution of protease (usually trypsin) at the surface of the sample, followed by 2–4 h of incubation and analysis of the resulting mix of peptides. The classical PMF method is then used to query protein databases for identification. Although this procedure presents some drawbacks (need for a microspotting system to deposit the enzyme solution as precisely as possible, time-consuming, and leads to mass spectra which can be complex and difficult to interpret), it has been extensively used since its implementation. MALDI-ISD, however, removes the need for in situ digestion when analyzing tissues.

4.1 MALDI-ISD on Tissue Slices

MALDI-ISD seems to be useful in this context, and Debois et al. developed an ISD-based method for protein identification, directly from a tissue section, without any further treatment [94]. Indeed, to be implemented, the technique requires only the use of an "ISD-favorable" matrix such as 2,5-DHB or 1,5-DAN as already mentioned. The analysis is then performed as a simple full scan MS acquisition, neither requiring an increase of the laser intensity nor the number of shots on the tissue. The resulting mass spectrum allows the creation of a sequence tag of amino acids and the original protein may be identified thanks to a query in protein databases, using a BLAST search.

Figure 3a shows an ISD mass spectrum recorded on a porcine eye lens with 1,5-DAN as a matrix [94]. Two ions series were detected, leading to two different tags: QPLNPKIIIF EKENFKGHSH ELNGPCPNLK ETGVEQAGSV and WQMDRIRRVS QVPQ, respectively. For the first tag, the interrogation of databases led to the identification of the Beta-Crystallin B2 for three different species (rabbit, bovine, and guinea pig). Nevertheless, some differences exist between the primary sequences of rabbit, bovine, and guinea pig Beta-Crystallins B2. To obtain the N-terminal sequence of the porcine Beta-Crystallin B2, a pMS^3-sequencing experiment was performed on the same tissue slice. The ion at *m*/*z* value 1,519.0 was selected and then fragmented. Figure 3b depicts the resulting MS/MS spectrum. Almost the entire *b*-ions and *y*-ions series were obtained and the acetylation of the Ala residue was also confirmed, giving access to the full sequence of the N-terminus part of the porcine Beta-Crystallin B2 (51 first amino acids), which is not referenced in databases (SwissProt and TrEMBL). These results open the way towards de novo sequencing, directly from tissue slices. The method was applied for biomarkers discovery [95, 96].

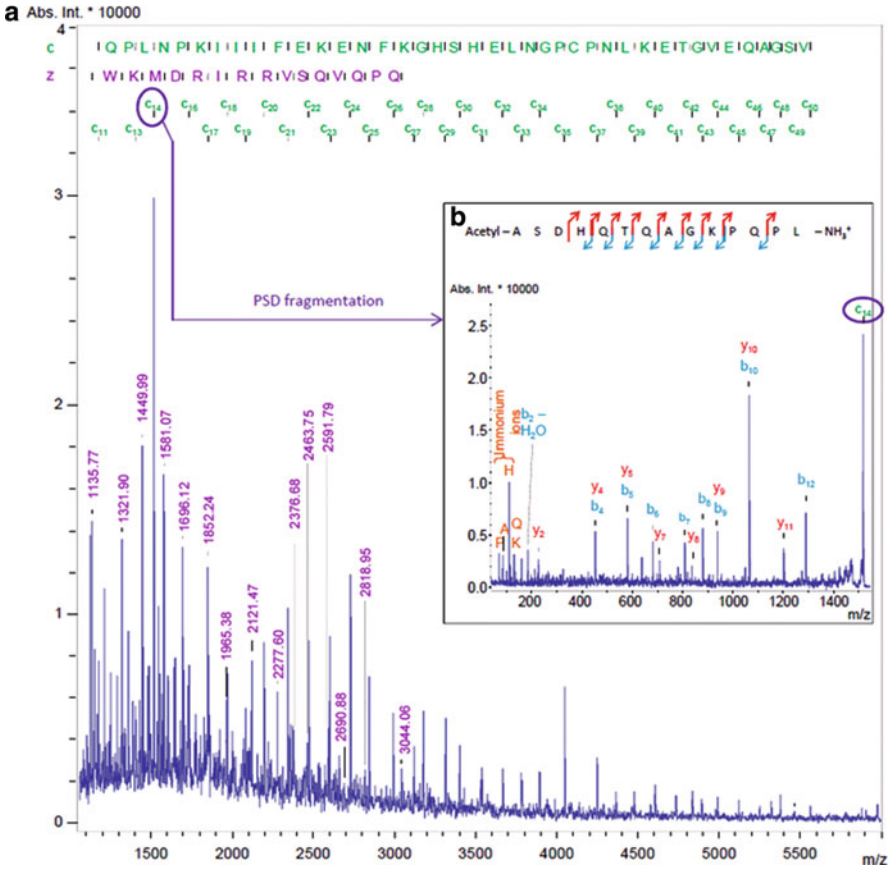


Fig. 3 (**a**) ISD mass spectrum recorded on a porcine eye lens slice with 1,5-DAN as a matrix. The tag and the masses indicated in *green* correspond to the N-terminus part of the Beta-Crystallin B2 (*c*-ions series), the *purple* ones to the C-terminus part (*z*-ions series). (**b**) T^3-sequencing mass spectrum obtained by isolating the ion at *m*/*z* 1,519.08 and by fragmenting it by PSD. Quasi-complete *b*- and *y*-ions series are observable, leading to the full sequencing of the N-terminus part of the protein

4.2 MALDI-ISD Imaging

ISD can also be performed during a tissue imaging experiment, allowing the localization of different fragments to be considered. Figure 4 presents MALDI images recorded on a porcine eye lens slice, with 1,5-DAN as the matrix [94]. Getting the distribution of ISD fragments provides two additional pieces of information. First, it becomes easy to find which fragments are coming from the same protein, as they all exhibit the same localization on the tissue surface. This can help to indicate real consecutive peaks, facilitating the protein sequencing. Second, by overlaying these images, it becomes possible to highlight high intensity pixel(s) and

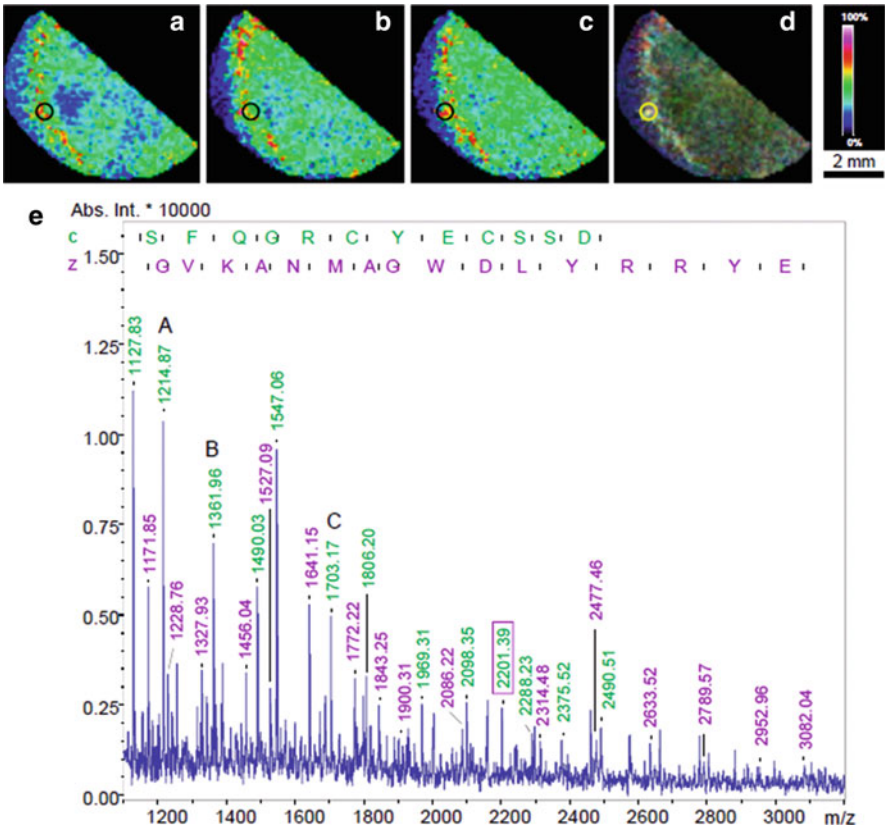


Fig. 4 MALDI images of ISD fragments at (**a**) *m*/*z* 1,214.9, (**b**) *m*/*z* 1,361.9, and (**c**) *m*/*z* 1,703.2, recorded on a porcine eye lens slice. (**d**) Overlay of the three previous ion images. The *black and the yellow circles* indicate the region where the mass spectrum has been extracted (high intensity pixel). (**e**) ISD mass spectrum extracted from a high intensity pixel. The tag and the masses indicated in *green* correspond to the N-terminus part of the Gamma-Crystallin B (*c*-ions series), the *purples* ones to the C-terminus part (*z*-ions series). The *green purple-framed annotation* indicates a peak for which *m*/*z* value corresponds to both *c*- and *z*-ion. The ion peaks which images are shown are labeled with letters A, B, C

specifically extract the corresponding mass spectrum. Of course, the more intense the mass spectrum and the more peaks present, the easier the interpretation. In Fig. 4, for example, the three fragments exhibit close localizations, suggesting their common origin, and the overlay of their images (Fig. 4d) shows an intense pixel (yellow-circled) which allows the extraction of the ISD mass spectrum shown in Fig. 4e. Two sequence tags were then established, leading to the identification of Gamma-Crystallin B.

The ISD process is often criticized, and the main denigrations are related to the presumed need for a pure sample, to its lower sensitivity, and to its reduced mass accuracy and resolution. In fact, some of these criticisms are based on false assumptions.

Concerning the kind of sample that can be analyzed by ISD, it is admitted that a tissue slice is a complex sample and that within one 50-μm-wide laser shot, several proteins may be desorbed and fragmented. To decrease the number of analyzed proteins, two approaches are considered. The first choice consists of decreasing the size of the laser shot (to 30 or even 20 μm). In this case, the question is: what about the sensitivity of the instrument? Would it be sensitive enough to obtain interpretable ISD mass spectra? How many molecules are present in a 20-μm-wide pixel? And, out of these, how many will fragment? This issue has to be investigated. The second choice for decreasing the number of analyzed proteins is to employ sample depletion methods to remove the high abundance proteins from the tissue slice [97]. Or, rather than depleting the abundant proteins, a variety of identifications can instead be obtained by imaging multiple yet slightly different tissue sections. Thus, in MALDI-ISD imaging, serial sections are still likely to result in detection of the same set of proteins, but the sensitivity of the ISD process to protein ionization efficiencies allows a larger variety of protein identifications to be made from slightly more different tissue sections. MALDI-ISD imaging therefore allows the variety of proteins identified and localized between different tissue sections to mirror more effectively the changes in sample composition. This feature of MALDI-ISD could be used in combination with sample depletion to reveal even more proteins.

Concerning the issue of the mass accuracy, this does not represent a severe problem. Indeed, during the data treatment, it is not the mass of peptides which is measured, but the mass difference between consecutive peaks that is used to obtain an amino acid residue identification. Accurate mass of peptides is more important in the PMF strategy where intact analyte masses are used to query protein databases, but for sequencing a good external calibration of the mass spectrum is sufficient to allow the mass spectrum interpretation. Moreover, to improve the quality of mass accuracy and calibration, it is best to use ISD fragments of a known protein (myoglobin for example) to create a calibration mass list. In this way, data used for calibration and data coming from the sample are recorded under the same conditions and the whole mass range is covered, far better than with a commercial peptides mixture. For mass resolution, as mentioned previously, ISD fragments are formed in the ionization source before the extraction of ions (contrary to PSD), then travel through the flight tube, and with the use of an electrostatic mirror the mass resolution obtained for ISD mass spectra is completely satisfactory.

Progress has recently occurred regarding data treatment for MALDI-ISD imaging. In fact, MALDI-ISD imaging presents a relatively new type of dataset, wherein a protein sequence can be found in each pixel, and this calls for the development of new software and data analysis methods. Zimmerman et al. have created an analytical pipeline suite of software that accomplishes two main goals: automated analysis of MALDI-ISD “top-down” spectra for protein sequencing and automated analysis of MALDI-ISD tissue imaging data [98]. These goals can be combined so that an imaging dataset is first searched by the software to find groups of highly correlated peaks that co-occur and that are likely to correspond to a single protein. Second, the software can then identify the highest quality spectra from the imaging dataset for

automated de novo sequencing. After a first protein is automatically sequenced or identified, another search is run in the imaging data to find spectral patterns of peaks that are different from that of the first sequenced protein, which are likely to correspond either to a second protein or to a post translationally modified form of the first protein. This semi-automated process is iterated for an exhaustive characterization all of the proteins in the MALDI-ISD imaging dataset. Ion images are subsequently plotted by the software showing the spatial localizations of the identified proteins over the tissue surface. The MALDI-ISD protein sequencing software is used either in combination with imaging datasets, or to sequence proteins from individual MALDI-ISD spectra [98–100].

As applied to individual spectra, the sequencing software uses mass differences calculated from spectral peak lists to perform protein sequencing, and the software is able to overcome interference due to unidentified fragmentation peaks or redundant ion series peaks. In addition, as it was shown that ISD patterns for up to three proteins can exist simultaneously within a single MALDI-ISD spectrum, the software can subtract the already known ISD peak pattern of a known protein out of the peak list. This isolates the remaining peaks, leaving the refined peak list to be passed through the software for identification of a second protein, and by repeating this process a third protein is identified. This analysis process for protein mixtures is advantageous for use with MALDI-ISD tissue imaging data, as MALDI-ISD mass peaks corresponding to more abundant proteins are likely to be already known and can be subtracted from the spectral peak lists to simplify the data analysis for the remaining proteins.

A future possible way to facilitate ISD analysis of complex mixtures is the generation of all possible sequence tags that fit the data within a tolerance window. At present it is difficult for software to interpret ISD mass spectra containing more than two or three proteins because of the large number of peaks present. In the rarer cases where more than three proteins exist in a single ISD spectrum, the idea is to have the software calculate any possible tag and then automatically submit them to query protein databases using a BLAST engine. After this first stage the user should validate (or not) the results and the software could then create an exclusion list for peaks already used, followed by identification of the remaining proteins. This could be seen as a “cleaning” of the ISD mass spectrum. This future direction for automated data treatment would greatly decrease the time needed for mass spectra interpretation and would not necessitate the presence of the user.

5 Conclusions

MALDI-ISD seems to have a promising future and will likely have increasing importance in mass spectrometry applications. Top-down sequencing of proteins seems to be the most promising application. Indeed, MALDI-ISD allows a rapid characterization of N-terminal and sometimes C-terminal sequences of a protein and from only a few picomoles of compounds. This technique could be employed

for the routine quality control of recombinant proteins. Another promising theme resides in MALDI-imaging with ISD and pMS^3 performed directly on tissue slices, which could drive the characterization of different proteins and, in the best cases, of pathological biomarkers. In parallel studies, MALDI-ISD has been performed with success on other biopolymers including oligonucleotides and therefore can be generalized as a fast sequencing method.

Acknowledgments The F.R.S.-FNRS (Fonds National de la Recherche Scientifique, Belgium) is acknowledged for postdoctoral fellowships to D.D., T.A.Z., and N.S. The Japanese Society for the Promotion of Science is acknowledged for the postdoctoral fellowship of D.A. The instrumentation was funded by the FNRS, The University of Liege, and the FEDER funds.

References

1. Rubino FM, Pitton M, Di Fabio D et al (2009) Toward an "omic" physiopathology of reactive chemicals: thirty years of mass spectrometric study of the protein adducts with endogenous and xenobiotic compounds. Mass Spectrom Rev 28:725–784
2. Cox J, Mann M (2011) Quantitative, high-resolution proteomics for data-driven systems biology. Annu Rev Biochem 80:273–299
3. Marchetti-Deschmann M, Allmaier G (2011) Mass spectrometry – one of the pillars of proteomics. J Proteomics 74:915–919
4. Griffiths WJ, Wang Y (2009) Mass spectrometry: from proteomics to metabolomics and lipidomics. Chem Soc Rev 38:1882–1896
5. Cramer R, Gobom J, Nordhoff E (2005) High-throughput proteomics using matrix-assisted laser desorption/ionization mass spectrometry. Expert Rev Proteomics 2:407–420
6. Knochenmuss R (2006) Ion formation mechanisms in UV-MALDI. Analyst 131:966–986
7. Gies AP, Vergne MJ, Orndorff RL et al (2007) MALDI-TOF/TOF CID study of polystyrene fragmentation reactions. Macromolecules 40:7493–7504
8. Konn D, Murrell J, Despeyroux D et al (2005) Comparison of the effects of ionization mechanism, analyte concentration, and ion "cool-times" on the internal energies of peptide ions produced by electrospray and atmospheric pressure matrix-assisted laser desorption ionization. J Am Soc Mass Spectrom 16:743–751
9. Spengler B (1997) Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules. J Mass Spectrom 32:1019–1036
10. Yoon S, Moon J, Kim M (2010) A comparative study of in- and post-source decays of peptide and preformed ions in matrix-assisted laser desorption ionization time-of-flight mass spectrometry: effective temperature and matrix effect. J Am Soc Mass Spectrom 21:1876–1883
11. Kelleher NL (2004) Top-down proteomics. Anal Chem 76:196A–203A
12. Spengler B, Kirsch D, Kaufmann R et al (1991) Metastable decay of peptides and proteins in matrix-assisted laser-desorption mass spectrometry. Rapid Commun Mass Spectrom 5:198–202
13. Karas M, Bahr U, Fournier I et al (2003) The initial-ion velocity as a marker for different desorption-ionization mechanisms in MALDI. Int J Mass Spectrom 226:239–248
14. Syka JEP, Coon JJ, Schroeder MJ et al (2004) Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. Proc Natl Acad Sci USA 101:9528–9533
15. Demeure K, Gabelica V, De Pauw E (2010) New advances in the understanding of the in-source decay fragmentation of peptides in MALDI-TOF-MS. J Am Soc Mass Spectrom 21:1906–1917

16. Schulz E, Karas M, Rosu F et al (2006) Influence of the matrix on analyte fragmentation in atmospheric pressure MALDI. J Am Soc Mass Spectrom 17:1005–1013
17. Gabelica V, Schulz E, Karas M (2004) Internal energy build-up in matrix-assisted laser desorption/ionization. J Mass Spectrom 39:579–593
18. Demeure K, Quinton L, Gabelica V et al (2007) Rational selection of the optimum MALDI matrix for top-down proteomics by in-source decay. Anal Chem 79:8678–8685
19. Quinton L, Demeure K, Dobson R et al (2007) New method for characterizing highly disulfide-bridged peptides in complex mixtures: application to toxin identification from crude venoms. J Proteome Res 6:3216–3223
20. Sakakura M, Takayama M (2010) In-source decay and fragmentation characteristics of peptides using 5-aminosalicylic acid as a matrix in matrix-assisted laser desorption/ionization mass spectrometry. J Am Soc Mass Spectrom 21:979–988
21. Smargiasso N, Quinton L, De Pauw E (2011) 2-Aminobenzamide and 2-aminobenzoic acid as new MALDI matrices inducing radical mediated in-source decay of peptides and proteins. J Am Soc Mass Spectrom. doi:10.1007/s13361-011-0307-5
22. Johnson RS, Martin SA, Biemann K et al (1987) Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. Anal Chem 59:2621–2625
23. Bache N, Rand KD, Roepstorff P et al (2008) Gas-phase fragmentation of peptides by MALDI in-source decay with limited amide hydrogen (1H/2H) scrambling. Anal Chem 80:6431–6435
24. Rand KD, Bache N, Nedertoft MM et al (2011) Spatially resolved protein hydrogen exchange measured by matrix-assisted laser desorption ionization in-source decay. Anal Chem 83:8859–8862
25. Takayama M (2001) N-Cα bond cleavage of the peptide backbone via hydrogen abstraction. J Am Soc Mass Spectrom 12:1044–1049
26. Takayama M (2001) In-source decay characteristics of peptides in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. J Am Soc Mass Spectrom 12:420–427
27. Köcher T, Engström Å, Zubarev RA (2005) Fragmentation of peptides in MALDI in-source decay mediated by hydrogen radicals. Anal Chem 77:172–177
28. Calba PJ, Muller JF, Hachimi A et al (1997) Spirooxazines as a molecular probe for the study of matrix-assisted laser desorption/ionization processes. Part I: study of the interaction effect between the molecular probe and the matrix. Rapid Commun Mass Spectrom 11:1602–1611
29. Calba PJ, Muller JF, Inouye M (1998) H-atom transfer following analyte photoionization in matrix-assisted laser desorption/ionization processes. Rapid Commun Mass Spectrom 12:1727–1731
30. Brown RS, Feng J, Reiber DC (1997) Further studies of in-source fragmentation of peptides in matrix-assisted laser desorption-ionization. Int J Mass Spectrom 169–170:1–18
31. Takayama M, Tsugita A (1998) Does in-source decay occur independent of the ionization process in matrix-assisted laser desorption? Int J Mass Spectrom 181:L1–L6
32. Takayama M, Tsugita A (2000) Sequence information of peptides and proteins with in-source decay in matrix assisted laser desorption/ionization-time of flight-mass spectrometry. Electrophoresis 21:1670–1677
33. Hardouin J (2007) Protein sequence information by matrix-assisted laser desorption/ionization in-source decay mass spectrometry. Mass Spectrom Rev 26:672–682
34. Paizs B, Suhai S (2005) Fragmentation pathways of protonated peptides. Mass Spectrom Rev 24:508–548
35. McCormack AL, Somogyi A, Dongre AR et al (1993) Fragmentation of protonated peptides: surface-induced dissociation in conjunction with a quantum mechanical approach. Anal Chem 65:2859–2872
36. Somogyi Á, Wysocki V, Mayer I (1994) The effect of protonation site on bond strengths in simple peptides: application of ab initio and modified neglect of differential overlap bond

orders and modified neglect of differential overlap energy partitioning. J Am Soc Mass Spectrom 5:704–717
37. Sachon E, Clodic G, Blasco T et al (2009) In-source fragmentation of very labile peptides in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. Anal Chem 81:8986–8992
38. Asakawa D, Takayama M (2011) Cα-C bond cleavage of the peptide backbone in MALDI in-source decay using salicylic acid derivative matrices. J Am Soc Mass Spectrom 22:1224–1233
39. Asakawa D, Takayama M (2011) Specific cleavage at peptide backbone Cα–C and CO–N bonds during matrix-assisted laser desorption/ionization in-source decay mass spectrometry with 5-nitrosalicylic acid as the matrix. Rapid Commun Mass Spectrom 25:2379–2383
40. Reiber DC, Brown RS, Weinberger S et al (1998) Unknown peptide sequencing using matrix-assisted laser desorption/ionization and in-source decay. Anal Chem 70:1214–1222
41. Brown RS, Lennon JJ (1995) Sequence-specific fragmentation of matrix-assisted laser-desorbed protein/peptide ions. Anal Chem 67:3990–3999
42. Lennon JJ, Walsh KA (1997) Direct sequence analysis of proteins by in-source fragmentation during delayed ion extraction. Protein Sci 6:2446–2453
43. Pfeifer T, Drewello M, Schierhorn A (1999) Using a matrix-assisted laser desorption/ ionization time-of-flight mass spectrometer for combined in-source decay/post-source decay experiments. J Mass Spectrom 34:644–650
44. Suckau D, Resemann A (2003) T3-sequencing: targeted characterization of the N- and C-termini of undigested proteins by mass spectrometry. Anal Chem 75:5817–5824
45. Resemann A, Wunderlich D, Rothbauer U et al (2010) Top-down de novo protein sequencing of a 13.6 kDa camelid single heavy chain antibody by matrix-assisted laser desorption ionization-time-of-flight/time-of-flight mass spectrometry. Anal Chem 82:3283–3292
46. Patterson SD, Katta V (1994) Prompt fragmentation of disulfide-linked peptides during matrix-assisted laser desorption ionization mass spectrometry. Anal Chem 66:3727–3732
47. Schnaible V, Wefing S, Resemann A et al (2002) Screening for disulfide bonds in proteins by MALDI in-source decay and LIFT-TOF/TOF-MS. Anal Chem 74:4980–4988
48. Wefing S, Schnaible V, Hoffmann D (2006) SearchXLinks. A program for the identification of disulfide bonds in proteins from mass spectra. Anal Chem 78:1235–1241
49. Lennon JJ, Walsh KA (1999) Locating and identifying posttranslational modifications by in-source decay during MALDI-TOF mass spectrometry. Protein Sci 8:2487–2493
50. Raska CS, Parker CE, Huang C et al (2002) Pseudo-MS3 in a MALDI orthogonal quadrupole-time of flight mass spectrometer. J Am Soc Mass Spectrom 13:1034–1041
51. van der Wel H, Fisher SZ, West CM (2002) A bifunctional diglycosyltransferase forms the Fucα1,2Galβ1,3-disaccharide on Skp1 in the cytoplasm of Dictyostelium. J Biol Chem 277:46527–46534
52. Hanisch F-G (2011) Top-down sequencing of O-glycoproteins by in-source decay matrix-assisted laser desorption ionization mass spectrometry for glycosylation site analysis. Anal Chem 83:4829–4837
53. Chaurand P, DaGue BB, Ma S et al (2001) Strain-based sequence variations and structure analysis of murine prostate specific spermine binding protein using mass spectrometry. Biochemistry-US 40:9725–9733
54. Harvey DJ, Hunter AP, Bateman RH et al (1999) Relationship between in-source and post-source fragment ions in the matrix-assisted laser desorption (ionization) mass spectra of carbohydrates recorded with reflectron time-of-flight mass spectrometers. Int J Mass Spectrom 188:131–146
55. Harvey DJ, Naven TJP, Küster B et al (1995) Comparison of fragmentation modes for the structural determination of complex oligosaccharides ionized by matrix-assisted laser desorption/ionization mass spectrometry. Rapid Commun Mass Spectrom 9:1556–1561
56. Naven TJP, Harvey DJ, Brown J et al (1997) Fragmentation of complex carbohydrates following ionization by matrix-assisted laser desorption with an instrument fitted with time-lag focusing. Rapid Commun Mass Spectrom 11:1681–1686

57. Cancilla MT, Penn SG, Carroll JA et al (1996) Coordination of alkali metals to oligosaccharides dictates fragmentation behavior in matrix assisted laser desorption ionization/Fourier transform mass spectrometry. J Am Chem Soc 118:6736–6745
58. Bashir S, Giannakopulos AE, Derrick PJ et al (2004) Matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. A comparison of fragmentation patterns of linear dextran obtained by in-source decay, post-source decay and collision-induced dissociation and the stability of linear and cyclic glucans studied by in-source decay. Eur J Mass Spectrom 10:109–120
59. Yang H, Yu Y, Song F et al (2011) Structural characterization of neutral oligosaccharides by laser-enhanced in-source decay of MALDI-FTICR MS. J Am Soc Mass Spectrom 22:845–855
60. Wuhrer M, Deelder AM (2006) Matrix-assisted laser desorption/ionization in-source decay combined with tandem time-of-flight mass spectrometry of permethylated oligosaccharides: targeted characterization of specific parts of the glycan structure. Rapid Commun Mass Spectrom 20:943–951
61. Yamagaki T, Suzuki H, Tachibana K (2005) In-source and postsource decay in negative-ion matrix-assisted laser desorption/ionization time-of-flight mass spectrometry of neutral oligosaccharides. Anal Chem 77:1701–1707
62. Suzuki N, Khoo K-H, Chen C-M et al (2003) N-glycan structures of pigeon IgG. J Biol Chem 278:46293–46306
63. Yu S-Y, Khoo K-H, Yang Z et al (2008) Glycomic mapping of O- and N-linked glycans from major rat sublingual mucin. Glycoconj J 25:199–212. http://www.ncbi.nlm.nih.gov/pubmed/17891558
64. Yu S-Y, Wu S-W, Khoo K-H (2006) Distinctive characteristics of MALDI-Q/TOF and TOF/TOF tandem mass spectrometry for sequencing of permethylated complex type N-glycans. Glycoconj J 23:355–369
65. Terada M, Khoo K-H, Inoue R et al (2005) Characterization of oligosaccharide ligands expressed on SW1116 cells recognized by mannan-binding protein. J Biol Chem 280:10897–10913
66. Klisch K, Jeanrond E, Pang P-C et al (2008) A tetraantennary glycan with bisecting N-acetylglucosamine and the Sda antigen is the predominant N-glycan on bovine pregnancy-associated glycoproteins. Glycobiology 18:42–52
67. Smargiasso N, De Pauw E (2010) Optimization of matrix conditions for the control of MALDI in-source decay of permethylated glycans. Anal Chem 82:9248–9253
68. Netz DJA, Pohl R, Beck-Sickinger AG et al (2002) Biochemical characterisation and genetic analysis of aureocin A53, a new, atypical bacteriocin from Staphylococcus aureus. J Mol Biol 319:745–756
69. Yoo C, Suckau D, Sauerland V et al (2009) Toward top-down determination of PEGylation site using MALDI in-source decay MS analysis. J Am Soc Mass Spectrom 20:326–333
70. Harris JM, Chess RB (2003) Effect of pegylation on pharmaceuticals. Nat Rev Drug Discov 2:214–221
71. Nordhoff E, Karas M, Cramer R et al (1995) Direct mass spectrometric sequencing of low-picomole amounts of oligodeoxynucleotides with up to 21 bases by matrix-assisted laser desorption/ionization mass spectrometry. J Mass Spectrom 30:99–112
72. Zhu L, Parr GR, Fitzgerald MC et al (1995) Oligodeoxynucleotide fragmentation in MALDI/TOF mass spectrometry using 355-nm radiation. J Am Chem Soc 117:6048–6056
73. Juhasz P, Roskey MT, Smirnov IP et al (1996) Applications of delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry to oligonucleotide analysis. Anal Chem 68:941–946
74. McLuckey S, Van Berkel G, Glish G (1992) Tandem mass spectrometry of small, multiply charged oligonucleotides. J Am Soc Mass Spectrom 3:60–70
75. Wu J, McLuckey SA (2004) Gas-phase fragmentation of oligonucleotide ions. Int J Mass Spectrom 237:197–241

