**University of Alberta**

# Development of Mass Spectrometric Methods for Membrane Proteome Analysis

by

Nan Zhang ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Department of Chemistry

Edmonton, Alberta

Fall 2003

*Your file Votre référence*
*ISBN: 0-612-88076-1*
*Our file Notre référence*
*ISBN: 0-612-88076-1*

# University of Alberta

## Library Release Form

**Name of Author:** Nan Zhang

**Title of Thesis:** Development of Mass Spectrometric Methods for Membrane Proteome Analysis

**Degree:** Doctor of Philosophy

**Year this Degree Granted:** 2003



September 26th, 2003

**University of Alberta**

**Faculty of Graduate Studies and Research**

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Development of Mass Spectrometric Methods for Membrane Proteome Analysis** submitted by Nan Zhang in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Dr. Liang Li (Supervisor)

Dr. Mark T. McDermott

Dr. John S. Klassen

Dr. Robert E. Campbell

Dr. Michele Barry

Dr. Ron Orlando, University of Georgia

Date: Sept. 26, 2003

To my parents and my husband, Guijun

## Abstract

While membrane proteins play critical roles in many biological functions and are often the molecular targets for drug discovery, their analysis presents a special challenge largely due to their highly hydrophobic nature. Recently, the shotgun proteomics approach has emerged as a powerful technique for proteome identification and quantitation. It is based on the digestion of proteins from extracts of whole cells, organelles, or specific fractions of thereof, followed by liquid chromatography tandem mass spectrometry analysis of the resulting peptides.

In this thesis, the shotgun proteomics approach is further developed to tackle membrane protein analysis. Because membrane proteins are not easily dissolved in water, surfactants are employed for membrane protein solubilization. However, surfactants, especially sodium dodecyl sulfate (SDS), have been found to be detrimental to mass spectrometric signal by some of the researchers, whereas others have found that SDS did not deteriorate the mass spectrometric signal. This controversy was resolved from part of the work described in this thesis. It was found that for SDS-containing samples, the intensity of the MALDI signals can be affected by the conditions of sample preparation: on-probe washing, choice of matrix, deposition method, solvent system, and protein-to-SDS ratio. However, under appropriate conditions, the two-layer method gave reliable MALDI signals for samples with levels of SDS up to ~1%. The mass resolution issue was also addressed and a method was developed to improve the resolution.

Furthermore, the surfactants' effects on protein digestion and mass spectrometric response were studied and a SDS-aided in-solution digestion for membrane protein identification method was developed. The developed method was validated by protein

identification from a membrane protein fraction of a human cell line, with two shotgun approaches, LC/ESI and LC/MALDI MS/MS techniques. The developed in-solution method was also compared with the conventional gel electrophoresis approach in terms of protein identification. It was found that the three techniques are complementary to each other and a preferred proteomics studies should include all of them. The in-solution method was further applied to the analysis of lipid raft proteins and was optimized with multi-dimensional separation technique to obtain a more comprehensive proteome map.

## Acknowledgements

First of all, I sincerely thank my supervisor, Dr. Liang Li, for giving me the opportunity to study in his research group. I would like to thank for his advice, encouragement, guidance, inspiration and enthusiasm during the course of my graduate study. He has been always available for discussion and help. What I have learned from him and his research group will benefit my life and career.

I would also like to thank my committee members, Dr. Mark M. McDermott, Dr. John S. Klassen, Dr. Robert E. Campbell, Dr. Charles A. Lucy, Dr. Michele Barry and Dr. Ron Orlando, for review and comments on this thesis.

I would like to thank the Department of Chemistry and University of Alberta for financial support and providing facilities. I thank all the kind help and support from personnel in the general office, copy room, machine shop and electronics shop.

My deep gratitude goes to members of Dr. Li's research group, past and present, whom I have worked with. I would like to thank Dr. Nan Li for her hard work in Chapter 6 and helpful and great discussions in Chapters 5 and 7. I am grateful to Dr. Alan A. Doucette, for his numerous work and help in Chapter 2. I would like to thank Mr. Chris J. McDonald and Mr. J. Bryce Young for spending their precious time to proofread my thesis. I would like to thank Bernd O. Keller for his endless help during the course of my graduate study. I also thank Dr. Randy Whittal for his support when I used the instruments in Mass Spectrometry Laboratory. Many appreciations also extend to Dr. Yuqin Dai, Dr. Truong Ta, Dr. Ken K.-C. Yeung, and Ms. Jing Zheng for their kind help and useful suggestion. I also thank Dr. H. Liu and Mr. C. Ji for their understanding of my demand to the QSTAR instrument time during the last four months.

I also thank Dr. Sandra Markus for cultured the cell in Chapter 5 and Paul Semchuk and Lorne Burke of the Institute for Biomolecular Design of the University of Alberta for generating some of the ESI data in Chapters 5 and 7.

Many thanks devoted to my friends, especially my special friend Mrs. Kathy Gottlob, for her care, passion and encouragement. The warmth of her family made my life in Edmonton joyful.

Finally, I would like to thank to my parents, sister and niece for their love, support and understanding. From bottom of my heart, I thank my dear husband, Guijun Dai, for his love and patience; especially that he has been so sensitive to my needs and helped me in every area, physically, emotionally and mentally in the past year. His sacrifice makes me focused on my studies and research, and I will be forever grateful for that.

# Table of Contents

# List of Tables

# List of Figures

## List of Abbreviations

| | |
|---|---|
| 2D LC | 2 Dimensional liquid chromatography |
| BRO | Bacteriorhodopsin |
| BSA | Bovine serum albumin |
| CE | Capillary electrophoresis |
| CHAPS | 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate |
| CID | Collision-induced dissociation |
| CMC | Critical micelle concentration |
| CNBr | Cyanogen bromide |
| DAHC | Diammonium hydrogen citrate |
| DC | Direct current |
| DHB | 2,5-dihydroxybenzoic acid |
| DTT | dithiothreitol |
| *E. coli* | *Escherichia coli* |
| EGTA | [Ethylenebis(oxyethylenenitrilo)] tetraacetic acid |
| ESI | Electrospray ionization |
| FBS | fatal bovine serum |
| FTICR | Fourier-transform ion cyclotron resonance |
| FWHM | Full width at half maximum |
| HCCA | α-cyano-4-hydroxycinnamic acid |
| HABA | 2-(4-hydroxyphenylazo)-benzoic acid |
| HEPES | N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid ( |
| HPLC | High performance liquid chromatography |

| | |
|---|---|
| IE | Ion exchange |
| IEF | Isoelectric focusing |
| m/z | mass to charge |
| MALDI | Matrix-assisted laser desorption/ionization |
| MES | 2-(4-Morpholino)ethane sulfonic acid |
| MS | Mass spectrometry |
| MW | Molecular weight |
| PBS | Phosphate-buffered saline |
| PI | Isoelectric point |
| PMSF | Phenylmethyl sulfonyl fluoride |
| ppm | part(s) per million |
| QqTOF | Quadruple time-of-flight |
| RCF | Relative centrifugal force (× g) |
| RF | Radio frequency |
| RP | Reversed-phase |
| SA | Sinapinic acid |
| SDS | Sodium dodecyl sulfate |
| SDS-PAGE | Sodium dodecyl sulfate- polyacrylamide gel electrophoresis |
| S/N | Signal to noise ratio |
| TFA | Trifluoroacetic acid |
| Tris | Tris (hydroxymethyl) aminomethane |
| Triton® X-100 | t-Octylphenoxypolyethoxyethanol |
| TOF | Time-of-flight |

| | |
|---|---|
| m | milli- ($10^{-3}$) |
| $\mu$ | micro- ($10^{-6}$) |
| n | nano- ($10^{-9}$) |
| p | pico- ($10^{-12}$) |

# Chapter 1

## Introduction to Mass Spectrometry for Proteomics

Understanding proteins and their functions is essential to a better and clearer understanding of biologically important processes and systems. It was hypothesized that deciphering the genetic codes of organism might enable such an understanding. However, it was soon realized that the genetic codes alone could not give all the information needed to understand the complex world of proteins and their functions [1]. This complexity is attributed to the complicated three-dimensional structures of proteins and the multifold possibilities of post-translational modifications leading to multiple functions of a single protein in different environments. To get an idea on functional correlations of proteins in a given tissue it is therefore imperative to identify all of them. The Global analysis of all expressed proteins in a cell, tissue or organelle is termed proteomics.

The science of proteomics involves the detection, identification, and characterization of the protein expression, function, activity, regulation and post-translational modification in post-genomic state [2, 3]. The technologies engaged in such analyses integrate separation science for separation/purification of proteins and peptides, analytical science for identification and quantification of the proteins or peptides of interest and bioinformatics for data management and analysis. Mass spectrometry (MS) is the center technology for proteomics studies, especially for the identification and sequencing of protein, or the characterization of the state of post-translational modifications and protein-protein or protein-complex interactions.

The beginning of the mass spectrometry technique dates back to the early days of the last century. However, in the last decade, mass spectrometry has played an important role in the identification of biological molecules. Technical breakthroughs in the 1980's lead to the development of two "soft" ionization methods, electrospray ionization (ESI) [4-6] and matrix-assisted laser desorption/ionization (MALDI) [7-9]. These two ionization methods make it possible to generate ions from large non-volatile analytes such as proteins and peptides without significant decomposition or fragmentation. And thus, the analysis of proteins has been changed fundamentally and this makes the analysis of large intact proteins possible [10]. Indeed, development in mass spectrometry has been the major driving force for the establishment and advances of the new field, proteomics [11-13].

The following two sections will be a brief overview of the principles on the mass spectrometry and protein identification methods.

## 1.1 Mass Spectrometry

The principle of mass spectrometry is based on the measurement of the mass to charge ratio of ions in gas phase. Hence, the analytes need to be ionized and converted to gas phase at the mass spectrometer source region. Those ions from the source region are brought into and separated by a suitable mass analyzer, and then recorded by a detector as ion currents and displayed as a mass spectrum. Three different principles may be applied to achieve the mass separation: separation on the basis of time-of-flight (TOF), separation by quadrupole electric field, or separation by selective trapping of ions from a three-dimensional trapping field (ion trap or FTICR). For peptide sequencing analysis, two

steps of mass spectrometry are performed in tandem (tandem mass spectrometry or MS/MS), which may be achieved by employing the same principle twice or by combination of the two different principles. Both MALDI and ESI can in principle be coupled with any one from these three separation methods. MALDI typically produces ions in pulsed intervals under vacuum, making it most suitable for coupling with TOF MS. Electrospray produces ions continuously at atmospheric pressure, making it amenable to coupling with ion trap or quadrupole MS.

### 1.1.1 MALDI-TOF MS

MALDI was introduced by two laboratories of Tanaka [9] and Hillenkamp [7, 8] independently as improvement to laser desorption/ionization in 1987. The MALDI event involves the analytes of interest, use of matrix, a suitable solvent, the probe surface and a laser. Figure 1.1 shows a schematic diagram of MALDI. The analyte is dissolved in a suitable solvent, typically methanol, acetonitrile, water or mixture of these. The matrix is also dissolved in an appropriate solvent. Once these solutions are prepared, they are mixed under an appropriate matrix-to-analyte ratio (typically >500:1) and deposited onto the sample probe surface allowing it to dry. The solvent evaporates and matrix and analyte co-crystallize on the target probe. The resulting solid is then irradiated by nanosecond laser pulses, typically a short pulse of ~3 ns duration and power of ~$10^6$ W/cm$^2$ with wavelength of 337 nm. The large amount of laser energy is absorbed efficiently by matrix molecules, which rapidly expand to gas phase. Energy is subsequently transferred to the analyte that becomes desorbed into the gas phase. Although the ionization mechanism is not fully understood and is still debated [14], it is

widely accepted that co-desorption of matrix and analyte enables the proton transfer from matrix molecules to the analyte molecules in the gas phase (i.e. chemical ionization).

In the MALDI process, the matrix is usually a small organic molecule with absorbance at the wavelength of the laser. Work involved in proteins and peptides almost exclusively uses α-cyano-4-hydroxycinnamic acid (HCCA) and 2,5-dihydroxybenzoic acid (DHB) as matrix. The commonly used matrices for MALDI MS are listed in Figure 1.2. The roles of matrix are: (1) to entrap analyte into the matrix-analyte co-crystallization, (2) to isolate analyte molecules which prevent the aggregation of the analyte, (3) to absorb energy from laser beam and easy to be evaporated into gas phase, and (4) to aid the analyte ionization process by the rapidly expended matrix plume.

Several features make MALDI suitable for biological sample analysis. It produces mainly singly charged ions, allowing analysis of very heterogeneous biological protein mixtures, such as protein digests. The devolvement of sample preparation and deposition methods makes MALDI an ionization technique that can tolerate certain amount of contaminants, such as salts or detergent [15]. These sample preparation methods include dried-droplet [8], vacuum drying [16], crushed-crystal [17], slow crystal growing [18], active film [19, 20], pneumatic spray [21], electrospray [22], fast solvent evaporation [23, 24], sandwich [25, 26], and two-layer method [27]. Choice of these methods can enable analyzing special and complex biological samples, such as hydrophobic proteins, or sample containing surfactants, in term of optimization and improvement of signal resolution or intensity.

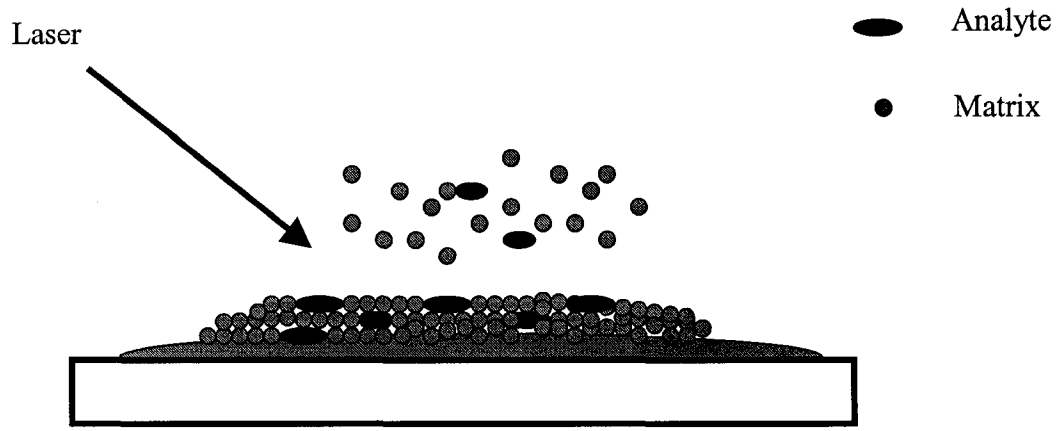
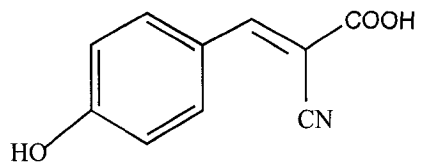
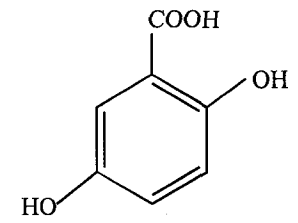


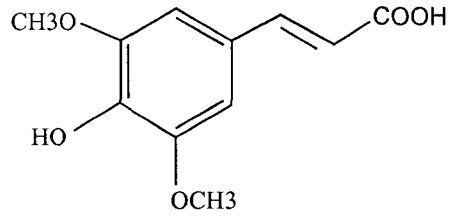

**Figure 1.1** Schematic diagram of MALDI.



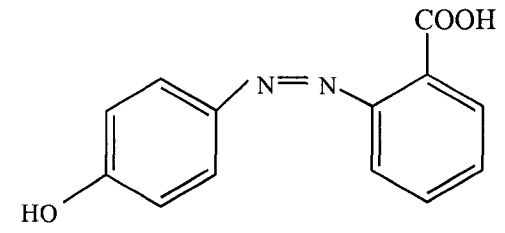

α-cyano-4-hydroxycinnamic acid (HCCA)




2,5-dihydroxybenzoic acid (DHB)




Sinapinic acid (SA)




2-(4-hydroxyphenylazo)-benzoic acid (HABA)

**Figure 1.2** Commonly used matrices in MALDI.

Because MALDI is a pulsed event, it usually combines with a pulsed mass analyzer such as TOF MS. The features of TOF analyzer such as theoretically unlimited detectable mass range, high ion transmission, short duty circle and high spectra recording speed make it highly desirable for high sensitivity MALDI analysis.

Figure 1.3 shows the basic principle of a MALDI TOF MS. Ions generated from source region are given a fixed amount of kinetic energy and accelerated by high voltage (HV) up to 30 kV before pass through a field free region. The energy obtained is:

$$zeV = \frac{mv^2}{2} \qquad (1.1)$$

where $z$ is the charge state of the ion, $e$ is the unit of elementary charge, $m$ is the mass of the ion, and $v$ is the linear velocity of the ion after acceleration.

Hence the ion crosses the field free region with velocity:

$$v = \left(\frac{2zeV}{m}\right)^{1/2} \qquad (1.2)$$

and flight time:

$$t = \left(\frac{m}{2zeV}\right)^{1/2} L \qquad (1.3)$$

In a mass spectrometer, mass resolution is defined as $m/\Delta m$. In a time-of-flight mass spectrometer, in which ions are accelerated by a constant energy, the resolution can be expressed as:

$$\frac{m}{\Delta m} = \frac{t}{\Delta t} \qquad (1.4)$$

where $\Delta m$ or $\Delta t$ is measured as the full width at half maximum (FWHM). Therefore, mass resolution depends upon the time resolution and upon the laser pulse width, detector response, initial kinetic energy, and velocity etc.

The most significant limit on the mass resolution in a TOF mass spectrometer is the broad range of ion initial velocities or energies as they are accelerated. MALDI produces a wide variation of ion kinetic energy, and the velocity spread produced by this variation has an effect on the mass resolution because ions with the same *m/z* acquired different velocity after acceleration, and therefore have different flight time, which results in peak broadening. However this effect of initial energy can be reduced with delayed extraction of the ions into the field free region (time-lag focusing) or an ion reflector.

Figure 1.4 shows the principle of a time-lag focusing in a linear TOF MS. In time-lag focusing, the ions generated from MALDI are allowed to separate according to their initial velocity or energy just after they are ionized. This is done in a field free region where the repeller and the first extraction grid hold the same potential (Figure 1.4 A-B). Ions with the same m/z but different axial velocities will be expanding away from the repeller, such that those ions with higher velocity or more energy will move further from the repeller than the initially less energetic ions (Figure 1.4B). After a certain time delay (hundreds of nanoseconds to several microseconds), an extraction pulse is applied to the repeller to extract ions into the flight tube. The extraction pulse imparts more energy to the ions closer to the repeller. The amplitude of the pulse is adjusted so that the ions with initially less energy will catch up to the initially more energetic ions at the detector. With time-lag focusing, the mass resolution is significantly improved.