76. Christian N, Reilly J, Mokler V et al (2001) Elucidation of the initial step of oligonucleotide fragmentation in matrix-assisted laser desorption/ionization using modified nucleic acids. J Am Soc Mass Spectrom 12:744–753
77. Chan T, Fung Y, Li Y (2002) A study of fast and metastable dissociations of adenine-thymine binary-base oligonucleotides by using positive-ion MALDI-TOF mass spectrometry. J Am Soc Mass Spectrom 13:1052–1064
78. Wang BH, Hopkins CE, Belenky AB et al (1997) Sequencing of modified oligonucleotides using in-source fragmentation and delayed pulsed ion extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry. Int J Mass Spectrom 169–170:331–350
79. Chifotides HT, Koomen JM, Kang M et al (2004) Binding of DNA purine sites to dirhodium compounds probed by mass spectrometry. Inorg Chem 43:6177–6187
80. Lemaire R, Menguellet SA, Stauber J et al (2007) Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker. J Proteome Res 6:4127–4134
81. Kang S, Shim HS, Lee JS et al (2010) Molecular proteomics imaging of tumor interfaces by mass spectrometry. J Proteome Res 9:1157–1164
82. Oppenheimer SR, Mi D, Sanders ME et al (2010) Molecular analysis of tumor margins by MALDI mass spectrometry in renal carcinoma. J Proteome Res 9:2182–2190
83. Pierson J, Norris JL, Aerni H-R et al (2004) Molecular profiling of experimental Parkinson's disease: direct analysis of peptides and proteins on brain tissue sections by MALDI mass spectrometry. J Proteome Res 3:289–295
84. Stauber J, Lemaire R, Franck J et al (2008) MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. J Proteome Res 7:969–978
85. Hsieh Y, Casale R, Fukuda E et al (2006) Matrix-assisted laser desorption/ionization imaging mass spectrometry for direct measurement of clozapine in rat brain tissue. Rapid Commun Mass Spectrom 20:965–972
86. Khatib-Shahidi S, Andersson M, Herman JL et al (2006) Direct molecular analysis of whole-body animal tissue sections by imaging MALDI mass spectrometry. Anal Chem 78:6448–6456
87. Cornett DS, Frappier SL, Caprioli RM (2008) MALDI-FTICR imaging mass spectrometry of drugs and metabolites in tissue. Anal Chem 80:5648–5653
88. Trim PJ, Henson CM, Avery JL et al (2008) Matrix-assisted laser desorption/ionization-ion mobility separation-mass spectrometry imaging of vinblastine in whole body tissue sections. Anal Chem 80:8628–8634
89. Atkinson SJ, Loadman PM, Sutton C et al (2007) Examination of the distribution of the bioreductive drug AQ4N and its active metabolite AQ4 in solid tumours by imaging matrix-assisted laser desorption/ionisation mass spectrometry. Rapid Commun Mass Spectrom 21:1271–1276
90. Sugiura Y, Setou M (2010) Imaging mass spectrometry for visualization of drug and endogenous metabolite distribution: toward in situ pharmacometabolomes. J Neuroimmune Pharm 5:31–43
91. Benabdellah F, Touboul D, Brunelle A et al (2009) In situ primary metabolites localization on a rat brain section by chemical mass spectrometry imaging. Anal Chem 81:5557–5560
92. Groseclose MR, Andersson M, Hardesty WM et al (2007) Identification of proteins directly from tissue: *in situ* tryptic digestions coupled with imaging mass spectrometry. J Mass Spectrom 42:254–262
93. Kaletas BK, van der Wiel IM, Stauber J et al (2009) Sample preparation issues for tissue imaging by imaging MS. Proteomics 9:2622–2633
94. Calligaris D, Villard C, Lafitte D (2011) Advances in top-down proteomics for disease biomarker discovery. J Proteomics 74:920–934
95. Calligari D, Villard C, Terras L et al (2010) MALDI in-source decay of high mass protein isoforms: application to alpha- and beta-tubulin variants. Anal Chem 82:6176–6184

96. Debois D, Bertrand V, Quinton L et al (2010) MALDI-in source decay applied to mass spectrometry imaging: a new tool for protein identification. Anal Chem 82:4036–4045
97. Bonnel D, Longuespee R, Franck J et al (2011) Multivariate analyses for biomarkers hunting and validation through on-tissue bottom-up or in-source decay in MALDI-MSI: application to prostate cancer. Anal Bioanal Chem 401:149–165
98. Zimmerman TA, Debois D, Mazzucchelli G et al (2011) An analytical pipeline for MALDI in-source decay mass spectrometry imaging. Anal Chem 83:6090–6097
99. Demine R, Walden P (2004) Sequit: software for de novo peptide sequencing by matrix-assisted laser desorption/ionization post-source decay mass spectrometry. Rapid Commun Mass Spectrom 18:907–913
100. Gao J, Tsugita A, Takayama M et al (2002) A programmable fragmentation analysis of proteins by in-source decay in MALDI-TOF mass spectrometry. Anal Chem 74:1449–1457

Top Curr Chem (2013) 331: 143–164
DOI: 10.1007/128_2012_383
© Springer-Verlag Berlin Heidelberg 2012
Published online: 25 October 2012

Advances of MALDI-TOF MS in the Analysis of Traditional Chinese Medicines

Minghua Lu and Zongwei Cai

Abstract Traditional Chinese medicines (TCMs) are attracting more and more attention because of their long historical clinical experience and reliable therapeutic efficacy for preventing and/or treating various human diseases. Many techniques and methods were developed for the analysis of TCMs to support new drug discovery and quality control. Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), a soft ionization mass spectrometric technique, has been widely used in the analysis of a wide variety of large molecular compounds including proteins, peptides, and polymers since it was introduced in the late 1980s. In the present chapter, advances of MALDI-TOF MS in TCMs analysis have been reviewed. The review covers MALDI-TOF MS applications in the identification of new bioactive ingredients, analysis of alkaloids, determination of small molecular compounds with new matrices, proteomics analysis associated with TCMs, direct analysis of plant tissue, and other applications in TCMs.

Keywords Chinese herbs · MALDI-TOF MS · Traditional Chinese medicines (TCMs)

Contents

M. Lu and Z. Cai (✉)
Department of Chemistry, Hong Kong Baptist University, Kowloon Tong, Kowloon, Hong Kong SAR, China
e-mail: zwcai@hkbu.edu.hk

Abbreviations

2-DE	Two-dimensional gel electrophoresis
α-CHCA	α-Cyano-4-hydroxycinnamic acid
DHB	2,5-Dihydroxybenzoic acid
ESI	Electrospray ionization
HPLC	High performance liquid chromatography
IEC	Ion exchange chromatography
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
TCMs	Traditional Chinese medicines
TLC	Thin-layer chromatography
TOF	Time-of-flight

1 Introduction

Traditional Chinese medicines (TCMs) have been widely used in China for the prevention and/or treatment of human diseases for thousands of years. Due to their long historical clinical experience and reliable therapeutic efficacy, TCMs applications have also been extended to other countries, including Korea, Japan, India, and even European [1] and North American countries [2]. Chinese people have accumulated rich clinical experimental and human clinical data with a long history of the development of theory and clinical practice. However, compared to modern drugs that contain one or two active compounds with known concentration, TCMs are complex mixtures which usually contain hundreds of chemically different constituents. Moreover, the content of the active ingredients in TCMs may be influenced by breed, region of growth, season of harvest, and processing procedures. So far, there have been 12,806 medical resources discovered in China, including 11,145 medicinal plants, 1,581 medicinal animals, and 80 medicinal minerals [3]. In the Pharmacopoeia of the People's Republic of China (2010 edition), a total of 2,136 Chinese medicine products have been compiled [4].

In the process of "modernization" and "globalization" of TCMs, quality control, identification of active ingredients, and study of the toxicology of TCMs, many methods were developed for the analysis of TCMs. Thin-layer chromatography (TLC) is a simple, low-cost, versatile, and specific method for the identification of

herbal medicines [5]. Due to its high resolution, selectivity, and sensitivity, gas chromatography-mass spectrometry (GC-MS) is the most commonly used technique for analysis of liposoluble constituents, especially volatile/semi-volatile compounds and their metabolites in TCMs [6]. Liquid chromatography-mass spectrometry (LC-MS) is another promising technique for the quantitative analysis of bioactive compounds and their metabolites of herbal extract and medicines in biological fluids because the technique can be directly applied for analysis of nonvolatile and low thermal stable compounds without derivatization [7]. Capillary electrophoresis (CE) and capillary electrophoresis-mass spectrometry (CE-MS), with the advantages of high resolution, minimal sample and solvents consumption, short analysis time, and high separation efficiency, are very effective tools for analyzing TCMs and related compounds [8, 9]. Multi-dimensional chromatographic separation systems have emerged with obvious advantages (e.g., strong separation ability, high resolution, and high peak capacity) in analysis of complex samples [10]. In addition to chromatographic and hyphenated techniques, spectroscopic methods, including Fourier transform infrared spectroscopy (FT-IR), near-infrared spectroscopy (NIR), and nuclear magnetic resonance (NMR) spectroscopy, have also been widely used in the quality analysis of TCMs [11].

Matrix-assisted laser desorption/ionization (MALDI), a powerful soft ionization technique proposed in the late 1980s [12, 13], has become a preeminent technique for the analysis of complex samples. Applications of MALDI coupled with time-of-flight mass spectrometry (TOF MS) have been extended to various fields, especially in identification of large molecular compounds, such as peptides/proteins [14–16], lipids [17, 18], polymers [19, 20], oligonucleotides [21], and oligosaccharides [22]. Compared to other mass analyzers such as Fourier transform ion cyclotron (FT-ICR) MS, ion trap MS, quadrupole-based MS, and magnetic-sector MS, TOF MS is the most frequently used detector in MALDI because it lends itself naturally to ion production by a short laser pulse [23, 24]. Therefore the configuration of MALDI-TOF MS is characterized by a large mass range, high sensitivity for ion detection, high tolerance to salts and buffers, and simple and fast analysis [25]. In recent years, MALDI-TOF MS was also applied to analyze small molecular compounds (MW $<$ 1,000) followed by the introduction of various new matrices [26, 27].

This chapter reviews the advances of MALDI-TOF MS in TCMs analysis in the last decade. MALDI-TOF MS as a very useful and powerful tool in the identification of new bioactive ingredients in TCMs is discussed in Sect. 2.1. Applications of MALDI-TOF MS to the analysis of alkaloids that are the large type of bioactive compounds existing in most TCMs are reviewed in Sect. 2.2. Following this, the determination of small molecular compounds in TCMs by MALDI-TOF MS with new matrices is discussed in Sect. 2.3. Proteomic analysis associated with TCMs and the direct analysis of plant tissue by MALDI-TOF MS are reviewed in Sects. 2.4 and 2.5, respectively. In the last section (Sect. 2.6), other applications of MALDI-TOF MS in TCMs are summarized.

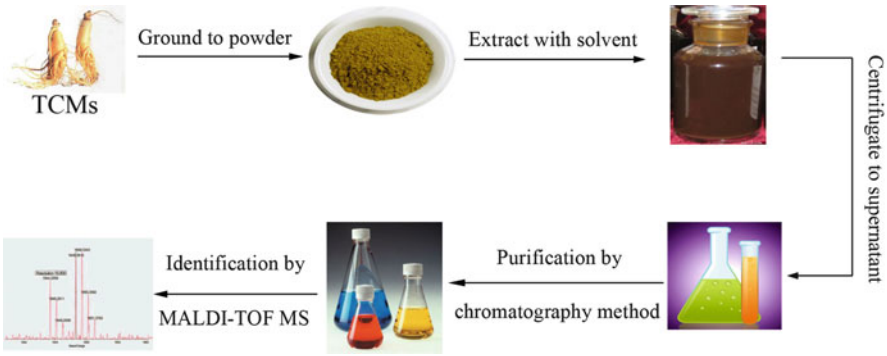


Fig. 1 Scheme for identification of new bioactive ingredients from TCMs by MALDI-TOF MS

2 Methods

2.1 Identification of New Bioactive Ingredients in TCMs

MALDI-TOF MS is a very useful and powerful tool that can be used to identify rapidly new bioactive ingredients in TCMs. Prior to the MALDI-TOF MS analysis, sample preparation procedures, including extraction, purification, and isolation, are usually required. To extract bioactive ingredients effectively from TCMs, herbal samples were usually cut into pieces and ground to powder, and then extracted with suitable extraction solvent. Supernatant obtained by centrifugation was then subjected to purification with chromatography methods, including gel filtration chromatography or size exclusion chromatography, gel electrophoresis, ion exchange chromatography (IEC), and RP-HPLC. The purified bioactive ingredients, such as peptides, proteins, anticoagulants, antifungals, and lectins, were identified by MALDI-TOF MS. A brief scheme for identification of bioactive ingredients in TCMs by MALDI-TOF MS is illustrated in Fig. 1.

Liu and coworkers [28] purified and identified three novel peptides with antioxidant properties from *Cornu Bubali* (water buffalo horn, WBH), a TCM that was used for dispelling heat, counteracting toxins, and relieving convulsions. In this research paper, consecutive chromatographic methods including gel filtration chromatography, IEC, and HPLC were applied to purify and isolate three novel antioxidant peptides from aqueous extract of WBH. The purified peptides were identified by MALDI-TOF/TOF MS in positive ion delayed extraction reflector mode with α-cyano-4-hydroxycinnamic acid (CHCA) as the matrix. The amino acid sequences of the peptides WBH-1, WBH-2, and WBH-3 were identified as hexapeptide Gln-Tyr-Asp-Gln-Gly-Val (708 Da), nonapeptide Tyr-Glu-Asp-Cys-Thr-Asp-Cys-Gly-Asn (1,018 Da), and dodecapeptide Ala-Ala-Asp-Asn-Ala-Asn-Glu-Leu-Phe-Pro-Pro-Asn (1,271 Da), respectively.

Two novel antifungal peptides named EAFP1 and EAFP2 were purified and identified by Huang and colleagues [29] in the bark of *Eucommia ulmoides* Oliv (Du-Zhong in Chinese). EAFPs were obtained from the isolation and purification of bark of *E. ulmoides* Oliv using chromatography methods. The molecular weight and sequence analysis were performed by MALDI-TOF MS combined with CPY time-dependent and concentration-dependent digestion. For MALD-TOF MS analysis, the matrix was prepared by dissolving CHCA in 50% ACN solution containing 0.1% TFA to a saturated solution. Two microliters (about 1–5 pmol/μL) of peptide dissolved in a 0.1% TFA solution in water was mixed with 20 μL of prepared matrix solution. Prior to the MALDI-TOF MS analysis, 1 μL of this mixture solution was deposited on the target and dried. The molecular weight of EAFP1 and EAFP2 were determined as 4,201.4 Da and 4,518.9 Da, respectively. A conclusion that both peptides contain ten cysteines forming five pairs of disulfides was obtained.

Cyclotides are macrocyclic peptides isolated from plants. The compounds typically contain 28–37 amino acids and are characterized by their head-to-tail cyclized peptide backbone and the interlocking arrangement of three disulfide bonds [30]. They were initially discovered based on the potent insecticidal activity, in addition to a range of other biological activities including anti-HIV, antimicrobial, and cytotoxic activities [31]. Wang and co-workers [32] isolated and identified five new and three known cyclotides that were shown to have anti-HIV activity from *Viola yedoensis*, an important Chinese herb from the Violaceae family that has been reported to contain potential anti-HIV agents. Before MALDI-TOF MS analysis, isolation and purification of cyclotides from *V. yedoensis* were performed. The desalted samples were mixed in a 1:1 ratio (v/v) with matrix consisting of a saturated solution of CHCA in 50% ACN (0.5% formic acid) and confirmed by MALDI-TOF MS.

Zhong et al. [33] developed a method for rapid isolation and purification of an anticoagulant from *Whitmania pigra*, a common traditional Chinese anticoagulant medicine in China. A new anticoagulant was isolated and purified from *W. pigra*. This novel component was named whitmanin and its molecular weight was determined as 8,608 Da by MALDI-TOF MS with sinapinic acid (SA) as matrix. *Rosa chinensis* (Yuejihua in Chinese) is a well-known ornamental plant, and its flowers are commonly used in TCM for treating catamenia disorder, trauma, and blood disorders. A total of 36 known and unknown phenolic antioxidants were simultaneously determined in methanolic crude extracts of dried *R. chinensis* flowers by LC coupled with single-quadrupole electrospray ionization MS (ESI-MS) and MALDI-quadrupole ion trap (QIT)-TOF MS [34]. MALDI-QIT-TOF MS was applied not only to confirm 31 of all 36 known and unknown phenolics isolated and identified by LC-ESI-MS, but also to identify tentatively two ellagitannins (rugosins B and C). Additionally, MALDI-QIT-TOF MS analysis only took a few minutes per run whereas LC-ESI-MS analysis took >100 min per run.

Clematis Montana lectin, a novel mannose-binding lectin with antiviral and apoptosis-inducing activities, was isolated from *Clematis montana* Buch.-Han stem (*Ranunculaceae*)[35]. *C. Montana* lectin was a homodimer of 11,968.9-Da subunits as determined by gel filtration and MALDI-TOF MS. The peptide mass

fingerprinting of *C. Montana* lectin was identified by MALDI-Q-TOF MS/MS, which indicated that *C. Montana* lectin may be a novel plant lectin. Yan and coworkers [36] isolated a novel homodimeric lectin (AMML) with antifungal activity from a Chinese herb, that is, the root of Astragalus mongholicus, by using a combination of ammonium sulfate fraction chromatography and IEC. The molecular mass of intact AMML was determined to be 66,396 Da by MALDI-TOF MS and 61.8 kDa by gel filtration, respectively. Experimental results also demonstrated that AMML was a dimeric protein composed of two identical subunits each with a molecular mass of 29.6 kDa.

Copper-zinc superoxide dismutase (Cu, Zn SOD) was extracted, purified, and characterized by Haddad and Yuan [37] from Radix lethospermi seed (RLS), a kind of TCM. Extraction and purification were carried out by consecutive chromatographic methods including ammonium sulfate fractionation, IEC, column chromatography, and a second step of IEC before MALDI-TOF MS analysis. The molecular weight of RLS Cu, Zn SOD was determined by following several methods. On gel filtration, a single peak was obtained with a molecular weight of 32 kDa for a whole enzyme. The rough size of the subunit determined by SDS-PAGE was 16 kDa. The exact mass of one subunit of SOD was determined as 15.166 kDa by MALDI-TOF MS with CHCA as matrix.

2.2 Analysis of Alkaloids in TCMs

Alkaloids, a group of compounds that mostly contain basic nitrogen atoms, are widely distributed in various living organisms including plants. Alkaloid-containing plants have been used by humans since ancient times for therapeutic and recreational purposes, especially TCM herbs. Therein, more than 2,000 alkaloids were identified since morphine, the first individual alkaloid, was isolated from the opium poppy plant by Friedrich Sertürner in 1804. Many alkaloids were discovered to have medicinal properties. TCMs containing tropane alkaloids have been used to treat asthma, chronic bronchitis, pain, and flu symptoms. However, not all TCM herbs can be directly used because they may contain toxics alkaloids, such as aconitum alkaloids in Fuzi. Thus, proper processing of the herb is usually required to reduce the amount of toxic alkaloids prior to use [38]. Many qualitative and quantitative methods have been developed for the purposes of quality control, forensic medicine, and therapeutic drug monitoring of TCMs [39–43]. MALDI-TOF MS, a relatively new analytical technique, with characteristics of easy sample preparation, high throughput, and strong identification ability, was also applied to the analysis of alkaloids in TCMs [44–50].

A method for direct analysis of alkaloid profiling in Chinese herbs tissue by using MALDI-TOF MS was established in our laboratory, which will be discussed in detail in Sect. 2.5 [45, 46]. Feng and Lu [47] developed a method for analyzing low molecular weight compounds and applied to the determination of carcinogenic

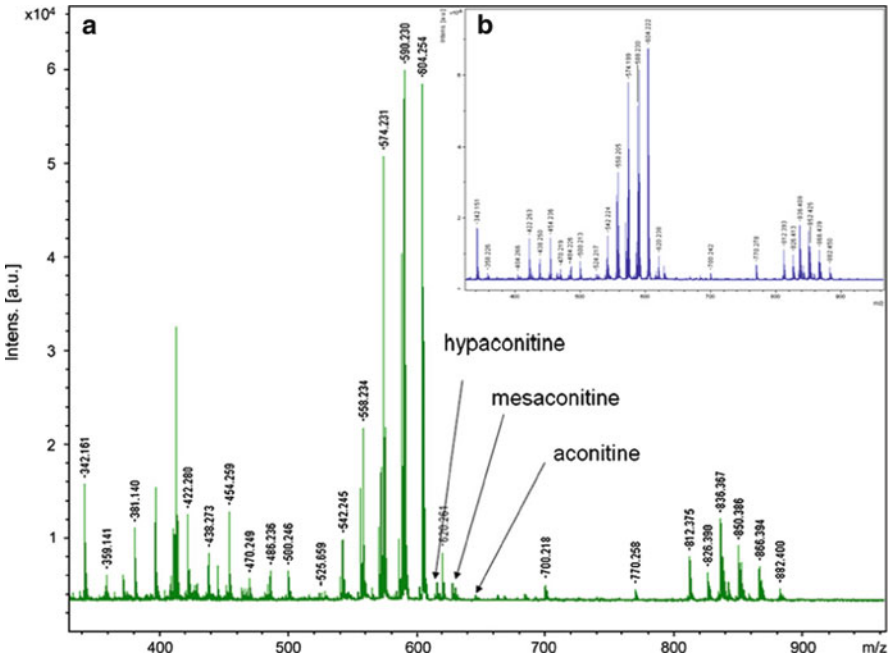


Fig. 2 (**a**) Typical MALDI-TOF MS spectrum of Aconitum alkaloids from *A. carmichaeli* root. (**b**) Typical MALDI-TOF MS spectrum of Aconitum alkaloids from direct tissue analysis. Copyright with permission from Elsevier Science B.V [49]

areca alkaloids by MALDI-TOF MS with a new matrix, which will be reviewed in the next section.

TLC combined with MALDI-TOF MS is a powerful technique for fast and high throughput analysis of compounds in complex samples. However, this technique usually suffers from dilution of the TLC bands resulting in decreased sensitivity and masking of signals in the low-mass region because of addition of matrix. A matrix-free TLC and laser desorption/ionization MS method was developed for separation and identification of medicinal alkaloids, berberine and palmatine from *Berberis barandana*, by Shariatgorji and coworkers [48]. Wang and coworkers [49] developed a high throughput and robust qualitative MALDI-TOF MS method for profiling alkaloids in Fuzi, the processed lateral roots of the TCM *Aconitum carmichaeli* Debx (*A. carmichaeli*). Under optimized conditions, a typical MALDI-TOF MS profiling spectrum of a Fuzi extract using DHB as matrix was obtained and is shown in Fig. 2. A solid sample analysis method was also investigated by applying the matrix solution onto the powdered samples directly on the sample plate. The ratio of sample mass to DHB matrix volume was proved to have a significant impact on the mass spectrum. When the ratio was 10 μg to 0.5 μL, similar mass spectra to that with the conventional solvent extraction method was achieved (Fig. 2a, b). Furthermore, the semi-quantitative potential of MALDI-TOF MS was studied and compared by using LC-MS as reference.

A MALDI-TOF MS method for rapid and direct profiling of alkaloids in medical herbs was developed by Lu and colleagues [50]. The dry herbs were ground to powder and passed through a stainless steel sieve, mixed with DHB matrix solution to form a homogeneous suspension, and directly subjected to MALDI-TOF MS analysis. DHB was considered as optimized matrix for analysis of alkaloids in TCM herbs by MALDI-TOF MS, which agreed with previous work published by Wang and coworkers [49].

2.3 *Determination of Small Molecular Weight Compounds in TCMs with New Matrices*

Since it was first introduced in the late 1980s [12, 13], MALDI-TOF MS has been successfully applied to the analysis of various types of large molecules such as peptides, proteins, and polymers. However, using conventional MALDI-TOF MS for the analysis of small molecular weight compounds is difficult since it suffers from strong interference in the low-mass range from the matrices. To overcome the drawbacks of MALDI-TOF MS in analysis of small molecular compounds (<1,000 Da), many methods and matrices, including various matrix-free desorption/ionization MS [26, 51, 52], various carbon nano-materials [27, 53, 54], as well as other inorganic [55, 56] and organic [57, 58] materials, were introduced in recent years. A method based on oxidized carbon nanotubes as matrix for analysis of a TCM *Psoralea corylifolia* by MALDI-TOF MS was developed by Chen and his coworkers [59]. By using oxidized carbon nanotubes as matrix, all 11 fractions were analyzed by MALDI-TOF MS in both negative and positive ionization modes. A large number of individual species of small molecular weight compounds were detected in the fractions with high detection accuracy and sensitivity by MALDI-TOF MS (Fig. 3). A total of more than 188 components were isolated and identified from the extract of *P. corylifolia* by integration of IEC fractionation with RPLC-APCI MS and MALDI-TOF MS. Among these compounds, a total of 92 dominant molecular ion peaks could only be detected by MALDI-TOF MS. Therefore, MALDI-TOF MS with matrix of oxidized carbon nanotubes is a complementary technique with LC-APCI MS for detection and analysis of small molecules in a complex sample. Chen and colleagues [60] also established a hyphenated method (including IEC, RPLC-APCI-MS, and MALDI-TOF MS) for isolation and identification of components in a TCM of Honeysuckle. For MALDI-TOF MS analysis, oxidized carbon nanotubes were used as matrix. A total of 262 components were detected from the extract of Honeysuckle by UV detector, APCI-MS, or MALDI-TOF MS. Among these compounds, 145 components only could be detected by MALDI-TOF MS alone. Compared with LC-UV and LC-APCI-MS, MALDI-TOF MS using oxidized carbon nanotubes as matrix exhibited a very powerful ability in the identification of low-mass compounds in complex samples, such as TCMs and biological samples. Using

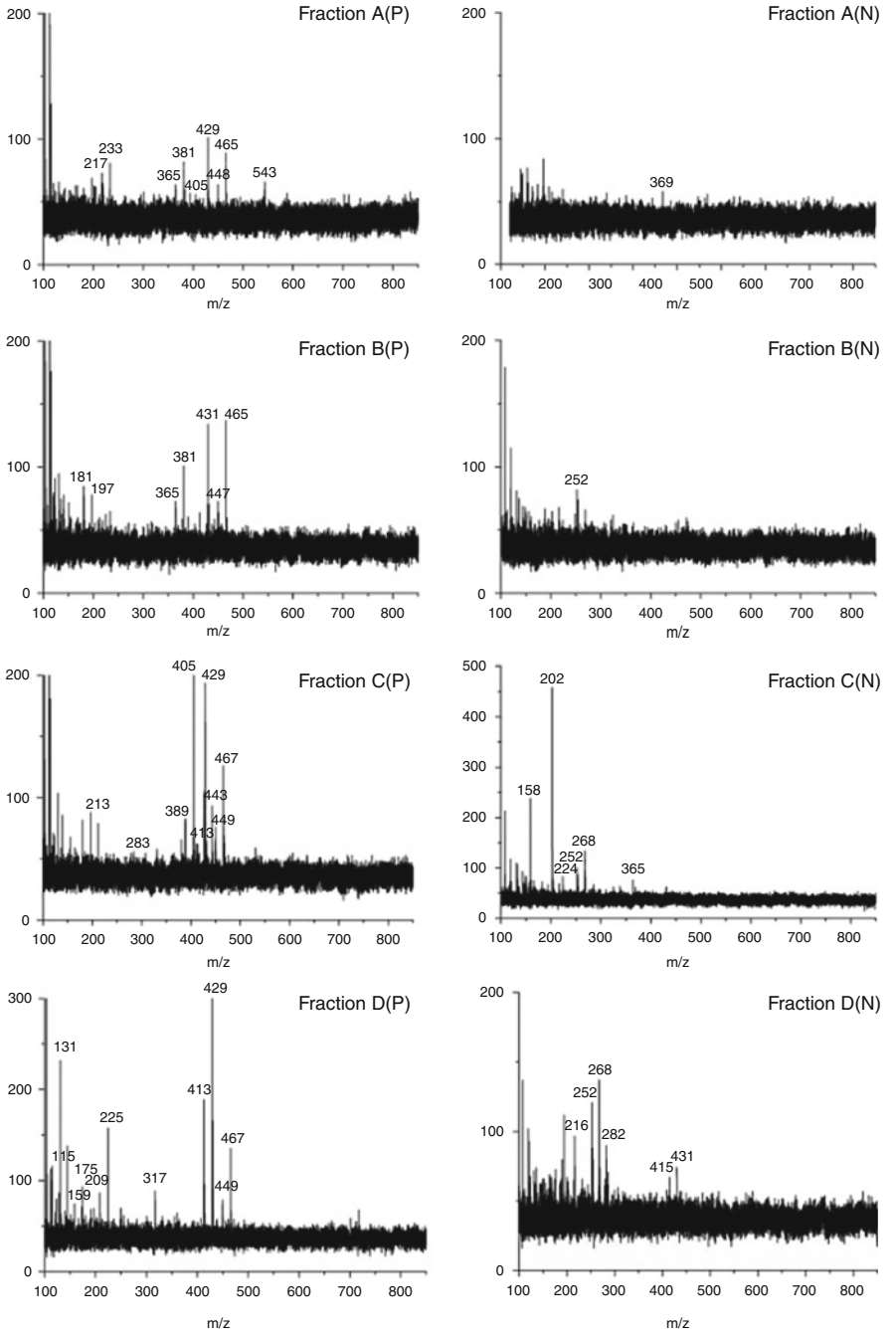


Fig. 3 (continued)

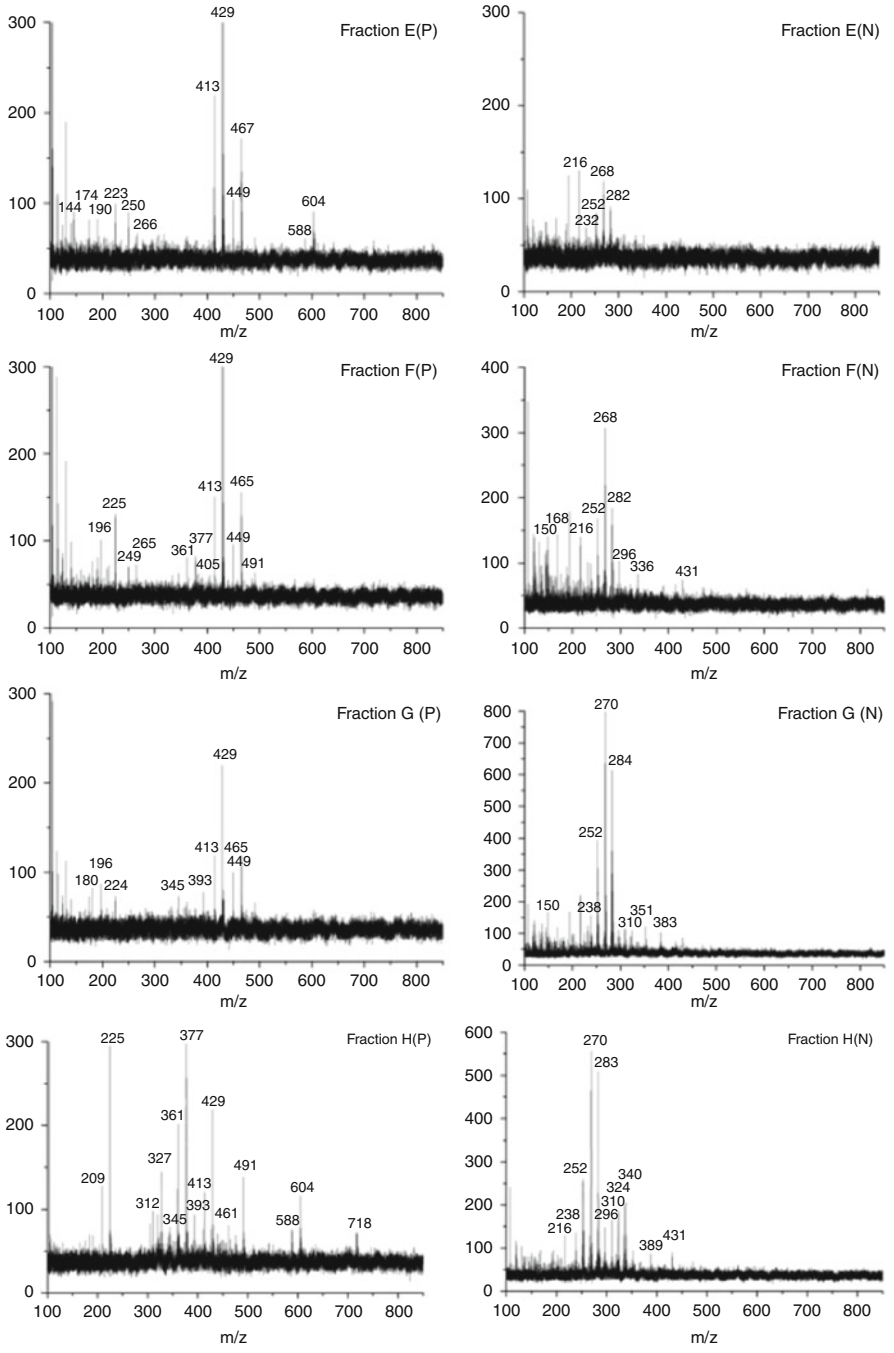


Fig. 3 (continued)

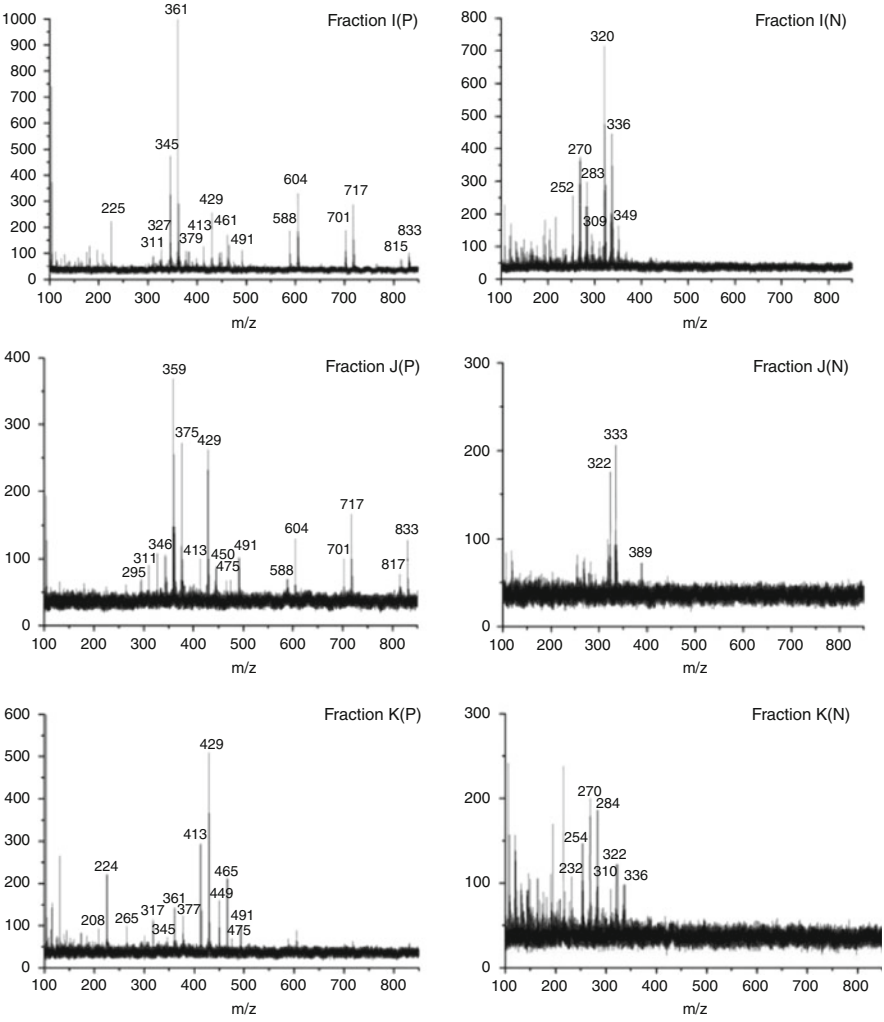


Fig. 3 MALDI-TOF MS spectra of the fractions A to K obtained from SCX column separation of *Psoralea corylifolia* extract. The mass spectra of all fractions were acquired with laser power adjusted to slightly above the threshold energy for all of the components with oxidized carbon nanotubes as matrix and detected in positive ion mode (P) and negative ion mode (N). Copyright with permission from Elsevier Science B.V [59]

oxidized carbon nanotubes as matrix, Pan et al. [53] developed a method for quantitative determination of the concentrations of jatrorrhizine (8.65 mg/mL) and palmatine (10.4 mg/mL) in an extract of *Coptis chinensis* Franch. Hu and his coworkers [61] established a MALDI-TOF MS using carbon nanotubes for the analysis of low-mass compounds in environmental samples. In this article, two arsenic speciations ($HAsO_2$ and H_3AsO_4) in the extracts of TCMs were directly

detected without any further separation and purification by MALDI-TOF MS with carbon nanotubes as matrix.

Feng and Lu [47] developed a MALDI-TOF MS method with a new matrix "7-mercapto-4-methylcoumarin" for analyzing low molecular mass compounds and it was applied to the determination of carcinogenic areca alkaloids. Furthermore, the new matrix was also used for determination of the signals of arecoline and arecaidine in the MALDI imaging experiment. The established method was succeeded to trace analysis of arecoline in human plasma at sub-micromolar level.

2.4 Proteomic Analysis Associated with TCMs

TCMs are great sources in which to discover new bioactive compounds with anticancer, antitumor, antifungal, and/or other biological activities [62–64]. To elucidate their working mechanism(s) in the treatment of various diseases, a comparative proteomics analysis was usually carried out by 2-DE and MALDI-TOF MS. Comparisons were made by gel images between TCMs treated samples and untreated controls. Differentially expressed proteins (usually greater than twofold difference) between treated samples and untreated controls that were excised from one or more gels were identified by MALDI-TOF MS.

Some TCMs, such as saponins, have great value as potent cancer prevention and chemotherapeutic agents. However, the active mechanisms of these compounds were not clear. Wang and colleagues [65] used a proteomic method to examine the cytotoxic effect of dioscin, a glucoside saponin, on human myeloblast leukemia HL-60 cells. Thirty-nine differentially expressed proteins after dioscin treatment for 24 h were identified by separation of the microsomal fraction and subsequently analyzed by MALDI-TOF MS. Ge and co-workers [66] applied proteomics to analyze the arsenic trioxide (ATO)-induced protein alterations in multiple myeloma (MM) cell line U266 and then investigated the molecular pathways responsible for the anticancer actions of ATO. It is noted that 84 differentially expressed proteins were excised from 2-DE and identified by MALDI-TOF/TOF MS followed by database interpretation in this study. Among them, 76 proteins including 29 up-regulated and 47 down-regulated proteins were identified successfully.

Tubeimoside-1 (TBMS1) is a triterpenoid saponin extracted from *Bolbostemma paniculatum* (Maxim.) Franquet (Cucurbitaceae), a Chinese herb with anticancer potential called "Tu Bei Mu". Xu et al. [67] studied the cytotoxic effects of TBMS1 on HeLa cells by a comparative proteomics approach to delineate the possible molecular basis of TBMS1-induced cancer cell death. 2-DE and MALDI-TOF MS/MS were applied to identify altered proteins related to energy metabolism and protein synthesis in this research work. To elucidate the anti-tumor mechanism of Rhizoma Paridis total saponin (RPTS) from the herb Rhizoma Paridis, a proteomic analysis for studying the change of proteins between the control and RPTS treated cells was carried out using MALDI-TOF MS by Cheng et al. [68]. The digests obtained were spotted onto the anchorchips with 1 μL of analyte in duplicate and

0.05 μL of 2 mg/mL CHCA in 0.1% TFA/33% ACN which contained 2 mM ammonium phosphate. The experiments were performed on a TOF Biflex IV mass spectrometer (Bruker Daltonics) in positive ion reflectron mode. The data were searched against NCBInr database by MASCOT search engine. More than 50 proteins showed a significant change between control (0.01% DMSO) and RPTS (IC50 approximately 10 μg/mL) treated cells after 48 h. Twelve proteins were identified by MALDI-TOF MS using peptide fingerprinting from 15 protein spots (density difference greater than twofold between the control and RPTS-treated groups).

Sanqi, also called tianqi, the root of *Panax notoginseng*, is an important medicinal substance in TCM with blood-stanching and blood-quickening actions. Notoginsengnosides (NG) isolated from Sanqi could inhibit ADP-induced platelet aggregation of rat washed platelets. To identify the possible target proteins of NG in platelets, 2-DE-based comparative proteomics was performed and proteins altered in expressional level after NG treatment were identified with MALDI-TOF MS/MS by Yao and coworkers [69]. The proteins were digested with trypsin and the peptides were analyzed using an ABI 4700 Proteomics Analyzer with delayed ion extraction. The MS data obtained were investigated using the MASCOT search engine against the NCBI protein sequence database with a score of more than 50.

To clarify the mechanism regarding the concomitant use of berberine (BBR) and fluconazole (FLC) could provide a synergistic action against FLC-resistant *Candida albicans* (C. albicans) clinical strains in vitro, a comparative proteomic study was performed in untreated control cells and cells treated with FLC and/or BBR in two clinical strains of *C. albicans* resistant to FLC by Xu et al. [70]. The peptides obtained were spotted onto a MALDI target and overlaid with 0.8 μL of matrix solution (CHCA in 0.1% TFA and 50% ACN). The samples were analyzed on an Applied Biosystems TOF-TOF Proteomics Analyzer in positive reflection mode. The MS and MS/MS spectra were searched using the MASCOT engine to identify the proteins. A total of 16 differentially expressed proteins, most of which were related to energy metabolisms (e.g., Gap1, Adh1, and Aco1) were identified by 2-DE and MALDI-TOF MS in this research work.

2.5 *Direct Analysis of Plant Tissue*

In recent years, direct tissue analysis has gained more and more attention because it does not require complicated and tedious extraction and purification procedures prior to MS analysis. Although new techniques have been introduced recently, including desorption electrospray ionization MS (DESI MS) [71] and surface-enhanced laser desorption ionization TOF MS (SELDI TOF MS) [72], most reports on direct analysis were focused on MALDI-TOF MS [14, 17, 22, 73–76]. Up to now there are two types of experimental methods using MALDI-TOF MS for direct tissue analysis, namely profiling and imaging. According to the type of experiment being performed, there are also two different methods of matrix deposition techniques used: spotting and coating. The profiling experiments focus on comparison of different regions in tissue by analyzing several spots (typically

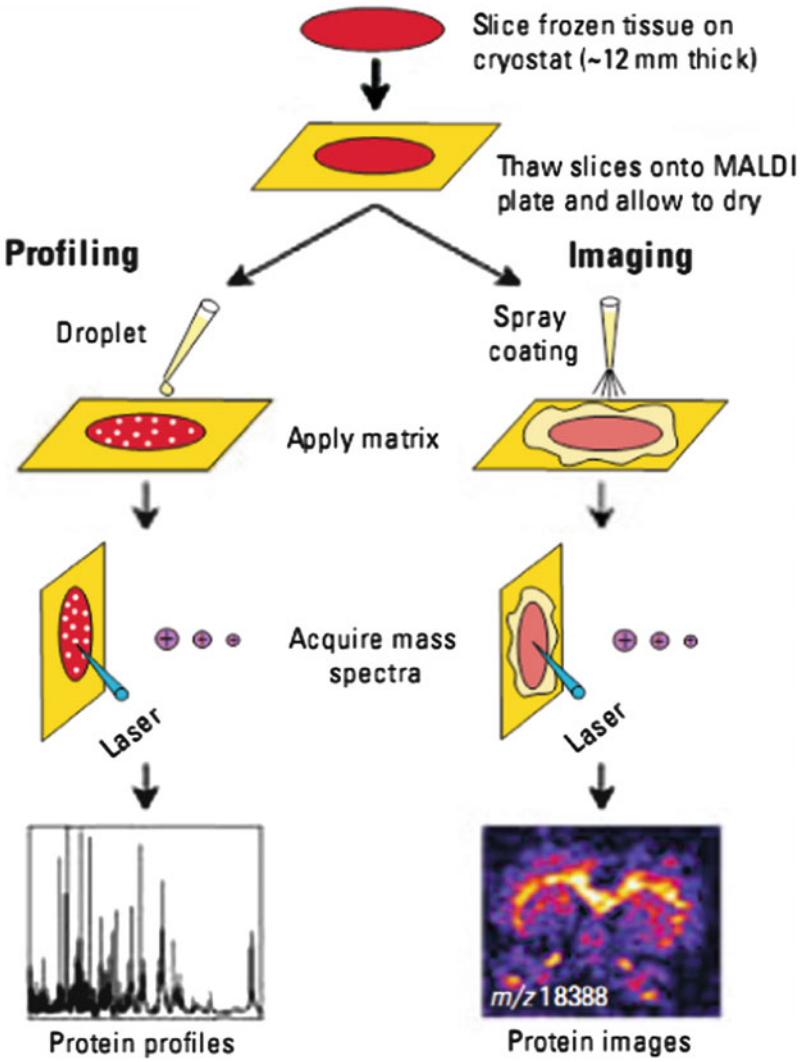


Fig. 4 Procedures involved in profiling and imaging MS of mammalian tissue samples. Copyright with permission from American Chemical Society [77]

ten spots) per tissue section from large numbers of samples. Thus, for profiling experiments, the matrix is deposited directly onto specific regions of interest in the tissue section. For imaging experiments, to obtain high-resolution two-dimensional images, the experiments are usually run on one or two prototypical samples. Therefore, the matrix must be homogeneously distributed over the entire tissue section and the individual mass spectra are automatically acquired in a raster pattern across the entire tissue section. A general scheme for profiling and imaging sample preparation is shown in Fig. 4 [77].

The application of MALDI-TOF MS to direct profiling and imaging has made great progress in tissue analysis. Most of the reports about direct profiling and imaging tissue were focused on animal tissue for proteins and peptides analysis. Only a few research papers for direct profiling and imaging plant tissues of TCMs were reported, probably because most of the compounds in TCMs have a molecular weight below 1,000 Da. In this mass range it is difficult to identify analytes by using traditional MALDI-TOF MS because the signals may be interfered with by matrix background ions. Direct analysis of plant tissue by MALDI-TOF MS has become more and more applicable with the development of many new matrices for analysis of small molecular compounds. A method for the direct determination of alkaloid profiling in plant tissue by MALDI-TOF MS was developed in our laboratory [45]. Four commonly used Chinese medicinal herbs including Aconitum Carmichaeli Debx. (Fuzi in Chinese), Processed Fuzi, *Coptis chinensis* Franch. (Huanglian in Chinese), and Corydalis yanhusuo W.T.Wang (Yanhusuo in Chinese) were studied for herb differentiation and explanation of the significant differences in their toxicities. Briefly, plant tissues were cut into slices with a thickness of 10–20 μm and then these slices were adhered to MALDI target plate. The matrix was deposited onto the tissue surface directly before the sample plant was placed in a vacuum desiccator to improve crystal homogeneity prior to the MALDI-TOF MS analysis. Among all commonly used matrices, CHCA and DHB were found to be effective for analyzing the low-molecular-weight compounds in complex samples [78, 79]. In optimized conditions, alkaloid profiles in Fuzi tissue were investigated by MALDI-TOF MS with CHCA as matrix and the profiling spectrum is illustrated in Fig. 5a. To study the differentiation of Fuzi and Processed Fuzi as well as their toxicity, the direct analysis of Processed Fuzi under the same experimental conditions was performed and the spectrum is shown in Fig. 5b. The usability of direct analysis on plant tissue for detecting the alkaloid profiles indicated that MALDI-TOF MS could be applicable for the discovery of new compounds. In addition, a rapid and straightforward method for direct alkaloid profiling in crude and processed Strychnos nux-vomica seeds by MALDI-TOF MS was also used in our laboratory [46]. Alkaloid profiles in tissues from different parts (endosperm and epidermis) of crude Semen Strychni and the tissues from different heating processes (sand and oil) were analyzed and differentiated by direct MALDI-TOF MS (data not shown here). The spectrum of alkaloid profiles obtained from MALDI-TOF MS could provide much valuable information for the differentiation of crude and processed Strychnos nux-vomica seeds and for explanation of the significantly different toxicity.

A method for direct spatial profiling of phytochemicals and secondary metabolites in integrated herbal tissue by MALDI-TOF MS was developed by Ng and coworkers [80]. Experiment demonstrated that among all the different matrices used, including CHCA, SA, and DHB, CHCA could form relatively uniform layers of crystals on the stem tissue and assist the desorption/ionization of the alkaloids effectively upon N_2 laser ablation. For example, the abundance of alkaloid ions desorbed with CHCA matrix is ten times that of SA and DHB matrices. The established method was applied to determine the spatial distributions

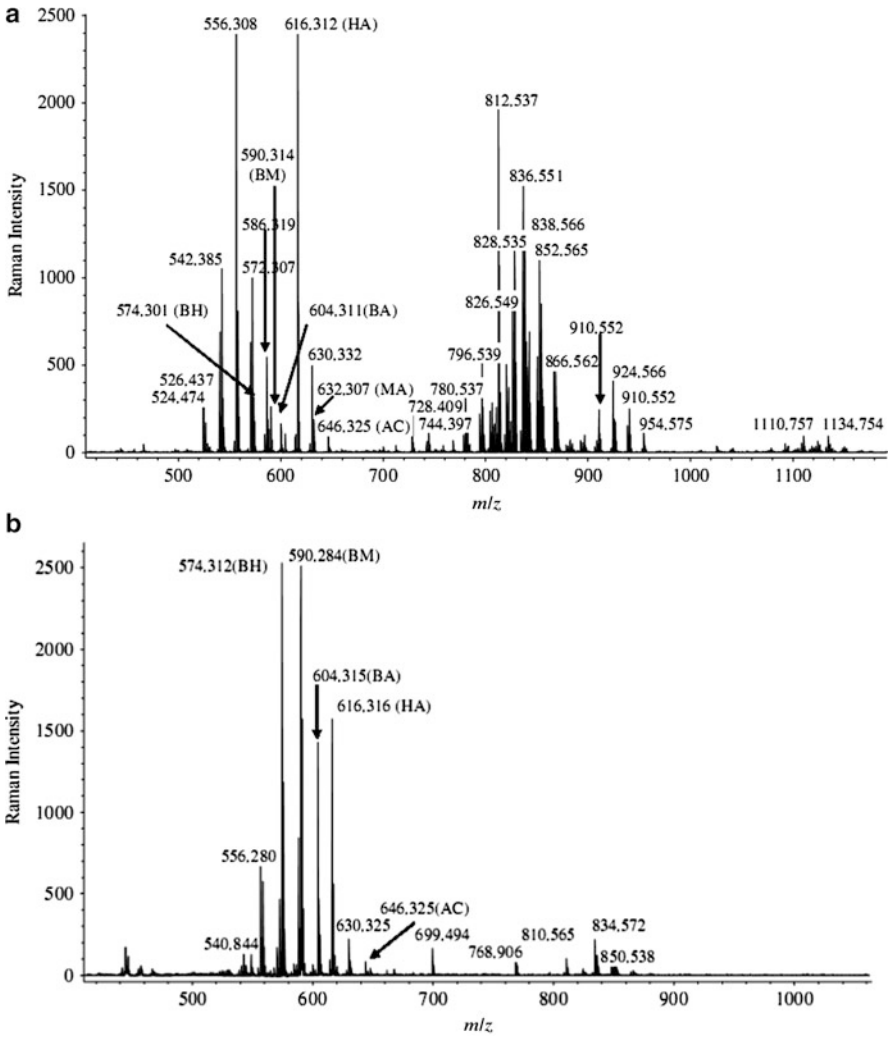


Fig. 5 MALDI-TOF MS spectra of alkaloids from the direct analysis of Fuzi root tissue (**a**) and Processed Fuzi (**b**) deposited with CHCA in 50:50 acetonitrile/0.1% TFA as matrix, about 20 μL/cm^2 under 22% laser energy. Copyright with permission from John Wiley and Sons Ltd [45]

of the metabolites and measure semiquantitatively their relative abundances in different tissue regions. A mass spectrum for direct desorption/ionization of secondary metabolites (alkaloids) from the cortex region of CHCA-coated stem tissue of Sinomenium acutum by MALDI-TOF MS is illustrated in Fig. 6. Compared with other techniques for analysis of Chinese herbs, this method showed some distinguished merits including simple, fast, clear, low sample consumption without any extraction process.