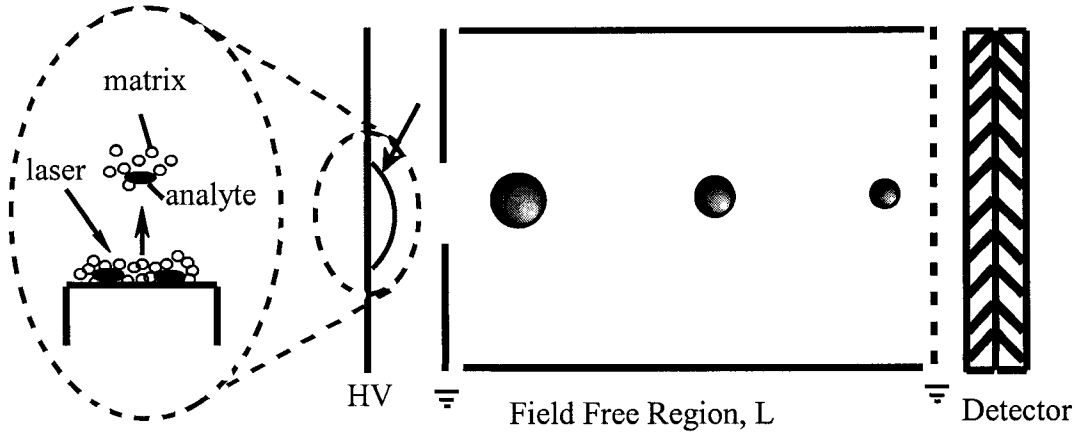



**Figure 1.3** Schematic graph of a linear MALDI-TOF mass spectrometer.

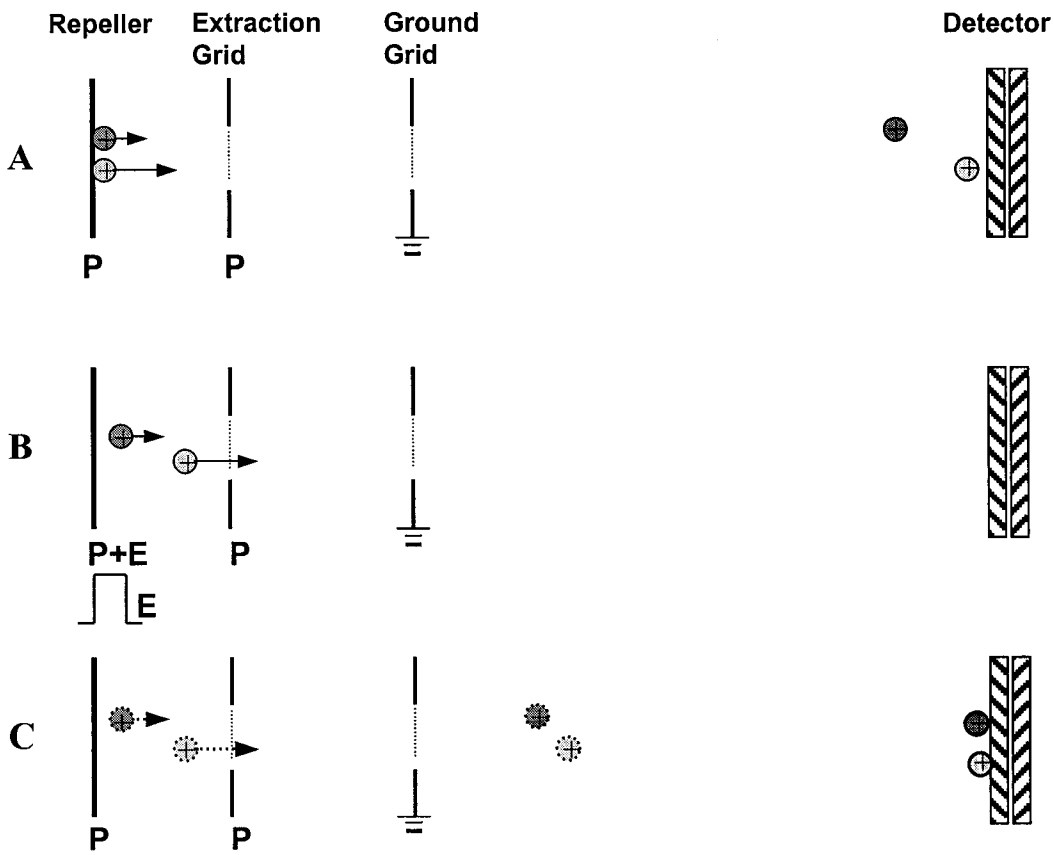



**Figure 1.4** Illustration of principle and the effect of time-lag focusing in a linear TOF MS. (A) Illustration of the flight of two ions with same m/z but different initial velocity without time-lag focusing. (B) In time-lag focusing, a pulse voltage is applied after ions are first separated by their initial velocity in the source region. (C) After the pulse voltage, the lower ion will catch up the faster one at detector.

Another method to improve the resolution is reflectron ion mirror. A schematic diagram of a TOF MS equipped with a reflectron flight tube is shown in Figure 1.5. In this type of instrument, ions are not detected by detector 1 at end of the field free region. Instead, they are reflected back to another field free flight tube by electric field (ion mirror) and detected by detector 2. Ions with higher initial velocity will penetrate deeper in the ion mirror than ions with same *m/z* but less initial velocity and take longer time to return. By properly arranging the voltage gradient on the reflectron, ions with the same *m/z* value can then be focused at the detector 2, and the initial energy spread are largely compensated and mass resolution is greatly improved up to 30,000 at m/z 3000, and typically > 5000 up to m/z 6000 [14].

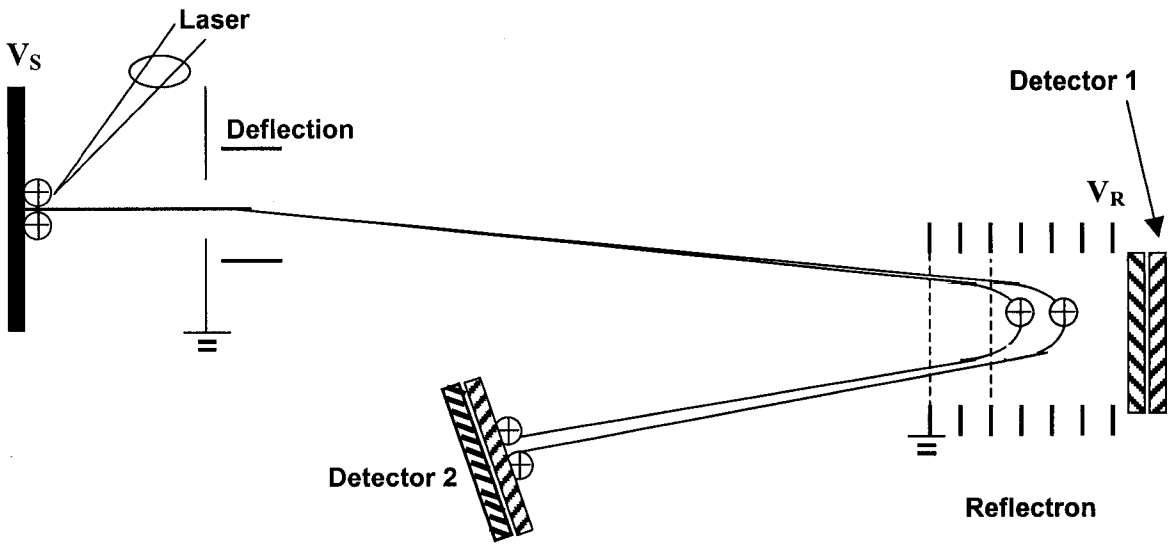



**Figure 1.5** Schematic diagram of a reflectron TOF mass spectrometer. Vs represents the voltage applied to repeller, and $V_R$ represents the voltage grid applied on the reflectron.

Besides the improvement of mass resolution, time-lag focusing and reflectron also provide a significant improvement of mass measurement accuracy and sensitivity because of peak sharpness. Mass accuracy of 50 ppm can be routinely obtained for ions below 5000 Da by internal calibration (analyte and calibrant located in the same spot). For complex sample mixture with a broad mass range or external calibration, mass accuracy of better than 500 ppm or 0.05% can normally be achieved.

### 1.1.2 ESI Ion Trap Mass Spectrometry

Electrospray is an atmospheric-pressure ionization technique that produces small charged droplets from a liquid medium under an electric field. The concept was first introduced by Dole and co-workers in 1968 [28]. Fenn and co-workers successfully combined ESI with mass spectrometry in late 1980s and first reported electrospray mass spectra from large biomolecules [4-6]. The ESI processes were described by Kebarle *et al* as shown in Figure 1.6 schematically [29]. There are four major processes involved. (1) Charged droplets were produced at ESI capillary tip by application of a high voltage. (2) Shrinkage of the charged droplets due to solvent evaporation accomplished by flowing dry nitrogen gas at moderate temperature. (3) As the droplets shrink, the coulombic repulsion forces overcome the surface tension forces where fission occurs and the charged droplets disintegrate into fine droplets. (4) Electric field at liquid surface becomes so high that the solute ions "escape" from the liquid phase to gas phase (Ion evaporation theory). Those ions in gas phase will be introduced to a mass analyzer.

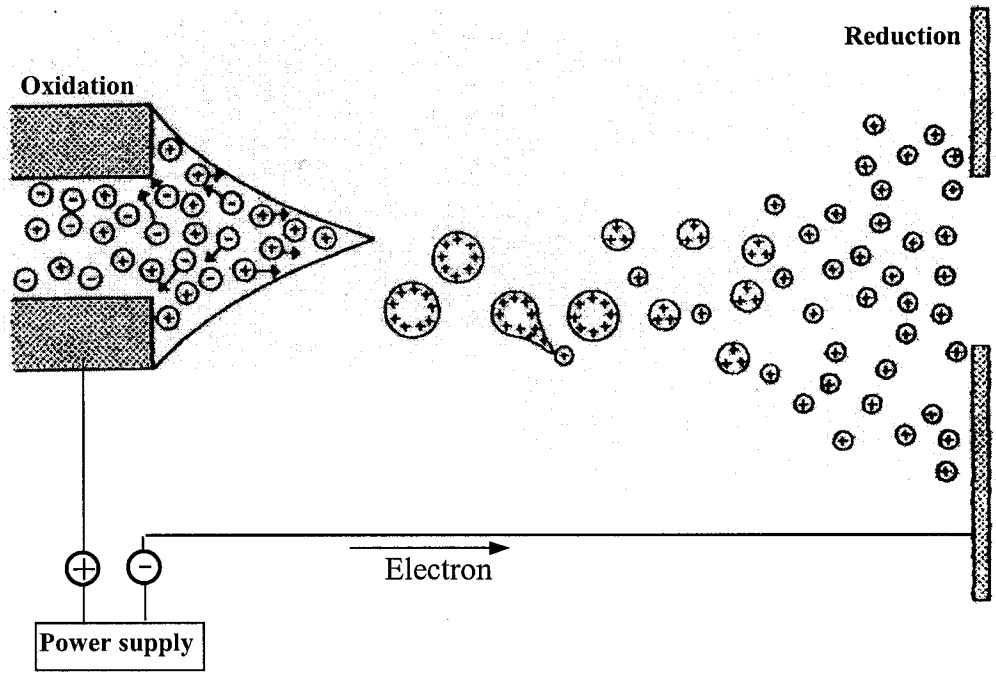



**Figure 1.6** Schematic representation of processes in ESI.

The electrospray ion source has gone through major developments after it was introduced. One of these is the development of ionspray as an improvement over electrospray in order to couple MS with HPLC systems. Conventional electrospray operates best at flow rate at 0.5~5 μl/min, but the coupling of liquid chromatography to mass spectrometry sometimes demands flow rate of up to 1 ml/min. The electrospray evaporation is facilitated by a coaxial gas flow (ionspray) [30]. Because electrospray signal intensity increases linearly with the analyte concentration over a wide range until its saturation, and almost independent of liquid flow rate, low flow rate electrospray ionization, nanospray, was developed by Wilm and Mann [31, 32]. This results in a flow rate of 20-50 nL/min, and hence small sample volumes are consumed.

A wide range of compounds can be analyzed by ESI MS; the only requirement is that analyte should be sufficiently polar to allow attachment of charges. Electrospray ionization commonly produces multiply charged ions, which are generally detected in the mass range below 2500 Da. ESI is less tolerant to salts and other contaminants because of competition for charge. Analytes must come to the droplet surface to desorb; contaminants compete with the analyte in this process, and also serve to disrupt the spray/evaporation processes [15]. Since ESI produces ions continuously, it usually combined with scanning mass analyzer such as quadrupole, ion trap or FTICR.

Quadrupole mass analyzer (or quadrupole mass filter) contains four cylindrical rods (Figure 1.7) with two opposite rods connected electrically. Direct current (DC) and radio frequency (RF) voltages are applied to the two pairs of rods with different polarities. The DC and RF values are set such that only ions of interest have stable trajectories through the quadrupole and reach the detector. All other ions will be deflected and annihilated upon incidence with the rods. As the voltages are scanned by increasing U and V (U is the DC potential applied between the electrodes, and V is the amplitude of the RF signal with frequency $\omega$) with U/V constant, ions with different m/z values are transmitted successively from low to high mass. Quadrupole instruments have two significant characteristics. First, at a given time, the vast majority of ions are discarded; thus, quadrupole is described as mass filter. Second, quadrupole can be set to contain and transmit ions of all m/z by applying RF-only field. This use is significant because it allows quadrupole to function as a sophisticated lens system in the instrument where ions transmission and focusing without filtering is needed. This feature makes quadrupoles useful as collision cells in tandem instruments that will be described later.

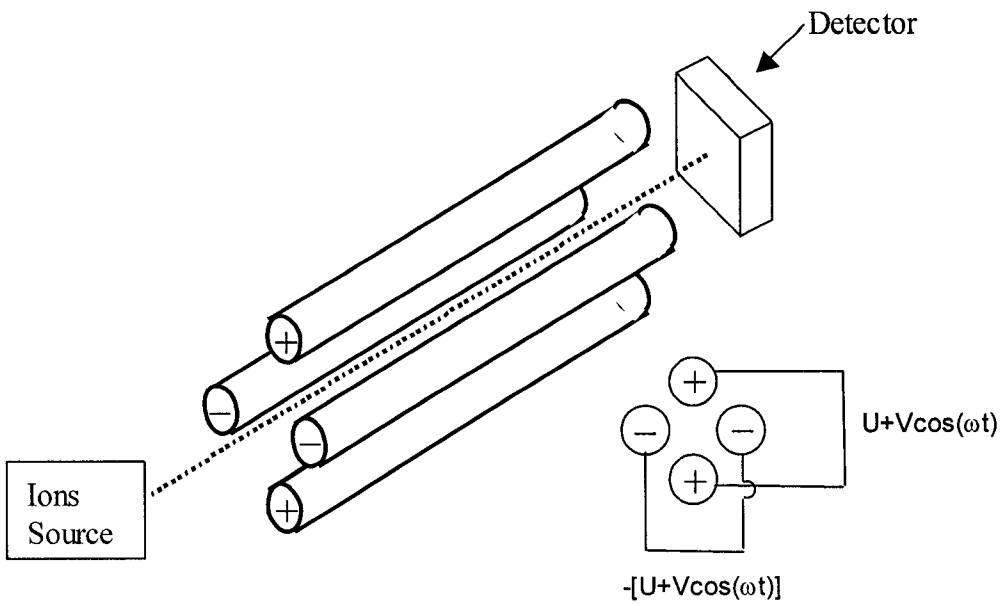



**Figure 1.7** Schematic representation of the quadrupole mass filter and basic principle.

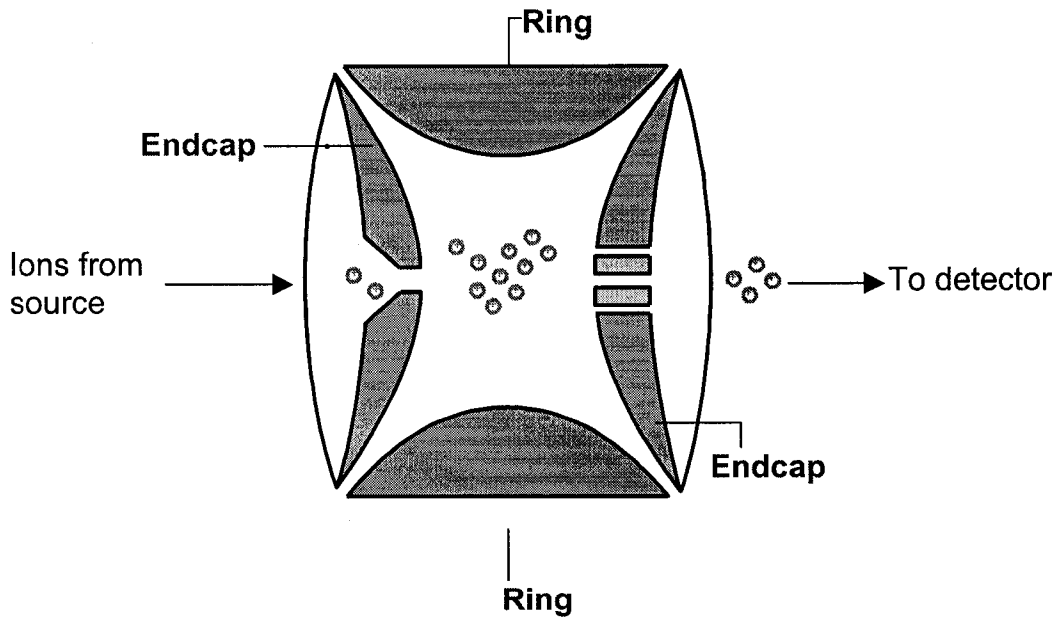



**Figure 1.8** Schematic diagram of ion trap mass analyzer.

An ion trap instrument is composed of two end-caps and a ring electrode (Figure 1.8). A hyperbolic electric field is created inside the chamber by applying a potential on the electrodes. After the ions are introduced from the source region, ions are subjected to the hyperbolic field, which sequentially ejects ions from the trap for detection to produce a mass spectrum. An important characteristic of ion traps is that ions can be accumulated, fragmented and analyzed in the same chamber. Thus tandem MS can be achieved in the volume of one ion trap mass analyzer. For MS/MS experiment, only desired ions are retained. The remaining ion species are fragmented by collision-induced dissociation (CID) and its products are analyzed. Several sequential steps of MS/MS in principle can be performed which allow detailed fragment studies ($MS^n$ ability).