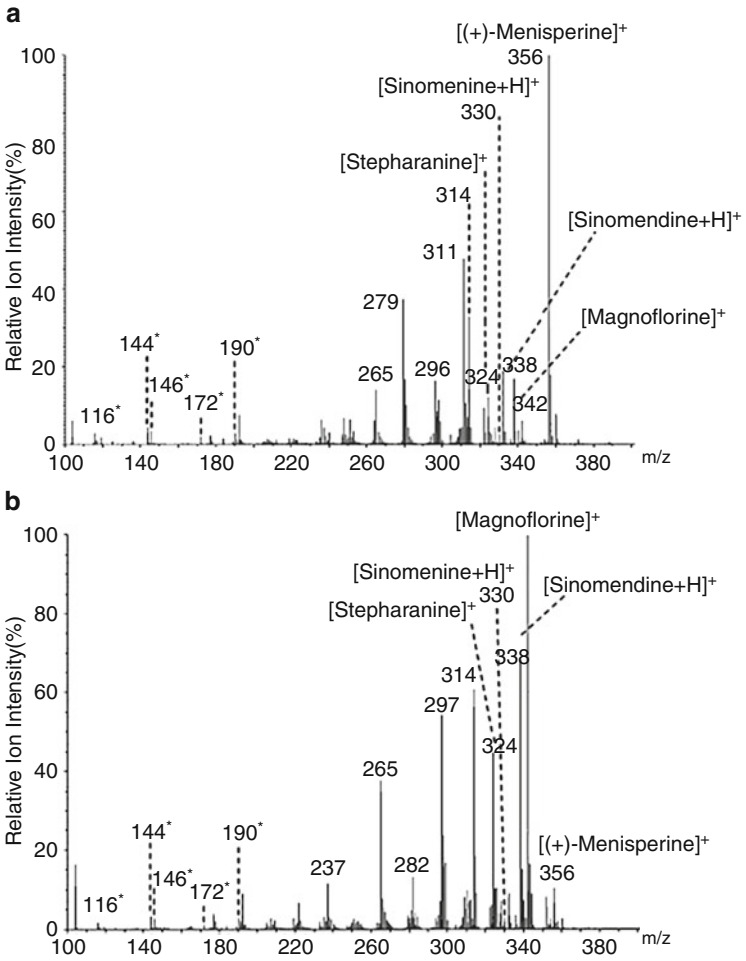


Fig. 6 Direct desorption/ionization of secondary metabolites (alkaloids) from the cortex region of CHCA-coated stem tissue of Sinomenium acutum. Samples were collected from Shaanxi (**a**) and Anhui (**b**) provinces, China. The solvent composition of CHCA solution (10 mg/mL) used for the matrix deposition was 80 vol.% acetone and 20 vol.% H_2O. (Asterisks denote ion peaks from CHCA matrix). Copyright with permission from The American Chemical Society [80]

2.6 Other Applications of MALDI-TOF MS in TCMs

Chinese gall (Wubeizi), a conventional TCM that contains high levels of gallotannins, has been mainly used as astringent, haemostatic, antiphlogistic, and antiseptic agents. Zhu and co-workers [81] demonstrated that MALDI-QIT-TOF MS

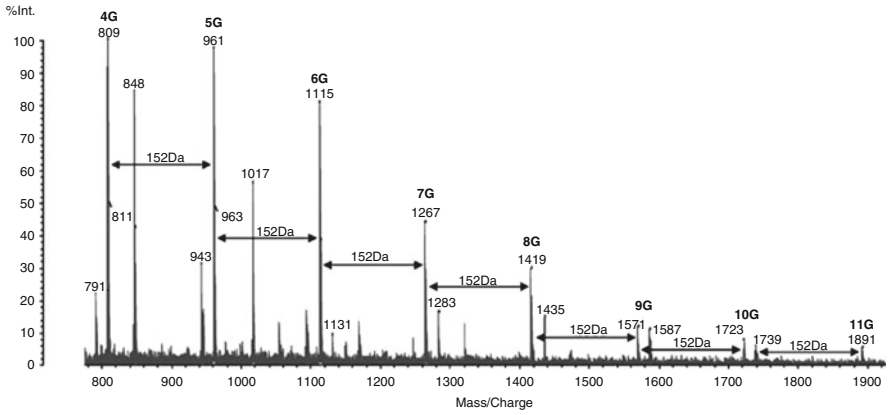


Fig. 7 MALDI-QIT-TOF MS positive ion spectrum of the gallotannins with DP of 4G to 11G in the crude extract of Chinese galls. Copyright with permission from John Wiley and Sons Ltd [81]

was a powerful and efficient technique for rapid direct analysis of the gallotannins in the crude extract of Chinese gall without any troublesome sample pretreatments. Several matrices such as DHB, 2',4',6'-trihydroxyacetophenone (THAP), and CHCA were initially examined for the measurement of the crude extract from Chinese galls by MALDI-QIT-TOF MS. Experimental results demonstrated that the use of THAP provided excellent peak intensity. A series of gallotannins with 4–11 galloyl units were identified in Chinese galls by MALDI-QIT-TOF MS. Figure 7 shows a typical MALDI-QIT-TOF positive ion spectrum of the gallotannins in the crude extract of Chinese galls by using THAP as matrix.

Glycyrrhizin (GC) is the main sweet tasting compound from liquorice root, which has protein kinase inhibitory, antiulcer, and antiviral activities. It is known that the hapten number in an antigen conjugate is important for immunization against small molecular compounds. Shan and coworkers [82] reported an analysis of hapten-carrier protein conjugates by MALDI-TOF MS that can directly describe the suitability of hapten number for immunization. By using BAS with molecular weight 66,433 Da as conjugate, a broad peak coinciding with the conjugate of GC and BSA appeared at *m*/*z* 70021 in the MALDI-TOF MS spectrum. Therefore the calculated value of the GC component (MW 823) is 3,588, indicating that at least four molecules of GC conjugated with the BAS.

Lin wa pi, the dried skin of the Heilongjiang brown frog, *Rana amurensis*, is commonly used as an ingredient of many medicines, as a general tonic, and as a topical antimicrobial/wound dressing. By using RP-HPLC and MALDI-TOF MS, Zhou and colleagues [83] identified components of the peptidome and transcriptome of the cutaneous granular gland, the source of skin-derived bioactive peptides in *Lin wa pi*. The molecular masses of polypeptides in fractionation of skin extract were detected by using MALDI-TOF MS in positive detection mode with CHCA as the matrix (data not shown).

3 Conclusions and Perspectives

MALDI-TOF MS, as a useful and powerful tool, has been successfully applied to solve a wide range of problems from TCMs analysis, such as rapid identification of new bioactive compounds, proteomics analysis to elucidate activity mechanism of TCMs, and direct tissue analysis with easy sample preparation. Compared with other techniques for analysis of TCMs, MALDI-TOF MS possesses some prominent advantages, including fast analysis, high sensitivity, low sample consumption, relative high tolerance towards salts and buffers, possibility to store sample on the target plate, and the ability of direct tissue analysis. Despite the higher salt and buffer tolerance of MALDI-TOF MS compared to ESI, the technique has limitations for qualitative and quantitative analysis of low molecular weight compounds (<1,000 Da). Poor reproducibility (sample to sample and shot to shot) is the main problem that hampers the application of the MALDI technique to the analysis of small molecular weight compounds. The presence of "hot spots" that comes from the co-crystallization process is the main cause of the poor reproducibility of MALDI-TOF MS. Various approaches have been developed to improve reproducibility, which can be divided into two main groups, namely the enhancement of homogeneous matrix crystallization and the averaging out of variations in instrument response [84]. Another challenge for analysis of low molecular weight compounds by MALDI-TOF MS is the interferences from matrix in low-mass range. To suppress or eliminate the interferences from matrix, various matrix-free methods and new matrices that are suitable for analysis of small molecular compounds were introduced, including desorption on porous silicon [51], sol-gel derived film [85], carbon nanostructure [54], and so on.

Direct analysis of plant tissue is another promising direction for analyzing TCMs by MALDI-TOF MS. The compositions of TCMs are very complicated and they usually contain hundreds of chemical constituents including flavonoids, alkaloids, phenols, and so on. The inherent advantages of MALDI-TOF MS, such as high throughput, easy sample preparation, and rapid analysis ability, make this technique most suitable for the direct analysis of complex samples. Although little work on the direct analysis of TCM tissue has been reported so far, it is believed that direct profiling or imaging of plant tissue by MALDI-TOF MS will be more and more applicable in future TCMs analysis.

Acknowledgements We are thankful for the financial support from the National Nature Sciences Foundation of China (No. 20928005 and 21175025)

References

1. Melchart D, Hager S, Hager U, Liao JZ, Weidenhammer W, Linde K (2004) Complement Ther Med 12:71–78
2. Bent S, Ko R (2004) Am J Med 116:478–485

3. Chen YZ, Chen SY (2003) Introduction of Chemical Methods in Study of Modernization of Traditional Chinese Medicines. Science Press, Beijing
4. Pharmacopoeia of the People's Republic of China (2010 edition) National Pharmacopoeia Committee Chemical Industry Publishing, Beijing
5. Yang L, Xu SJ, Feng QR, Liu HP, Tian RT, Xie PS (2009) J Liq Chromatogr Relat Technol 32:2893–2905
6. Ye JS (2009) Chem Pap 63:506–511
7. Gray MJ, Chang D, Zhang Y, Liu JX, Bensoussan A (2010) Biomed Chromatogr 24:91–103
8. Ganzera M (2008) Electrophoresis 29:3489–3503
9. Fu XF, Liu Y, Li W, Pang NN, Nie HG, Liu HW, Cai ZW (2009) Electrophoresis 30:1783–1789
10. Gao MX, Deng CH, Lin S, Hu FL, Tang J, Yao N, Zhang XM (2007) J Sep Sci 30:785–791
11. Jiang Y, David B, Tu PF, Barbin Y (2010) Anal Chim Acta 657:9–18
12. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T (1988) Rapid Commun Mass Spectrom 2:151–153
13. Karas M, Hillenkamp F (1988) Anal Chem 60:2299–2301
14. Qiao L, Roussel C, Wan JJ, Yang PY, Girault HH, Liu BH (2007) J Proteome Res 6:4763–4769
15. Karbassi ID, Nyalwidhe JO, Wilkins CE, Cazares LH, Lance RS, Semmes OJ, Drake RR (2009) J Proteome Res 8:4182–4192
16. Debois D, Bertrand V, Quinton L, De Pauw-Gillet MC, De Pauw E (2010) Anal Chem 82:4036–4045
17. Jackson SN, Wang HYJ, Woods AS (2005) Anal Chem 77:4523–4527
18. Hidaka H, Hanyu N, Sugano M, Kawasaki K, Yamauchi K, Katsuyama T (2007) Ann Clin Lab Sci 37:213–221
19. Trimpin S, Wijerathne K, McEwen CN (2009) Anal Chim Acta 654:20–25
20. Sachon E, Matheron L, Clodic G, Blasco T, Bolbach G (2010) J Mass Spectrom 45:43–50
21. Kong XL, Huang LCL, Liau SCV, Han CC, Chang HC (2005) Anal Chem 77:4273–4277
22. Laremore TN, Zhang FM, Linhardt RJ (2007) Anal Chem 79:1604–1610
23. Fenyo D, Beavis RC (2008) Mass Spectrom Rev 27:1–19
24. Zenobi R (2010) Ionization processes and detection in MALDI-MS of polymers, 1st edn. Wiley, New York
25. Gross J, Strupat K (1998) Trac Trends Anal Chem 17:470–484
26. Cohen LH, Gusev AI (2002) Anal Bioanal Chem 373:571–586
27. Rainer M, Najam-ul-Haq M, Huck CW, Vallant RM, Heigl N, Hahn H, Bakry R, Bonn GK (2007) Recent Pat Nanotechnol 1:113–119
28. Liu R, Wang M, Duan JA, Guo JM, Tang YP (2010) Peptides 31:786–793
29. Huang RH, Xiang Y, Liu XZ, Zhang Y, Hu Z, Wang DC (2002) FEBS Lett 521:87–90
30. Craik DJ, Daly NL, Bond T, Waine C (1999) J Mol Biol 294:1327–1336
31. Daly NL, Rosengren KJ, Craik DJ (2009) Adv Drug Deliv Rev 61:918–930
32. Wang CKL, Colgrave ML, Gustafson KR, Ireland DC, Goransson U, Craik DJ (2008) J Nat Prod 71:47–52
33. Zhong S, Cui Z, Sakura N, Wang D, Li JL, Zhai Y (2007) Biomed Chromatogr 21:439–445
34. Cai YZ, Xing J, Sun M, Zhan ZQ, Corke H (2005) J Agric Food Chem 53:9940–9948
35. Peng H, Lv H, Wang Y, Liu YH, Lia CY, Meng L, Chen F, Bao JK (2009) Peptides 30:1805–1815
36. Yan QJ, Jiang ZQ, Yang SQ, Deng W, Han LJ (2005) Arch Biochem Biophys 442:72–81
37. Haddad NIA, Yuan QS (2005) J Chromatogr B Analyt Technol Biomed Life Sci 818:123–131
38. Chen JH, Lee CY, Liau BC, Lee MR, Jong TT, Chiang ST (2008) J Pharm Biomed Anal 48:1105–1111
39. Yu LJ, Xu Y, Feng HT, Li SFY (2005) Electrophoresis 26:3397–3404
40. Chen JH, Zhao HQ, Wang XR, Lee FSC, Yang HH, Zheng L (2008) Electrophoresis 29:2135–2147

41. Ranieri TL, Ciolino LA (2008) Phytochem Anal 19:127–135
42. Li HJ, Jiang Y, Li P (2009) J Chromatogr A 1216:2142–2149
43. Zhang J, Jin Y, Dong J, Xiao YS, Feng JT, Xue XY, Zhang XL, Liang XM (2009) Talanta 78:513–522
44. Sun WX, Liu SY, Liu ZQ, Song FR, Fang SP (1998) Rapid Commun Mass Spectrom 12:821–824
45. Wu W, Liang ZT, Zha ZZ, Cai ZW (2007) J Mass Spectrom 42:58–69
46. Wu W, Qiao CF, Liang ZT, Xu HX, Zhao ZZ, Cai ZW (2007) J Pharm Biomed Anal 45:430–436
47. Feng CH, Lu CY (2009) Anal Chim Acta 649:230–235
48. Shariatgorji M, Spacil Z, Maddalo G, Cardenas LB, Ilag LL (2009) Rapid Commun Mass Spectrom 23:3655–3660
49. Wang J, van der Heijden R, Spijksma G, Reijmers T, Wang M, Xu G, Hankemeier T, van der Greef J (2009) J Chromatogr A 1216:2169–2178
50. Lu L, Yue H, Song FR, Tsao R, Liu ZQ, Liu SY (2010) Chem Res Chin Univ 26:11–16
51. Wei J, Buriak JM, Siuzdak G (1999) Nature 399:243–246
52. Peterson DS (2007) Mass Spectrom Rev 26:19–34
53. Pan CS, Xu SY, Hu LG, Su XY, Ou JJ, Zou HF, Guo Z, Zhang Y, Guo BC (2005) J Am Soc Mass Spectrom 16:883–892
54. Najam-ul-Haq M, Rainer M, Szabo Z, Vallant R, Huck CW, Bonn GK (2007) J Biochem Biophys Methods 70:319–328
55. Castro AL, Madeira PJA, Nunes MR, Costa FM, Florencio MH (2008) Rapid Commun Mass Spectrom 22:3761–3766
56. Lorkiewicz P, Yappert MC (2009) Anal Chem 81:6596–6603
57. Woldegiorgis A, von Kieseritzky F, Dahlstedt E, Hellberg J, Brinck T, Roeraade J (2004) Rapid Commun Mass Spectrom 18:841–852
58. Zhang S, Liu JA, Chen Y, Xiong SX, Wang GH, Chen J, Yang GQ (2010) J Am Soc Mass Spectrom 21:154–160
59. Chen XG, Kong L, Su XY, Pan CS, Ye ML, Zou HF (2005) J Chromatogr A 1089:87–100
60. Chen XG, Hu LH, Su XY, Kong L, Ye ML, Zou HF (2006) J Pharm Biomed Anal 40:559–570
61. Hu LG, Xu SY, Pan CS, Yuan CG, Zou HF, Jiang GB (2005) Environ Sci Technol 39:8442–8447
62. Huang XD, Kong LA, Li X, Chen XG, Guo M, Zou HF (2004) J Chromatogr B Analyt Technol Biomed Life Sci 812:71–84
63. Su XY, Kong L, Li X, Chen XG, Guo M, Zou HF (2005) J Chromatogr A 1076:118–126
64. Zhang H, Hu CX, Liu CP, Li HF, Wang JS, Yuan KL, Tang JW, Xu GW (2007) J Pharm Biomed Anal 43:151–157
65. Wang Y, Che CM, Chiu JF, He QY (2007) J Proteome Res 6:4703–4710
66. Ge F, Lu XP, Zeng HL, He QY, Xiong S, Jin L (2009) J Proteome Res 8:3006–3019
67. Xu Y, Chiu LF, He QY, Chen F (2009) J Proteome Res 8:1585–1593
68. Cheng ZX, Liu BR, Qian XP, Ding YT, Hu WJ, Sun J, Yu LX (2008) J Ethnopharmacol 120:129–137
69. Yao Y, Wu WY, Guan SH, Jiang BH, Yang M, Chen XH, Bi KS, Liu X, Guo DA (2008) Phytomedicine 15:800–807
70. Xu Y, Wang Y, Yan L, Liang RM, Dai BD, Tang RJ, Gao PH, Jiang YY (2009) J Proteome Res 8:5296–5304
71. Talaty N, Takats Z, Cooks RG (2005) Analyst 130:1624–1633
72. Bonam-Rani A, Ternier J, Ratel D, Benabid AL, Issartel JP, Brambilla E, Berger F (2006) Clin Chem 52:2103–2106
73. Jackson SN, Wang HYJ, Woods AS (2005) J Am Soc Mass Spectrom 16:133–138
74. Khatib-Shahidi S, Andersson M, Herman JL, Gillespie TA, Caprioli RM (2006) Anal Chem 78:6448–6456

75. Lemaire R, Wisztorski M, Desmons A, Tabet JC, Day R, Salzet M, Fournier I (2006) Anal Chem 78:7145–7153
76. Walch A, Rauser S, Deininger SO, Hofler H (2008) Histochem Cell Biol 130:421–434
77. Chaurand P, Schwartz SA, Caprioli RM (2004) Anal Chem 76:86A–93A
78. McCombie G, Knochenmuss R (2004) Anal Chem 76:4990–4997
79. Sleno L, Volmer DA (2005) Rapid Commun Mass Spectrom 19:1928–1936
80. Ng KM, Liang ZT, Lu W, Tang HW, Zhao ZZ, Che CM, Cheng YC (2007) Anal Chem 79:2745–2755
81. Zhu F, Cai YZ, Xing J, Ke JX, Zhan ZQ, Corke H (2009) Rapid Commun Mass Spectrom 23:1678–1682
82. Shan SJ, Tanaka H, Shoyama Y (2001) Anal Chem 73:5784–5790
83. Zhou M, Liu Y, Chen TB, Fang XX, Walker B, Shaw C (2006) Peptides 27:2688–2694
84. van Kampen JJA, Burgers PC, de Groot R, Luider TM (2006) Anal Chem 78:5403–5411
85. Lin YS, Chen YC (2002) Anal Chem 74:5793–5798

Top Curr Chem (2013) 331: 165–192
DOI: 10.1007/128_2012_364
© Springer-Verlag Berlin Heidelberg 2012
Published online: 6 September 2012

Chemical and Biochemical Applications of MALDI TOF-MS Based on Analyzing the Small Organic Compounds

Haoyang Wang, Zhixiong Zhao, and Yinlong Guo

Abstract This review focuses on the recent applications of matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI TOF-MS) technologies for analysis of small molecular compounds (SMCs) on qualitative and quantitative levels to provide valuable information for the relative chemical and biochemical researches. We summarized the approaches to minimize the interference in low *m*/*z* range and ion suppression effects by adding chemical reagents, sample cleaning and preparation, matrix selection and performing chromatographic separation before MALDI TOF-MS analysis. Meanwhile, we discussed the strategies to enhance the MALDI TOF-MS detection sensitivity and selectivity by derivatization to attach the "charge tags" to SMCs. In addition, the mass spectrometric imaging (MSI) methods for locating the target SMCs in bio-tissues by MALDI TOF-MS were reviewed. Furthermore, the applications of MALDI TOF-MS for monitoring enzyme reactions and the screening of their inhibitors were also presented. Finally, the chemical applications of MALDI TOF-MS on characterization of small molecular transition-metal complexes and monitoring the organic reactions, especially the polymerization reactions, were discussed.

Keywords MALDI TOF-MS • Mass spectrometry • Small molecule • Transition-metal organometallic compounds

H. Wang (✉), Z. Zhao and Y. Guo
Shanghai Mass Spectrometry Center, Shanghai Institute of Organic Chemistry, Chinese Academy of Science, LingLing Road 345, 200032 Shanghai, People's Republic of China
e-mail: haoyangwang@sioc.ac.cn; ylguo@sioc.ac.cn

Contents

Abbreviations

AChE	Acetylcholinesterase
CAD	Collision-activated dissociation
CHCA	α-Cyano-4-hydroxycinnamic acid
CID	Collision-induced dissociation
CNTs	Oxidized carbon nanotubes
CT	Charge-transfer ionization
DCTB	2-[(2*E*)-3-4-(*tert*-butylphenyl)-2-methylprop-2-enylidene] malononitrile
DHB	2,5-Dihydroxybenzoic acid
DIOS	Desorption/ionization on silicon
ESI	Electrospray ionization
FTICR-MS	Fourier transform ion cyclotron resonance mass spectrometry
LC	Liquid chromatography
MALDI	Matrix assisted laser desorption ionization
MS	Mass spectrometry or mass spectrometer
MS/MS	Tandem mass spectrometry
MSI	Mass spectrometric imaging
NALDI	Nano-assisted laser desorption–ionization
NAPIQ	*N*-alkylpyridinium isotope quaternization
SALDI	Surface-assisted laser desorption/ionization
SMCs	Small molecular compounds
SRM	Selected ion reaction monitoring
TLC	Thin layer chromatography
TOF-MS	Time-of-flight mass spectrometer

1 Introduction

Nowadays, matrix-assisted laser desorption/ionization [1, 2] time-of-flight mass spectrometry (MALDI TOF-MS) has become a versatile and important soft ionization technique in mass spectrometry for the determination of molecular masses of various fragile and nonvolatile samples, including biopolymers [3] such as proteins [4, 5], peptides [6], and saccharides [7, 8], and large organic molecules such as polymers [9], dendrimers [10], and other macromolecules [11], as well as small molecular compounds (SMCs) [12]. The ability to obtain accurate and high resolution mass information rapidly with relative ease of sample preparation suggests the potential of MALDI TOF-MS to become a powerful tool in the analysis of nearly all types of organic molecules [13]. MALDI has some similar characteristics with electrospray ionization (ESI) [14] both in relative softness and the ions produced. However, multiply charged ions always appear in ESI-MS analysis. In traditional MALDI analysis the sample is dispersed in solid matrices. The matrices are normally organic aromatic compounds, which are able to efficiently absorb the laser radiation, transfer laser energy to analytes and achieve rapid desorption/ionization of analytes into positively or negatively charged ions by using a pulse of radiation from an ultraviolet/infrared laser. Therefore, mass analyzer of TOF-MS is well suited to MALDI experiments because TOF-MS is also working in pulse mode rather than in continuous operation. An MALDI TOF-MS spectrum could be generated with an individual laser shot or accumulation of several laser shots [15]. Until now, MALDI has also been coupled with TOF/TOF-MS (tandem TOF/TOF mass spectrometer) [16], Q-TOF-MS (quadrupole-time of flight mass spectrometer) [17], or Q-IT-TOF-MS (quadrupole ion trap time-of-flight mass spectrometer) [18]. The introduction of the post source dissociation (PSD) and collision induced dissociation (CID) function for the ions generated by MALDI ion source helps the ion structural assignment and expands the analytic power of the MALDI technique [19]. Today, commercial reflection TOF-MS instruments and Q-TOF-MS instruments could reach the resolving power $m/\Delta m$ of about 10,000–20,000 FWHM (full-width half-maximum, Δm defined as the peak width at 50% of peak height) [20, 21], or obtain even higher resolution values depending on the operation mode.

Coupling MALDI with TOF-MS instrumentation allows a "virtually unlimited" mass range to be monitored; therefore MALDI TOF-MS applications have been focused on the analysis of large biomolecules. However, in recent years reports of the characterization of SMCs by MALDI TOF-MS have been emerging in large numbers. Recently some nice reviews have summarized the important and interesting applications of MALDI TOF-MS using novel strategies to improve the performance of the MALDI-MS technique in the analysis of SMCs. Cohen reviewed the applications of MALDI-MS to provide both qualitative and quantitative determination of low molecular weight compounds [12, 13]. Just recently Luider reviewed the biomedical application of MALDI-MS for the analysis of small molecules and discussed its favorable properties [22]. Higashi reviewed the derivatization of neutral steroids to enhance their detection characteristics in mass spectrometry [23].

Wang and Guo summarized the recent applications of MALDI Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) for qualitative and quantitative analysis of low molecular weight compounds [24]. Schiller reviewed the applications of MALDI TOF-MS in lipidomics [25] and recent developments of useful MALDI matrices for the MS characterization of apolar compounds [26]. McIndoe discussed the utility of MALDI TOF-MS in the characterization of metal clusters [27]. Montaudo reviewed the studies of synthetic polymers by MALDI TOF-MS [28] and Schubert reviewed the structural assignment of synthetic polymers by tandem mass spectrometry [29]. Fogg reviewed the utility of inert-atmosphere MALDI TOF-MS for study of transition metal complexes/catalysts [30].

These published reviews are an excellent background to this account and the overlapping field will not be discussed again. This review summarizes the most recent chemical and biochemical applications of MALDI TOF-MS based on analyzing the SMCs. Besides the technologies for straightforward analysis of SMCs by MALDI TOF-MS, the corresponding suitable sample preparation [31], tandem mass spectrometric methods [32], detection-oriented derivatizations [33, 34], and the related chemical treatments [35, 36] for this purpose are discussed. These strategies often help to pre-concentrate target analytes before MALDI TOF-MS analysis to reduce the matrix effects and ion suppression effects, and also to translate the compounds into MS acceptable and easily MS detectable species to enhance MS detection selectivity and sensitivity [23]. Further use of MALDI TOF-MS to monitor enzyme-catalyzed reactions and screen their inhibitors are also discussed. Meanwhile, the characterization of small organometallic complexes and monitoring organic reaction by MALDI TOF-MS are presented systemically. Thus, the present review highlights the unique advantages of MALDI TOF-MS, as a powerful tool, to perform chemical and biochemical research by analyzing SMCs.

2 Basic Aspects for Analysis of SMCs by MALDI TOF-MS

MALDI TOF-MS experiments can be performed in vacuum or atmospheric pressure [37]. Atmospheric pressure (AP)-MALDI has a larger tolerance for laser frequency variations and shows reduced fragmentation due to the collision cooling of the expanding plume. Therefore, clustering ions between matrix and analyte sometime could be clearly observed. Dissociation of these clusters can be achieved by applying higher laser frequency or adapting the parameters of the atmospheric pressure interface [38]. The ionization of the MALDI process is triggered by a laser beam, typically using ultraviolet (UV) lasers such as nitrogen lasers (337 nm) or frequency-tripled and frequency-quadrupled Nd:YAG lasers (355 and 266 nm, respectively) [39]. Infrared (IR) lasers are also used due to their softer mode of ionization and greater material removal (useful for biological samples), less low-mass interference, and compatibility with other laser desorption/ionization methods [40].

The matrices generally used have some common characteristics, such as separating the molecules into the matrix solid solution, protecting the molecules from being destroyed by direct laser beams, and facilitating vaporization and ionization [41]. The introduction of organic matrix reduces the ion fragmentation in the ionization process, which is the basic characteristic of MALDI. The top three commonly used are 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid), α-cyano-4-hydroxycinnamic acid (CHCA), and 2,5-dihydroxybenzoic acid (DHB) [42]. Normally MALDI is able to create singly-charged ions. "Cationized molecules" are the dominant ionic species in positive ion mode, such as $[M+H]^+$ and $[M+Na]^+$, while the deprotonated molecule $[M-H]^-$ is the main ionic species in negative ion mode. When the matrices, such as *para*-nitroaniline (PNA) or 9-aminoacridine (9-AA), are used, multiply charged ions $[M+nH]^{n+}$ might also be observed [43]. Ion signals of radical cations can also be observed in the analysis of organometallic compounds and aromatic compounds due to the charge-transfer process in MALDI-MS ionization [44]. However, the high levels of noise from the background ions of organic matrices in the low m/z range restricted SMCs analysis by MALDI TOF-MS [12, 45].

Laser desorption/ionization (LDI) is a mass spectrometric approach normally performed with TOF-MS but without applying any traditional organic matrices. In these methods, sample was loaded onto the photoactive but non-desorbable supports or surfaces. The most attractive advantages of LDI-MS are little or no signals from the photoactive and modified surfaces observed in the mass spectrum, as well as the simplified sample preparation without mixing or co-crystallizing of the analyte with organic matrices [12, 24]. Several LDI-MS techniques have emerged as important alternative strategies to classic MALDI-MS methods, such as desorption/ionization on porous silicon mass spectrometry (DIOS) [46], desorption/ionization on metal films [47, 48], desorption/ionization on oxide plates [49], nano-assisted laser desorption–ionization (NALDI) [50], and surface-assisted laser desorption/ionization (SALDI) [51]. Nowadays, LDI methods are still widely applied in crude oil analysis [52], atmospheric aerosol analysis [53], surface analysis [54], tissue imaging [55], and inorganic compounds analysis [56].