### 1.1.3 MALDI QqTOF Mass Spectrometer

The tandem mass spectrometer MALDI QqTOF MS as shown in Figure 1.9 was recently developed in Standing's lab and commercialized in Sciex [33-35]. Ions generated from the MALDI source first pass through q0, a quadrupole with RF only as ions focusing device in source region, and enter the quadrupole mass analyzer (Q1). In the MS mode, the quadrupole q2 is used as an ion transmission device (RF only) which allows all the ions from Q1 to reach the Modulator after focusing. A pulsed voltage is applied orthogonally to the Q direction, and forces the ions to enter the reflectron TOF and analyzed by detector. In the MS/MS mode, ions of certain m/z are first selected by Q1 (used as mass filter) and enter the collision cell (q2) where collision induced dissociation occurs with a collision gas ($N_2$ or Ar). Product ions from CID are analyzed by the reflectron TOF. The key advantage of a QqTOF instrument (also called QSTAR as commercial name) is that both the mass spectra and products spectra are recorded by a

reflectron TOF mass analyzer with all the advantages of TOF analysis such as high mass accuracy (10 ppm) and high resolution (typical 25000 at m/z 2500). This allows very high specificity in database searching and is very suitable for detection of post-translational modifications.

Compared with conventional MALDI TOF MS instrument, such as Bruker Reflex, in-source fragmentation and peptide *in vitro* modification are usually observed in the QSTAR for peptides with certain composition, and fragment ions are displayed in MS spectra [36]. The laser beam for this QSTAR is usually used with higher intensity compared to the conventional MALDI TOF MS and the higher laser power imparts more energy onto the desorbed ions, which potentially leads to fragmentation. The QSTAR has quadrupole plus TOF geometry, whereas conventional MALDI only has TOF. The ion transmission time in quadrupole (quadrupole scanning speed: one second from 0 to 4000 Da) is much longer than that in TOF analyzer (~ ns), therefore, ions have more time to fragment after they gain energy. The second generation QSTAR, Sciex introduced a cone in front of the MALDI plate towards q0 — this increases the pressure in the region where the target plate sits and thus desorbed ions are collisionally cooled down before they go further down the q0 region — thus the increased pressure, although it leads to more collisions with gas molecules (nitrogen), does not lead to more fragmentation but just the opposite to actually reduce the in-source fragmentation. Table 1.1 lists the commonly observed modifications in the MALDI QqTOF instrument. These modifications make the spectra complex which complicates the database search. However, by understanding the origin of the in-source fragmentation, pseudo-$MS^3$ of the fragment ions displayed in MSmode can be performed which leads to a database search

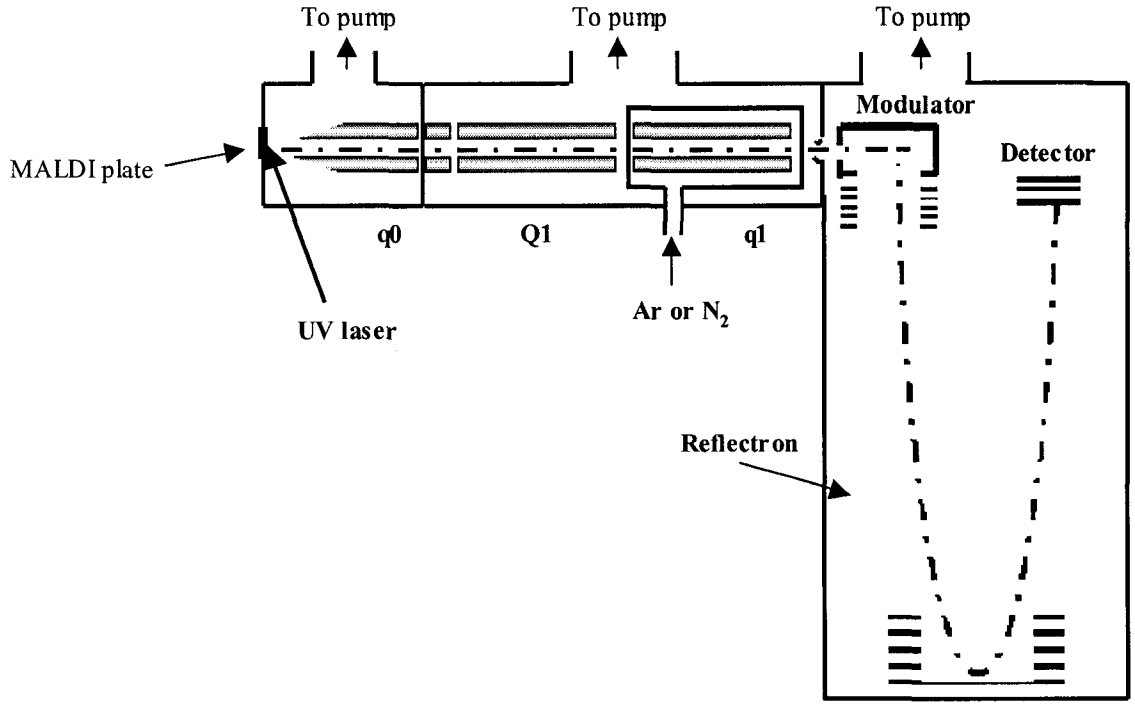



**Figure 1.9** Schematic diagram of a hybrid quadrupole time-of-flight mass spectrometer.

**Table 1.1 Commonly Observed peaks in MALDI QqTOF Instrument.**

| Observed Mass Da | Origin | Detection mode |
|---|---|---|
| -42 | Ornithine from Arginine, loss of $CN_2H_2$ | MS/MS |
| -44 | decarboxylation of gamma carboxy Glutamate, loss of $CO_2$ | MS/MS |
| -17 | loss of $NH_3$ | MS or MS/MS |
| -18 | loss of $H_2O$ | MS or MS/MS |
| N/A | Fragmentation between DP or PD in a peptide | MS |
| 175.1 & 158 | From C-terminal Arginine | MS/MS |
| $[MH]^+$–64 (Dominant) | Peptide that has oxidation of methionine to loss of $SOCH_4$ | MS/MS |

with high confidence [37, 38]. Because of in-source fragmentation, the best way to use MALDI QqTOF for peptide sequence analysis is to obtain MS mode spectra from both conventional MALDI TOF (e.g. Bruker Reflex III) and MALDI QqTOF, and perform MS/MS experiment by MALDI QqTOF with ions of known origin (tryptic digest or in-source fragmentation etc).

It should be mentioned for this instrument is the observation of matrix cluster in MS/MS spectra, especially in the case where the analyte concentration is low or matrix cluster has the same mass as the parent ion. Every single amino acid has a mass with decimal about 0.05. As the number of amino acid in the peptide increase, the mass decimal has a latter increase manner. Such as peptide LKHMS has a mass of 615.3288 Da, and peptide of LKHMSTYWR has a mass of 1221.6203 Da as well. However, matrix cluster always has a decimal of 0.09 to 0.1. For example, ion of $[3*HCCA + K]^+$ has m/z of 606.0915 Da and $[6*HCCA - 3H + 3K + Na]^+$ has m/z of 1271.1130 Da. Therefore, for the purpose of spectra interpretation or database searching it is important to distinguish the matrix cluster from peptide fragment ions co-existed in the MS/MS spectra and pick out the matrix cluster information from the spectra.

## 1.2 Protein Identification

### 1.2.1 Analysis of Intact Proteins

The mass measurement of the intact protein is the first essential step for protein characterization. Accurate measurement of molecular masses can verify the correctness of the translated sequences. In addition, accurate molecular weight measurement of known proteins can provide information about post-translational modifications, such as phosphorylation or the number of cysteines and disulfide bridge in a protein. Both

MALDI MS and ESI MS can be used for determination of an intact protein molecular weight with an accuracy of better than 0.05%, which is much higher compared to the conventional SDS-PAGE analysis. For proteins separated by gel electrophoresis, MALDI MS is commonly used for protein molecular weight measurement after protein extracted out of the gel piece or electroblotted onto a membrane. However, ESI MS analysis is problematic due to the observation of SDS-protein adducts, and additional effort must be used to eliminate these adducts [39].

### 1.2.2 Peptide Mass Mapping

A set of peptide masses obtained by proteolysis or chemical cleavage can be used as a unique fingerprint, allowing a protein to be identified in a database. The strategy is outlined in Figure 1.10. First, a protein is digested by a chemical or protease with high sequence specificity (such as CNBr or trypsin) to produce a set of peptides. The peptide map is usually determined by MALDI TOF MS or ESI MS. MALDI TOF MS is more favorable for a peptide mapping experiment because MALDI produces exclusively singly charged ions for low mass peptides and, in addition, MALDI can tolerate buffers and salts in the samples. This set of masses is then compared to the mass profile generated from a theoretical digestion of a protein in the protein database in order to find the protein generating the most similar pattern.

Five independent laboratories published methods for database searching using peptide mass mapping [40-44]. In these algorithms, the scoring schemes are either to order the proteins according to decreasing numbers of the matched peptides [40, 42, 44] or based on probability to prevent large random scores caused by large proteins [43]. Search engines on the Internet include Mascot (http://www.matrix-science.com)

PeptIdent (http://ca.expasy.org/tools/peptident.html), MOWSE (http://srs.hgmp.mrc.ac.uk/cgi-bin/mowse), MS-FIT (http://prospector.ucsf.edu/ucsfhtml4.0/msfit.htm) and PeptideSearch (http://www.mann.embl-heidelberg.de/GroupPages/PageLink/peptidesearchpage.html).

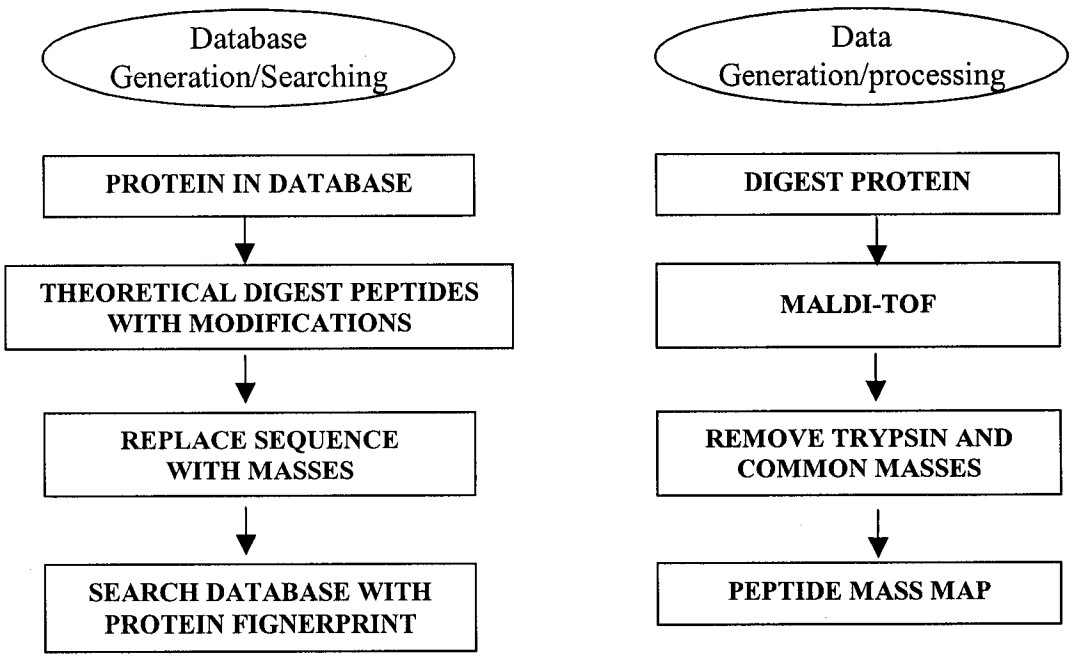



**Figure 1.10** An outline of the process of the peptide mass mapping.

Reliable protein identification is based on several parameters, such as the accuracy of peptide mass determination, the number of masses submitted for query, the mass distribution of the query masses in the candidate protein, the number of matched peptides and the size of the sequence database. In many cases, peptide mass mapping alone is not sufficient for reliable identification of a protein. This is particularly true with protein mixtures or when dealing with small amounts of protein in which only a few peptides can be detected because of the MS detection sensitivity issue. In such cases, additional information, such as peptide fragment information, is required for a confident identification.

### 1.2.3 Peptide Sequence by Tandem MS and Database Searching

Proteins can also be identified by tandem MS analysis of peptides. Because tandem mass spectra contain sequence information of the peptide, rather than only its mass, those searches are generally more unambiguous and discriminative than peptide mass mapping.

Algorithms using uninterpreted tandem spectra for database searching have been developed, such as SEQUEST (http://fields.scripps.edu/sequest/) or Mascot. In the searching process, experimental fragment information is matched against predicted fragment information for all the peptides in the databases which have parent ions of the same m/z. Although this method is highly automated, the sequences need to be manually examined by comparison of fragment ion masses of interest with theoretical ones of the candidate (e.g. MS-Product program http://prospector.ucsf.edu/ucsfhtml4.0/msprod.htm) unless the score of candidate is significantly higher than the next possible identity.

The most widely used method for peptide fragmentation in tandem mass spectrometry is collision-induced dissociation (CID) [45, 46]. When ions possessing high translational energy collide with neutral molecules or atoms, a small amount of their translational energy is converted into internal energy. This can cause decomposition of the ions. CID can be performed at high or low collision energies. Low energy CID (10-100 eV) is widely used for most instruments (triple quadrupole, ion trap or QqTOF) in protein laboratories. Under CID, fragment ions are mostly produced by the dissociation of peptide backbone. Figure 1.11 describes the nomenclature for peptide fragmentation pattern [47]. The fragments will only be detected if they carry at least one charge. If the charge is retained on the C-terminal fragment, the ion is classed as either *x, y* or *z*. If the charge is retained on the N-terminal fragment, the ion type is either *a, b* or *c*. The subscript indicates the number of residues in the fragment. For low CID, the most abundant fragments are neutral loss and *y/b* ions which are formed by dissociation of the peptide backbone. *a*-type ions can be formed by losing of carbonyl group from *b* ions. The less frequent fragments are *c* and *z* ions.

It is believed that the fragmentation reaction of peptides by low energy CID is initiated by a mobile proton and directed by a charge-site. In the absence of strongly basic residue (Arg), the migration of a mobile proton to carbonyl oxygen or amide nitrogen initiates the cleavage of various peptide bonds via a cyclic intermediate. Residues that tend to localize mobile protons, for example, His, or donate protons (Asp, Glu) will have a significant effect on peptide fragmentation. Due to charge localization or donation, a selective cleavage of the protonated peptide bonds are often observed in the N-terminal side of Pro, C-terminal side of Asp, Glu [36, 48-50]. Knowledge of the

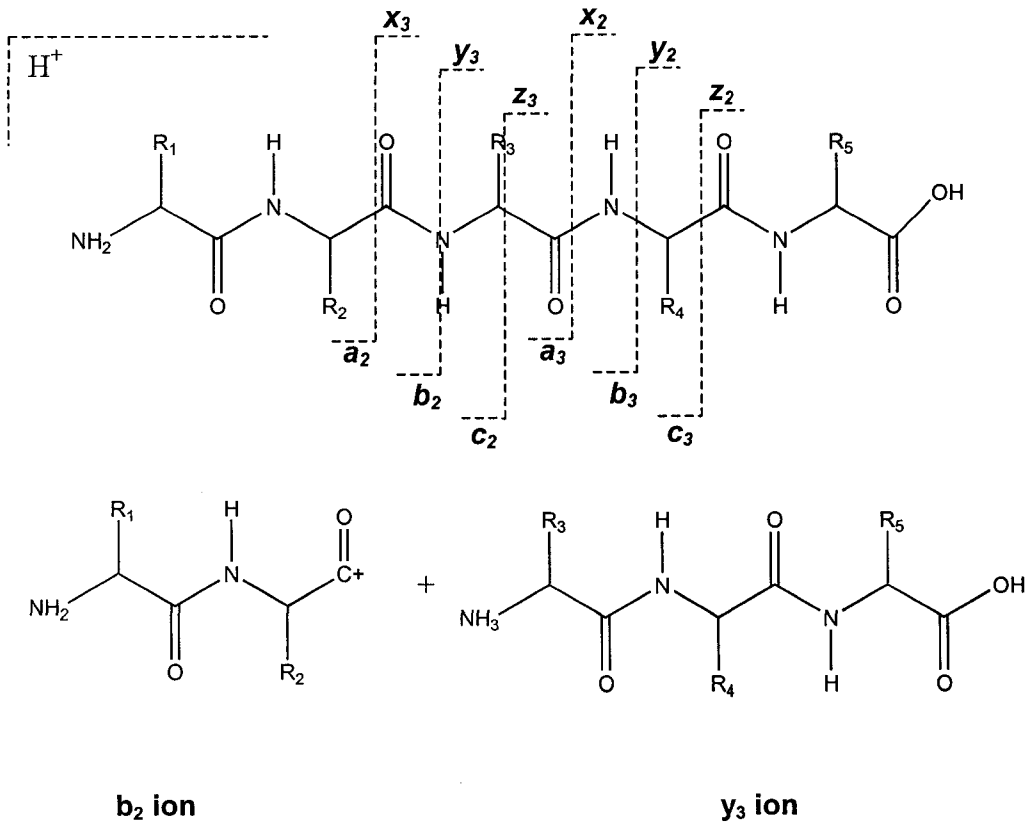



**Figure 1.11** (A) Nomenclature for peptide fragmentation pattern under low CID. (B) Representation of *b* and *y* series ions.

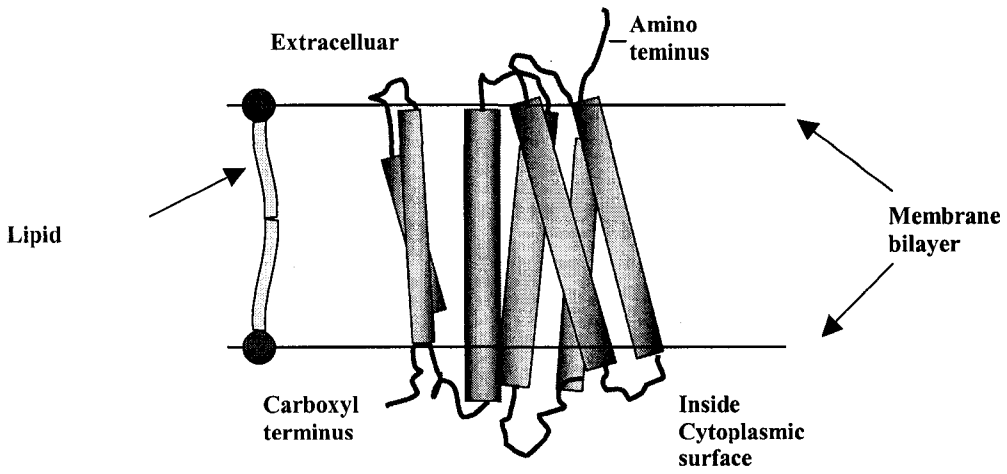



**Figure 1.12** Structure model of hydrophobic membrane protein Bacteriorhodopsin

selective cleavage site is very important for predicting the fragment pattern of the MS/MS spectra of a peptide, and can increase the protein identification confidence where a manual search is required because most algorithms available today do not consider the abundance of a particular product ion.

If positive search results cannot be obtained by using peptide mass mapping and fragment ions searching or if the proteins are not in the database, *de novo* sequencing can be pursued. In this approach, the sequence information can be "read' out manually from the MS/MS spectrum [33, 51, 52]. After obtaining several short sequences from the protein of interest, a cross species database search (BLAST) is performed to find similar proteins from other species. Because the MS/MS spectra are composed of different types of fragment ions, the ions type assignment can be simplified using isotope-labeling techniques, such as *y* ions can be assigned by $^{16}O/^{18}O$ labeling technique [33].

## 1.3 MS for Membrane Protein Analysis

### 1.3.1 Membrane Protein Separation

Membrane proteins are either extrinsic or intrinsic. Extrinsic membrane proteins are outside of the biological membrane, but are bound to it by weak molecular attractions. Intrinsic membrane proteins are embedded in the membrane. Many of them extend from one side of the membrane to the other and are referred to as transmembrane proteins. Figure 1.12 represents the structure of bacteriorhodopsin (BRO), a well-known transmembrane protein commonly used as a model for hydrophobic membrane protein study. Many transmembrane proteins contain a hydrophobic domain that is anchored to the non-polar tails of the lipid membrane, (e.g. the seven helices in Figure 1.12). Transmembrane proteins also have two hydrophilic domains, one at the N-terminal that

extends out into the extracellular medium and one at C-terminal that protrudes into the cytoplasm. Membrane proteins are very important in biological systems because they play a role in cellular processes, such as ion transport, signal transduction, cell adhesion etc. It is estimated that transmembrane proteins correspond to 30% of the total cell proteins based on the already completed sequences of several genomes [53, 54]. Therefore, membrane proteins draw much attention and interest for proteomics researchers.

However, characterization of membrane proteins encounters several difficulties. Generally speaking, protein characterization by MS involves two major issues – separation/purification and analysis [55]. Two-dimensional-gel electrophoresis is one of the most employed techniques for protein separation/purification. After 2D gel electrophoresis was first described [56], it has become a general method for protein separation because it provides the highest resolution. Proteins are separated according to their isoelectric point (PI) by means of isoelectric focusing (IEF) in the first dimension, and according to molecular weight by sodium dodecyl sulfate polyarylamide electrophoresis (SDS-PAGE) in the second dimension. Because these two parameters are distinct, it is very likely that each spot across the 2D gel contains only one protein conformer. However, the separation of membrane proteins by 2D-gel electrophoresis is underrepresented mainly because membrane proteins are poorly soluble in the media used for IEF [57-61]. The membrane proteins are intended to be soluble in lipid bilayers and not in water, and some of them are difficult to dissolve in a water-based environment even though chaotropes and mild detergents (such as urea and CHAPS) are employed in the IEF separation. On the other hand, the surfactant SDS has almost perfect solubilizing

power, and remains used as one of the most popular method for membrane proteins solubilization. One-dimensional SDS-PAGE is often used for membrane protein separation although the method sacrifices separation resolution.

### 1.3.2 Membrane Protein Analysis by MS

Because membrane protein requires surfactants to be present in the biological sample work up for extraction, solubilization and separation, surfactants have always co-existed in the membrane protein samples for mass spectrometrists. However, the analysis of membrane proteins solubilized with surfactants by MS is a challenging task because mass spectrometry requires relatively pure samples to achieve good sensitivity and accuracy, and surfactants are contaminants to MS. Since the MALDI technique was first introduced in 1987, researchers have examined the effects of surfactants on MALDI MS, and many conflicting literature has been presented since then. Some researchers have found that surfactants, especially SDS, were detrimental to the MS signal and must be removed from the sample prior to MS analysis, or avoided altogether [60, 62-69]. Therefore, much effort was directed towards the removal of surfactants [70, 71]. However, the removal of surfactants usually results in loss of proteins and adds undesirable steps to MS analysis [72, 73]. On the other hand, other researchers have found that surfactants did not deteriorate the MALDI MS signal, and could be used for sample preparation [74, 75]. The two camps remained separate until the studies in Chapter 2 were published. It was found that for SDS-containing samples, the intensity of the MALDI signals can be affected by the conditions of sample preparation: on-probe washing, choice of matrix, deposition method, solvent system, and protein-to-SDS ratio.

However, under appropriate conditions, the two-layer method gave reliable MALDI signals for samples with levels of SDS up to ~1% [76].

## 1.4 Summary of the Thesis

In this thesis, I first investigated the effect of SDS on MALDI MS analysis of proteins and peptides. The results showed that sample preparation methods were crucial for analyzing protein samples containing SDS, and proved that by using two-layer sample preparation method, MALDI MS could analyze protein samples containing SDS up to 1%. Then I further explored the cause of SDS' effect on MALDI MS resolution and proposed an alternative surfactant to improve MS measurement accuracy. Furthermore, the effects of commonly used biological surfactants on proteolysis and mass spectrometric response were studied in detail. And based on this study, a methodology for membrane protein identification by SDS-aided in-solution digestion HPLC-MALDI MS/MS was developed and the results obtained from this method were compared to those obtained from the conventional protein identification method, i.e., gel electrophoresis-MS. It was shown that this new method could significantly extend the proteome coverage, resulting in a more comprehensive proteome map. This method was then applied to the analysis of a real sample, membrane protein identification in lipid raft extract. Finally, the in-solution method was further optimized by employment of the multi-dimensional separation technique. This work illustrated that future method development in separation would likely further improve the proteome coverage.

## 1.5 Literature Cited

(1) Fisher, E. H. In *Proteome Research: New Frontiers in Functional Genomics*; Wilkins, M. R., Williams, K. L., Appel, R. D., Hochstrasser, D. F., Eds.; Springer: Berlin, 1997, Preface.

(2) Gabor Miklos, G. L.; Maleszka, R. *Proteomics* **2001,** *1*, 30-41.

(3) Anderson, N. G.; Matheson, A.; Anderson, N. L. *Proteomics* **2001,** *1*, 3-12.

(4) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989,** *246*, 64-71.

(5) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984,** *88*, 4451-4459.

(6) Yamashita, M.; Fenn, J. B. *J. Phys. Chem.* **1984,** *88*, 4671-4675.

(7) Karas, M.; Bachmann, D.; Bahr, U.; Hillenkamp, F. *Int. J. Mass Spectrom. Ion Proces.* **1987,** *78*, 53-68.

(8) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988,** *60*, 2299-2301.

(9) Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988,** *2*, 151-153.

(10) Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001,** *101*, 269-295.

(11) Blackstock, W. P.; Weir, M. P. *Trends Biotechnol.* **1999,** *17*, 121-127.

(12) Dove, A. *Nature Biotechnol.* **1999,** *17*, 233-236.

(13) Pandey, A.; Mann, M. *Nature* **2000,** *405*, 837-846.

(14) Cotter, R. J. In *Time of Flight Mass Spectrometry: Instrumentation and Application in Biological Research*; Am. Chem. Soc.: D. C., 1997. pp 311-314.

(15) Kinter, M.; Sheerman, N. E. In *Protein Sequencing and Identification Using Tandem Mass Spectrometry*; Wiley-Interscience: New York, 2000. pp301.

(16) Weinberger, S. R.; Boernsen, K. O.; Finchy, J. W.; Robertson, V.; Musselman, B. D. *Proceedings of the 41st ASMS Conference on Mass Spectrometry and Allied Topics*, San Francisco, CA,, May 31- June 4 1993; pp 775a-b.

(17) Xiang, F.; Beavis, R. C. *Rapid Commun. Mass Spectrom.* **1994,** *8*, 199-204.

(18) Xiang, F.; Beavis, R. C. *Organic Mass Spectrom.* **1993,** *28*, 1424-2429.

(19) Mock, K. K.; Sutton, C. W.; Cottrell, J. S. *Rapid Commun. Mass Spectrom.* **1992,** *6*, 233-238.

(20) Bai, J.; Liu, Y. H.; Cain, T. C.; Lubman, D. M. *Anal. Chem.* **1994,** *66*, 3423-3430.

(21) Köchling, H. J.; Biemann, K. *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*, Atlanta, Georgia, May 21- 26 1995; p 1225.

(22) Hensel, R. R.; King, R.; Owens, K. G. *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*, Atlanta, Georgia, May 21- 26 1995; p 947.

(23) Vorm, O.; Roepstorff, P.; Mann, M. *Anal. Chem.* **1994,** *66*, 3281-3287.

(24) Edmondson, R. D.; Campo, K. K.; Russell, D. H. *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*, Atlanta, Georgia, May 21-26 1995; p 1246.

(25) Li, L.; Golding, R. E.; Whittal, R. M. *J. Am. Chem. Soc.* **1996,** *118*, 11662-11663.

(26) Kussmann, M.; Nordhoff, E.; Rahbek-Nielsen, H.; Haebel, S.; Rossel-Larsen, M.; Jakobsen, L.; Gobom, J.; Mirgorodskays, E.; Kroll-Kristensen, A.; Palm, L.; Roepstorff, P. *J. Mass Spectrom.* **1997,** *32*, 593-601.

(27) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1996,** *68*, 2494-2500.

(28) Dole, M.; Mach, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B. *J. Chem. Phys* **1968,** *49*, 2240-2247.

(29) Kebarle, P.; Tang, L. *Anal. Chem.* **1993,** *65*, 972A-986A.

(30) Bruins, D. R. *J. Chromatogr.* **1991,** *554*, 39-46.

(31) Wilm, M.; Mann, M. *Int. J. Mass Spectrom. Ion Proces.* **1994,** *136*, 167-180.

(32) Wilm, M.; Mann, M. *Anal. Chem.* **1996,** *68*, 1-8.

(33) Shevchenko, A.; Chernushevich, I.; Ens, W.; Standing, K. G.; Thomson, B.; Wilm, M.; Mann, M. *Rapid Commun. Mass Spectrom.* **1997,** *11*, 1015-1024.

(34) Loboda, A.; Krutchinsky, A.; Loboda, O.; McNabb, J.; Spice, V.; Ens', W.; Standing, K. G. *IM Publications* **2000,** *6*, 531-536.

(35) Loboda, A. V.; Krutchinsky, A. N.; Bromirski, M.; Ens, W.; Standing, K. G. *Rapid Commun. Mass Spectrom.* **2000,** *14*, 1047-1057.

(36) Sullivan, A. G.; Brancia, F. L.; Tyldesley, R.; Bateman, R.; Sidhu, K.; Hubbard, S. J.; Oliver, S. G.; Gaskell, S. J. *Int. J. Mass Spectrom.* **2001,** *210/211*, 665-676.

(37) Raska, C. S.; Parker, C. E.; Huang, C.; Han, J.; Glish, G. L.; Pope, M.; Borchers, C. H. *J. Am. Soc. Mass Spectrom.* **2002,** *13*, 1034-1041.

(38) Zhang, N.; Li, N.; Quach, T.; Keller, B. O.; Li, L. *Proceedings of the 50th Conference of the American Society for Mass Spectrometry and Allied Topics*, Orlando, Florida, June 2-6 2002.

(39) Fridriksson, E. K.; Baird, B.; F.W., M. *J. Am. Soc. Mass Spectrom.* **1999,** *10*, 453-455.

(40) Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. F.; Grimley, C.; Watanabe, C. *Proc. Natl. Sci. USA* **1993,** *90*, 5011-5015.

(41) James, P.; Quadroni, M.; Carafoli, E.; Gonnet, G. *Biochem. Biophys. Res. Commun.* **1993,** *195*, 58-64.

(42) Mann, M.; Hojrup, P.; Roepstorff, P. *Biol. Mass Spectrom.* **1993,** *22*, 338-345.

(43) Pappin, D. J. C.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993,** *3*, 327-332.

(44) Yates III, J. R.; Speicher, S.; Griffin, P. R.; Hunkapiller, T. *Anal. Biochem.* **1993,** *214*, 397-408.

(45) Jennings, K. R. *Int. J. Mass Spectrom. Ion Phys.* **1968,** *1*, 227-235.

(46) Hunt, D. F.; Buko, A. M.; Ballard, J. M.; Shabanowitz, J.; Giordani, A. B. *Biomed. Mass Spectrom.* **1981,** *8*, 397-408.

(47) Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984,** *11*, 601.

(48) Wang, Y.; Johansson, J.; Griffiths, W. J. *Rapid Commun. Mass Spectrom.* **2000,** *14*, 2182-2202.

(49) G., S. S.; Whiting, A.; Gaskell, S. J. *Int. J. Mass Spectrom. Ion Proces.* **1997,** *162*, 149-161.

(50) Gu, C.; Tsaprailis, G.; Breci, L.; Wysocki, V. H. *Anal. Chem.* **2000,** *72*, 5804-5813.

(51) Taylor, J. A.; Johnson, R. S. *Anal. Chem.* **2001,** *73*, 2594-2604.

(52) Keough, T.; Lacey, M. P.; Youngquist, R. S. *Rapid Commun. Mass Spectrom.* **2000,** *14*, 2348-2356.

(53) Paulsen, I. T.; Sliwinski, M. K.; Nelissen, B.; Goffeau, A.; Saier Jr, M. H. *FEBS lett.* **1998,** *430*, 116-125.

(54) Wallin, E.; Von Heijne, G. *Prot. Sci.* **1998,** *7*, 1029-1038.

(55) Tomer, K. B. *Chem. Rev.* **2001,** *101*, 297-328.

(56) O'Farrell, P. H. *J. Biol. Chem.* **1975,** *250*,4007-4021.

(57) Wilkins, M. R.; Gasteriger, E.; Sanchez, J.-C.; Bairoch, A.; Hochstrasser, D. F. *Electrophoresis* **1998,** *19*, 1501-1505.

(58) Santoni, V.; Molloy, M.; Rabilloud, T. *Electrophoresis* **2000,** *21*, 1054-1070.

(59) Adessi, C.; Miege, C.; Albrieux, C.; Rabilloud, T. *Electrophoresis* **1997,** *18*, 127-135.

(60) Maire, M. L.; Champeil, P.; Moller, J. V. *Biochim. Biophys. Acta* **2000,** *1508*, 86-111.

(61) Fountoulakis, M.; Takacs, B. *Electrophoresis* **2001,** *22*, 1593-1602.

(62) Tomer, K. B.; Deterding, L. J.; Parker, C. E. In *Advances in Chromatography*; Brown, P. R., Grushka, E., Eds.; Marcel Dekker: New York, 1995; Vol. 35, Charpter 2.

(63) Ogorzalek Loo, R. R.; Dales, N.; Andrews, P. C. *Protein Sci.* **1994,** *3*, 1975-1983.

(64) Rosinke, B.; Strupat, K.; Hillenkamp, F.; Rosenbusch, J.; Dencher, N.; Krueger, U.; Galla, H.-J. *J. Mass Spectrom.* **1995,** *30*, 1462-1468.

(65) Amini, A.; Dormady, S.; Riggs, L.; Regnier, F. E. *J. Chromagr. A* **2000,** *894*, 345-355.

(66) Karas, M.; Bahr, A.; Ingendoh, E.; Nordhoof, B.; Stahl, B.; Strupat, K.; Hillenkamp, F. *Anal. Chim. Acta* **1990,** *241*, 175-185.

(67) Bornsen, K. O.; Gass, M. A. S.; Bruin, G. J. M.; von Adrichem, J. H. M.; Biro, M. C.; Kresbach, G. M.; Ehrat, M. *Rapid Commun. Mass Spectrom.* **1997,** *11*, 603-609.

(68) Rundlett, K. L.; Armstrong, D. W. *Anal. Chem.* **1996,** *68*, 3493-3497.

(69) Jensen, C.; Haebel, S.; Anderson, S. O.; Roepstorff, P. *Int. J. Mass Spectrom. Ion Proces.* **1997,** *160*, 339-356.

(70) Schaegger, H. In *A practical guide to membrane protein purification*; von Jagow, G., Schaegger, H., Eds.; Academic Press: San Diego, CA, 1994, pp 31-34.

(71) Immler, D.; Mehl, E.; Lottspeich, F.; Meyer, H. E. In *Microcharacterization of proteins*; Kellner, R., Lottspeich, F., Meyer, H. E., Eds.; Wiley-VCH: New York, 1999, pp 87-93.

(72) Galvani, M.; Hamdan, M. *Rapid Commun. Mass Spectrom.* **2000,** *14*, 721-723.

(73) Eckerskorn, C.; Strupat, K.; Schleuder, D.; Dennis Hochstrasser, D.; Sanchez, J.-C.; Lottspeich, F.; Hillenkamp, F. *Anal. Chem.* **1997,** *69*, 2888-2892.

(74) Amado, F. M.; Santana-Marques, M. G.; Ferrer-Correia, A. J.; Tomer, K. B. *Anal. Chem.* **1997**, *69*, 1102-1106.

(75) Breaux, G. A.; Green-Church, K. B.; France, A.; Limbach, P. A. *Anal. Chem.* **2000**, *72*, 1169-1174.

(76) Zhang, N.; Doucette, A.; Li, L. *Anal. Chem.* **2001**, *73*, 2968-2975.

# Chapter 2

# Two-Layer Sample Preparation Method for MALDI Mass Spectrometric Analysis of Protein and Peptide Samples Containing Sodium Dodecyl Sulfate[a]

## 2.1 Introduction

Matrix-assisted laser desorption ionization (MALDI) and electrospray ionization (ESI) mass spectrometry (MS) are widely used for the analysis of biomolecules such as proteins and peptides. MALDI MS is particularly attractive for the direct analysis of complex biological samples such as cell extracts. It can tolerate relatively high levels of contaminants that are often present in a real world sample [1]. Such contaminants are often difficult to remove, either completely or partially, prior to MS analysis without an extensive effort on sample cleanup or separation. Detergents are a class of contaminants that are often added to a sample during workup, with sodium dodecyl sulfate (SDS) being most commonly employed. SDS is used in protein separation by polyacrylamide gel electrophoresis (SDS-PAGE). It is also used for improving solubilization of hydrophobic proteins, reducing protein aggregation and minimizing sample loss due to adsorption to sample containers [2,3].

Several reports have shown that SDS is detrimental to the MALDI signal of peptides and proteins and should therefore be removed from the sample prior to MS

[a] A form of this chapter is published as:
N. Zhang, A. Doucette and L. Li, "Two-Layer Sample Preparation Method for MALDI Mass Spectrometric Analysis of Protein and peptide Samples Containing Sodium Dodecyl Sulfate" *Anal. Chem.*, **2001**, *73*, 2968-2975. Dr. A. Doucette and N. Zhang contributed equally to this work because of the bulk results of the data collected.

analysis, or avoided altogether [4-9]. Nonionic detergents such as β-D-octylglucoside cause less signal degradation than SDS [5-7]. On the other hand, a few recent studies indicated that SDS could help the MALDI analysis. For example, Amado *et al.* showed that if the SDS concentration was increased above 0.23%, the ion signal was seen to improve [10]. They reported this phenomenon to be correlated with protein molecular weight, as well as with protein concentration. The partial signal recovery was reported to consistently occur at higher SDS levels, and the level of SDS at which signal recovery occurred was very reproducible. Limbach *et al.* showed that the ion signal of hydrophobic peptides improved in the presence of SDS at levels above 0.5% [11]. While it is known that the addition of a detergent to the sample can reduce mass discrimination of peptide mixtures [9], SDS was recently shown to improve sequence coverage of proteins from enzymatic digests by improving the relative ionization of the more hydrophobic peptides [12,13].