The ion suppression effect could decrease the sensitivity of target SMCs and reduce the accuracy and reproducibility of MALDI TOF-MS analysis, which limits its application to quantification analysis. The ion suppression effect might come from salts and buffers, ion-pairing agents, ionic detergents, endogenous compounds, drugs, metabolites, and proteins in biomedical research [57]. When MALDI TOF-MS is used for the quantification analysis of SMCs, many strategies could be applied to minimize these problems, such as desalting and sample clean-up, increasing homogeneous sample during matrix crystallization or adding some additive reagents, applying chromatographic separation, modification or derivatization of target compounds, using ionic liquid matrices (ILMs), choosing high molecular matrix or LDI approaches to reduce the chemical interferences caused by organic matrices. Moreover, applying internal standards and averaging results within spot and spot-to-spot were also effective methods to enhance the precision of MALDI TOF-MS analysis [12, 22]. In some cases, the difficulties to analyze SMCs by

MALDI TOF-MS might cause mass discrimination and low mass cut off effects of ion transfer optics, such as quadrupole, hexapole, or octopole mass filters, in mass spectrometers. Under these conditions the parameters of ion transfer optic have to be optimized to guarantee ion accumulation and the successful passing of ions in the low mass range through ion transfer optics to the mass detector [13, 24].

The combination of various column and planar separation techniques prior to MALDI or DIOS TOF-MS analysis could significantly reduce signal suppression of low-abundance interference species from complicated mixtures and enhance detection sensitivity [58]. The separation-coupling MALDI methods can be classified as off-line and on-line modes, respectively. MALDI and relative MS methods share the advantages of speed, efficiency, and high tolerance to nonvolatile buffers and impurities [59, 60]; the samples are first separated by chromatographic methods, such as HPLC [61], GPC [62], CE [63], TLC [64], and gel separations [65], and are then transferred and applied to solid supports in off-line mode. At the same time, the on-line MALDI analysis of flowing liquid samples was developed rapidly due to its high speed, high throughput and lower sample consumption. In on-line mode, liquid fractions containing analytes from chromatographic separations and solution of matrices are transported into the TOF-MS by means of capillaries and deposited on a rotating wheel or ball [66, 67], where the solvent evaporates to form a thin sample trace containing matrices and analytes on the position in vacuum condition for further laser desorption/ionization MS analysis when the wheel or ball rotates.

3 MALDI TOF-MS Analysis of SMCs

3.1 Direct Analysis of SMCs by MALDI TOF-MS

For a typical MALDI TOF-MS, the low-mass range ($m/z < 1{,}000$) is dominated by small molecular matrix ions. Consequently, the most significant problem for MALDI-MS to analyze SMCs directly is the interference from matrix signals in low mass range areas. However, in some cases MALDI TOF-MS could be used for the analysis of SMCs directly. Zou established a MALDI TOF-MS method for suppressing the matrix related ions background and improving the analysis of SMCs substantially by adding the surfactant of cetrimonium bromide (CTAB) to the conventional CHCA matrix [68]. MALDI TOF-MS has proven to be a reliable and rapid tool to detect and identify microcystin variants in very small amounts of samples such as single microcystis colonies. Mass signals of microcystin variants can be characterized and identified by post source dissociation (PSD) fragmentation [69]. Moreover, quantification of the glucose produced by enzyme hydrolysis of starch was achieved by MALDI TOF-MS protocol, by using sorbitol as an internal standard [70].

Just recently, Liu et al. applied a novel matrix of isoliquiritigenin (ISL), a flavonoid with a chalcone structure (4,2′,4′-trihydroxychalcone), to the analysis of neutral oligosaccharides by MALDI TOF-MS [71]. Compared with widely used matrices, 2,5-dihydroxybenzoic acid (DHB) and 2,4,6-trihydroxyacetophenone (THAP), ISL showed several advantages, such as higher homogeneity of crystallization, higher sensitivity, better spectral quality, better salt tolerance, and enough fragmentation yield to get richer structural information. Moreover, quantitative analysis of oligosaccharides was achieved by using ISL matrix.

A fast-drying sample preparation method, using CHCA as a matrix in an acetonitrile/tetrahydrofuran solvent system, has been used to analyze commercial canola, castor, and olive oils by MALDI TOF-MS and the resulting MS spectra showed sodium adduct ions of oil molecules [72]. Furthermore, the relative percentages of the fatty acids in olive oil were detected and such a MALDI TOF-MS method is simpler and less time-consuming than the established transesterification method by GC/MS [73]. MALDI-MS/MS could be used for the studies of oils and lipids [74] due to the abundance of matrix-derived chemical noise in the mass range of interest. MALDI has also been extensively used in the analysis of phospholipids [75]. Al-Saad have published an interesting study on the "prompt" or "in-source" fragmentation and the PSD of MALDI ionized phospholipids [76]. With the development of MALDI TOF/TOF and MALDI Q-TOF instruments [16], further applications in studies of lipophilic molecules are expected [77]. Griffiths reviewed the application of tandem mass spectrometry in the study of fatty acids, bile acids, steroid conjugates, and neutral steroids [78].

3.2 MALDI TOF-MS Analysis of SMCs with Specific Matrices

Various matrix substances, such as porphyrins, inorganic materials, fullerenes, porous silicon, carbon nanotubes, ionic (liquid) matrix [79], and nanoparticles have also been used to eliminate or reduce matrix ion interference during MALDI TOF-MS analysis [80]. The most important and attractive advance in this research area is the introducing of specific materials as MALDI matrices, which could absorb and transfer laser energy to analytes to facilitate ionization in MALDI TOF-MS analysis.

Luider applied a high molecular weight matrix, mesotetrakis(pentafluorophenyl) porphyrin, to eliminate chemical noise in the low-mass range and quantify small molecular drugs, lopinavir and ritonavir (HIV protease inhibitors), by MALDI TOF-MS. A "brushing" spotting technique in combination with pre-structured target plates allowed fast preparation of homogeneous matrix crystals and the addition of Li^+ gave rise to intense cationized drug species. These approaches were applied to quantify opinavir in extracts of small numbers of peripheral blood mononuclear cells [81]. Ayorinde used meso-tetrakis(pentafluorophenyl) porphyrin (F20TPP) as a matrix in analysis of some commercial nonylphenol ethoxylates with sodium ion dopant by MALDI TOF-MS [82]. A comparison of

the mass spectrometric data with those obtained with CHCA showed less matrix interference in the low-mass range. Recently, Xiong analyzed small molecules with metal-phthalocyanines as matrices, which were even capable of forming matrix-analyte adducts in either the positive or negative ion mode [83].

ILMs were also tested as MALDI matrices for quantification of analysis of SMCs. Good calibrations with high linearity and reproducibility were achieved over a broad concentration range for all the tested ILMs in spite of their different physical states. The experimental results indicated that various ILMs had different sensitivities owing to changes in their cation components and the hydrophobicity of target compounds [84]. Fast screening of SMCs was performed by thin-layer chromatography (TLC) followed by direct on-spot MALDI TOF-MS identification using a UV-absorbing ILM with minimal background ions from the proton donor triethylamine/CHCA ILM. Three arborescidine alkaloids – anesthetics levobupivacaine and mepivacaine and the antibiotic tetracycline – were characterized by MALDI TOF-MS [85]. Moreover, aflatoxins B_1, B_2, G_1, and G_2 have been analyzed by MALDI TOF-MS using a UV absorbing ILM (Et_3N-α-CHCA) and addition of NaCl enhanced the detection sensitivity of aflatoxins to an LOD of 50 fmol [86].

An ionic crystal MALDI matrix, which combined the lipid response enhancing UV-absorber *p*-nitroaniline with the protic agent butyric acid, was synthesized and applied in MALDI TOF-MS analysis of the zwitterionic phosphatidylcholine (PC) headgroup-containing lipids [87]. Polyanionic oligosaccharides, such as dermatan sulfate (DS) and chondroitin sulfate (CS), exhibit poor ionization efficiencies and tend to undergo thermal fragmentation by loss of SO_3 under conventional UV-MALDI conditions. A new ILM, a guanidinium salt of CHCA, facilitates direct UV-MALDI TOF-MS analysis of underivatized DS and CS oligosaccharides up to a decasaccharide in their common form as sodium salts. The resulting mass spectra show a very low extent of fragmentation by loss of SO_3 [88].

Buriak first developed the surface-supporting laser desorption method and reported the use of the pulsed LDI from a porous silicon surface for analysis of small biomolecules by MALDI TOF-MS [46]. Such a method uses porous silicon to trap analytes deposited on the surface, and laser irradiation to vaporize and ionize them, with detection limits at femtomole or attomole levels, inducing little or no fragmentation as well as exhibiting compatibility with silicon-based microfluidics and microchip technologies. Later such a method was called desorption/ionization on porous silicon mass spectrometry (DIOS-MS) [89], which could be applied to the quantitative analysis of SMCs, organic reaction monitoring, PSD, and for combining with various chromatographic technologies. Quinoline functionalized mesoporous silica was applied as a matrix for the MALDI-MS analysis of small molecules [90], showing several advantages in analysis of small dye molecules, such as less background interference, high homogeneity, and better reproducibility.

Carbon nanotubes (CNTs) have large surface areas to disperse the analyte molecules by preventing sample aggregation. The strong ultraviolet laser absorption characteristics of CNTs allow easy laser energy transfer to the analytes, facilitating the ionization of analytes. Zou used CNTs as an alternative to conventional organic matrices for SMCs with MALDI TOF-MS [91]. Guo reported a simple method,

which immobilized carbon nanotube with a type of polyurethane adhesive (NIPPOLAN-DC 205), for analysis of SMCs [92]. The strong bonding of CNTs on the target facilitated experiments such as PSD by providing longer-lasting signals and even had potential application in urine glucose analysis due to low interference signals in the low *m*/*z* range when using CNTs as the matrix. Moreover, oxidized CNTs which can form a stable homogeneous suspension in water close to a solution phase were synthesized and used for analysis of small biomolecules by MALDI TOF-MS [93] and MALDI FTICR-MS [94, 95]. Impurities including amorphous carbon, which is one of the main reasons for ion source contamination, were destroyed by the oxidization. Furthermore, the 2,5-dihydroxybenzoyl hydrazine functionalized CNTs [96] were synthesized and used as both pH adjustable enriching reagent and matrix in MALDI TOF-MS analysis of trace peptides. The high efficiency of adsorption and enrichment towards trace peptides can be achieved by adjusting pH value of the functionalized CNT dispersion.

Wang used graphene as a matrix for the analysis of SMCs, including: amino acids, polyamines, anticancer drugs, nucleosides, and nonpolar steroids by using MALDI TOF-MS [97]. The use of graphene as a matrix avoided the fragmentation of analytes and provided good reproducibility and high salt tolerance. Compared with conventional organic matrices, graphene was not only an effective matrix for nonpolar compounds, but also an adsorbent for solid-phase extraction to improve the detection limit. Just recently, Cai reported a new interesting approach for the analysis of SMCs, such as peptides, amino acids, fatty acids, nucleosides, and nucleotides, with direct negative ion LDI-TOF-MS on graphene flakes [98]. Deprotonated monomeric species $[M-H]^-$ ions were achieved on uniform graphene flake films with better sensitivity and reproducibility in negative ion mode. However in positive ion mode the analytes could be detected as various adduct ions, such as $[M+H]^+$, $[M+Na]^+$, $[M+K]^+$, $[M+2Na-H]^+$, $[M+2K-H]^+$, and even $[M+Na+K-H]^+$. Thus the matrix interference free negative ion LDI on graphene flakes may be expanded for LDI-MS analysis of various SMCs.

3.3 MALDI TOF-MS Analysis of SMCs with Derivatization

The main objects of analytical derivatization in MALDI TOF-MS analysis for SMCs were: (1) making the derivatizated SMCs to higher *m*/*z* range to reduce the interference of matrix signals in the low *m*/*z* range, which could increase detection sensitivity; (2) selectively attaching "charge tags" or "easily ionized tags" to the neutral small molecules, which would significantly increase detection sensitivity and selectivity; (3) introducing "isotopic tag" compounds as internal standards to quantify target compounds. Therefore, when MALDI TOF-MS was combined with suitable derivatization, the whole method would be more powerful to solve the problems concerning analysis of SMCs.

Heinzle investigated the application of straightforward one-pot derivatization procedures to analyze alcohols, aldehydes and ketones, carboxylic acids,

ketocarboxylic acids, and amines by MALDI TOF-MS [99]. A pair of isotopically coded light and heavy reagents, tris(2,4,6-trimethoxyphenyl)phosphonium acetic acid *N*-hydroxysuccinimide esters, were synthesized and used to derivatize small molecular compounds with primary or secondary amine groups for MALDI TOF-MS analysis [100, 101]. Such derivatization greatly facilitated MALDI analysis of SMCs and significantly improved the sensitivity of analysis, allowing a limit of detection in the low femtomole range and the direct quantification of SMCs without sample clean-up. Recently Denekamp reported a rapid derivatization reaction of *para*-methoxy substituents of tris(2,4,6-trimethoxyphenyl)methyl carbenium ion with primary and secondary alkyl amines for MALDI TOF-MS analysis of amines, amino acids, and small C-protected peptides [102]. Meanwhile, a universal ionization label derivatization reagent containing an anthracene moiety was developed by Schmitz for the atmospheric pressure laser ionization APLI-(TOF)-MS analysis of SMCs and small molecular polymers [103].

Guo developed a pyrimidine-based stable-isotope labeling reagent, $[d_0]/[d_6]$ 4,6-dimethoxy-2-(methylsulfonyl) pyrimidine (DMMSP) [104] for comparative quantification of protein [105] and small peptides [106]. This was further used as a derivatization agent in studies of different expression levels of keratins in tongue coating samples of Hepatitis B patients by MALDI TOF-MS [107]. In addition, Guo and Wang developed a simple derivatization reaction based on direct *N*-alkylpyridinium isotope quaternization (NAPIQ) for mild derivatization of cholesterol and fatty alcohols [108]. Different from the conventional quaternary reagents with cations on themselves, the "*N*-cationic tag" was introduced onto the target compounds by derivatization reaction with neutral pyridine in the presence of trifluoromethanesulfonic anhydride (Tf_2O). The derivatization significantly improved the detection limits of analytes by 10^3-fold in MALDI FTICR-MS analysis. The NAPIQ method was used in analyses of cholesterol and fatty alcohols in small amounts of human hair samples (<0.5 mg). Moreover, recently NAPIQ was further proven to be an efficient method for analyzing steroids and carbohydrates [109].

MALDI TOF-MS has several advantages over other mass spectrometric methods in analysis of oligosaccharides. A number of publications have tried to incorporate derivatization steps prior to MALDI TOF-MS experiments for increasing detection sensitivity of oligosaccharides [110]. Here we selectively discussed some widely used methods. In MALDI analysis of small oligosaccharides a very large increase in sensitivity may be achieved due to introduction of a quaternary ammonium center ("quaternization"). Such a quaternary ammonium center may be introduced into the saccharide by reaction with commercially available glycidyltrimethylammonium chloride (GTMA) [111] or by using Girard's reagent T. GTMA reacts with alcohol functionalities, whereas Girard's reagent T is specific for aldehyde and keto groups. Reducing saccharides can be derivatized by both GTMA and Girard's reagent T. Non-reducing saccharides as well as sugar alcohols can be derivatized using GTMA. Although sucrose, raffinose, and sorbitol do not react with Girard's reagent T, they all produce intense signals after derivatization with GTMA.

Small oligosaccharides could be derivatized into hydrazone with Girard's T reagent in order to introduce a cationic site for detection by MALDI TOF-MS [112]. The derivative was prepared in high yield and did not require extensive clean-up prior to MS examination, unlike the products of the more commonly used reductive amination derivatization. The derivatives gave a tenfold increase in detection sensitivity over those afforded by the underivatized oligosaccharides. In addition, the use of these derivatives removed ambiguities caused by the presence of $[M+Na]^+$ and $[M+K]^+$ from the underivatized sugars. Furthermore, Girard's T derivatization overcame problems associated with the reducing-terminal and *N*-acetylamino groups, which were introduced when the oligosaccharides were prepared by cleavage from glycoproteins with hydrazine.

Okamoto demonstrated the usefulness of derivatization to enhance sensitivity of MALDI TOF-MS analysis of oligosaccharides by tagging maltopentaose with 2-aminopyridine (PA), 4-aminobenzoic acid ethyl ester (ABEE), and trimethyl (*p*-aminophenyl)ammonium chloride (TMAPA). Among the derivatives, the sensitivity of the PA-tagged maltopentaose showed a 100-fold improvement over the native one with DHB matrix, while the oligosaccharide derivatized with ABEE and TMAPA gave 30- and 10-fold increases in sensitivity over the underivatized one. Structural information of these derivatized oligosaccharides was obtained by post source decay (PSD) [113]. Predictable and reproducible fragmentation patterns could be obtained in all cases and matrix-dependence fragmentation with the PA-labeled oligosaccharide was observed; with CHCA a simple spectrum ascribable to Y series ions was obtained, while both B and Y series ions were clearly observed in the DHB case. Oligosaccharides could also be derivatized by reductive amination with benzylamine followed by *N*,*N*-dimethylation with methyl iodide and analyzed by MALDI-MS and PSD [114] and the approximate detection limit for the resulting carbohydrate derivatives was 50 fmol. When the derivatives were analyzed by MALDI-PSD, the fragmentation pattern observed was dominated by fragment ions retaining the modified reducing terminus, thus simplifying the interpretation of the mass spectra.

3.4 MALDI TOF-MS for Metabolites and Lipidomics

In the last 10 years, many "omics" techniques have been developed for biomarker discovery and early diagnosis of human cancers [115]. Considering the relation between MALDI TOF-MS analysis for SMCs and "omics" techniques, here we have only discussed the application of MALDI TOF-MS on metabolomics [116] and lipidomics [117].

Metabolomics is defined as "the quantitative measurement of the dynamic multiparametric metabolic response of living systems to pathophysiological stimuli or genetic modification" [118]. In metabolomics research, MS analysis is normally performed to identify and quantify metabolites after chromatographic separation [119, 120]. The metabolites could be identified by mass spectral libraries or

according to their fragmentation patterns. However, the MALDI matrices significantly give background chemical noise at $<1{,}000$ m/z that complicates analysis of the low-mass range for metabolites. Thus the use of MALDI TOF-MS as a standalone method for metabolomics was limited. However, MALDI FTICR-MS with the higher resolution showed better performance in analysis and quantification of small molecular metabolites and these applications have been reviewed by Brown [121]. Due to these limitations, some LDI approaches have been applied to the analysis of biofluids and tissues. Among the technologies being developed to address this challenge, nanostructure-initiator MS (NIMS) [122, 123] does not require the application of traditional matrices and thereby facilitates small molecular metabolites identification.

Lipidomics is a subdivision of metabolomics and could be defined as a large-scale study of pathways and networks of cellular lipids in biological systems [124]. The progress of modern lipidomics has been greatly accelerated by the development of the general and soft ionization MS techniques [116, 125]. Comprehensive detections of an entire range of lipids within a complex mixture can be correlated to experimental conditions or disease states by MALDI TOF-MS. Schiller recently stated that MALDI has become increasingly popular for the analysis of nonpolar lipids [126]. Nowadays, MALDI TOF-MS has become a very promising approach for lipidomics studies, particularly for the imaging of lipids from tissue slides [127, 128].

Recent developments in MALDI TOF-MS methods have enabled direct detection of lipids in situ and collision-activated dissociation (CAD) of the molecular ions can be used to determine the lipid family and often structurally define the molecular species. Such a technique enables MS detection of phospholipids, sphingolipids, and glycerolipids in tissues such as heart, kidney, and brain. Furthermore, the distribution of many different lipid molecular species often define anatomical regions within these tissues [129]. Integration of experimental data of MALDI TOF-MS and relative databases of lipids, as well as with metabolic networks, offers an opportunity to devise therapeutic strategies to prevent or reverse these pathological states involving dysfunction of lipid-related processes [130].

4 Monitoring Enzyme Activities and Screening Inhibitors by MALDI TOF-MS

Screening for inhibitors of pharmacologically-relevant enzymes is one of main aims nowadays during drug discovery, and mass spectrometric methods have increasingly been developed for such purposes by measuring enzyme activities and kinetics [131]. Such methods could be divided into two categories: (1) direct screening methods isolate and detect the enzyme-inhibitor complexes or the inhibitors to prove activities on the drug target; (2) indirect screening methods rely on detection of reporter molecules to measure enzyme-inhibitory activities of analytes and the assay provides a functional response rather than relying on

detection of inhibitor binding alone, which is performed by chromatographic and MS methods [132]. Here, we would focus on discussing indirect methods for studying enzyme activities and screening of enzyme inhibitors by MALDI TOF-MS. This could be regarded as the extensive application of MALDI TOF-MS to the analysis of SMCs, because here the analysis target is not the enzymes or enzyme-inhibitor complexes but the small molecular substrates and the products from the enzyme-catalyzed reactions.

The key point for this application is characterizing the signal ratios of substrate/product by MS; thus MS-based screening methods could screen various kinds of samples or compound libraries and provide mass information, structural information, enzyme activity information, and enzyme-inhibitor affinity data, as well as achieving high throughputs. Compared with the traditional fluorescence approach, the mass spectrometric label-free method offers direct identification of the substrates and products in enzyme-catalyzed reactions. Furthermore, MALDI-MS does not require complicated sample pretreatment procedures and even the enzyme reaction solution could be directly analyzed by MALDI-MS without chromatographic separation due to its high tolerance to many buffer salts and chemical reagents.

As an early report, MALDI TOF-MS is combined off-line with rapid chemical quench-flow methods to investigate the pre-steady-state kinetics of a protein-tyrosine phosphatase (PTPase) [133]. Furthermore, a digital micro-fluidic system based on electrowetting has been developed to facilitate the investigation of pre-steady state reaction kinetics using rapid quenching and MALDI TOF-MS for the analysis of reaction kinetics that were previously too rapid to analyze by MALDI TOF-MS [134]. The different effects of chiral isomer inhibitors on the enzyme could also be studied by MALDI TOF-MS, which showed that butyrylcholinesterase (BChE) inhibited by (1*R*)-isomalathions readily reactivated, while enzyme inactivated by (1*S*)-isomers did not. The results indicated that EBChE inhibition by (1*R*)-isomalathions proceeded with loss of diethyl thiosuccinate, but inactivation by (1*S*)-isomers occurred with loss of thiomethyl as the primary leaving group, followed by rapid expulsion of diethyl thiosuccinate to yield an aged enzyme [135].

The MALDI TOF-MS platform for quantitatively monitoring enzyme activities of acetylcholinesterase and screening enzyme inhibitors has been demonstrated by Zou and the method described employed a matrix of oxidized carbon nanotubes [136]. The activity of acetylcholinesterase was quantitatively monitored and the acetylcholinesterase inhibitors could be screened. Karas developed a MALDI TOF-MS method for the determination of acetylcholine (ACh) and choline (Ch) in mouse brain microdialysis samples [137]. An optimized dried droplet preparation with the CHCA was used with good accuracy and precision. Despite the high salt content of the perfusion fluid, a direct measurement was also successfully performed and the experimental results showed that no other compounds, which were naturally dialyzed from the brain in conjunction with ACh and Ch, caused interferences with the analysis. Another important advantage of this method was that no LC preparation of the sample was required.

Due to the characteristics of high resolution and sensitivity to distinguish the signals of small molecules with the signals of matrices and other interferences, a MALDI FTICR-MS-based enzyme activity assay method was also developed for enzyme kinetic measurements and inhibitor screening of acetylcholinesterase (AChE) without specific matrix requirement [138]. Recently Guo improved the MALDI FTICR-MS method to obtain low detection limits of organophosphorus pesticides (OPs) through the use of external reagents, *n*-octylphosphonic acid, which helped to enhance AChE inhibition by OPs. The detection limits were improved significantly by 10^2-fold to 10^3-fold in comparison with conventional enzyme-inhibited methods [139].

Siuzdak described a nanostructure-initiator mass spectrometry (NIMS) enzymatic (Nimzyme) assay, in which enzyme substrates were immobilized on the TOF-MS surface by using fluorous-phase interactions. This "soft" immobilization allowed efficient desorption/ionization while also allowing the use of surface-washing steps to reduce signal suppression from complex biological samples. The Nimzyme assay was sensitive to subpicogram levels of enzyme, detecting both addition and cleavage reactions (sialyltransferase and galactosidase), was applicable over a wide range of pHs and temperatures, and could measure activity directly from crude cell lysates [140]. Villanueva performed a comprehensive detection and functional analysis of pre-existent peptides and small proteins with the capability of binding to trypsin-like proteases related to blood coagulation by "intensity fading" MALDI TOF-MS. Combining "intensity fading MS" and off-line LC prefractionation allowed the detection of more than 75 molecules present in the leech extract, which interacted specifically with a trypsin-like protease over a sample profile of nearly 2,000 different peptides/proteins [141].

MALDI TOF-MS-based methods could also be used for rapid and versatile characterization of protein kinases and their inhibitors. New kinase substrates were designed by the modification of common synthetic peptides with mass tag technology by *N*-derivatization through stable isotope labeling and C-terminal conjugation with tryptophanylarginine. Results showed that C-terminal conjugation with the tryptophanylarginine moiety enhanced the ionization potency of these new substrates and this radioactive isotope-free quantitative kinase assay would greatly accelerate the discovery of a new generation of potential kinase inhibitors that exhibited high selectivity or unique inhibitory profiles [142]. The concept of indirect screening methods for studying enzyme activity has been applied for tackling the trace amount of β-lactamase in milk, which showed another promising development orientation of this research filed. A selective, fast, and effective enzyme assay based on MALDI FTICR-MS for quantifying β-lactamase, an illegal additive in milk products, has been reported [143]. The amount of β-lactamase that could be determined in milk samples is 6×10^{-3} U mL^{-1} by this approach. It is assumed that this method might be useful for detection of illegal added enzyme in foodstuff and other matrices.

5 MALDI TOF-MS Imaging for SMCs

The concept of MALDI mass spectrometric imaging (MSI) was introduced in 1997 by Caprioli [144] for rapid and direct profiling of the endogenous and exogenous compounds in tissues by MALDI TOF-MS to obtain their spatial distribution and orientation. This operation mode was often referred to as the mass microprobe mode of MALDI TOF-MS imaging with a spatial resolution of 50–200 mm, which was limited by the laser spot size/resolution available for sample target movement. The mass microscope approach is the other MALDI TOF-MSI method developed by Heeran. In this approach, rather than the laser beam being highly focused, a mass spectrometer that accepted a 150–300 mm diameter ion beam was used to map a magnified image of the spatial distribution of a selected m/z value onto a two-dimensional detector [145]. Using such instrumentation a spatial resolution of 4 mm has been achieved but this technique is not commercially available. Thus, we would focus on discussing the mass microprobe mode of MALDI TOF-MS imaging.

Mass spectrometric imaging is a powerful tool for localizing compounds of biological interest with molecular specificity and relatively high resolution. The determination of the localization of various target compounds in a whole animal is valuable for many applications, including pharmaceutical absorption, distribution, metabolism, and excretion (ADME) studies, and biomarker discovery [146]. As for its applications in SMCs, the two-dimensional visualization of the distribution of a drug and the relative metabolites in whole body sections of animals could be achieved [147]. Utilizing MALDI TOF-MSI for whole-body animal sections offered considerable advantages compared to traditional methods, such as sample preparation, matrix application, signal normalization, and image generation, and provided a molecular ex vivo view of organs or whole-body sections from an animal. This made possible the label-free tracking of both endogenous and exogenous compounds with spatial resolution and molecular specificity [148].

The Caprioli group studied drug distribution of anti-tumor drugs in mouse tumor tissue and rat brains. In these experiments matrix was applied to intact tissue by either spotting small volumes of the matrix in selected areas, or by coating the entire surface by air-spraying. MALDI images were created by using the selected ion reaction monitoring (SRM) technique to profile the drug under study specifically. Such an approach minimized the potential for ions arising from either endogenous compounds or the MALDI matrix and this image indicated that the drug was present over most of the tumor but was concentrated in the outer periphery [149]. The Caprioli group has recently reported an interesting sample preparation method involving the use of matrix pre-coated MALDI targets before analysis for tissue imaging of small molecules [150]. Pre-coated targets were coated with tiny crystals of approximately 1–2 μm; thus the tissue sections needed only to be transferred onto the pre-coated target for fast and easy sample preparation without using solvents, which might lead to analyte delocalization within a tissue section.

Drug distribution and individual metabolite distribution within whole-body tissue sections were detected for the first time simultaneously at various time points following drug administration by Khatib-Shahidi [151]. MALDI TOF-MS/MS imaging analysis of tissues from 8 mg/kg olanzapine dosed rats revealed temporal distribution of the drug and metabolites. The emergence of its metabolites were also detected in tissue and correlated to the loss of parent drug signals. Moreover, the MALDI-MS/MS imaging data of the parent drug and its metabolites compared well with published quantitative whole body autoradiography data. In order to detect SMCs in complex biological tissue sections using MALDI, a tandem mass spectrometer was required in order to separate the analyte ions from the background interference ions from the matrix [152]. For example, a hybrid Q-TOF-MS was utilized to resolve some of the mass spectrometric interference from the matrix so that CAD fragmentation of protonated drug molecules could be performed [153].