We believe the wide range of results on the effects of SDS on MALDI MS analysis could be explained by differences in the sample preparation methods used by various groups. It has been shown that pH and on-probe washing are crucial to the success of MALDI analysis of samples containing SDS, and that good spectra could be obtained for solutions containing up to 0.2% SDS [14]. However, it is not known if any other controllable variables for sample preparation are important when dealing with samples containing SDS. One of the most influential variables is the technique of sample/matrix deposition to the MALDI sample plate. There are a number of sample preparation methods developed for MALDI applications. They include dried-droplet [15], vacuum drying [16], crushed-crystal [17], slow crystal growing [18], active film

[19,20], pneumatic spray [21], electrospray [22], fast solvent evaporation [23,24], sandwich [25,26], and two-layer method [27]. We have recently demonstrated that the two-layer method can provide much improved performance in analyzing complex protein and peptide mixtures [28-30], compared to the conventional dried-droplet method and the fast evaporation method. The two-layer method involves the use of fast solvent evaporation to form the first layer of small matrix crystals, followed by deposition of a *mixture* of matrix and analyte solution on top of the crystal layer. With this method, the matrix and analyte solution conditions for preparing the second layer can be readily altered and fine-tuned for specific applications [28].

In this work, we demonstrate that the two-layer sample preparation method is remarkably robust in handling protein and peptide samples containing SDS for MALDI analysis. A detailed study of the effects of SDS, under different preparation conditions involving the use of the two-layer method as well as the dried-droplet method, is performed. In addition, we present two new applications where the inclusion of SDS in the sample workup may be beneficial and where successful MALDI analysis was still possible in the presence of SDS using the two-layer method.

## 2.2 Experimental

### 2.2.1 Chemicals and Reagents

Bovine serum albumin (BSA), equine cytochrome *c*, myoglobin, ubiquitin, insulin chain B oxidized, and bradykinin were from Sigma Aldrich Canada (Oakville, ON). Sinapinic acid (SA), 2,5-dihydroxybenzoic acid (DHB), 2-(4-hydroxyphenylazo)-benzoic acid (HABA), α-cyano-4-hydroxycinnamic acid (HCCA), and sodium dodecyl sulfate

were received from Aldrich Canada. HCCA was purified by recrystallization from ethanol prior to use. Analytical grade acetone, methanol, acetonitrile, formic acid, isopropanol, and trifluoroacetic acid (TFA) were purchased from Caledon Laboratories (Edmonton, AB). Water used in all experiments was from a NANOpure water system (Barnstead/Thermolyne).

### 2.2.2 MALDI Sample Preparation

Protein samples were prepared in SDS solutions and were thoroughly vortexed. The SDS concentrations (w/v) indicated in this work are for the protein samples prior to mixing with the MALDI matrices.

Two MALDI sample deposition methods were employed in this study, namely the dried-droplet, and two-layer method. For the dried-droplet method, the samples were mixed with one of the following four matrices on a stainless steel target: HCCA, HABA, sinapinic acid and DHB (0.5μL sample + 0.5μL matrix). With HCCA, HABA [31] and sinapinic acid, the matrix was saturated in 50% (v/v) acetonitrile/water unless otherwise indicated. With DHB, the matrix was dissolved in 9:1 (v/v) water/ethanol at a concentration of 20 g/L. For the two-layer method, 1 μL of a 12-mg/mL solution of HCCA in 80% acetone/methanol was deposited on the target and was allowed to dry, and form a very thin layer of fine crystals. A 1 μL aliquot of the protein solution was mixed with 1 μL of HCCA solution (saturated in 40% methanol/ water) and a 0.4 μL portion of this mixture was deposited on top of the first layer and allowed to dry.

For each method, an on-probe washing step is performed in cases where the sample spot is not water-soluble. The washing step involves adding 1 μL of water to the

sample spot and then absorbing the water with a Kimwipe. This procedure was performed twice on each sample spot.

### 2.2.3 Bacteria Extraction

The *Escherichia coli* used in this study (*E. coli* ATCC 9637) was cultured at the Edgewood RDE Center (EDREC). A 1 mg sample of *E. coli* was mixed with 500 μL of water. The sample was briefly vortexed to suspend the cells, and the suspension was divided evenly among 5 separate vials. To remove excess salts, the suspensions were vortexed for a few seconds, and the supernatant was discarded after centrifugation. The *E. coli* samples were extracted with 0.1% TFA by vortexing for 3 minutes followed by centrifugation. The pellet was then further extracted with 0.1% TFA containing various concentrations of SDS. The supernatant of the extracts was collected for MALDI analysis after mixing 1:1 with saturated HCCA (in 1:2:3 formic acid: isopropanol: water) solution. Samples were stored in ice prior to MALDI analysis.

### 2.2.4 Instrumentation

MALDI experiments were carried out on an Applied Biosystems Voyager Elite laser desorption/ionization time-of-flight (TOF) mass spectrometer (Framingham, MA) or a HP Model G2025A MALDI-TOF MS (Hewlett-Packard, Reno, NV). The linear mode of operation was used for protein detection while the reflectron mode was for peptide detection. Ionization was performed with a 337-nm pulsed nitrogen laser. The laser power was adjusted slightly above the threshold of the desorption/ionization process and was kept constant for a given set of protein samples (i.e., one protein at a given concentration, with various concentrations of SDS). All spectra resulted from signal

averaging of 50 to 64 shots. For each sample, the laser was moved throughout the sample spot in order to generate the best possible spectrum in each case.

All data were processed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR). The signal to noise ratio (S/N) was calculated from the peak height and the standard deviation of the baseline, determined at a region near the peak of interest. The peak height was determined from the difference between the baseline, calculated as the average over the region used to calculate S/N, and the highest point of a modified Gaussian fitted curve of the peak. The error bars in the figures represent the calculated standard deviation of the signal to noise ratios of the spectra. At least three spectra, recorded from independent sample spots, were averaged for each data point in the figures.

## 2.3 Results and Discussion

### 2.3.1 Effect of on-Probe Sample Washing

The sample/matrix crystals prepared by the two-layer method can be readily washed with water. The effect of sample washing on MALDI analysis of proteins containing different amounts of SDS is illustrated in Figure 2.1. In this experiment, samples consisting of 500 nM BSA in 0%, 0.05%, or 1% SDS were prepared. The samples were mixed in a 1:4 ratio with saturated HCCA matrix solution and spotted on a stainless steel probe coated with a thin layer of matrix crystals. The samples were initially analyzed by MALDI MS without the on-probe washing step. Following acquisition of the spectra, the same sample spots were then washed on-target and re-analyzed by MALDI MS. Figure 2.1 displays some representative MALDI spectra

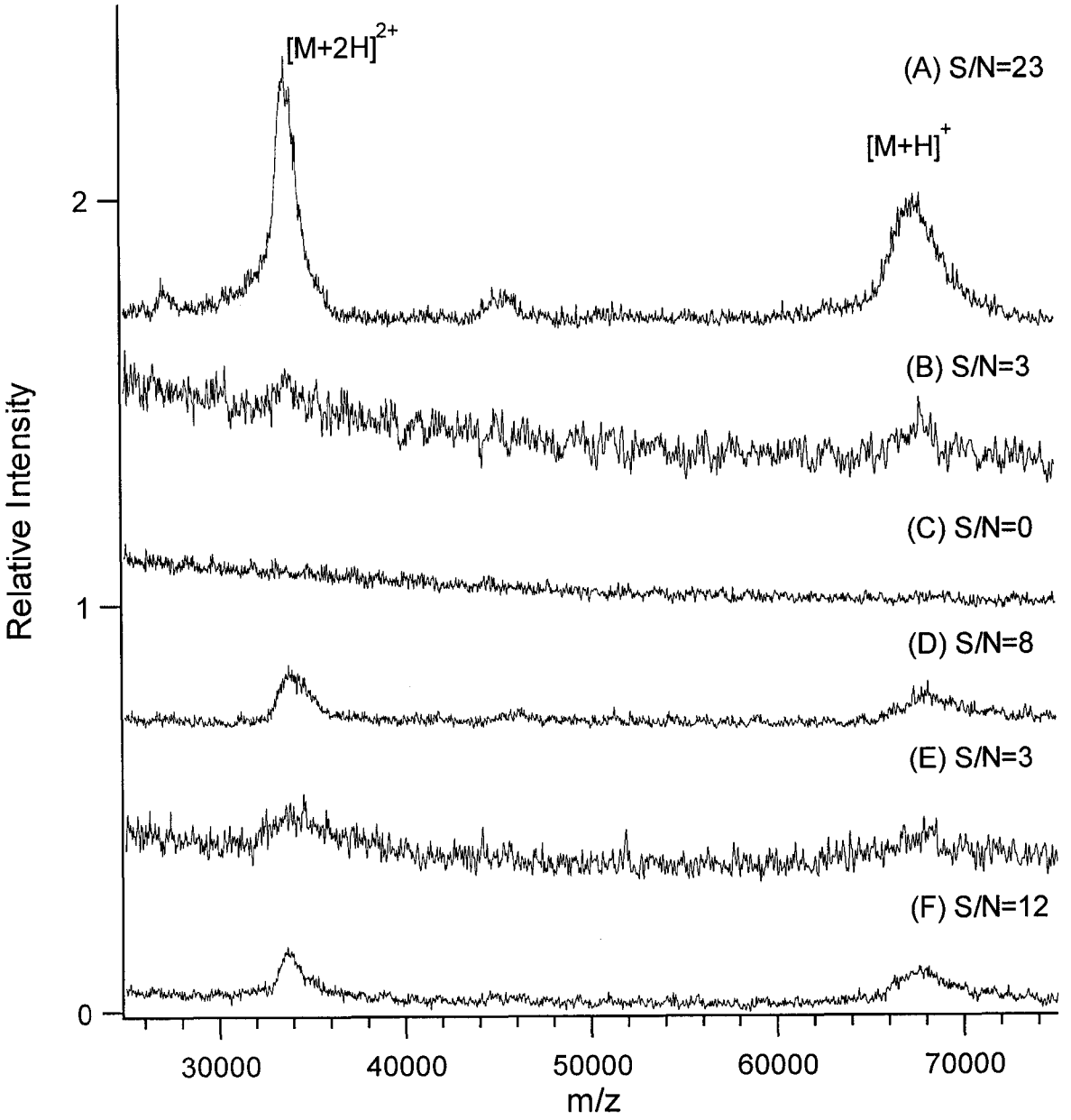



**Figure 2.1** MALDI spectra obtained from samples of 500 nM BSA prepared in (A) 0% SDS, (B) 0.05% SDS, (C) 1% SDS, without on-probe washing; and in (D) 0% SDS, (E) 0.05% SDS, (F) 1% SDS, with the addition of on-probe washing. (A) & (D), (B) & (E), and (C) & (F) were from same sample spot. The calculated signal to noise of the molecular ion peak is indicated in the figure.

obtained from 500 nM BSA samples with and without on-probe washing.

As shown in Figure 2.1A, a strong $[M+H]^+$ BSA signal was observed for the sample prepared without SDS, having a signal to noise (S/N) ratio of ~23. However, the presence of 0.05% SDS in the sample resulted in a weakly detected peak with a calculated S/N of ~3 (see Figure 2.1B). At levels of SDS higher than 0.1%, the BSA peak was undetectable (see Figure 2.1C). When the same sample spots were washed, a remarkable difference was observed. Figure 2.1D shows that on-probe washing reduces the signal intensity for the sample that does not contain SDS. This is due to the loss of sample during the process of on-probe washing. Figure 2.1E shows that a weak BSA peak was again observed at the 0.05% SDS level. However, unlike the sample analyzed without the washing step (see Figure 2.1C), a strong BSA signal (S/N=12) can now be detected for the sample containing 1% SDS with on-probe washing (see Figure 2.1F). Similar results were obtained with various protein samples analyzed by MALDI with HCCA as matrix. This deterioration in signal at low levels of SDS and signal recovery at higher SDS levels is in agreement with earlier work [10]. In summary, the washing step is crucial in order to obtain a signal at high SDS concentrations (>0.1%) with the two-layer sample preparation method. It should be noted that although the washing step appears to improve MALDI signals at high SDS concentrations, not all sample preparation conditions are tolerant to on-probe washing. It was observed that as the SDS level exceeded 1% in the sample, the sample spots became increasingly soluble. In addition, the water-soluble matrix DHB would completely dissolve if subjected to such washing. In general, however, a washing step should be performed on all samples if the deposited spot can tolerate this procedure.

### 2.3.2 Effect of Matrix

Four commonly used matrices, namely HCCA, DHB, HABA, and SA, were employed in this study. Solutions of 1 μM cytochrome *c* prepared at various concentrations of SDS were analyzed using these four matrices. We note the two-layer method is not applicable to DHB – this matrix does not dissolve in an appropriate organic solvent such as acetone, which is required to form a thin layer by fast evaporation. To include DHB in this comparative study, the dried droplet deposition method was used for all matrices. On-probe washing was performed on the samples, except for those prepared with DHB or with HABA at SDS levels ≥1%, as the sample spots were water soluble. Figure 2.2 summarizes the average signal to noise ratios of the singly charged cytochrome *c* peaks obtained from the spectra using each matrix at varying concentrations of SDS.

Figure 2.2 shows that the type of matrix used plays a significant role in determining spectral quality at varying concentrations of SDS. Sinapinic acid was least tolerant towards the presence of SDS in the sample. Figure 2A shows that at most levels of SDS either very weak signals or no signals were observed. The spectra acquired using HCCA as matrix also display generally weak signals when SDS was added to the samples (Figure 2.2B). The matrix HABA seemed to be slightly more tolerant to SDS (Figure 2.2C), with good S/N at all levels of SDS ≤1%. A decrease in signal intensity was seen for HABA at high levels of SDS (2 and 5%). Finally, samples prepared using DHB as matrix yielded much stronger signals in the presence of SDS as compared to the other matrices (Figure 2.2D). However, DHB tends to form a characteristic ring of crystals at the edge of the sample spot giving poor spot-to-spot spectral reproducibility. It was

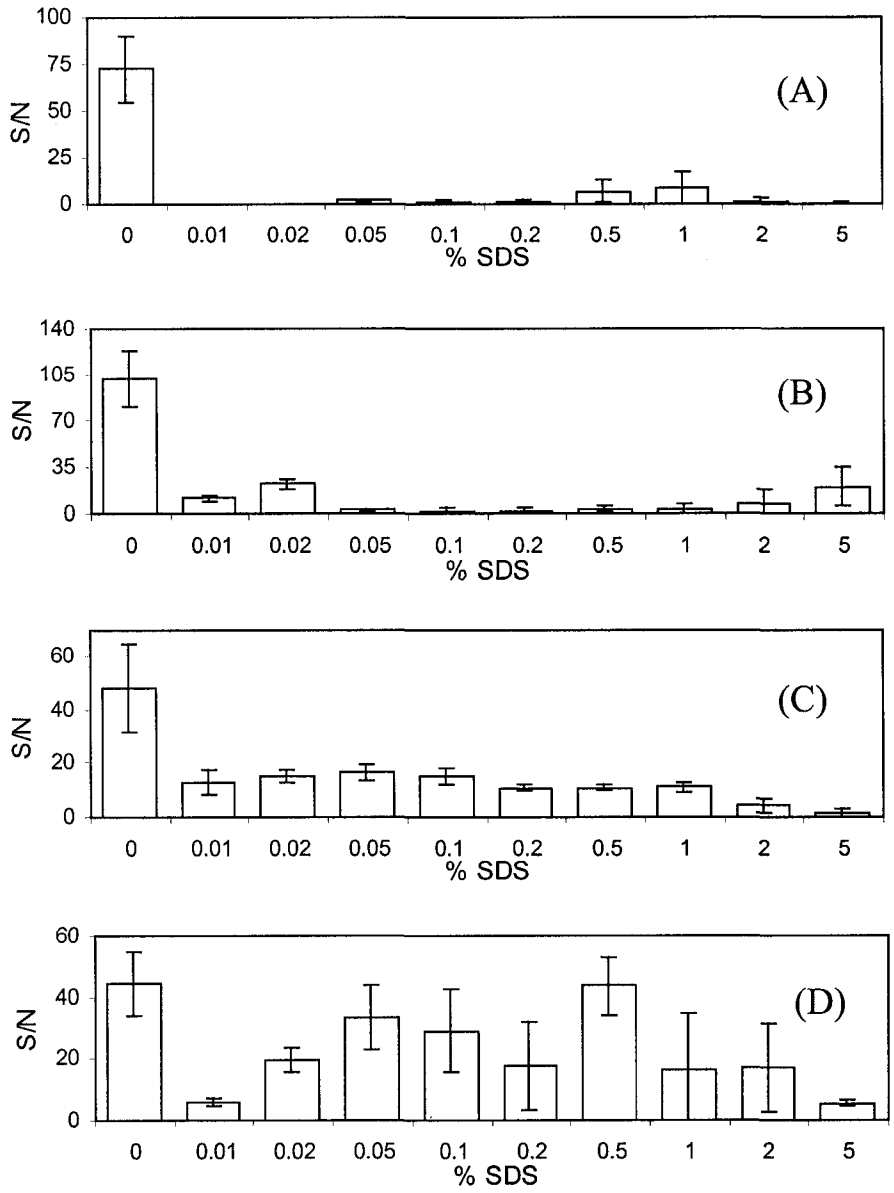



**Figure 2.2** Average signal to noise ratios (S/N) for the molecular ion signals calculated from the MALDI spectra of 1 μM cytochrome *c* at various concentrations of SDS. The matrices used for the analysis are (A) sinapinic acid, (B) HCCA, (C) HABA, (D) DHB. Error bars represent the standard deviations of S/N from 9 spectra recorded from 3 independent sample spots (A, C, D) or 3 independent spectra (B).

necessary to search for "hot" spots in order to obtain any detectable signal.

It was not surprising that the choice of matrix would play a role in obtaining MALDI spectra for SDS-containing samples. The success of protein ionization and subsequent detection depends extensively on co-crystallization of sample and matrix [27]. The addition of SDS to the sample is expected to change the solubility of the protein and likely affects the extent of analyte incorporation as well as analyte distribution in the matrix crystals. However, the results presented in Figure 2.2 cannot be exclusively correlated to the choice of matrix, since other factors are expected to influence the trends. In particular, the solvent system was not identical for all matrices used. Sinapinic acid, HABA, and HCCA were prepared as saturated solutions in 50% acetonitrile/water, while the DHB solution was prepared as a 20-mg/mL solution in 9/1 water/ethanol. Therefore, we investigated the effects of solvent and sample preparation method on SDS-containing samples using HCCA as matrix. HCCA was chosen since it provides superior detection sensitivity and higher tolerance to salts for protein samples, particularly high molecular weight proteins [32], compared to SA and HABA.

### 2.3.3 Effects of Solvent and Sample Preparation Method

Samples of 1 μM cytochrome *c* were prepared and analyzed by MALDI MS using two different sample preparation methods with HCCA as the matrix. In the dried droplet method, the sample was mixed with HCCA (saturated in either 50% acetonitrile/water or 40% methanol/water) in a 1:1 ratio. For the two-layer method, the sample was mixed in a 1:1 ratio with HCCA (saturated in either 50% acetonitrile/water or 40% methanol/water) and 0.4 μL of the mixture was deposited on top of a seed layer of matrix crystals. All samples were washed on-probe and then analyzed by MALDI MS. Figure 3 displays the

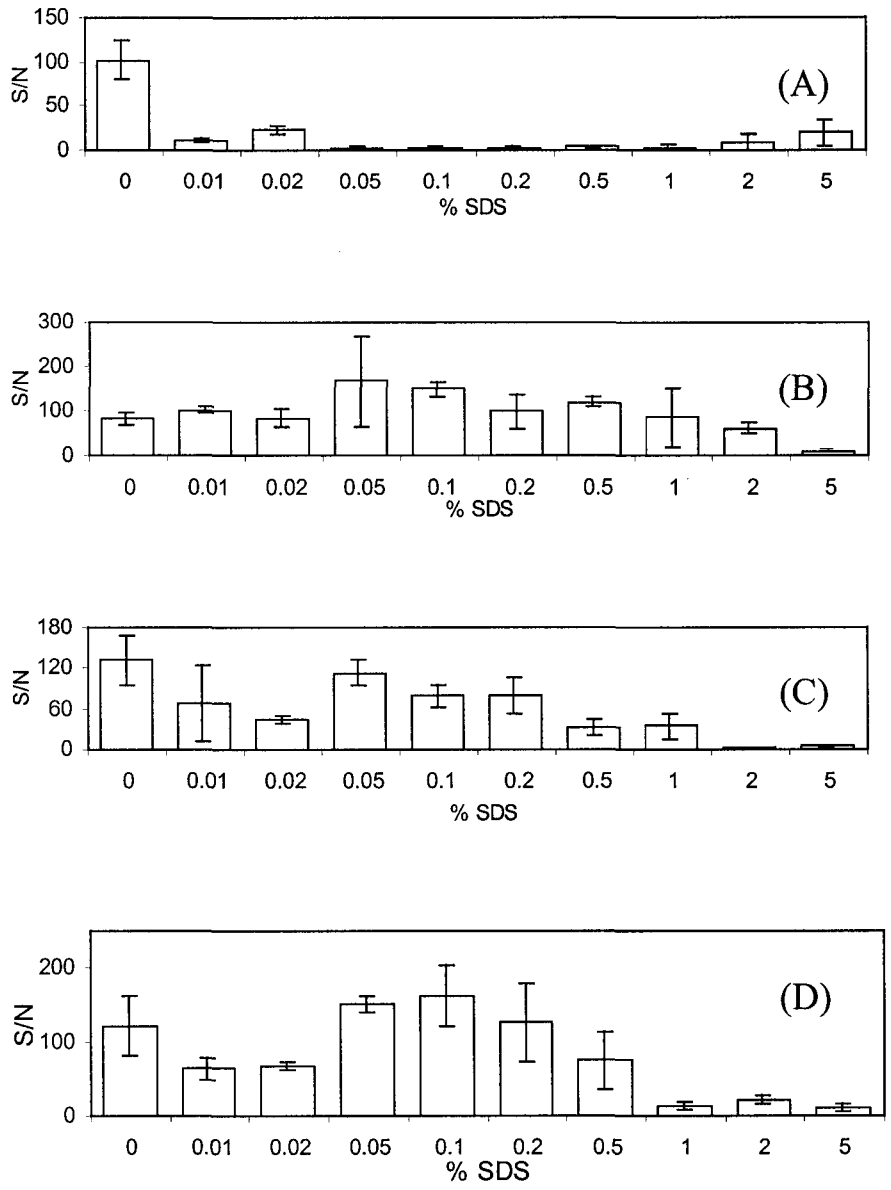



**Figure 2.3** Average signal to noise ratios (S/N) for the molecular ion signals calculated from the MALDI spectra of 1 μM cytochrome *c* at various concentrations of SDS using HCCA as matrix with different sample deposition methods and solvent systems. The dried droplet method was used with HCCA saturated in (A) 50% acetonitrile/water and (B) 40% methanol/water. The two-layer method was used with the second layer HCCA saturated in (C) 50% acetonitrile/water and (D) 40% methanol/water. Error bars represent the standard deviations of S/N from 3 independent spectra.

averaged signal to noise ratios calculated from the MALDI spectra collected under the various conditions.

The comparison of the results between Figure 2.3A and 2.3B indicates that, for the dried-droplet method, the solvent system has a significant effect on the protein signals at different levels of SDS. Figure 2.3A was obtained by using the sample prepared with HCCA in 50% acetonitrile/water and it shows that the analyte signals were reduced drastically for the samples containing different levels of SDS. In contrast, Figure 2.3B illustrates that when the sample was prepared in the methanol/water solvent system, at all levels of SDS up to 2%, the S/N ratio was equal to or greater than that of the sample prepared without SDS. The cause of this strong solvent effect is unknown. It might be related to the extent of analyte incorporation into matrix crystals prepared under different solvent conditions.

Although the choice of solvent has a significant effect on signal response for SDS-containing samples deposited using the dried droplet method, this effect was not observed when a two-layer sample preparation method was used. This can be seen from the comparison of panels C and D of Figure 2.3. For either of the two solvent systems, the addition of 0.01 or 0.02% SDS was seen to reduce the observed S/N ratio of the cytochrome *c* peak. The presence of 0.05% SDS in the sample resulted in signal recovery with each solvent system. When the SDS concentration was further increased, the S/N ratio of the cytochrome *c* peak deteriorated.

The results of Figure 2.3 also show that, under identical solvent conditions, the sample deposition method has an effect on the signal intensity of SDS-containing samples. This is most clearly seen for the samples prepared using the acetonitrile-water

solvent system, shown in panels A and C of Figure 2.3. Large differences are noted at levels of SDS in the range of 0.05% to 0.2%. At these levels of SDS, very weak signals were obtained with the dried droplet method (Figure 2.3A). However, using the two-layer method, the addition of 0.05 to 0.2% SDS resulted in signals with intensities that were comparable to the signal for the sample without SDS (Figure 2.3C). The samples prepared using the methanol-water solvent system, shown in Figure 2.3B and 2.3D, also show the dependence of signal intensity on sample deposition method. Significant differences are noted for the samples prepared in 1 and 2% SDS. While the dried droplet method still revealed relatively strong signals at these high levels of SDS, the samples prepared using the two-layer method displayed only weak signals as compared to the samples prepared without SDS.

It should be noted that, when the dried droplet method was used for the samples prepared at SDS levels above 1%, it became increasingly difficult to find the hot spots that produced the signals shown in Figure 2.3A or 2.3B. The fact we observed analyte signals for the samples containing 2 or 5% SDS can be contributed to the inhomogeneous analyte distribution in the matrix/analyte crystals prepared by using the dried droplet method. Some spots, albeit difficult to find, contain the proper crystals that give arise the MALDI analyte signals. For general applications in analyzing samples containing SDS up to ~1%, the two-layer sample preparation is preferred. With this method, the analyte signals can be easily observed without the need of searching for the hot spots.

### 2.3.4 Effect of Protein Concentration

Another important variable in MALDI sample preparation is the protein concentration, which determines the analyte to matrix ratio in the co-crystals formed on

the probe. For SDS-containing samples, protein concentration also governs the relative amount of SDS to protein for a given concentration of SDS. Figure 2.4 displays the average signal to noise ratios for the singly charged molecular ion signals obtained from the MALDI spectra of 500 nM, 1.5 μM, 5 μM, and 15 μM BSA with various concentrations of SDS. The sample spots were prepared using the two-layer method with the second layer containing analyte/matrix in a 1:4 ratio.

Figure 2.4 shows different trends for the low and high concentrations of BSA. For the 500-nM solution (Figure 2.4A), one can see that the MALDI signal intensity dropped only slightly at low amounts of SDS, and then fully recovered as the SDS concentration was increased to 0.5%. As the SDS concentration was increased above 0.5%, the signal intensity again dropped. Figure 2.4D shows that, for a 15 μM solution of BSA, the signal intensity decreased significantly as the concentration of SDS in the sample increased from 0% to 0.05%. The signal intensity did not return to any significant extent as the SDS concentration was further increased. At intermediate levels between these low and high concentrations of BSA, one can observe that for a 1.5 μM solution of BSA (Figure 2.4B), the trend most resembled that of the 500 nM BSA solution, with only a slight drop in signal intensity as the SDS was increased to 0.1%. The 5 μM BSA samples (Figure 2.4C) displayed a trend that was very similar to the 15 μM BSA solution. These results demonstrate that protein concentration in the sample has an effect on the observed signal trends at different levels of SDS concentrations.

The differences observed in the trends in Figure 2.4 may result from variations in one, or a combination of several variables between each set. The most notable variations

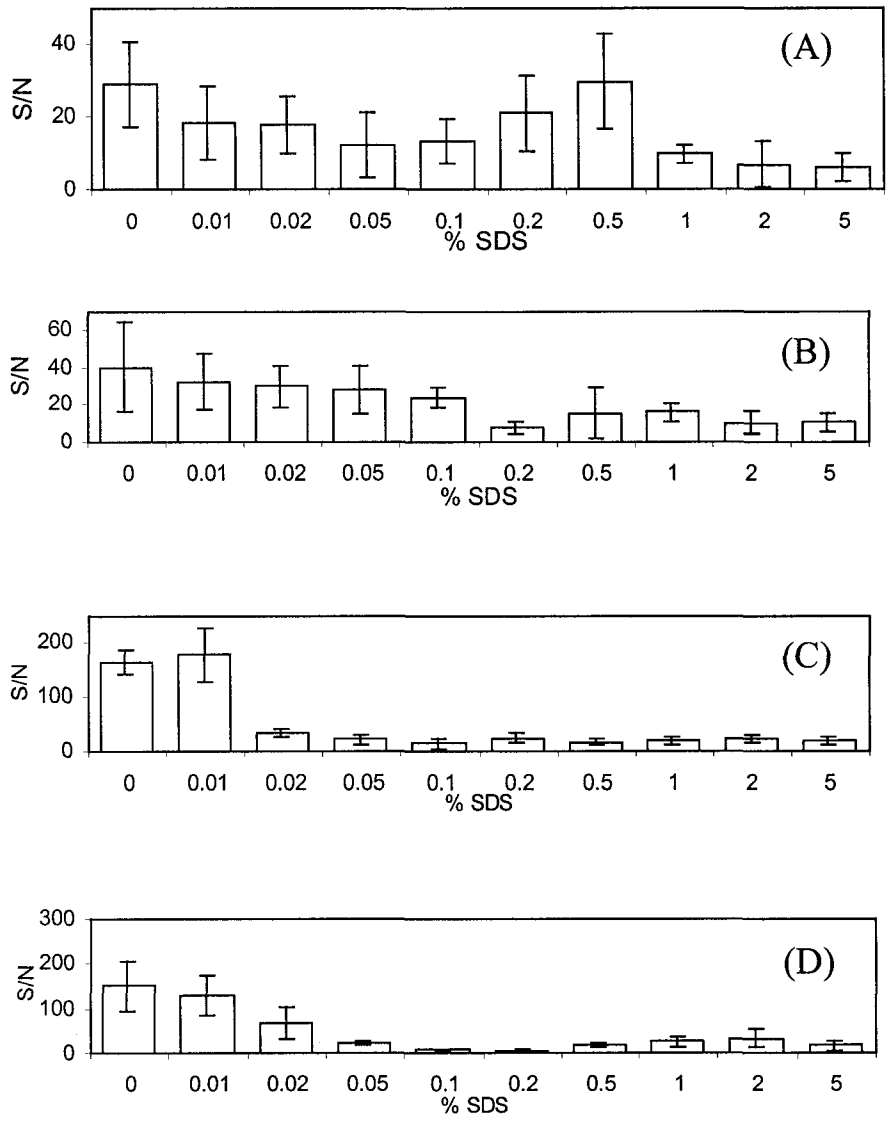



**Figure 2.4** Average signal to noise ratios (S/N) for the molecular ion signals obtained from the MALDI spectra of BSA at various concentrations of SDS. (A) 500 nM, (B) 1.5 μM, (C) 5 μM, or (D) 15 μM BSA was mixed with the second-layer matrix solution in a ratio of 1:4. The second-layer matrix solution consisted of 40% methanol/water HCCA matrix. Error bars represent the standard deviations of S/N from 3 independent spectra.

include the ratios between protein and matrix, as well as protein to SDS. These variations are further explored in the following section.

### 2.3.5 Effect of Protein to Matrix Ratio

To address how the protein to matrix ratio in the presence of different amounts of SDS can affect the protein signal intensity, solutions of 15 μM BSA were prepared with various concentrations of SDS (0%, 0.5%, 1%, and 5%). The samples were mixed with matrix solutions at ratios of 1:1, 1:4, 1:10, 1:50, and 1:100. All samples were analyzed by MALDI-TOF MS using identical laser power. The average signal to noise ratios of the singly charged BSA peaks were calculated and the results are summarized in Table 2.1.

From Table 2.1, the effect of varying the protein to matrix ratio on the signal intensity is clearly observed. For the 15-μM BSA sample prepared without SDS (0%), the strongest signals were observed for the samples that were diluted 1:4 or 1:10 with matrix. As might be expected, at low dilutions of sample to matrix (1:1), there is insufficient matrix on the target to effectively ionize the high quantity of BSA deposited on the probe (3 pmol). At high dilutions of sample to matrix (1:50, 1:100), the decreased amount of BSA present on the target results in a loss in signal intensity.

For the SDS-containing samples in Table 2.1, we can also observe a dependence of signal intensity on protein to matrix ratio; however, this dependence is less pronounced than that of the sample without SDS. By comparing the S/N ratios at a given level of SDS, only a small difference in S/N is incurred by changing the ratio of protein to matrix. Table 2.1 shows that the ideal protein to matrix ratio is dependent on the amount of SDS

**Table 2.1 Signal to Noise Ratios (S/N) for 15 μM BSA Samples Prepared with Various Concentrations of SDS at Different Protein to Matrix Ratios.**

| | Calculated S/N | | | |
|---|---|---|---|---|
| Protein-to-matrix ratio | 0% SDS | 0.5% SDS | 1% SDS | 5% SDS |
| 1:1 | 29 ± 11 | 85 ± 20 | 68 ± 24 | 21 ± 4 |
| 1:4 | 215 ± 39 | 55 ± 28 | 55 ± 15 | 44 ± 5 |
| 1:10 | 240 ± 14 | 32 ± 10 | 35 ± 9 | 36 ± 16 |
| 1:50 | 107 ± 22 | 44 ± 14 | 37 ± 8 | 15 ± 3 |
| 1:100 | 27 ± 6 | 55 ± 15 | 39 ± 4 | 7 ± 5 |

in the sample. At 0.5% and 1% SDS, the strongest signals were obtained for the 1:1 dilution of analyte with matrix. At 5% SDS, the strongest signal was at the 1:4 dilution.

Another important observation can be made by comparing the S/N ratios obtained for a fixed ratio of protein to matrix. From Table 2.1, we can see that for the 1:1 dilution, the signal to noise ratio was seen to increase from 29 at 0%, to 85 at 0.5% SDS; in other words, an improvement in signal intensity was achieved by adding SDS to the sample. When 1% SDS was added to the sample, the S/N ratio still remained high, and at 5% SDS, the signal decreased to a similar intensity as for the 1:1 dilution without SDS. A very similar trend was also obtained for the 1:100 dilution of sample with matrix, with signal improvement being achieved with the addition of 0.5% and 1% SDS to the sample. However, from the results in Table 2.1, a significantly different trend was obtained for the 1:4 and 1:10 dilutions of sample with matrix. From these dilutions, the strongest signal was observed for the sample containing no SDS. When SDS was added to the sample at the 1:4 and 1:10 dilutions, a marked decrease in signal intensity was observed. For a given dilution, the S/N ratio obtained at the 0.5%, 1%, and 5% levels of SDS were essentially identical. A similar trend is also apparent at the 1:50 dilution, where again the addition of 0.5% SDS or more to the sample significantly decreased the signal intensity. The type of trend that is obtained by varying the SDS in the sample will therefore depend on the protein to matrix ratio used for the two-layer method.

It is noted that, the preparations that yield a trend of decreasing signal intensity as SDS is added to the sample, occurred when the protein-matrix ratio was at or near the optimal ratio of 1:10. Such a trend is in agreement with the theory that SDS will decrease signal intensity, due to interference in crystallization and ionization. For the samples that

were prepared with a 1:1 or a 1:100 dilution of sample to matrix, adding SDS to the sample results in an improvement in signal over the initial preparations (without SDS). However, These improved S/N ratios are still significantly lower than the S/N of 240, which was obtained at the optimization ratio of sample to matrix of 1:10. The trend of improved signal intensity as SDS is added to the sample was only obtained for the preparation with poor ratios (too low or too high) of protein to matrix. These results therefore indicate that the ratio of sample to matrix will affect the type of trend obtained by adding SDS to the sample.

We note that the reason(s) as to how the presence of different amounts of SDS in a protein or peptide sample can affect MALDI signals is unknown. SDS can affect sample preparation and/or ionization. In the two-layer method, the analyte/matrix crystals are of submicrometer sizes; there are no visible differences (under 50x magnification) between the morphology of crystals prepared with varying amounts of SDS. One can speculate that proteins are still associated with SDS in the matrix crystals; but exactly how the matrix, protein and SDS in the crystals are associated is unknown. The actual conformer of the protein in the crystal is also unknown. SDS, at a sufficient amount in a solution, is well known to cause protein denaturing. But once the SDS/protein sample is mixed with a matrix and the mixture is allowed to dry to form co-crystals, one really does not know the state of the protein in the crystals. It is thus difficult to ascertain if protein denaturation by SDS plays any role on MALDI signal detectability. Nevertheless, based on the above study, it is quite clear that the two-layer method, using well-controlled sample preparation conditions, can be used to handle protein samples containing SDS. In the following we demonstrate some applications of

protein sample preparations involving the use of SDS and study the effect of SDS on MALDI signals in these applications.

### 2.3.6 Hydrophobic Protein

Many biologically significant membrane-associated proteins are hydrophobic [33]. Surfactants are often added to solutions of hydrophobic proteins to facilitate solubilization, as well as to avoid sample loss due to adsorption to the walls of the container. In this study, bacteriorhodopsin was used as a model hydrophobic protein. In Figure 2.5, the MALDI spectra obtained from the analysis of 0.25 μg/μL bacteriorhodopsin solutions containing various levels of SDS are shown. The two-layer method was employed, using HCCA as matrix. Figure 2.5A corresponds to the sample prepared in aqueous solution without SDS. It shows the bacteriorhodopsin peak near m/z 27000, along with a second peak approximately 1200 Da higher than the bacteriorhodopsin peak, which likely corresponds to an unprocessed precursor of bacteriorhodopsin [34]. When 0.1% SDS was added to the sample, the resulting spectrum (Figure 2.5B) gave a stronger signal. At the 1% SDS level, the bacteriorhodopsin peak is still clearly detected. A small increase in signal intensity obtained by adding SDS to the sample is perhaps due to the improved solubility of the protein with the SDS solvent system. This representative example demonstrates that hydrophobic proteins can be analyzed by the two-layer sample preparation MALDI MS in the presence of SDS.