One favorable feature of MALDI TOF-MSI over traditional drug image techniques is the capability to detect a non-radiolabeled molecule with molecular specificity. Therefore MSI has the ability to differentiate between the intact drug and its metabolites which may be present within a single tissue section collected at various time points following drug administration. MALDI signals were found to be proportional to the densities of pharmaceuticals in tissues although different regions within the same organ section and different types of tissues might demonstrate different surface properties [154]. The linearity of analyte responses and the ionization suppression degree due to the heterogeneity within a tissue section was addressed by making a calibration curve for the semi-quantitation of parent drug and its metabolite by depositing several droplets with increasing concentrations on a blank tissue section.

DIOS strategy was also applied in MALDI TOF-MSI. A comparison of the use of DIOS and 9-aminoacridine for metabolite profiling in *Escherichia coli* demonstrated that the sensitivity for such analyses obtainable using DIOS was superior owing to the reduced chemical background [155]. The use of ion mobility (IM) has also been reported in conjunction with MALDI TOF-MS as a means to fractionate lipids [156]. In this technique, ions were first separated by IM and then analyzed by TOF-MS [157]. The desorbed lipid ions fell on a trend line that was separate from those of oligonucleosides, peptides, proteins, and drugs and metabolites having the same nominal mass [158]. This allowed ions originating from lipids to be distinguished from other small biomolecules. The potential of MALDI-IMS based on MALDI TOF-MS is great, and advances in the instrumentation and operating protocols will bring new applications and insights into molecular processes involving health and disease [159, 160].

Although the important recent developments were achieved in the field of MALDI TOF-MSI in tissue, the precise identification of compounds still needs improvement. Fournier developed an on-tissue N-terminal peptide derivatization strategy to enhance protein identification in MALDI TOF-MSI [161]. Derivatizations made the MS/MS spectra easily interpretable, leading to precise identification and easy manual reading of sequences for de novo sequencing. MALDI-MSI studies were limited due to the poor sensitivity of some SMCs.

The chemical derivatization on tissues for SMCs has rarely been applied. Boutaud used on-tissue chemical derivatization of 3-methoxysalicylamine (3-MoSA), a scavenger of gamma-ketoaldehydes, with 1,1′-thiocarbonyldiimidazole (TCDI) to form an oxothiazolidine derivative with much greater sensitivity in MALDI-MS than 3-MoSA [162]. The 3-MoSA spatial distribution and its pharmacokinetic profile in different organs were obtained by on-tissue chemical derivatization with TCDI onto tissues from mice. These results showed that on-tissue chemical derivatization could be used to improve MALDI TOF-MSI performance for SMCs.

6 Organic Chemical Applications of MALDI TOF-MS

Nowadays, MALDI TOF-MS has an important application and contribution for research in organic chemistry. Organometallic compounds are challenging analytes for MS methods. At the same time, these complexes have attracted extensive research interests of chemists for the discovery of efficient and selective transition-metal catalysts. The widely used high-energy MS ionization methods always induced decomposition and dissociation of sensitive organic and organometallic molecules. However the structural elucidation of the sensitive organic and organometallic complexes could be greatly simplified by applying "soft" ionization mass spectrometric methods, such as ESI and MALDI, providing their *m*/*z* information and their isotope patterns [163, 164].

In ESI-MS analysis, compounds first have to be prepared into solution, which would cause the decomposition of some air or water sensitive organic and organometallic complexes. High contamination and residues of transition metal complexes in the ESI ion source required extensive washing and cleaning. These limitations restricted the extensive applications of ESI-MS for studying sensitive organic compounds and organometallic complexes. However the easy cleaning steps and low or non-existent solvent usage characteristics of MALDI TOF-MS are advantages in analysis of organometallic complexes. Fogg recently gave a nice review concerning MALDI TOF-MS analysis of reactive organometallic molecules for use in studying organic reactions [29]. Here we have summarized some new advances in this area. This includes MALDI TOF-MS methods for analysis of fragile organometallic complexes and some air and moisture sensitive organic compounds, as well as monitoring the process of the chemical transformation.

6.1 DCTB: The Star Matrix for Small Molecular Organometallics

The selection of the appropriate matrix for analysis of organometallic compounds is important. Most routine matrices, such as DHB and CHCA, often provide plenty of protons in the MALDI ionization process to form $[M+H]^+$. However, these protic matrices are poorly suited for analysis of organometallic compounds sensitive to

protons. Protonolysis of metal–ligand bonds limited the observation of organometallic species by MALDI-MS since many of the ligands offering lone pair electrons to form the metal–ligand bonds were basic. The dissociation of ligand and ligand exchange reactions with the matrix may occur in MALDI TOF-MS. Thus, some aprotic and neutral matrices were developed and tested. Among them, 2-[(2*E*)-3-4-(*tert*-butylphenyl)-2-methylprop-2-enylidene]malononitrile (DCTB) as a neutral matrix seems to be a promising and excellent matrix for analysis of organic and organometallic compounds by MALDI TOF-MS.

DCTB is not only a neutral matrix, which maintains intact acid-sensitive organometallic compounds, but is also an effective electron-transfer agent, which promotes ion formation at considerably reduced threshold laser influence, leading to very "clean" mass spectra often entirely free of unwanted decomposition of the analyte. Luftmann found DCTB required extremely low laser intensity to produce positive and negative ions and it was helpful to retain the integrity of fragile compounds [165]. Moreover, the Drewello group found that DCTB was the best suited matrix for fullerene derivatives analysis, providing analyte signals in both positive and negative ion modes at comparatively lower threshold laser influence [166]. The Lou group characterized some synthetic Ru and Ir complexes by MALDI TOF-MS and found that DCTB was the best matrix among the ten tested for the complexes that were prone to ligand exchange by matrix [167]. Among the ten matrices investigated, DCTB was the only one that gave no detectable signals of substitution products by ligand exchange reaction with DCTB matrix. This was probably because DCTB was an aprotic matrix and contains no electron-donating group. It demonstrated that matrix substitution to the reactive compounds could also occur in the gas phase initiated by laser irradiation.

The Wyatt group characterized various analytes, especially some organometallics compounds, by using DCTB as matrix with MALDI TOF-MS [168]. Moreover, they recently studied the solvent-free MALDI TOF-MS sample preparation methods for analysis of organometallic and coordination compounds [169]. These procedures comprised two distinct steps: (1) the "solids mixing" of sample with matrix; (2) transferring solid sample/matrix mixture to the MALDI target plate. Such methods were desirable for insoluble materials, compounds that are only soluble in disadvantageous solvents, or complexes that dissociate in solution. Such situations presented major "difficulties" for most mass spectrometric techniques in sensitive organic compounds and organometallics analysis. Although DCTB has been considered to be an excellent matrix, it provided troublesome results for compounds containing aliphatic primary or secondary amino groups. On the basis of the possible mechanisms proposed, the unknown strong extra ions in the MALDI mass spectra were the products of nucleophilic adduction reactions between analyte amino groups and DCTB molecules or radical cations. Thus, care should be taken in MALDI TOF-MS when DCTB is used as the matrix for compounds containing amino groups [170].

6.2 *Inert-Atmosphere CT MALDI TOF-MS*

Another major problem for the MS analysis of organometallic complexes is caused by decomposition of analytes during the sample preparation, introduction, and analysis when the analytes are air/moisture sensitive. Sample preparation for some extremely air/moisture sensitive small organic compounds or organometallic complexes is ideally carried out under inert conditions, generally in a glovebox under dry and oxygen-free conditions [171]. The matrix must be kept dry and saturated by inert gas. Sample introduction is the most difficult process to avoid sample decomposition owing to most MS systems having no facility for protecting samples from the air during sample transfer. For these reason, various methods have been proposed to minimize sample decomposition. The Scott group developed a capillary-seal apparatus to aero-spray the matrix solution and sample onto probe tips under a nitrogen stream [172]. Other methods generally involve the use of extra sacrificial additives or thin coating layers of matrix. These methods were helpful for the relatively robust complexes, but extremely reactive species still remained susceptible to decomposition.

To solve this problem ideally, the Fogg group connected a MALDI TOF-MS instrument directly to a glovebox so as to bring both sample preparation and introduction inside the inert-atmosphere glovebox [173]. Another urgent requirement of sample solubility imposes limitations on MALDI-MS analysis. Direct analysis of solid samples is desirable as a means of characterizing insoluble compounds or metal complexes. Solvent-free analysis of polymeric solids has been achieved by grinding with solid matrix and analyzing the powder [174]. Fogg modified this approach by adding paraffin oil (Nujol) to promote matrix-analyte mixing and subsequently applying a thin layer of the amorphous mull to the MALDI target plate [173].

Inert-atmosphere charge-transfer ionization (CT) MALDI TOF-MS is strongly complementary to ESI-MS, enabling analysis of neutral, charged, insoluble metal complexes with minimal perturbation. Such an idea might be applied to the study of thermal and thermoxidative decomposition processes of poly(bisphenol A carbonate) (PC) under inert-atmosphere by MALDI TOF-MS as used by Montaudo [175]. Inert-atmosphere CT-MALDI TOF-MS methods also eliminate the danger of deactivation of air/water sensitive sample prior to analysis and promises to provide more structure insights of organometallic complexes. Till now inert-atmosphere CT MALDI TOF-MS has been widely used to characterize the Ru metathesis catalysts containing aryloxide ligands [176] and Ru–OAr bonding in a five-coordinate Ru^{II} complex [177], as well as Au^{I} *N*-heterocyclic carbene complexes bearing biologically compatible moieties [178].

The CT-MALDI TOF-MS method could exploit the rich redox chemistry of the transition metals. Juhasz and Costello first performed such a study on analyzing ferrocene and ruthenocene oligomers by using 2-(4-hydroxyphenylazo) benzoic acid, quinizarin, dithranol, and 9-nitroanthracene matrices [179]. They noted that positive radical ions and not protonated species were produced, even with acidic

polar matrices. These observations were not so surprising given that similar observations were made when compounds of this type were analyzed by FAB/LSI MS. They proposed two mechanisms for their generation: (1) the ion may be generated in the gas phase involving the charge exchange from a matrix radical cation to the analyte; (2) the direct photoionization of the analyte on the target surface and the assistance by the matrix to transfer the analyte ions to the gas phase. Later, Duncan reported observation of the radical cations from low molecular weight metal macrocyclic complexes [180].

Limbach investigated CT MALDI TOF-MS for analysis of metallocene and found that the gas phase electron-transfer (ET) mechanism between radical cation of matrix and analyte to form the radical cation of analyte in MALDI was more reasonable than the direct photoionization mechanism [181]. They also studied the influence of ionization energy (IE) on CT ionization and proposed that the condition for the presence of an analyte radical cation was that analyte had lower IE than matrix [182]. Vasil'ev reported the determination of the IE of DCTB (8.54 $\pm$ 0.05 eV) by applying photoelectron (PE) spectroscopy, which was in excellent agreement with the theoretical value of 8.47 eV, obtained by AM1 calculations. The same level of theory determines the electron affinity (EA) as 2.31 eV. Model analytes of known thermochemistry (phenanthrene, anthracene, and fluorofullerene) were used to bracket the CT reactivity in DCTB-MALDI. The formation of radical cations of the analytes could be expected within the thermochemical framework of DCTB [183]. The Wyatt group also discussed the influence of the IE values for the ET ionization mechanism and found that periodic trends of the metal center could also help to predict the formation of the radical cation of analytes [184].

Furthermore, based on successful detection of the radical cations of the metal complexes by CT MALDI TOF-MS methods, the accurate mass measurement helped to identify the organometallic compounds effectively by obtaining their typical elemental compositions. Wyatt group performed accurate mass measurements for radical ions of organometallics compounds by MALDI TOF-MS with standard reference materials [185, 186]. Just recently Wyatt made another breakthrough of analysis of various organic and organometallic compounds using NALDI-TOF-MS. The experiments showed the NALDI surface of silicon nanowires was effective for several nonpolar organic, organometallic, and ionic compounds in positive ion mode, as well as fluorinated compounds in positive and negative ion mode. NALDI data were compared with MALDI data for the same compounds, and the higher sensitivity of NALDI was highlighted by the successful characterization of two porphyrins for a sample amount of 10 amol per spot [187].

Kerton probed the reactions of lanthanide amide reagents $Ln(N(SiMe_3)_2)_3$ (where Ln = Sm, Gd, Ho, or Yb) with amine-bis(phenol) ligands by using inert-atmosphere MALDI TOF-MS with anthracene as matrix. This technique rapidly confirmed ligand coordination and gave an excellent agreement with theoretical isotope patterns for lanthanide (amine-phenolate) fragments [188]. Karlin used MALDI TOF-MS to characterize a high-spin, five coordinated peroxo adduct $[(^6L)Fe^{III}–(O_2^{2-})–Cu^{II}]^+$ of an iron^{II}–copper^{I} complex, obtained by reaction of $[(^6L)Fe^{II}Cu^{I}]^+$ with O_2 [189]. Such a complex was believed to have a relevant

role in the heme-copper dioxygen reactivity relevant to cytochrome-*c* oxidase O_2-reduction chemistry. Thus, based on characterization of the transition metal complexes by inert-atmosphere CT MALDI TOF-MS, the chemical transformations of these complexes could also be monitored and studied.

Beside the aromatic and fused-ring aromatic compounds used as the matrix in CT MALDI TOF-MS analysis, there were reports about the utilization of other charge-transfer matrices, especially in the characterization of neutral organometallics, which were susceptible to degradation in the presence of protons and were hard to ionize and be characterized by other MS methods. Michalak developed C_{60} as CT matrix for MALDI TOF-MS analysis of biomolecules and obtained their radical ions in both positive and negative ion mode [190]. Liu studied several series of tungstate ion clusters formed in MALDI FTICR-MS analysis of Keggin-type silicopolyoxotungstate anions [191]. The experimental results showed that matrix applied in MALDI FTICR-MS analysis has an influence on the cluster ion production. Brune applied elemental sulfur as a matrix for analysis of photosynthetic pigments and fullerenes by MALDI TOF-MS [192]. Recently sulfur was further used as CT matrix in MALDI TOF-MS to characterize various neutral Grubbs catalysts and some ferrocene derivatives to give their cationic radical ions [193]. Petković investigated Pt(II), Pt(IV), Pd(II), and Ru(III) complexes by using flavonoids as matrices for MALDI TOF-MS [194]. Use of inert-atmosphere CT MALDI-MS as a time-resolved method offers new ways to simplify the study of transition metal catalyst activation and deactivation pathways. Such studies deserved attention and usage in the challenging research on transition metal catalysis.

6.3 Probing Chemical Transformations by MALDI TOF-MS

As an extension of the use of MALDI TOF-MS in analyzing small molecular organometallics and polymer, some organic reactions involving these species nowadays could also be monitored and studied by MALDI TOF-MS, especially polymerization reactions, which were difficult to study by ESI due to high contamination and blocking of capillaries in ESI ion source by newly formed polymer. Thus the propagation steps of the polymerization reaction could be monitored by MALDI TOF-MS. Moreover, MALDI TOF-MS spectra give information not only about the polymers' molecular weight distribution and polydispersity index (PDI) but also about the nature of their repeat units and end groups [195, 196]. However, polymer analysis by MALDI TOF-MS was usually limited to structures containing heteroatoms or unsaturated functionalities as ionization sites in MS analysis. In addition, MALDI did not provide accurate determination of the molecular weight distribution of polydisperse samples since it usually underestimated the molecular weight, because heavier species are harder to vaporize and detect. Coupling chromatography (GPC) with MALDI mass spectrometry overcame this limitation and was widely used for polymer characterization [13, 28].

Harruna studied several polystyrene polymers, which were prepared by reversible addition-fragmentation chain-transfer (RAFT) polymerization of styrene

with two different RAFT agent-initiator systems, by nuclear magnetic resonance (NMR) and size exclusion chromatography (SEC), as well as MALDI TOF-MS techniques. The structures arising from the intermediate RAFT radicals and their cross-termination adducts were detected by MALDI TOF-MS, showing narrow molecular weight distribution and confirming the operation of the Rizzardo mechanism including the Monteiro intermediate radical termination model for RAFT polymerization [197]. Kostjuk studied an anionic ring-opening polymerization of hexafluoropropylene oxide using the conventional alkali metal fluorides/tetraglyme catalytic system in the presence of different fluorinated solvents at various temperatures by MALDI TOF-MS. Polymers with chain ends of methyl esters were fully characterized by gas chromatography, NMR, and MALDI TOF-MS [198]. Kamigaito reported the simultaneous chain-growth and step-growth polymerization via the metal-catalyzed radical copolymerization of conjugated vinyl monomers and designed monomers containing unconjugated C═C and active C–Cl bonds. The polymerization mechanism was studied in detail by NMR and MALDI TOF-MS for the polymerizations and the model reactions [199].

MALDI TOF-MS has also been applied to the study of ring-opening metathesis polymerization (ROMP) of 1,3,5,7-cyclooctatetraene (COT) in the presence of a chain transfer agent with Ru-olefin metathesis catalyst [200]. Astruc used MALDI TOF-MS to study polycycles, cyclophanes, and capsules, generated by metathesis reactions of perallylated arenes [201]. Wagener employed MALDI TOF-MS to examine polymers, generated via acyclic diene metathesis (ADMET) polymerization using a Ru-metathesis catalyst, for analysis of ADMET polymers having amino acid pendant groups placed at specific positions along the polyolefin backbone. The MALDI spectra clearly showed that olefin isomerization competes with propagation when different catalysts were employed, which results in the loss of precise control of polymer structures [202]. MALDI TOF-MS analysis identified the mass of the oligomeric chains, allowing for increased confidence in assignment of the possible chain structures. Furthermore, Wagener found competing metathesis and isomerization during ADMET depended on the structural properties; the amino acid functionalized olefins and the olefins displayed different selectivity toward main metathesis or isomerized products. In addition, this study detected dependence of the product selectivity on the olefin functional group properties. Two possible product distributions caused by different olefin activities were related, based on previous literature reports, to different kinetics of the reaction due to different coordination of the olefins to the catalyst center [203].

Meanwhile, the microfluidic system [204, 205] could be used together with MALDI TOF-MS to screen and optimize organic reaction conditions on a submicrogram scale. Ismagilov developed a system, which used discrete droplets (plugs) as microreactors separated and transported by a continuous phase of a fluorinated carrier fluid, for performing submicrogram or nanoliter scale reactions [206]. The evaluation of the reaction efficiency by MALDI-MS could be achieved by comparing the fraction of peak area of products for each reagent and reaction condition because MALDI TOF-MS can be used to characterize reactions quantitatively or semiquantitatively.

7 Conclusion and Perspectives

This review describes the recent advances and applications of MALDI TOF-MS analysis for small organic, biochemical and organometallic compounds. The MALDI TOF-MS analysis strategies for SMCs in a direct manner or in other methods, such as applying specific matrices, coupling to chromatographic separation, and chemical derivatization, are highlighted. The tolerance of salt and contamination, as well as the potential automated procedure, represented the orientation for high throughput characteristics of MALDI TOF-MS technology with easy sample preparation. Meanwhile the high resolution measurement power and the convenient tandem mass spectrometric function of modern MALDI TOF-MS methods provided high quality and accurate MS characterization for target SMCs, which even allowed the locating of SMCs on bio-tissues by mass spectrometric imaging. The introduction of derivatization significantly increased the detection sensitivity and selectivity, especially by attaching "charge tags" or "isotopically labeled tags," to profile and quantify the target SMCs by MALDI TOF-MS. Furthermore, MALDI TOF-MS methods to study the kinetics of enzyme reactions and to screen enzyme inhibitors became valuable tools for discovering potential drugs. In addition, chemical reactions could be monitored by MALDI TOF-MS for evaluating reaction conditions, screening catalysts, and revealing the details of reaction processes. Therefore, MALDI TOF-MS technology could provide valuable information and deep knowledge to explore and imagine new fields of the chemical world, especially relating to SMCs analysis, as well as offering new opportunities for innovations in analytical, biochemical, and organic chemistry research.

Acknowledgments The authors express great thanks to Prof. Dr. Deryn E. Fogg in University of Ottawa and Jürgen Schiller in Universität Leipzig for their generosity and for kindly providing their publications. The authors acknowledge financial support from NSFC (20902104, 21072215, and 21172250), Innovation Method Fund of China (2010IM030900), and CAS (YZ200938, YG2010056).

References

1. Karas M, Bachman D, Bahr U, Hillenkamp F (1987) Int J Mass Spectrom Ion Process 78:53–68
2. Tanaka K, Waki H, Ido Y, Akita S, Yoshida Y, Yoshida T, Matsuo T (1988) Rapid Commun Mass Spectrom 2:151–153
3. Castro JA, Köster C, Wilkins C, Cotter R (1992) Rapid Commun Mass Spectrom 6:239–241
4. Karas M, Hillenkamp F (1988) Anal Chem 60:2299–2301
5. Mustafa DAN, Burgers PC, Dekker LJ, Charif H, Titulaer MK, Sillevis Smitt PAE, Luider TM, Kros JM (2007) Mol Cell Proteomics 6:1147–1157
6. Dekker LJ, Burgers PC, Gűzel C, Luider TM (2007) J Chromatogr B 847:62–64
7. Franz AH, Molinski TF, Lebrilla CB (2001) J Am Soc Mass Spectrom 12:1254–1261
8. Mizuno Y, Sasagawa T, Dohmae N, Takio K (1999) Anal Chem 71:4764–4771
9. Dey M, Castoro JA, Wilkins CL (1995) Anal Chem 67:1575–1579

10. Li QB, Li FX, Jia L, Li Y, Liu YC, Yu JY, Fang Q, Cao AM (2006) Biomacromolecules 7:2377–2387
11. Senko MW, McLafferty FW (1994) Annu Rev Biophys Biomol Struct 23:763–785
12. Cohen LH, Gusev AI (2002) Anal Bioanal Chem 373:571–586
13. Hillenkamp F, Peter-Katalinić J (2007) MALDI MS – a practical guide to instrumentation, methods and application. Wiley-VCH, Weinheim
14. Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM (1989) Science 246:64–71
15. Hager JW (2004) Anal Bioanal Chem 378:845–850
16. Yergey AL, Coorssen JR, Backlund PS, Blank PS, Humphrey GA, Zimmerberg J, Campbell JM, Vestal ML (2002) J Am Soc Mass Spectrom 13:784–791
17. Hunnam V, Harvey DJ, Priestman DA, Bateman RH, Bordoli RS, Tyldesley R (2001) J Am Soc Mass Spectrom 12:1220–1225
18. Suzuki Y, Suzuki M, Ito E, Goto-Inoue N, Miseki K, Lida J, Yamazaki Y, Yamada M, Suzuki A (2006) J Biochem 139:771–777
19. Vestal ML, Campbell JM (2005) Methods Enzymol 402:79–108
20. Gross JH (2004) Mass spectrometry: a textbook. Springer, Berlin
21. Ingendoh A, Karas M, Hillenkamp F, Giessmann U (1994) Int J Mass Spectrom Ion Process 131:345–354
22. van Kampen JJA, Burgers PC, de Groot R, Gruters RA, Luider TM (2011) Mass Spectrom Rev 30:101–120
23. Higashi T, Shimada K (2004) Anal Bioanal Chem 378:875–882
24. Wang HY, Chu X, Zhao ZX, He XS, Guo YL (2011) J Chromatogr B 879:1166–1179
25. Fuchs B, Schiller J (2009) Eur J Lipid Sci Technol 111:83–98
26. Fuchs B, Schiller J (2009) Curr Org Chem 13:1664–1681
27. Johnson BFG, McIndoe JS (2000) Coord Chem Rev 200:901–932
28. Montaudo G, Samperi F, Montaudo MS (2006) Prog Polym Sci 31:277–357
29. Crecelius AC, Baumgaertel A, Schubert US (2009) J Mass Spectrom 44:1277–1286
30. Eelman MD, Moriarty MM, Fogg DE (2006) Educ Adv Chem 10:213–234
31. Mitra S (2003) Sample preparation techniques in analytical chemistry. Wiley, Hoboken
32. Prakash C, Shaffer CL, Nedderman A (2007) Mass Spectrom Rev 26:340–369
33. Toyo'oka T (1999) Modern derivatization methods for separation sciences. Wiley, Chichester
34. Zaikin V, Halket JM (2009) A handbook of derivatives for mass spectrometry. IM Publications, Chichester
35. Chapman JR (1995) Encyclopedia of analytical science. Academic, London
36. Raynie DE (2004) Anal Chem 76:4659–4664
37. Laiko VV, Baldwin MA, Burlingame AL (2000) Anal Chem 72:652–657
38. Doroshenko VM, Laiko VV, Taranenko NI, Berkout VD, Lee HS (2002) Int J Mass Spectrom 221:39–58
39. Su AK, Lin CH (2006) Talanta 68:673–678
40. Berkenkamp S, Kirpekar F, Hillenkamp F (1998) Science 281:260–262
41. Zenobi R, Knochenmuss R (1998) Mass Spectrom Rev 17:337–366
42. Fitzgerald MC, Parr GR, Smith LM (1993) Anal Chem 65:3204–3211
43. Knochenmuss R, Zenobi R (2003) Chem Rev 103:441–452
44. Knochenmuss R (2006) Analyst 131:966–986
45. Bailes J, Vidal L, Ivanov DA, Soloviev M (2009) J Nanobiotechnology 7:10
46. Wei J, Buriak JM, Siuzdak G (1999) Nature 399:243–246
47. Nayak R, Knapp DR (2010) Anal Chem 82:7772–7778
48. Tarui A, Kawasaki H, Taiko T, Watanabe T, Yonezawa T, Arakawa R (2009) J Nanosci Nanotechnol 9:159–164
49. Bi H, Qiao L, Busnel JM, Devaud V, Liu B, Girault HH (2009) Anal Chem 81:1177–1183
50. Daniels RH, Dikler S, Li E, Stacey C (2008) J Assoc Lab Automation 13:314–321
51. Yao T, Kawasaki H, Watanabe T, Arakawa R (2010) Int J Mass spectrom 291:145–151