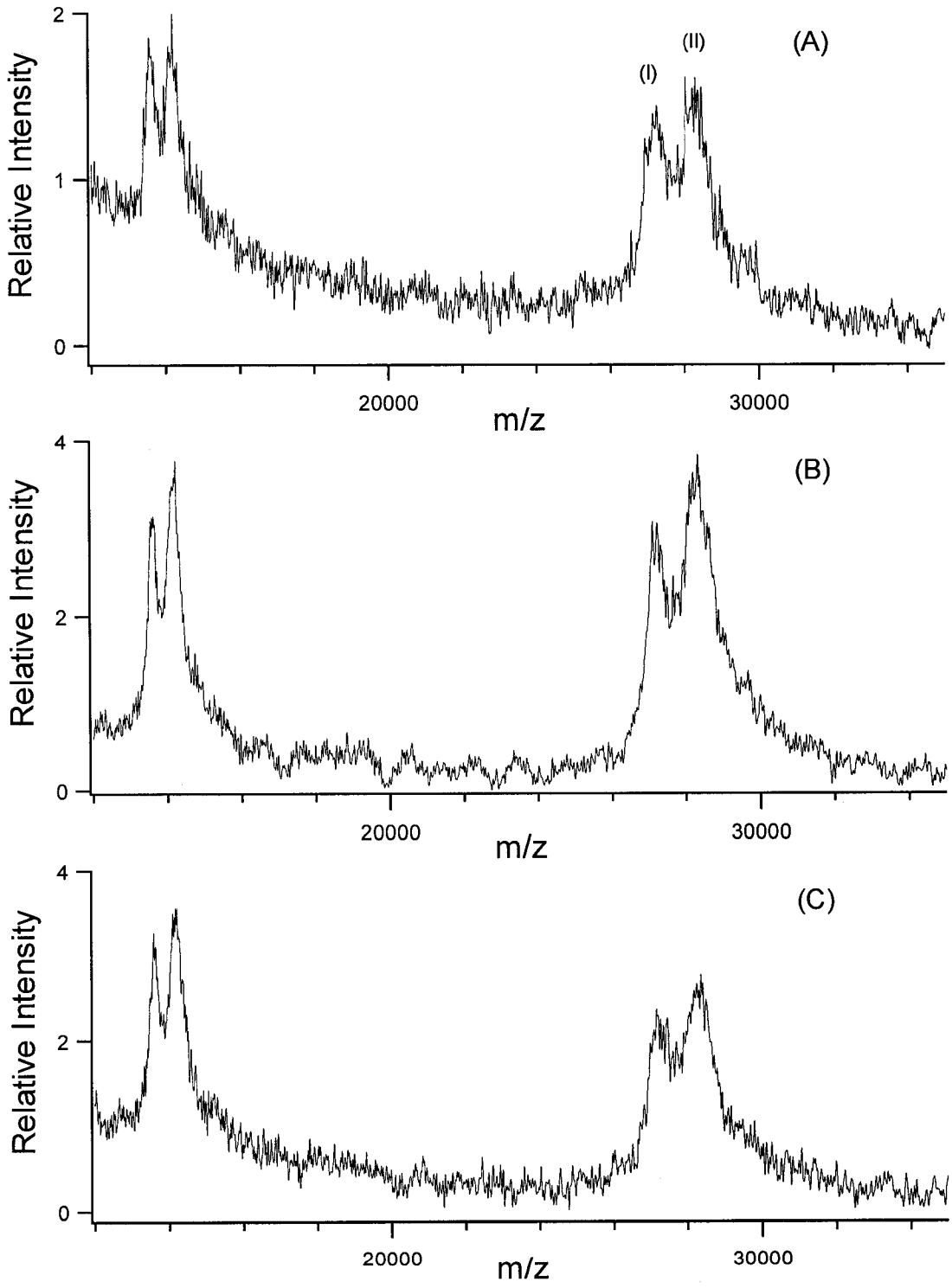



**Figure 2.5** MALDI spectra obtained from samples of 0.25 μg/μL bacteriorhodopsin with (A) 0% SDS, (B) 0.1% SDS and (C) 1% SDS using HCCA matrix and a two-layer deposition method. The labeled peak I indicates the bacteriorhodopsin molecular ion signal and the peak labeled as II is likely from an unprocessed precursor of bacteriorhodopsin.

### 2.3.7 Bacterial Protein Extraction

Detergents such as SDS are often incorporated in extraction procedures to solubilize certain types of proteins from cells, such as the more hydrophobic membrane proteins [2]. Here we demonstrate that the addition of SDS to *E. coli* protein extraction aids in the detection of additional proteins by MALDI. In this case, a 2 mg/ml sample of *E. coli* was first extracted with 0.1% aqueous TFA in order to solubilize the more water-soluble proteins. Figure 2.6A depicts the MALDI spectra obtained from the analysis of this extract. Several peaks are visible in the spectrum and they are mainly from proteins [29,30]. The *E. coli* cell pellet was further extracted using an identical solvent. This extract was then analyzed by MALDI MS and the spectrum is shown in Figure 2.6B. Very weak signals were obtained, indicating that most of the water-soluble proteins were already removed by the first extraction. The pellet was then further extracted using a solution containing 0.2% SDS and 0.1% aqueous TFA. Figure 2.6C shows the MALDI spectrum from this extract. A strong peak corresponding to a protein with MW~7000 Da is visible, along with the doubly charged molecular ion peak. This indicates that SDS can solubilize some proteins that were not solubilized by the 0.1% aqueous TFA solvent.

It is interesting to note that the peak shown in Figure 2.6C is unusually broad, suggesting that it is from a mixture of proteins such as those with different modifications. Tryptic digestion of this sample revealed several peptide ion peaks in the low mass region, but peptide mass fingerprinting alone was not sufficient to identify the protein(s). Unfortunately, ESI MS/MS could not be performed on this sample because of the presence of SDS. MALDI MS/MS could potentially be used to sequence the tryptic peptides and therefore identify the protein(s), although MALDI instruments capable of

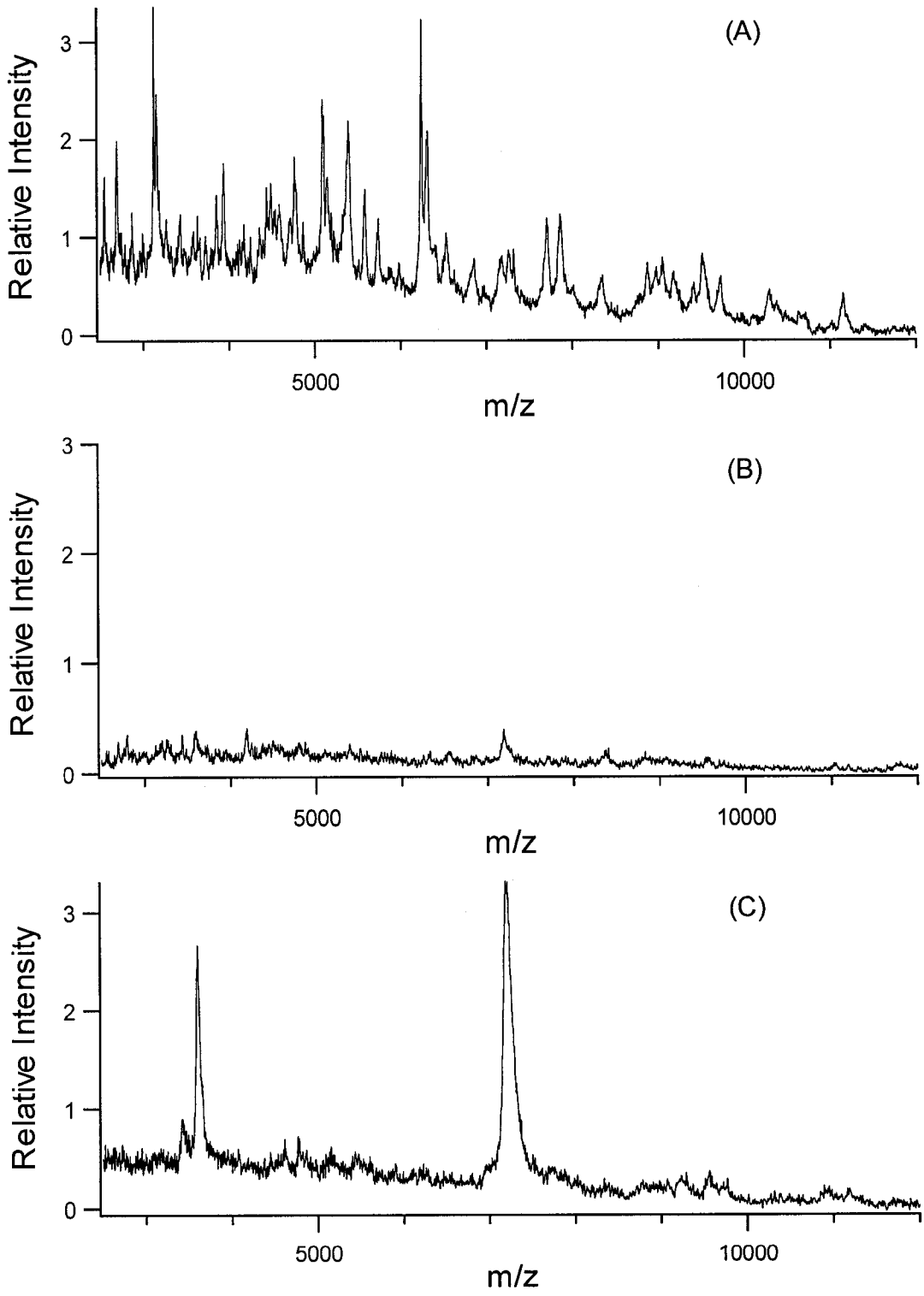



**Figure 2.6** MALDI spectra obtained from sequential extractions of an *E. coli* sample: (A) first extract with 0.1% aqueous TFA, (B) second extract with 0.1% aqueous TFA, and (C) third extract with 0.2% SDS in 0.1% TFA.

generating routine MS/MS spectra are not widely available at present. But it should be noted that a MALDI MS/MS based quadrupole/time-of-flight system is now commercially available.

## 2.4 Conclusions

We have demonstrated that a two-layer sample preparation method can be very useful in analyzing protein and peptide samples containing SDS by MALDI MS. This method involves first the formation of a thin layer of matrix crystals, followed by depositing a mixture of sample and matrix dissolved in a proper solvent and with an adjustable sample to matrix ratio. Several variables related to the sample preparation are shown to affect the intensity of the MALDI signal. With this method, an on-probe washing step should be performed, if the deposited spot can tolerate this procedure. Commonly used solvents (i.e., 50% acetonitrile/water or 40% methanol/water) for handling protein samples in MALDI can be utilized to prepare the second-layer matrix/analyte mixture. The matrix to SDS ratio is not critical, but the protein to SDS ratio can have a significant effect on the MALDI signal response. Thus dilution or concentration of the sample solution may be required when this method is applied to analyze samples with unknown protein concentration. The level of SDS should be kept at $\leq \sim 1\%$. With the two-layer method, there is no need of searching for hot spots in the process of acquiring MALDI spectra. If no analyte signals are observed in the first few shots, it usually indicates that the sample preparation conditions are not optimized and need to be adjusted.

The applications of the two-layer method for analyzing SDS-containing protein and peptide samples are illustrated. We demonstrate that hydrophobic protein samples in the presence of SDS can be readily detected by MALDI. For direct analysis of cell extract by MALDI, it is now possible to incorporate solvent extraction involving SDS to examine more hydrophobic proteins. Since the two-layer sample preparation method is tolerant towards a certain percentage of SDS in peptide samples, we can readily analyze protein digests containing SDS (data not shown). We envision that sample preparation methods, such as the two-layer method shown here, that are capable of handling impure samples including those containing SDS, will greatly benefit the MALDI MS and MS/MS approach in proteome identification.

## 2.5 Literature Cited


(1) Patterson, S. D.; Aebersold, R. *Electrophoresis*, **1995**, *16*, 1791-1814.

(2) Rabilloud, T. *Electrophoresis*, **1996**, *17*, 813-829.

(3) Beavis, R. C.; Chait, B, T. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6873-6877.

(4) Vorm, O.; Chait, B. T.; Roepstorff, P. *Proceedings of the 41st ASMS Conference on Mass Spectrometry and Allied Topics*, San Francisco, CA, May 31-June 4, **1993**; pp 621.

(5) Rosinke, B.; Strupat, K.; Hillenkamp, F.; Rosenbusch, J.; Dencher, N.; Kruger, U.; Galla, H. *J. Mass Spectrom.* **1995**, *30*, 1462-1468.

(6) Schnaible, V.; Michels, J.; Zeth, K.; Freigang, J.; Welte, W.; Bühler, S.; Glocker, M. O.; Przybylski, M. *Int. J. Mass Spectrom. Ion Proces.* **1997**, *169/170*, 165-177.

(7) Jeannot, M. A.; Jing, Z.; Li, L. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 512-520.

(8) Puchades, M.; Westman, A.; Blennow, K.; Davidsson, P. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 344-349.

(9) Galvani, M.; Hamdan, M. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 721-723.

(10) Amado, F. M.; Santana-Marques, M. G.; Ferrer-Correia, A. J.; Tomer, K. B. *Anal. Chem.* **1997**, *69*, 1102-1106.

(11) Breaux, G. A.; Green-Church, K. B.; France, A.; Limbach, P. A. *Anal. Chem.* **2000**, *72*, 1169-1174.

(12) Tummala, R.; Ballard, L. M.; Breaux, G. A.; Green-Church, K. B.; Limbach, P. A. *Proceedings of the 48th Conference of the American Society for Mass Spectrometry and Allied Topics*, Long Beach, CA, June 11-15, **2000**; pp481-483.

(13) Saunders, C. J.; Frankel, L. K.; Green-Church, K. B.; Tummala, R.; Bricker, T.; Limbach, P. A. *Proceedings of the 48th Conference of the American Society for Mass Spectrometry and Allied Topics*, Long Beach, CA, June 11-15, **2000**; pp 489-490.

(14) Jensen, C.; Haebel, S.; Andersen, S. O.; Roepstorff, P. *Int. J. Mass Spectrom. Ion Proces.* **1997**, *160*, 339-356.

(15) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.

(16) Weinberger, S.R.; Boernsen, K.O.; Finchy, J.W.; Robertson, V.; Musselman, B.D. *Proceedings of the 41st ASMS Conference on Mass Spectrometry and Allied Topics*; San Francisco, CA, May 31- June 4, **1993**, pp 775a-b.

(17) Xiang, F.; Beavis, R.C. *Rapid Commun. Mass Spectrom.* **1994**, *8*, 199-204.

(18) Xiang, F.; Beavis, R.C. *Org. Mass Spectrom.* **1993**, *28*, 1424-1429.

(19) Mock, K.K.; Sutton, C.W.; Cottrell, J.S. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 233-238.

(20) Bai, J.; Liu, Y.H.; Cain, T.C.; Lubman, D.M. *Anal. Chem.* **1994**, *66*, 3423-3430.

(21) Köchling, H.J.; Biemann, K. *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*; Atlanta, Georgia, May 21- 26, **1995**, pp 1225.

(22) Hensel, R.R.; King, R.; Owens, K.G. *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*; Atlanta, Georgia, May 21- 26, **1995**, pp 947.

(23) Vorm, O.; Roepstorff, P.; Mann, M. *Anal. Chem.* **1994**, *66*, 3281-3287.

(24) Edmondson, R.D.; Campo, K.K.; Russell, D.H. *Proceedings of the 43rd ASMS Conference on Mass Spectrometry and Allied Topics*; Atlanta, Georgia, May 21-26, **1995**, pp 1246.

(25) Li, L.; Golding, R.E.; Whittal, R.M. *J. Am. Chem. Soc.* **1996**, *118*, 11662-11663.

(26) Kussmann, M.; Nordhoff, E.; Rahbek-Nielsen, H.; Haebel, S.; Rossel-Larsen, M.; Jakobsen, L.; Gobom, J.; Mirgorodskays, E.; Kroll-Kristensen, A.; Palm, L.; Roepstorff, P. *J. Mass Spectrom.* **1997**, *32*, 593-601.

(27) Dai, Y.Q.; Whittal, R.M.; Li, L. *Anal. Chem.* **1996**, *68*, 2494-2500.

(28) Dai, Y.Q.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999**, *71*, 1087-1091.

(29) Wang, Z.; Russon, L.; Li, L.; Roser, D.C.; Long, S.R. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 456-464.

(30) Dai, Y.Q.; Li, L.; Roser, D.C.; Long, S.R. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 73-78.

(31) Juhasz, P.; Costello, C. E.; Biemann, K. *J. Am. Soc. Mass Spectrom.* **1993**, *4*, 399-409.

(32) Keller, B. O.; Li, L. manuscript in preparation.

(33) Santoni, V.; Molloy, M.; Rabilloud, T. *Electrophoresis* **2000**, *21*, 1054-1070.

(34) Barnidge, D. R.; Dratz, E. A.; Jesaitis, A. J.; Sunner, *J. Anal. Biochem.* **1999**, *269*,1-9.

# Chapter 3

# Ammonium Dodecyl Sulfate as an Alternative to Sodium Dodecyl Sulfate for Protein Sample Preparation with Improved Performance in MALDI Mass Spectrometry [a]

## 3.1 Introduction

Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) is widely used for protein analysis. Compared to other methods such as electrospray ionization, MALDI can tolerate a relatively high level of contaminants present in a sample [1]. The MALDI spectrum consists of peaks from mainly singly and doubly charged protein molecular ions. These attributes make the MALDI technique particularly useful for direct analysis of complex protein mixtures, such as those obtained from cell extracts or affinity purification. Samples obtained from these preparation methods may contain surfactants, which are commonly used reagents in protein sample workup. Surfactants help to solubilize hydrophobic proteins, reduce protein aggregation and protein adsorption on the container walls [2, 3]. Because of the possible adverse effect on MALDI performance caused by various types of surfactants, special care must be exercised in dealing with protein samples containing surfactants.