52. Dutta TK, Harayama S (2001) Anal Chem 73:864–869
53. Kalberer M, Paulsen D, Sax M, Steinbacher M, Dommen J, Prevot ASH, Fisseha R, Weingartner E, Frankevich V, Zenobi R, Baltensperger U (2004) Science 303:1659–1662
54. Gnaser H, Savina MR, Calaway WF, Tripa CE, Veryovkin IV, Pellin MJ (2005) Int J Mass spectrom 245:61–67
55. Trimpin S, Herath TN, Inutan ED, Wager-Miller J, Kowalski P, Claude E, Walker JM, Mackie K (2009) Anal Chem 82:359–367
56. Becker JS, Dietze HJ (2000) Int J Mass spectrom 197:1–35
57. Annesley TM (2003) Clin Chem 49:1041–1044
58. Fuchs B, Süß R, Nimptsch A, Schiller J (2009) Chromatographia 69:95–105
59. Hillenkamp F, Karas M, Beavis RC, Chait BT (1991) Anal Chem 63:1193A–1203A
60. Hardouin J (2007) Mass Spectrom Rev 26:672–682
61. McComb ME, Perlman DH, Huang H, Costello CE (2007) Rapid Commun Mass Spectrom 21:44–58
62. Dunphy JC, Busch KL, Hettich RL, Buchanan MV (1993) Anal Chem 65:1329–1335
63. Bogan MJ, Agnes GR (2004) Rapid Commun Mass Spectrom 18:2673–2681
64. Therisod H, Labas V, Caroff M (2001) Anal Chem 73:3804–3807
65. Rejtar T, Hu P, Juhasz P, Campbell JM, Vestal ML, Preisler J, Karger BL (2002) J Proteome Res 1:171–179
66. Preisler J, Foret F, Karger BL (1998) Anal Chem 70:5278–5287
67. Ørsnes H, Graf T, Degn H, Murray KK (2000) Anal Chem 72:251–254
68. Guo Z, Zhang QC, Zou HF, Guo BC, Ni JY (2002) Anal Chem 74:1637–1641
69. Welker M, Fastner J, Erhard M, Von Dohren H (2002) Environ Toxicol 17:367–374
70. Grant GA, Frison SL, Yeung J, Vasanthan T, Sporns P (2003) J Agric Food Chem 51:6137–6144
71. Yang HM, Wang JW, Song FR, Zhou YH, Liu SY (2011) Anal Chim Acta 701:45–51
72. Ayorinde FO, Elhilo E, Hlongwane C (1999) Rapid Commun Mass Spectrom 13:737–739
73. Ayorinde FO, Garvin K, Saeed K (2000) Rapid Commun Mass Spectrom 14:608–615
74. Kaltashov IA, Doroshenko V, Cotter RJ, Takayama K, Qureshi N (1997) Anal Chem 69:2317–2322
75. Harvey DJ (1995) J Mass Spectrom 30:1333–1346
76. Al-Saad KA, Zabrouskov V, Siems WF, Knowles NR, Hannan RM, Hill HH Jr (2003) Rapid Commun Mass Spectrom 17:87–96
77. Griffiths WJ, Liu S, Alvelius G, Sjövall J (2003) Rapid Commun Mass Spectrom 17:924–935
78. Griffiths WJ (2003) Mass Spectrom Rev 22:81–152
79. Tholey A, Heinzle E (2006) Anal Bioanal Chem 386:24–37
80. Pan CS, Xu SY, Zhou H, Fu Y, Ye ML, Zou HF (2007) Anal Bioanal Chem 387:193–204
81. van Kampen JJA, Burgers PC, de Groot R, Luider TM (2006) Anal Chem 78:5403–5411
82. Ayorinde FO, Hambright P, Porter TN, Keith QL Jr (1999) Rapid Commun Mass Spectrom 13:2474–2479
83. Zhang S, Liu JA, Chen Y, Xiong SX, Wang GH, Chen J, Yang GQ (2010) J Am Soc Mass Spectrorm 21:154–160
84. Li YL, Gross ML (2004) J Am Soc Mass Spectrom 15:1833–1837
85. Santos LS, Haddad R, Höehr NF, Pilli RA, Eberlin MN (2004) Anal Chem 76:2144–2147
86. Catharino RR, de Azevedo Marques L, Santos LS, Baptista AS, Glória EM, Calori-Domingues MA, Facco EMP, Eberlin MN (2005) Anal Chem 77:8155–8157
87. Ham BM, Jacob JT, Cole RB (2005) Anal Chem 77:4439–4447
88. Laremore TN, Zhang FM, Linhardt RJ (2007) Anal Chem 79:1604–1610
89. Shen ZX, Thomas JJ, Averbuj C, Broo KM, Engelhard M, Crowell JE, Finn MG, Siuzdak G (2001) Anal Chem 73:612–619
90. Li XH, Wu X, Kim JM, Kim SS, Jin M, Li DH (2009) J Am Soc Mass Spectrom 20:2167–2173
91. Xu SY, Li YF, Zou HF, Qiu JS, Guo Z, Guo BC (2003) Anal Chem 75:6191–6195

92. Ren SF, Zhang L, Cheng ZH, Guo YL (2005) J Am Soc Mass Spectrom 16:333–339
93. Ren SF, Guo YL (2005) Rapid Commun Mass Spectrom 19:255–260
94. Zhang J, Wang HY, Guo YL (2005) Chin J Chem 23:185–189
95. Wang CH, Li J, Yao SJ, Guo YL, Xia XH (2007) Anal Chim Acta 604:158–164
96. Ren SF, Guo YL (2006) J Am Soc Mass Spectrom 17:1023–1027
97. Dong XL, Cheng JS, Li JH, Wang YS (2010) Anal Chem 82:6208–6214
98. Lu MH, Lai YQ, Chen GN, Cai ZW (2011) Anal Chem 83:3161–3169
99. Tholey A, Wittmann C, Kang MJ, Bungert D, Hollemeyer K, Heinzle E (2002) J Mass Spectrom 37:963–973
100. Adamczyk M, Gebler JC, Wu J (1999) Rapid Commun Mass Spectrom 13:1413–1422
101. Lee PJ, Chen WB, Gebler JC (2004) Anal Chem 76:4888–4893
102. Denekamp C, Lacour J, Laleu B, Rabkin E (2008) J Mass Spectrom 43:623–627
103. Schiewek R, Mönnikes R, Wulf V, Gäb S, Brockmann KJ, Benter T, Schmitz OJ (2008) Angew Chem Int Ed 47:9989–9992
104. Zhang J, Guo YL (2005) Rapid Commun Mass Spectrom 19:2461–2464
105. Zhang J, Zhang L, Zhou Y, Guo YL (2007) J Mass Spectrom 42:1514–1521
106. Fang F, Zhang J, Zhang L, Guo YL (2009) Chin J Chem 27:2397–2404
107. Fang F, Liu P, Wang HY, Zhang L, Zhang J, Gao YP, Zeng LM, Guo YL (2009) Rapid Commun Mass Spectrom 23:1703–1709
108. Wang H, Wang HY, Zhang L, Zhang J, Guo YL (2011) Anal Chim Acta 690:1–9
109. Wang H, Wang HY, Zhang L, Zhang J, Leng JP, Cai TT, Guo YL (2011) Anal Chim Acta 707:100–106
110. Zaia J (2004) Mass Spectrom Rev 23:161–227
111. Gouw JW, Burgers PC, Trikoupis MA, Terlouw JK (2002) Rapid Commun Mass Spectrom 16:905–912
112. Naven TJP, Harvey DJ (1996) Rapid Commun Mass Spectrom 10:829–834
113. Okamoto M, Takahashi K, Doi T, Takimoto Y (1997) Anal Chem 69:2919–2926
114. Broberg S, Broberg A, Duus JØ (2000) Rapid Commun Mass Spectrom 14:1801–1805
115. Zhang XW, Wei D, Yap YL, Li L, Guo SY, Chen F (2007) Mass Spectrom Rev 26:403–431
116. Dettmer K, Aronov PA, Hammock BD (2007) Mass Spectrom Rev 26:51–78
117. Han XL, Gross RW (2005) Mass Spectrom Rev 24:367–412
118. Nicholson JK, Lindon JC, Holmes E (1999) Xenobiotica 29:1181–1189
119. Smith CA, O'Maille G, Want EJ, Qin C, Trauger SA, Brandon TR, Custodio DE, Abagyan R, Siuzdak G (2005) Ther Drug Monit 27:747–751
120. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G (2006) Anal Chem 78:779–787
121. Brown SC, Kruppa G, Dasseux JL (2005) Mass Spectrom Rev 24:223–231
122. Northen TR, Yanes O, Northen MT, Marrinucci D, Uritboonthai W, Apon J, Golledge SL, Nordström A, Siuzdak G (2007) Nature 449:1033–1036
123. Woo HK, Northern TR, Yanes O, Siuzdak G (2008) Nat Protoc 3:1341–1349
124. Wenk MR (2005) Nat Rev Drug Discov 4:594–610
125. Murphy RC, Fiedler J, Hevko J (2001) Chem Rev 101:479–526
126. Fuchs B, Schiller J (2008) Subcell Biochem 49:541–565
127. Gross RW, Han X (2007) Methods Enzymol 433:73–90
128. Schiller J, Suss R, Fuchs B, Muller M, Zschornig O, Arnold K (2007) Front Biosci 12:2568–2579
129. Murphy RC, Hankin JA, Barkley RM (2008) J Lipid Res 50:S317–S322
130. Yetukuri L, Ekroos K, Vidal-Puig A, Oresic M (2008) Mol Biosyst 4:121–217
131. Greis KD (2007) Mass Spectrom Rev 26:324–339
132. de Boer AR, Lingeman H, Niessen WMA, Irth H (2007) TrAc, Trends Anal Chem 26:867–883
133. Houston CT, Taylor WP, Widlanski TS, Reilly JP (2000) Anal Chem 72:3311–3319
134. Nichols KP, Gardeniers HJGE (2007) Anal Chem 79:8699–8704

135. Doorn JA, Schall M, Gage DA, Talley TT, Thompson CM, Richardson RJ (2001) Toxicol Appl Pharmacol 176:73–80
136. Hu LG, Jiang GB, Xu SY, Pan CS, Zou HF (2006) J Am Soc Mass Spectrom 17:1616–1619
137. Persike M, Zimmermann M, Klein J, Karas M (2010) Anal Chem 82:922–929
138. Xu Z, Yao SJ, Wei YL, Zhou J, Zhang L, Wang CH, Guo YL (2008) J Am Soc Mass Spectrom 19:1849–1855
139. Cai TT, Zhang L, Wang HY, Zhang J, Guo YL (2011) Anal Chim Acta 706:291–296
140. Northen TR, Lee JC, Hoang L, Raymond J, Hwang DR, Yannone SM, Wong CH, Siuzdak G (2008) Proc Natl Acad Sci USA 105:3678–3683
141. Yanes O, Villanueva J, Querol E, Aviles FX (2005) Mol Cell Proteomics 4:1602–1613
142. Kondo N, Nishimura SI (2009) Chem Eur J 15:1413–1421
143. Xu Z, Wang HY, Huang SX, Wei YL, Yao SJ, Guo YL (2010) Anal Chem 82:2113–2118
144. Caprioli RM, Farmer TB, Gile J (1997) Anal Chem 69:4751–4760
145. Luxemborg SL, Mize TH, McDonnell LA, Heeren RMA (2004) Anal Chem 76:5339–5344
146. Reyzer ML, Chaurand P, Angel PM, Caprioli RM (2010) Methods Mol Biol 656:285–301
147. Reyzer ML, Caprioli RM (2007) Curr Opin Chem Biol 11:29–35
148. Hsieh YS, Korfmacher WA (2008) MALDI imaging mass spectrometry for direct tissue analysis of pharmaceuticals. In: Ramanathan R (ed) Mass spectrometry in drug metabolism and pharmacokinetics. Wiley, Hoboken, pp 359–382
149. Reyzer ML, Hsieh Y, Ng K, Korfmacher WA, Caprioli RM (2003) J Mass Spectrom 38:1081–1092
150. Grove KJ, Frappier SL, Caprioli RM (2011) J Am Soc Mass Spectrom 22:192–195
151. Khabit-Shahidi S, Andersson M, Herman JL, Gillespie TA, Caprioli RM (2006) Anal Chem 78:6448–6456
152. Puolitaival SM, Burnum KE, Cornett DS, Caprioli RM (2008) J Am Soc Mass Spectrom 19:882–886
153. Garden RW, Sweedler JV (2000) Anal Chem 72:30–36
154. Hsieh Y, Casale R, Fukuda E, Chen JW, Knemeyer I, Wingate J, Morrison R, Korfmacher W (2006) Rapid Commun Mass Spectrom 20:965–972
155. Amantionico A, Zenobi R (2006) 17th international mass spectrometry conference, Prague, 27th August to 1st September 2006
156. Jackson SN, Wang HY, Woods AS, Ugarov M, Egan T, Schultz JA (2005) J Am Soc Mass Spectrom 16:133–138
157. Woods AS, Ugarov M, Jackson SN, Egan T, Wang HY, Murray KK, Schultz JA (2006) J Preoteome Res 5:1484–1487
158. McLean JA, Ridenour WB, Caprioli RM (2007) J Mass Spectrom 42:1099–1105
159. Korfmacher WA (2005) Bioanalytical assays in a drug discovery environment. In: Using mass spectrometry for drug metabolism studies. CRC, Boca Raton, pp 1–34
160. Korfmacher WA (2010) Using mass spectrometry for drug metabolism studies, 2nd edn. CRC, Boca Raton, p 429
161. Franck J, El Ayed M, Wisztorski M, Salzet M, Fournier I (2009) Anal Chem 81:8305–8317
162. Chacon A, Zagol-Ikapitte I, Amarnath V, Reyzer ML, Oates JA, Caprioli RM, Boutaud O (2011) J Mass Spectrom 46:840–846
163. Traeger JC (2000) Int J Mass Spectrom 200:387–401
164. Henderson W, Mcindoe SJ (2005) Mass spectrometry of inorganic and organometallic compounds: tools techniques tips. Wiley, Chichester
165. Ulmer L, Mattay J, Torres-Garcia HG, Luftmann H (2000) Eur J Mass Spectrom 6:49–52
166. Brown T, Clipston NL, Simjee N, Luftmann H, Hungerbühler H, Drewello T (2001) Int J Mass Spectrom 210:249–263
167. Lou X, van Buijtenen J, Bastiaansen JJAM, de Waal BFM, Langeveld BMW, van Dongen JLJ (2005) J Mass Spectrom 40:654–660
168. Wyatt MF, Stein BK, Brenton AG (2006) Anal Chem 78:199–206
169. Hughes L, Wyatt MF, Stein BK, Brenton AG (2009) Anal Chem 81:543–550

170. Lou XW, de Waal BFM, van Dongen JLJ, Vekemans JAJM, Meijer EW (2010) J Mass Spectrom 45:1195–1202
171. Lubben AT, McIndoe JS, Weller AS (2008) Organometallics 27:3303–3306
172. Ham JE, Durham B, Scott JR (2003) J Am Soc Mass Spectrom 14:393–400
173. Eelman MD, Blacquiere JM, Moriarty MM, Fogg DE (2008) Angew Chem Int Ed 47:303–306
174. Trimpin S (2007) In: Gross ML, Caprioli RM (eds) Encyclopedia of mass spectrometry, vol 6. Elsevier, New York, p 683
175. Montaudo G, Carroccio S, Puglisi C (2002) J Anal Appl Pyrol 64:229–247
176. Monfette S, Fogg DE (2006) Organometallics 25:1940–1944
177. Monfette S, Duarte Silva JA, Gorelsky SI, Dalgarno SJ, dos Santos EN, Araujo MH, Fogg DE (2009) Can J Chem 87:361–367
178. de Frémont P, Stevens ED, Eelman MD, Fogg DE, Nolan SP (2006) Organometallics 25:5824–5828
179. Juhasz P, Costello CE (1993) Rapid Commun Mass Spectrom 7:343–351
180. Lidgard R, Duncan MW (1995) Rapid Commun Mass Spectrom 9:128–132
181. McCarley TD, McCarley RL, Limbach PA (1998) Anal Chem 70:4376–4379
182. Macha SF, McCarley TD, Limbach PA (1999) Anal Chim Acta 397:235–245
183. Vasil'ev YV, Khvostenko OG, Streletskii AV, Boltalina OV, Kotsiris SG, Drewello T (2006) J Phys Chem A 110:5967–5972
184. Wyatt MF, Havard S, Stein BK, Brenton AG (2008) Rapid Commun Mass Spectrom 22:11–18
185. Wyatt MF, Stein BK, Brenton AG (2006) J Am Soc Mass Spectrom 17:672–675
186. Griffiths NW, Wyatt MF, Kean SD, Graham AE, Stein BK, Brenton AG (2010) Rapid Commun Mass Spectrom 24:1629–1635
187. Wyatt MF, Ding S, Stein BK, Brenton AG, Daniels RH (2010) J Am Soc Mass Spectrom 21:1256–1259
188. Ikpo N, Butt SM, Collins KL, Kerton FM (2009) Organometallics 28:837–842
189. Ghiladi RA, Huang H, Moenne-Laccoz P, Stasser J, Blackburn NJ, Woods AS, Cotter RJ, Incarvito CD, Rheingold AL, Karlin KD (2005) J Biol Inorg Chem 10:63–77
190. Michalak L, Fisher KJ, Alderdice DS, Jardine DR, Willett GD (1994) Org Mass Spectrom 29:512–515
191. Bai YP, Liu S, Song FR, Liu ZQ, Liu SY (2012) Rapid Commun Mass Spectrom 26:715–718
192. Brune DC (1999) Rapid Commun Mass Spectrom 13:384–389
193. Zhu W, Wang HY, Guo YL (2012) J Mass Spectrom 47:352–354
194. Petković M, Petrović B, Savić J, Bugarčić ŽD, Dimitrić-Marković J, Momić T, Vasić V (2010) Int J Mass Spectrom 290:39–46
195. Liu JS, Loewe RS, McCullough RD (1999) Macromolecules 32:5777–5785
196. Puglisi C, Samperi F, Alicata R, Montaudo G (2002) Macromolecules 35:3000–3007
197. Zhou GC, Harruna II (2007) Anal Chem 79:2722–2727
198. Kostjuk SV, Ortega E, Ganachaud F, Améduri B, Boutevin B (2009) Macromolecules 42:612–619
199. Mizutani M, Satoh K, Kamigaito M (2010) J Am Chem Soc 132:7498–7507
200. Scherman OA, Rutenberg IM, Grubbs RH (2003) J Am Chem Soc 125:8515–8522
201. Martinez V, Blais JC, Bravic G, Astruc D (2004) Organometallics 23:861–874
202. Petkovska VI, Hopkins TE, Powell DH, Wagener KB (2005) Macromolecules 38:5878–5885
203. Petkovska VI, Hopkins TE, Powell DH, Wagener KB (2006) Anal Chem 78:3624–3631
204. Mason BP, Price KE, Steinbacher JL, Bogdan AR, McQuade DT (2007) Chem Rev 107:2300–2318
205. McMullen JP, Jensen KF (2010) Ann Rev Anal Chem 3:19–42
206. Hatakeyama T, Chen DL, Ismagilov RF (2006) J Am Chem Soc 128:2518–2519

Top Curr Chem (2013) 331: 193–210
DOI: 10.1007/128_2012_365
© Springer-Verlag Berlin Heidelberg 2012
Published online: 14 September 2012

Bioinformatic Analysis of Data Generated from MALDI Mass Spectrometry for Biomarker Discovery

Zengyou He, Robert Z. Qi, and Weichuan Yu

Abstract In this chapter we first describe the applications of matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (MS) in biomarker discovery. After a summary of the general analysis pipeline of MALDI MS data, each step of the pipeline will be elaborated in detail. In particular we try to provide a categorization of existing solutions with the hope that the reader can obtain a global picture on this topic. In addition we show how to apply such an analysis pipeline in protein and glycan profiling for biomarker discovery and for a deeper understanding of diseases. Finally we discuss the limitations of current analysis methods and the perspectives of future research.

Keywords Biomarker Discovery • Data Mining • Feature Selection • MALDI MS Data • Peak Detection

Contents

Z. He (✉)
School of Software, Dalian University of Technology, Dalian, China
e-mail: zyhe@dlut.edu.cn

R.Z. Qi
Division of Life Science and State Key Laboratory of Molecular Neuroscience, The Hong Kong University of Science and Technology, Hong Kong, China
e-mail: qirz@ust.hk

W. Yu
Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, China
e-mail: eeyu@ust.hk

1 Introduction

In order to facilitate diagnosis and prognosis of diseases, biomarker identification by proteomic profiling of human tissue or cell specimens has become popular in clinical proteomics [1]. Mass spectrometry (MS) has proved to be the most promising tool for generating protein profiles of tissue, serum, and urine samples. In particular, the matrix-assisted laser desorption/ionization (MALDI) MS and its variants (e.g., SELDI and MALDI-TOF/TOF) have often been employed for such a purpose [2].

Ideally, the MALDI device should only generate signals that correspond to ionized proteins or peptides. However, the acquired spectra contain not only "true" peaks that represent proteins or peptides of scientific interest but also baseline drift and substantial background noise. The baseline is formed by the matrix materials and impurities that enter the analyzer as by-products of ionization. As the matrix materials are often fragmented to have small masses, the baseline shows a slowly decreasing trend along the mass-over-charge axis. The background noises are produced by electronic disturbances and impurities from the samples, with rapid fluctuations randomly varying over small mass ranges.

As a consequence, the identification of biomarkers from MALDI MS data becomes a very challenging task. It requires the management and analysis of large amounts of data in quite complex ways. One common practice is to divide the analysis efforts into different stages and thus to reduce the complexity. In general, it consists of the following three steps:

1. *Feature extraction*: Fundamental to any analysis of MS data is the extraction of "signals" or "features" from each spectrum.
2. *Feature alignment*: Feature alignment establishes the correspondence among biological features extracted from different spectra. In other words, the aim of feature alignment is to produce a two-dimensional table that can be used in biomarker selection.
3. *Feature selection*: Feature selection is to find a set of features as biomarkers to distinguish specimens from health and disease.

Existing review articles on MALDI MS data analysis either cover every analysis steps (e.g., [3]) or focus on particular methods (e.g., [4, 5]). This chapter concentrates on the categorization of recent developments and methodological

rationale. With the understanding of discussed methods, one may rapidly construct a precise computational pipeline for particular applications. We also discuss the limitations of current analysis methods and explore the use of case-based reasoning (CBR) as a remedy to facilitate data analysis.

MALDI MS produces singly charged ions from target samples, thus simplifying the interpretation of the spectra. Note that some analysis methods such as baseline removal are unique to MALDI MS data, while other methods are applicable to different kinds of MS data. The scope of the present chapter focuses on the bioinformatic methods used for analysis of MALDI MS data.

2 Overview of Currently Used Methods

2.1 Framework

A schematic overview of MALDI MS data analysis steps is given in Fig. 1. In feature extraction, each mass spectrum is preprocessed with smoothing and baseline removal before peak detection and quantification. Feature alignment takes features from multiple spectra as inputs to generate a set of consensus features. Feature selection reports a subset of common features with high classification accuracy as biomarkers.

Furthermore, both statistical and biological validation need to be performed on reported biomarkers. First, the feature selection results should be able to pass the statistical test. Next, a new set of samples should be measured independently to test the marker sensitivity. Finally, we need to conduct experiments to determine biological validity. Since the current chapter focuses on data analysis, we will concentrate on statistical validation in the forthcoming sections.

2.2 Feature Extraction

All current MALDI MS feature extraction methods share the same goal: to extract and quantify peaks of interest accurately and make the result applicable for further analysis [6]. In general, a feature extraction procedure consists of four steps: smoothing, baseline removal, peak detection, and peak quantification. Note that some steps such as smoothing and baseline removal may switch their locations in the pipeline [4]. Furthermore, it is also possible that some steps are merged or some additional steps such as calibration are included.

Feature extraction is the most important step and has a great impact on the accuracy of biomarker identification. This is because all subsequent analysis steps have to utilize the output of feature extraction as input. In other words, this is probably the only opportunity to recover real signals from noisy raw MS data in the whole analysis flow.

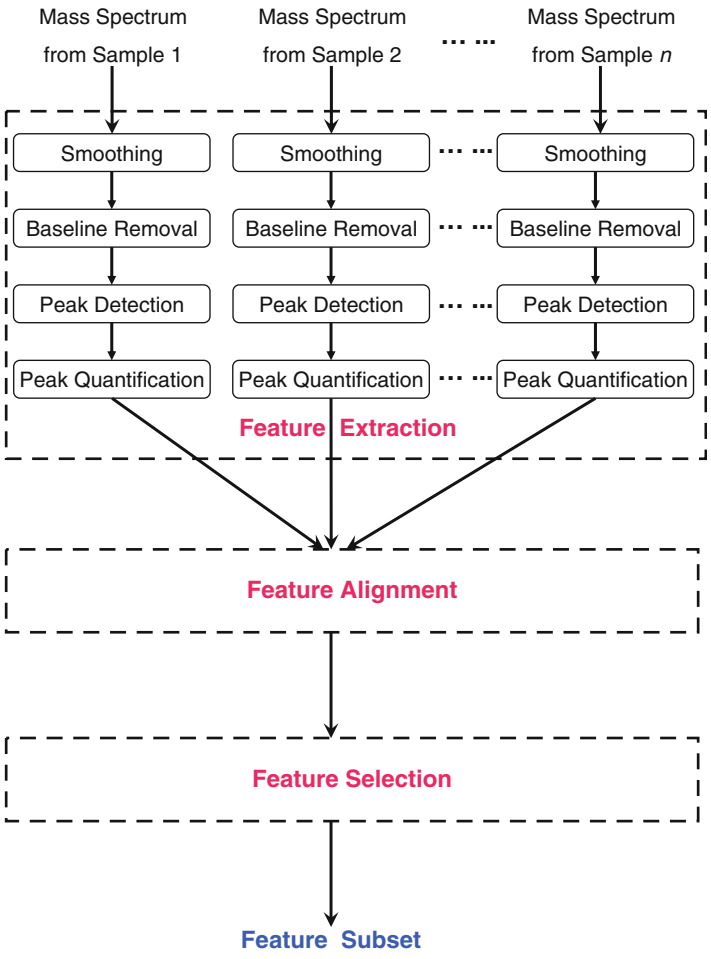


Fig. 1 A typical MALDI MS data analysis pipeline and its major components. This includes feature extraction, feature alignment, and feature selection. The reported markers will be subject to both statistical and biological validation. Commonly used feature extraction methods are intensity-based approaches, wavelet-based approaches, model-based approaches, and dimension reduction approaches. Feature alignment methods take either raw MS data or pre-processed MS data as input. Various clustering algorithms are the most widely methods for feature alignment. Existing feature selection approaches include filter methods, wrapper methods, and embedded methods

For each sub-procedure in feature extraction, people have developed many computational methods. These methods can be categorized differently according to different dimensions. Here we wish to create multiple hierarchical frameworks that are systematic and expandable, and attempt to cover all existing methods. With these considerations in mind, we propose two hierarchical frameworks in Figs. 2 and 3. Note that these two categorization methods are independent of each other. In each framework we use one specific classification criterion to organize all existing feature extraction methods.

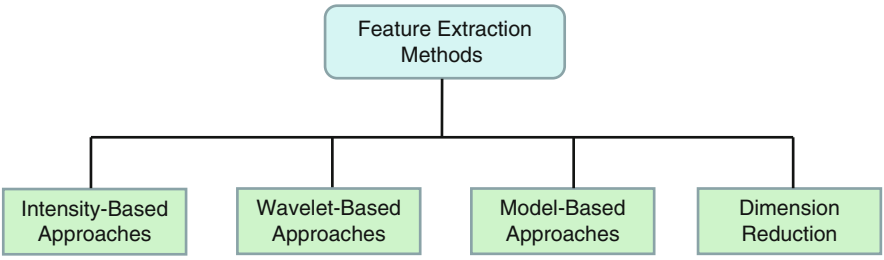


Fig. 2 The first framework for categorizing feature extraction methods. The classification criterion is the underlying algorithmic techniques used

In the first framework for categorizing available feature extraction methods, we use the underlying algorithmic technique as the classification criterion. As shown in Fig. 2, existing feature extraction methods may be broadly classified into four categories:

1. *Intensity-based approaches*: These approaches use certain thresholds to filter weak peaks and keep the most intense peaks [7]. Such intensity-based approaches are probably the most widely used feature extraction methods, e.g., [8–10]. Their main advantage is that numerous signal processing techniques are available to facilitate feature extraction. However, real peptide signals may be weaker than spectrum noise. Thus, they sometimes fail to extract true features, leading to a significant decrease in the performance of downstream analysis.
2. *Wavelet-based approaches*: These approaches can be further divided into two categories – continuous wavelet (CWT) methods (e.g., [11, 12]) and discrete wavelet (DWT) methods (e.g., [13–16]). Here we shall use CWT as an example to illustrate the basic idea of wavelet-based approach. Mathematically, the CWT can be expressed as [11]

$$C(a,b) = \int_R S(t)\Psi_{a,b}(t)\,\mathrm{d}t, \qquad \Psi_{a,b}(t) = \frac{1}{\sqrt{a}}\Psi\left(\frac{t-b}{a}\right), \tag{1}$$

where $S(t)$ is the signal, a is the scale ($a > 0$), b is the translation, $\Psi(t)$ is the mother wavelet, $\Psi_{a,b}(t)$ is the scaled and translated wavelet, and C is a two-dimensional matrix of wavelet coefficients. The wavelet coefficients can be used to measure the fitness between signal $S(t)$ and wavelet $\Psi_{a,b}(t)$. Higher coefficients reflect better matching and hence indicate the possible existence of biologically meaningful features. As shown in a recent comparison study [4], the wavelet-based approach generally exhibits good performance in practice. However, it should be noted that there is no model-based approach in the performance comparison. In this regard, a more comprehensive comparison study is still necessary.

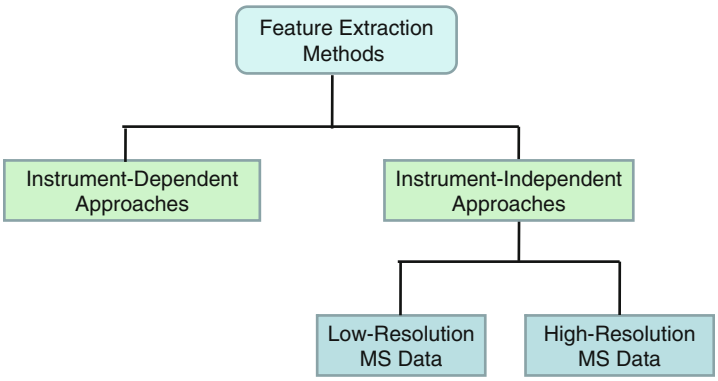


Fig. 3 The second framework for categorizing feature extraction methods. The classification criterion is the dependence of extraction algorithms on the MS instruments. Instrument-independent approaches are applicable to any spectra. In contrast, instrument-dependent approaches utilize some auxiliary information that is specific to the MS platform. Low resolution mass spectrometry data record masses to only unit mass resolution. This means that ions of the same nominal mass would be inseparable. In high resolution MS data, coincident ions are fully resolved

3. *Model-based approaches*: These approaches provide a better representation of the MS signals by incorporating information about isotopic distributions and peak shapes [17]. In high-resolution MS, the model-based approach is more appealing since peaks and wavelet functions may not directly correspond to the underlying molecules of interest. Recently, more and more research efforts have been conducted in this direction (e.g., [17–20]). The essential idea of such a model-based approach is to represent the observed spectrum as a linear combination of theoretical isotope distribution patterns. Therefore most model-based techniques are computationally expensive since they usually need to solve a large-scale optimization problem. We believe that the model-based approach will become the most reliable feature extraction method since it incorporates prior knowledge from sequence database into the detection process.
4. *Dimension reduction*: This approach utilizes dimension reduction techniques to extract features. For instance, the independent component analysis (ICA) has been applied to MALDI mass spectra in [21]. This method is less generally adopted since the generated features are hard to interpret biologically.

The second framework shown in Fig. 3 is application-oriented. To date, most feature extraction methods are independent of the MS instruments used. Such independence is plausible if we are able to reliably extract true biological signals from the data. However, when these instrument-independent approaches cannot achieve satisfactory performance, we may switch to instrument-dependent methods.

Due to the different properties of spectra from different types of MS instruments, it is very difficult to provide a feature extraction method that can always achieve good performance on all MS devices. Therefore, one alternative strategy is to use

algorithms that are designed for specific platforms [22]. For instance, the prOTOF 2000 (PerkinElmer), an orthogonal MALDI-TOF MS, has high mass accuracy and resolution over a wide mass range. The prOTOF MS spectra show a unique noise pattern in the background [23], making it possible to design more effective feature extraction methods specific to the instrument.

2.3 Feature Alignment

All alignment approaches fall into one of two very broad categories [24, 25]: they are based either on raw MS data or on pre-processed MS data. In the first alignment strategy, one attempts to compare the complete MS spectra directly so as to find an alignment with the minimal overall difference between intensities of all MS spectra and the reference spectrum. Alternatively, the second strategy utilizes only features from the feature extraction step as input. As a result, the performance of the overall alignment process strongly depends on the performance of the feature extraction [25]. To overcome this limitation, one may combine feature extraction and alignment in a feedback loop [26]. That is, when aligning all the identified features (including false-positive ones) from multiple spectra together, those false-positive features are not as consistent as true features. Thus, iteratively performing feature alignment and removing false-positive features could improve the quality of feature extraction and alignment simultaneously. More importantly, the discovery of cancer biomarkers also benefits from this improvement [26].

Despite their differences on the target data, both alignment strategies follow the same computational principle: formulating the alignment problem as a specific optimization problem and solving it using techniques from different domains. For instance, this problem has been addressed from the viewpoint of cluster analysis [27] and scale-space representation [28], respectively.

2.4 Feature Selection

Feature selection has been widely studied in statistics and machine learning for a long time. In the context of biomarker discovery, the objective of feature selection is to find a small set of features (markers) that best explains the difference between the disease and the control samples. Existing feature selection approaches can be organized into three categories [29]: filter methods, wrapper methods, and embedded methods. The filter method assesses the relevance of features by looking only at the intrinsic properties of the data. The wrapper method evaluates the goodness of feature subsets using the performance of a classification algorithm. The embedded method combines feature selection and classifier construction into an integrated procedure.

To date, various feature selection algorithms have been exploited for identifying biomarkers from MALDI data [30–34]. For a recent review of feature selection techniques for proteomic biomarker studies, one may refer to [35].

As biomarker discovery applications pose many new challenges, traditional feature selection approaches often fail to detect true markers consistently. This is because only classification accuracy is used as the selection criteria for marker identification. In high dimensional MS data, one may find many different subsets of features that can achieve the same or similar predictive accuracy. If there is only one real marker, it is obvious that such accuracy-based strategy cannot distinguish true marker from false ones effectively.

Therefore, the non-reproducibility of reported markers has become one major obstacle in biomarker discovery. Surprisingly, the analysis of the stability or robustness of feature selection techniques is only a topic of recent interest, and has not yet made it into mainstream methodology for biomarker discovery [36]. There are three main sources of instability in biomarker discovery [37]:

1. *Algorithm design without considering stability*: Traditional feature selection approaches concentrate on discovering a minimum subset of features that can achieve the best predictive accuracy, but often ignore "stability" in designing the algorithm.
2. *The existence of multiple sets of true markers*: It is possible that there are multiple sets of true markers in the MS data.
3. *Small number of samples in high dimensional data*: In the analysis of MALDI MS data, there are typically only hundreds of samples but thousands of features.

To improve the marker reproducibility, stable feature selection methods begin to receive more and more attention. Existing stable feature selection algorithms fall into four categories [37]:

1. *Ensemble feature selection*: This first uses different feature selectors and then aggregates the results of component selectors to generate the final output.
2. *Feature selection with prior feature relevance*: In most biomarker discovery applications, we assume that all features are equally relevant. Actually, some prior knowledge may be available to differentiate their relevance.
3. *Group feature selection*: This treats feature cluster as the basic unit in the selection process to improve the robustness.
4. *Sample injection*: This tries to increase the sample size to address the small-sample size vs large-feature-size issue.

Here we argue that future biomarker applications should incorporate "stability" into the feature selection process.

After feature selection, statistical and biological tests should be conducted to validate the identified markers. To reduce validation cost, a good feature selector is expected to report only a small subset of candidate markers without missing the true target.

2.5 Statistical Validation

To assess the statistical significance of a single biomarker, a permutation test permits us to choose the test statistic best suited to the task at hand [38]. In a permutation test, one basic operation is to reassign randomly class labels to samples so as to generate an uninformative data set of the same size as the original data. It repeats the above procedure multiple times to create many permutations of the data. Testing biomarkers on these permutations provides a null distribution of the performance found by chance, to which the performance on the original data can be compared. More precisely, we first compute the test statistic based on original data: t_{obj}. Then we generate B independent new data sets by randomly permuting the class label. We obtain the permuted statistic $T_{(b)}$ of target marker using the same procedure from the permuted data sets, where $1 \leq b \leq B$. Finally, the permutation p-value of the biomarker is calculated as

$$1/B \sum_{b=1}^{B} I(T_{(b)} > t_{\text{obj}}), \tag{2}$$

where $I(\cdot)$ is the indicator function.

In fact, the biomarker discovery investigation unusually generates multiple markers for validation. Suppose we need to validate m biomarkers independently and the p-value of each marker is s_i. If $s_i < \alpha$ (α is a pre-defined significance threshold), one can reject the null hypothesis that the ith marker has the same distribution for cases and controls and argue that there is a detectable or significant difference between the compared groups with respect to marker i. However, note that every time a test is conducted independently on each biomarker, there is a possibility that an error will be made. For instance, if $\alpha = 0.01$ and $m = 100$, one can expect that at least one claimed biomarker is actually not significantly differently expressed. In other words, if m biomarkers are tested simultaneously, some markers that are detected as "significant" could actually be false positives. Therefore, one has to address such "multiple hypothesis testing" problems in the validation of biomarkers. To date, the most effective strategy is to control the false discovery rate (FDR) [39].

In the context of biomarker discovery, the FDR is the expected rate of false positives among the markers that are deemed to be true. To control FDR, many methods and tools have been proposed after Benjamini and Hochberg's landmark paper in 1995 [39], as summarized in [40]. However, there is still no consensus on the best FDR control method and one has to choose the tool according to the analysis requirements and data characteristics.

Cross-validation is also widely used for validating biomarkers. It splits the data into different parts with approximately equal size. Each part is masked in turn as the test set, while the remaining parts are used to as the training set. The classifier built on the training set is then applied to predict the test set. This process is repeated until all parts have been masked once, and then the error made in all blinded test sets is combined to give an independent performance estimate.

There are many variants of cross-validation. Leave-one-out cross-validation uses only one sample as the test set. The k-fold cross-validation partitions the data into approximately k parts. The stratified cross-validation preserves the ratio of the class sizes in the training and test sets so as to make them accurate representations of the original data. These cross-validation procedures are available in many data mining tools (e.g., Weka [41]).

Note that a statistically valid biomarker should always be subject to biological validation since it can be biologically irrelevant. Even the most thorough statistical procedure cannot safeguard against this type of findings [42].

3 Case Studies

3.1 Protein/Peptide Profiling

Most clinical proteomics studies are designed to measure proteins and peptides in human body fluids. The analysis of protein profiling data uses the pipeline we have discussed in the previous section. In practice, the user needs to specify the algorithm used in each step. We use the work by Ressom et al. [31] as an example to elaborate on the data analysis process.

The objective in [31] is to identify candidate markers that distinguish hepatocellular carcinoma (HCC) from cirrhosis through analysis of enriched low-molecular-weight (LMW) serum samples. Through the removal of proteins >50 kDa, it improves quality of the spectra and allows the analysis of around 300 peptides. There are 215 non-replicate spectra, which were generated using 84 sera from HCC patients, 51 sera from cirrhotic patients, and 80 sera from healthy individuals.

Before actual data analysis, the authors used outlier screening to exclude spectra whose total ion current (TIC) differed by more than two standard deviation from the median TIC. In addition, each remaining spectrum was binned with a size of 100 ppm to reduce the dimension size. The mean of the intensities within each bin was used as the expression variable.

In feature extraction, spline approximation was applied to regress the baseline and the regressed baseline was smoothed using the lowess smoothing method. The resulting baseline was subtracted from the spectrum and each spectrum was normalized by dividing by its total ion current. Then the slope of the peaks is used as the criterion for peak detection.

In feature alignment, the list of detected peaks and combined peaks that differed in location by no more than seven bins or in relative mass no more than 0.9% was pooled.

In feature selection, a wrapper approach that combines ant colony optimization and support vector machine (SVM) was exploited. The proposed algorithm selected

a panel of eight peaks as markers. An SVM classifier built with these peaks achieved 94% sensitivity and 100% specificity in distinguishing HCC from cirrhosis in a blind validation set of 69 samples. Area under the receiver operating characteristic (ROC) curve was 0.996.

3.2 Glycan Profiling

For decades cancer has been known to be associated with changes in glycosylation at different levels [43]. Glycan profiling has advantages over traditional peptide or protein profiling since it only focuses on glycosylated proteins, reducing the potential number of biomarkers that need to be examined [44]. In fact, protein profiling data analysis and glycan profiling data analysis are very similar. Here we will use [44] as an example to illustrate this point.

By profiling oligosaccharides cleaved from glycosylated proteins shed by tumor cells into the blood stream, it is hoped that the effect analysis of glycan profiles will help identify cancer patients using a simple blood test [44]. The analysis procedure consists of six steps: baseline correction, data transformation, peak location, peak selection, normalization, and statistical analysis. Note we use the original terms and notations in [44] for each step. Actually, the first three steps correspond to feature extraction, peak selection corresponds to feature alignment, and statistical analysis corresponds to feature selection.

The analysis reported several masses that are significantly different between cancer and control patients. The significantly different masses form two isotope series representing known glycans ($Hex_3HexNAc_4Fuc_1$ and $Hex_5HexNAc_4Fuc_1$) that are structurally related by the addition of two hexose groups.

Further investigation on the detected glycans showed that there is a difference in the responses of men and women to cancer [44]. The relative levels of the discovered glycans in prostate cancer are approximately the same, while in breast and ovarian cancers they switched from highly over-expressed to highly under-expressed in cancer as the glycans become more massive. Hence, the analysis of glycan profiling data is able to discover new glycan markers, which are complementary to peptide markers.

3.3 Integrated Peptide and Glycan Profiling

It is also feasible to generate peptide and glycan profiling simultaneously. In [45], the serum samples of 203 participants from Egypt were used: 73 HCC cases, 52 patients with chronic liver disease consisting of cirrhosis and fibrosis cases,

and 78 population controls. Two complementary sample preparation methods were applied before generating mass spectra:

1. LMW enrichment was carried out for MALDI-TOF quantification of peptides.
2. Glycans were enzymatically released from proteins and permethylated for MALDI-TOF quantification of glycans.

In data analysis, the same procedure was applied to identify biomarkers from both peptide and glycan profiling. The analysis steps are similar to those in [31], which have been discussed in the previous section.

In this study, the selection of peptide and glycan features was performed separately. A very interesting result is that the combination of peptide features and glycan features together leads to a slight improvement in diagnostic capability. This suggests that it is beneficial to perform an integrated data analysis procedure in biomarker discovery. Here the "integrated" analysis is conducted on MALDI MS data from the same serum with respect to peptides and glycans. Potentially, it is also feasible to perform an "integrated" analysis using data generated from samples of different body fluids such as serum and urine.

4 Limitation and Perspective

MALDI MS has been widely used in proteomics for high-throughput biomarker discovery and drug target identification. Such applications require the management and analysis of data in large amounts and in quite complex ways. To date, both commercial and free academic software are provided to perform data analysis in different contexts. Though existing algorithms and software toolkits have been provided to analyze MALDI MS data, they still deserve attention regarding certain drawbacks:

- While some steps in the analysis pipeline are common to all applications, the organization of these steps is highly dependent on the kind of experiments being conducted. Hence, the overall data analysis pipeline is often subject to large variations. The design of an optimal pipeline is critical to the success of data analysis. However, the currently available methods are far from being optimized.
- In each step of the data analysis pipeline, various algorithms can be employed for the task. However, the performance of different algorithms can vary significantly and such uncertainty will propagate to subsequent analysis modules. For example, we have compared four public peak detection algorithms and found that even the same peak picking method can exhibit significant performance variations under different parameter settings [4]. To obtain good performance, users are required to be familiar with algorithmic details. This is unfeasible for most users.

In summary, most users cannot easily use data analysis tools to solve their biological problems due to the complexity of the data analysis process and software toolkits. Currently, it still often depends on data analysis professionals with relevant experience to provide successful solutions. In order to alleviate these issues, it would be desirable to have a flexible software platform that enables new users to construct rapidly a precise computational workflow for the analysis of MALDI MS data.

To overcome current limitations, we suggest the development of a knowledge-based data analysis platform using case-based reasoning (CBR). CBR is the process of solving new problems based on the solutions of similar past problems [46]. Such CBR-based data analysis toolkits may enable users to conduct their own data analysis pipeline easily and quickly using similar past cases of proteomics applications. The knowledge-based data analysis follows the normal CBR process: retrieve, reuse, revise, and retain:

- *Retrieve*: Given a new MS data analysis problem, we retrieve similar cases from the repository. We assume there are enough cases in the repository.
- *Reuse*: Map the solution from the previous case to the target problem. The user can check retrieved candidate cases manually and select one as template to solve the new analysis problem.
- *Revise*: The previous solution needs to be modified in order to fit the target situation better.
- *Retain*: After the solution has been successfully adapted to the target problem, we add this new solution to case repository for future reuse.

In addition, we would like to discuss several key issues in the implementation of such data analysis platform: case design, case initialization, and process optimization.

From the viewpoint of CBR, a case is a knowledge container. In our context, we may model each MALDI MS data analysis problem and its solution pipeline as a case. As shown in Fig. 4, each case consists of four basic elements: task, data, operator, and process.

The task element defines the objective of our data analysis. For instance, the "problem type" could be "protein biomarker discovery" or "glycan biomarker discovery." The data element includes information about MS data storage and metadata in different stages. The operator element contains atomic algorithms used in the pipeline, e.g., feature alignment, feature selection. Finally, the process element describes the analysis workflow, i.e., the sequential relations of data/operators.

The initialization of case-base includes two tasks: the first is to implement various basic operators and the second is to provide some typical MS data analysis cases.

To fulfill the first task, one may use available data analysis algorithms as the basic operators in the initialization stage. Any other newly proposed algorithms could also be plugged into the system.

To build high-quality MALDI MS analysis cases, the analysis pipeline can be optimized at two levels: operator and parameter.

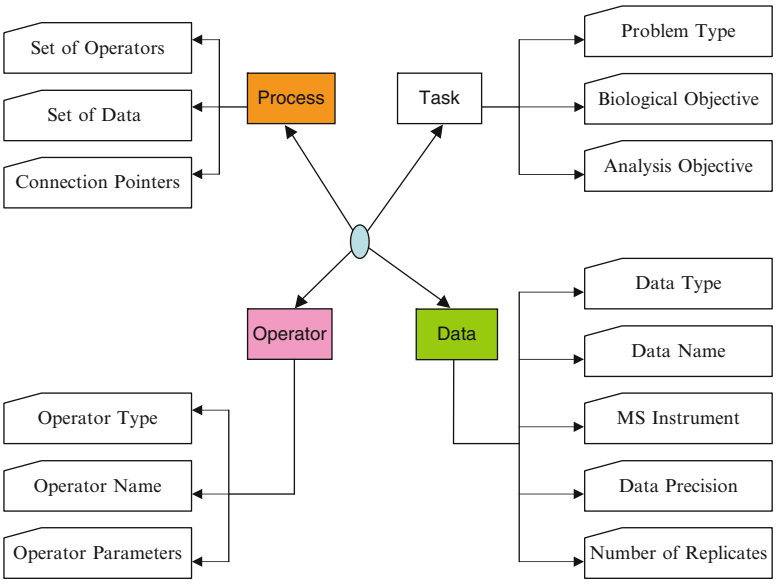


Fig. 4 The basic components of an MS data analysis case

Since there are multiple algorithms available for each type of operator, the exhaustive search may be used to test all potential combinations in the system so as to obtain a better configuration.

Since the parameter space is much larger than the operator space, it is infeasible to test all parameter settings. To optimize parameters, it would be a better practice to use heuristic methods to find local optimal parameters efficiently.

Such a case-based platform can simplify the data analysis process and reduce the dependency on expert. More precisely, it may provide the following salient features:

- It possesses a knowledge-base of domain specific MS data analysis processes. Such a knowledge-base contains typical analysis workflows of different biomarker discovery applications on different MS instruments.
- It makes use of the knowledge captured in past data analysis cases to formulate semi-automatic solutions for typical MS data analysis problems. Knowledge reuse is key to this case-based analysis platform, making it possible to obtain very good data analysis results without the intervention of MS experts and statisticians.
- The case-based system design enables us to easily integrate externally developed algorithms into the software.

We believe future biomarker discovery studies will benefit from the development of such a case-based MS data analysis platform.

Acknowledgements This work was partially supported by the Natural Science Foundation of China under Grant No. 61003176, the General Research Fund 621707 and 662509 from the Hong Kong Research Grant Council, the Area of Excellence Scheme and Special Equipment Grant from the University Grants Committee of Hong Kong, and the Research Proposal Competition Awards RPC07/08.EG25 and RPC10.EG04 from the Hong Kong University of Science and Technology.

References

1. Rifai N, Gillette M, Carr S (2006) Protein biomarker discovery and validation: the long and uncertain path to clinical utility. Nat Biotechnol 24(8):971–983
2. Koomen J, Shih L, Coombes K, Li D, Xiao L, Fidler I, Abbruzzese J, Kobayashi R (2005) Plasma protein profiling for diagnosis of pancreatic cancer reveals the presence of host response proteins. Clin Cancer Res 11(3):1110–1118
3. Roy P, Truntzer C, Maucort-Boulch D, Jouve T, Molinari N (2011) Protein mass spectra data analysis for clinical biomarker discovery: a global review. Brief Bioinform 12(2):176–186
4. Yang C, He Z, Yu W (2009) Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis. BMC Bioinformatics 10(1):4
5. Liu Q, Sung A, Qiao M, Chen Z, Yang J, Yang M, Huang X, Deng Y (2009) Comparison of feature selection and classification for MALDI-MS data. BMC Genomics 10(suppl 1):S3
6. Chen S, Li M, Hong D, Billheimer D, Li H, Xu B, Shyr Y (2009) A novel comprehensive wave-form MS data processing method. Bioinformatics 25(6):808–814
7. Wang P, Yang P, Arthur J, Yang JYH (2010) A dynamic wavelet-based algorithm for preprocessing tandem mass spectrometry data. Bioinformatics 26(18):2242–2249
8. Katajamaa M, Miettinen J, Oresic M (2006) MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. Bioinformatics 22(5):634–636
9. Yu W, Wu B, Lin N, Stone K, Williams K, Zhao H (2006) Detecting and aligning peaks in mass spectrometry data with applications to MALDI. Comput Biol Chem 30(1):27–38
10. Mantini D, Petrucci F, Pieragostino D, DelBoccio P, Nicola MD, Ilio CD, Federici G, Sacchetta P, Comani S, Urbani A (2007) LIMPIC: a computational method for the separation of protein MALDI-TOF-MS signals from noise. BMC Bioinformatics 8:101
11. Du P, Kibbe W, Lin S (2006) Improved peak detection in mass spectrum by incorporating continuous wavelet transform-based pattern matching. Bioinformatics 22(17):2059–2065
12. Wee A, Grayden D, Zhu Y, Petkovic-Duran K, Smith D (2008) A continuous wavelet transform algorithm for peak detection. Electrophoresis 29(20):4215–4225
13. Coombes K, Tsavachidis S, Morris J, Baggerly K, Hung M, Kuerer H (2005) Improved peak detection and quantification of mass spectrometry data acquired from surface-enhanced laser desorption and ionization by denoising spectra with the undecimated discrete wavelet transform. Proteomics 5(16):4107–4117
14. Kwon D, Vannucci M, Song J, Jeong J, Pfeiffer R (2008) A novel wavelet-based thresholding method for the pre-processing of mass spectrometry data that accounts for heterogeneous noise. Proteomics 8(15):3019–3029
15. Alexandrov T, Decker J, Mertens B, Deelder A, Tollenaar R, Maass P, Thiele H (2009) Biomarker discovery in MALDI-TOF serum protein profiles using discrete wavelet transformation. Bioinformatics 25(5):643–649
16. Mostacci E, Truntzer C, Cardot H, Ducoroy P (2010) Multivariate denoising methods combining wavelets and principal component analysis for mass spectrometry data. Proteomics 10(14):2564–2572
17. Noy K, Fasulo D (2007) Improved model-based, platform-independent feature extraction for mass spectrometry. Bioinformatics 23(19):2528–2535

18. Samuelsson J, Dalevi D, Levander F, Rognvaldsson T (2004) Modular, scriptable, and automated analysis tools for high-throughput peptide mass fingerprinting. Bioinformatics 20(18):3628–3635
19. Renard B, Kirchner M, Steen H, Steen J, Hamprecht F (2008) NITPICK: peak identification for mass spectrometry data. BMC Bioinformatics 9:355
20. Wang Y, Zhou X, Wang H, Li K, Yao L, Wong S (2008) Reversible jump MCMC approach for peak identification for stroke SELDI mass spectrometry using mixture model. Bioinformatics 24(13):i407–i413
21. Mantini D, Petrucci F, Boccio P, Pieragostino D, Nicola M, Lugaresi A, Federici G, Sacchetta P, Ilio C, Urbani A (2008) Independent component analysis for the extraction of reliable protein signal profiles from MALDI-TOF mass spectra. Bioinformatics 24(1):63–70
22. McLerran D, Feng Z, Semmes O, Cazares L, Randolph T (2008) Signal detection in high-resolution mass spectrometry data. J Proteome Res 7(1):276–285
23. Zhang S, DeGraba T, Wang H, Hoehn G, Gonzales D, Suffredini A, Ching W, Ng M, Zhou X, Wong S (2009) A novel peak detection approach with chemical noise removal using short-time FFT for prOTOF MS data. Proteomics 9(15):3833–3842
24. Vandenbogaert V (2008) Alignment of LC-MS images, with applications to biomarker discovery and protein identification. Proteomics 8(4):650–672
25. Kong X, Reilly C (2009) A Bayesian approach to the alignment of mass spectra. Bioinformatics 25(24):3213–3220
26. Yu W, He Z, Liu J, Zhao H (2008) Improving mass spectrometry peak detection using multiple peak alignment results. J Proteome Res 7(1):123–129
27. Tibshirani R, Hastie T, Narasimhan B, Soltys S, Shi G, Koong A, Le Q (2004) Sample classification from protein mass spectrometry by 'peak probability contrasts'. Bioinformatics 20(17):3034–3044
28. Yu W, Li X, Liu J, Wu B, Williams KR, Zhao H (2006) Multiple peak alignment in sequential data analysis: a scale-space-based approach. IEEE/ACM Trans Comput Biol Bioinform 3(3):208–219
29. Guyon I, Elisseeff A (2003) An introduction to variable and feature selection. J Mach Learn Res 3:1157–1182
30. Ressom H, Varghese R, Abdel-Hamid M, Eissa S, Saha D, Goldman L, Petricoin E, Conrads T, Veenstra T, Loffredo C et al (2005) Analysis of mass spectral serum profiles for biomarker selection. Bioinformatics 21(21):4039–4045
31. Ressom H, Varghese R, Drake S, Hortin G, Abdel-Hamid M, Loffredo C, Goldman R (2007) Peak selection from MALDI-TOF mass spectra using ant colony optimization. Bioinformatics 23(5):619–626
32. Park Y, Downing SR, Kim D, Hahn WC, Li C, Kantoff PW, Wei L (2007) Simultaneous and exact interval estimates for the contrast of two groups based on an extremely high dimensional variable: application to mass spec data. Bioinformatics 23(12):1451–1458
33. Oh J, Kim Y, Gurnani P, Rosenblatt K, Gao J (2008) Biomarker selection and sample pre diction for multi-category disease on MALDI-TOF data. Bioinformatics 24(16):1812–1818
34. Oh J, Gurnani P, Schorge J, Rosenblatt K, Gao J (2009) An extended Markov blanket approach to proteomic biomarker detection from high-resolution mass spectrometry data. IEEE Trans Inf Technol Biomed 13(2):195–206
35. Hilario M, Kalousis A (2008) Approaches to dimensionality reduction in proteomic biomarker studies. Brief Bioinform 9(2):102–118
36. Abeel T, Helleputte T, Van de Peer Y, Dupont P, Saeys Y (2010) Robust biomarker identification for cancer diagnosis with ensemble feature selection methods. Bioinformatics 26(3):392–398
37. He Z, Yu W (2010) Stable feature selection for biomarker discovery. Comput Biol Chem 34(4):215–225
38. Good P (2005) Permutation, parametric and bootstrap tests of hypotheses. Springer, Heidelberg

39. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Stat Methodol 57(1):289–300
40. Benjamini Y (2010) Discovering the false discovery rate. J R Stat Soc Ser B Stat Methodol 72(4):405–416
41. Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten I (2009) The WEKA data mining software: an update. SIGKDD Explor 11(1):10–18
42. Smit S, Hoefsloot H, Smilde A (2008) Statistical data processing in clinical proteomics. J Chromatogr B 866(1–2):77–88
43. Wuhrer M (2007) Glycosylation profiling in clinical proteomics-heading for glycan biomarkers. Expert Rev Proteomics 4(2):135–136
44. Barkauskas D, An H, Kronewitter S, De Leoz M, Chew H, de Vere White R, Leiserowitz G, Miyamoto S, Lebrilla C, Rocke D (2009) Detecting glycan cancer biomarkers in serum samples using MALDI FT-ICR mass spectrometry data. Bioinformatics 25(2):251–257
45. Ressom HW, Varghese RS, Goldman L, An Y, Loffredo CA, Abdel-Hamid M, Kyselova Z, Mechref Y, Novotny M, Drake SK, Goldman R (2008) Analysis of MALDI-TOF mass spectrometry data for discovery of peptide and glycan biomarkers of hepatocellular carcinoma. J Proteome Res 7(2):603–610
46. Aamodt A, Plaza E (1994) Case-based reasoning: foundational issues, methodological variations, and system approaches. Artif Intell Commun 7(1):39–59

Index