Sodium dodecyl sulfate (SDS) is a strong solubilization reagent that is commonly used for extracting proteins remaining in cells or tissues after they have been subjected to sequential extractions by non-ionic detergents [4-6]. SDS is also used as a denaturing

[a] A portion of this chapter is published as: N. Zhang, and L. Li, *"Ammonium Dodecyl Sulfate as an Alternative to Sodium Dodecyl Sulfate for Protein Sample Preparation with Improved Performance in MALDI Mass Spectrometry" Anal. Chem.*, **2002**, *74*, 1729-1736. The cell lysis was done in Dr. Shaw's lab.

reagent for SDS-polyacrylamide gel electrophoresis (PAGE). Early reports showed that SDS was detrimental to MALDI signals and should be avoided or removed from the samples prior to MS analysis [7-10]. Electroelution or electroblotting of proteins, solvent extraction, and protein precipitation are several techniques that have been developed to remove the SDS interference [11-15]. These techniques obviously add an undesirable step for MALDI analysis. The sample-cleaning step can result in protein loss, especially for hydrophobic or membrane proteins [16, 17]. An alternative approach is to develop new surfactants that have the similar properties as SDS, but will be MS-compatible. For example, an acid-labile surfactant was developed to replace SDS that was found to be compatible with MS analysis [18, 19]. However, this surfactant at present has only found limited use and its overall performance as a viable alternative to SDS in protein chemistry remains to be determined.

The effect of SDS on MALDI MS has been recently studied in more details. Recent studies have indicated that the negative effect by SDS on protein analysis was strongly dependent on sample preparation conditions [20-25]. For example, Roeptorff and coworkers reported that useful MALDI signals could be obtained for protein samples containing 0.2% SDS when the pH of the sample solution was kept under 2 [20]. They also indicated that SDS affected the mass resolution with unknown reason. Limbach and coworkers recently illustrated that SDS could assist in analyzing membrane proteins and improve ionization of hydrophobic peptides [21-23]. Amado and coworkers showed that the MALDI signal was improved if the SDS concentration was above 0.23% [24]. In our recent study, we presented a practical and robust technique based on a two-layer matrix/sample deposition method for the analysis of protein and peptide samples

containing SDS by MALDI MS [25]. The two-layer method involves the deposition of a mixture of sample and matrix on top of a thin layer of matrix crystals. We found that for SDS-containing samples, the intensity of the MALDI signals could be affected by the conditions of sample preparation, such as on-probe washing, choice of matrix, deposition method, solvent system and protein to SDS ratio. However, we found that under appropriate conditions, the two-layer method gave consistent MALDI signals for samples with levels of SDS up to ~1%.

While the two-layer method can readily produce MALDI signals from protein samples containing SDS, the protein peaks observed are generally broad and the peak centroids are shifted to higher masses. In this work, we address these issues and present a method using ammonium dodecyl sulfate as a viable alternative to SDS for protein sample preparation with much improved MALDI MS performance.

## 3.2 Experimental

### 3.2.1 Chemicals and Reagents

Bovine insulin, equine cytochrome *c*, equine myoglobin, bovine trypsinogen, bovine carbonic anhydrase II, bovine serum albumin (BSA), lactoferrin, α-cyano-4-hydroxycinnamic acid (HCCA), diammonium hydrogen citrate (DAHC), sulfuric acid and sodium dodecyl sulfate (SDS) were purchased from Sigma Aldrich Canada (Oakville, ON). Ammonium dodecyl sulfate (ADS) and dodecyl sulfate tris (hydroxymethyl) aminomethane salt (TDS) were purchased from Fluka, Canada (Oakville, ON). Tris (hydroxymethyl) aminomethane (Tris) base was purchased from Bio-Rad (Mississauga, ON). Dodecyl sulfuric acid (HDS) was generated in our

laboratory by exchanging $Na^+$ of SDS for $H^+$ cations using AG 50W-X8 ion-exchange resin purchased from Bio-Rad. Approximately 450 grams of resin was packed into a column. The hydrogen form of the resin was generated by washing the column with 300 mL of water, followed by 400 mL of 3 M HCl and flushed with water until the effluent from the column was neutral [26]. About 50 mL of 1% SDS (i.e., ~35 mM) was then loaded onto the column and the effluent was monitored for acidity. The first 20 mL of acidic effluent was discarded and the following 10 mL was collected for MALDI sample preparation. The molar concentration of HDS collected was approximately 35 mM. HCCA was purified by recrystallization from ethanol prior to use. Analytical grade acetone and methanol were purchased from Caledon Laboratories (Edmonton, AB). Water used in all experiments was from a Nanopure water system (Barnstead/Thermolyne).

### 3.2.2 Sample Preparation

Standard protein samples containing different types of dodecyl sulfate (DS) compound were prepared by mixing the protein solution with the dodecyl sulfate solution, followed by vortexing. The concentration of the DS compound (w/v) indicated in this work is for the protein samples prior to mixing with the MALDI matrix.

The two-layer method was employed for MALDI sample deposition [27, 28]. Briefly, 1 μL of a 12-mg/mL solution of HCCA in 80% acetone/methanol was deposited onto the target and was allowed to dry, forming a very thin layer of fine crystals. A 1-μL aliquot of the protein sample was mixed with HCCA second-layer solution (saturated in 40% methanol/ water) and a 0.5 μL aliquot of this mixture was deposited on top of the first layer and allowed to air dry. On-probe washing is always performed. The washing

step involves adding 2 μL of water to the sample spot and then blowing away the water with air. This procedure was performed twice on each sample spot.

### 3.2.3 SDS-PAGE

SDS-PAGE was carried out in a Bio-Rad mini-Protein III system using 12% polyacrylamide mini-gels. The SDS sample buffer contains 2% mercaptoethanol (v/v), 1% SDS, 12% glycerol, 50 mM Tris-HCl pH 6.8 and trace amount of bromophenol blue. Prior to electrophoresis, the protein sample was mixed in the sample buffer and heated at ~95°C for 4 min. Visualization of protein bands was done by using Bio-Rad's Biosafe Coomassie blue stain reagent or silver stain-kit.

### 3.2.4 Cell Lysis

Lymphoblastoid cell lines were grown in RPMI-1640 culture, buffered with 25 mM N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid (HEPES) and supplemented with 10% fatal bovine serum (FBS), 2 mM L-glutamine, 100 units penicillin and 100 μg/mL streptomycin. Cell concentration was kept in logarithmic growth phase at 2 ~ 5 × $10^5$ cells/mL medium. Approximately 3.5 × $10^8$ cells were spun down in a Beckman CS-6KR swing-bucket centrifuge, rinsed twice in ice-cold PBS, and lysed in 5 mL ice-cold lysing buffer (1% Triton X-100 RipA buffer). The composition of the RipA buffer is: 50 mM Tris-HCl, pH 7.4, 1% Triton X-100, 0.25% sodium deoxycholate, 150 mM NaCl, 1 mM EGTA plus Aprotinin, leupeptin and pepstatin at 10 μg/ml and 1 mM phenylmethyl sulfonyl fluoride (PMSF). The detergent insoluble fraction (DIF) is the detergent insoluble material after the cell lysates were extracted with RipA buffer. The pelleted DIF was finally rinsed in distilled water to reduce the salt concentration.

### 3.2.5 Instrumentation

MALDI experiments were carried out on a Bruker Reflex III time-of-flight mass spectrometer (Bremen/Leipzig, German) using the linear mode of operation. Ionization was performed with a 337-nm pulsed nitrogen laser. The laser power was adjusted slightly above the threshold of the desorption/ionization process and was kept constant for a given set of protein samples (i.e., one protein at a given concentration with different DS salts). All data were processed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

## 3.3 Results and Discussion

### 3.3.1 Effect of SDS on Mass Resolution

In our previous work, we described the two-layer sample preparation method and discussed how SDS affected the MALDI signal intensity [25]. With this method, good signal-to-noise ratios in MALDI spectra can be obtained from low micromolar concentrations of protein samples containing SDS up to 1%. However, the molecular ion peak is usually quite broad, compared to that obtained from a sample containing no SDS. Analysis of low mass proteins by this MALDI method reveals that peak broadening is mainly due to the formation of multiple sodium adduct ions with the proteins during the MALDI process. Figure 3.1 shows the MALDI mass spectrum obtained from a 2-μM bovine insulin solution containing 2% of SDS. Several partially resolved peaks are observed in the molecular ion region of the protein. They are from the protonated molecular ions, $[M+H]^+$ and sodium adduct ions, $[M-(n-1)H+nNa]^+$ where n = 1, 2, 3. In this case, the instrument resolving power is sufficiently high to resolve these adduct

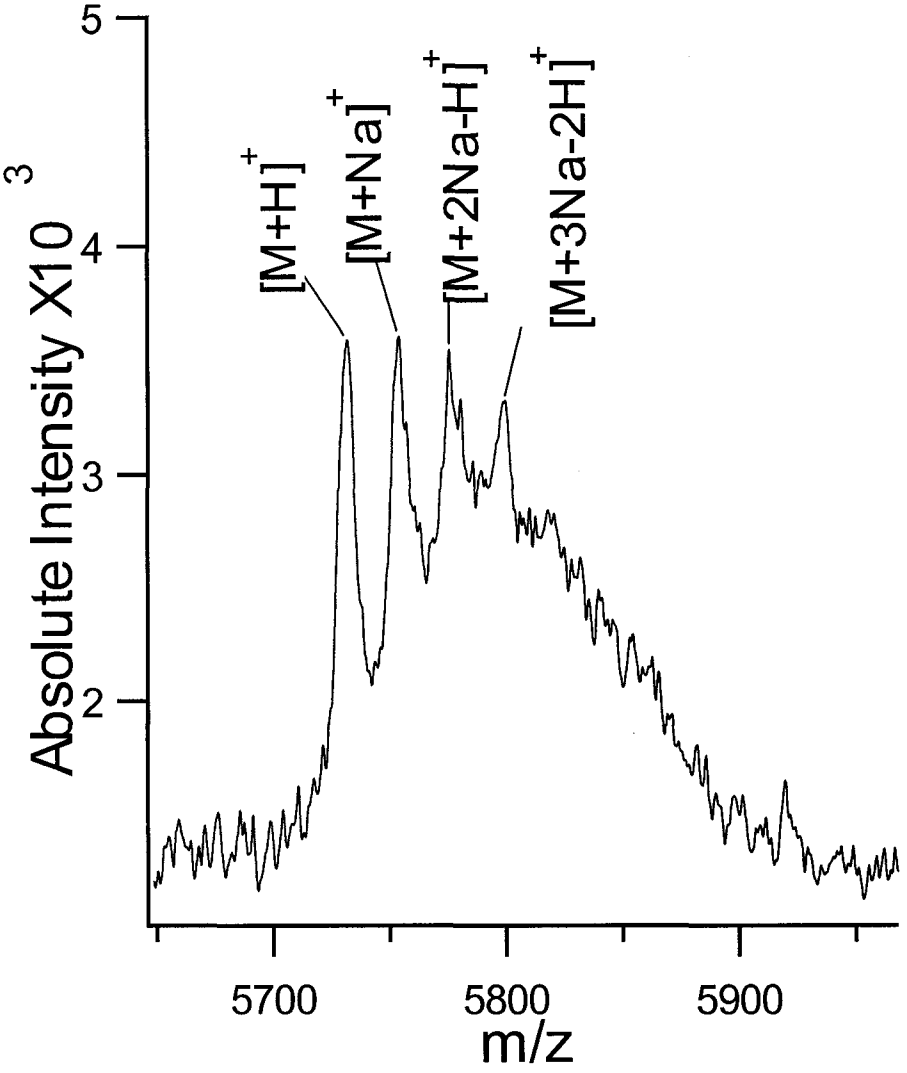



**Figure 3.1** MALDI spectrum of 2 μM bovine insulin with 2% SDS. Analyte to matrix, HCCA (saturated in 40% MeOH/$H_2O$), ratio is 1:1.

peaks, even though after n=3, the peaks are partially resolved because the multi-attachment of the potassium and sodium mixture. The protonated molecular ion can be readily identified, even though the peaks are not fully baseline-resolved. However, for higher mass proteins, sodium adduct peaks are not well resolved and can become the dominant peaks in the molecular ion region, resulting in poor mass resolution and a mass shift of the peak centroid to a higher mass. Since the extent of adduct formation cannot be predicted in the protein MALDI spectrum, the degree of mass shift is unknown. As a consequence, the protein mass cannot be accurately defined.

### 3.3.2 Effects of ADS, TDS and HDS on Resolution

SDS is commonly used to extract hydrophobic or membrane proteins from cells [5, 6]. Most biopolymer samples such as cell lysates contain a lot of salts that can be the potential source of sodium ions in subsequent MALDI analysis. However, prior to the use of SDS for protein extraction, the water-soluble proteins are commonly first extracted by using a buffer system that may or may not contain a mild surfactant or detergent such as Triton. The detergent insoluble component can be washed by using an aqueous solution to remove the salts. But, when SDS is used as a strong surfactant to solubilize any hydrophobic proteins remaining in the detergent insoluble component, a large quantity of sodium ions from SDS are also introduced to the sample. Thus the major source of sodium ions in the SDS-containing sample is from the SDS reagent. Since the surfactant property of SDS is mainly related to the dodecyl sulfate structure, other type of dodecyl sulfate salts still function like SDS for protein solubilization [29]. We have chosen two dodecyl sulfate salts, ammonium dodecyl sulfate (ADS) and tris (hydroxymethyl) aminomethane (Tris) dodecyl sulfate (TDS), for this study. Ammonium

and Tris were expected to interact less strongly with proteins in the gas phase, and to form fewer adducts compared to the sodium ion. In addition, these two compounds are commercially available and relatively inexpensive – an important consideration for the wide use of these reagents in protein biochemistry and in mass spectrometry laboratories. For comparison, we have also examined the effect of hydrogen dodecyl sulfate (HDS) in MALDI analysis.

Figure 3.2, from the left panel to the right, shows the MALDI spectra obtained from 2 μM of insulin, cytochrome *c*, myoglobin and trypsinogen, respectively, without the addition of any surfactant and with the addition of 0.5% various dodecyl sulfate surfactants. The molecular peaks for these four proteins are labeled in Figure 3.2A. In the MALDI spectra obtained from the samples containing 0.5% SDS (Figure 3.2B), the sodium adduct ion peaks are resolved from the protonated molecular ion peak for insulin and cytochrome c. For insulin the sodium adduct ion peaks are much less intense than the protonated molecular ion peak, whereas, in Figure 3.1, the sodium adduct ion peaks are comparable to the protonated one (2 μM insulin containing 2% SDS). It is generally the case that, as the concentration of SDS in a protein solution increases, the sodium adduct peaks in the MALDI spectrum become more intense. In the spectrum of cytochrome *c* with 0.5% SDS, the sodium adduct peaks become dominant. Even though the peaks are only partially resolved, the nominal molecular mass of cytochrome c still can be accurately determined. However, as the protein molecular mass further increases, the protonated and sodium adduct ion peaks are no longer well resolved, as illustrated in the spectra of myoglobin and trypsinogen containing 0.5% SDS. The peak centroids for these two proteins shift to higher masses than the molecular masses.

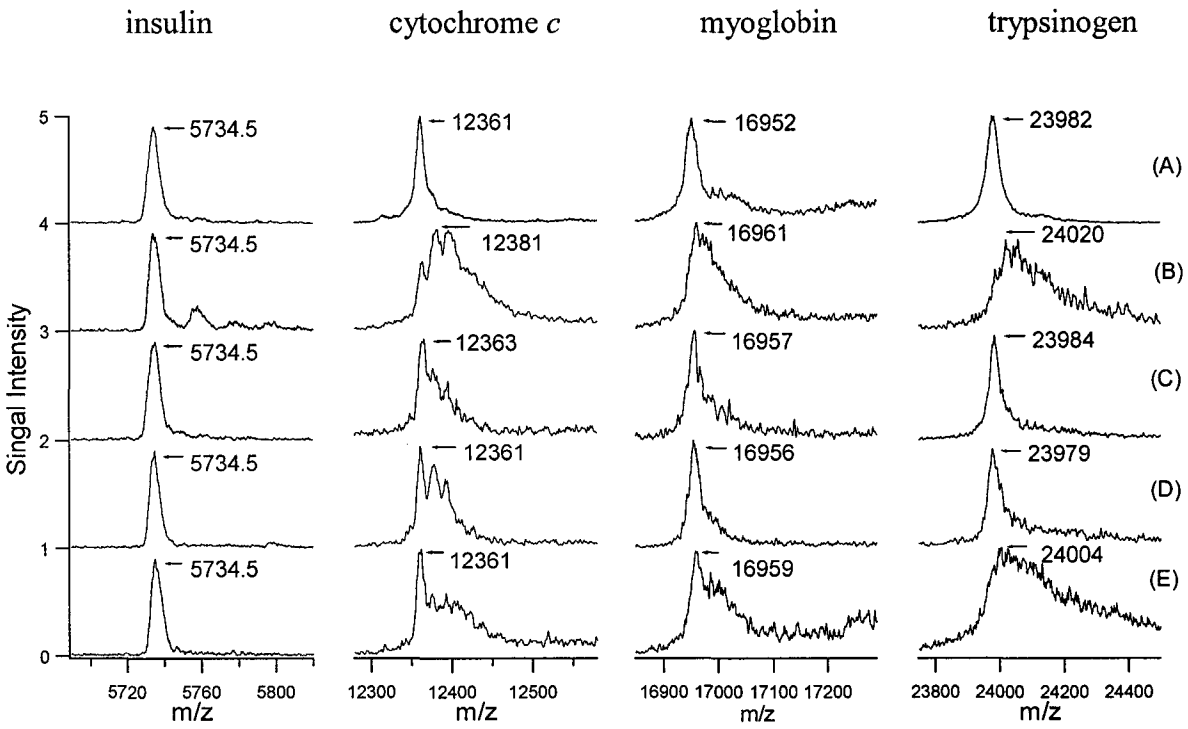



**Figure 3.2** MALDI spectra of 2 μM insulin, cytochrome c, myoglobin, and trypsinogen: (A) without surfactant, (B) with 0.5% SDS, (C) with 0.5% HDS, (D) with 0.5% ADS, and (E) with 0.5% TDS. Analyte to matrix ratio varies from 1:1 to 1:2.

It is worthwhile to comment on the differences in running protein samples containing common salts such as NaCl and those containing an equivalent amount of SDS by MALDI. In the case of protein samples containing NaCl, most salts will likely be excluded from the incorporation into the co-crystals of proteins and matrix molecules during the sample/matrix preparation process. A significant fraction of salts present in the sample/matrix layer can be washed away by the on-probe washing procedure in the two-layer sample preparation method. As a result, the protonated protein molecular ions are dominant in the MALDI spectrum of protein samples containing NaCl. However, in the case of protein samples containing SDS, dodecyl sulfate interacts strongly with the protein molecule. Three species, namely the protein, a protein-dodecyl sulfate complex, and dodecyl sulfate, can be incorporated into the matrix crystals. The presence of the latter two species in the matrix crystals would require counter ions that would mainly be the sodium ions in a SDS-containing protein sample. Thus, matrix/analyte co-crystallization and on-probe washing can only reduce a certain amount of sodium ions. Many remaining sodium ions residing in the co-crystals will be ionized in MALDI, which increases the possibility of forming adduct ions with the proteins in the gas phase.

The Ziptip sample cleaning technique is commonly used to reduce salt contaminants. We find it particularly effective for cleaning in-gel digests of proteins prior to MS analysis. However, we find that desalting the SDS-containing protein sample by using either ion exchange or C18 Ziptip does not result in any improvement in MALDI analysis. This is likely due to the difficulty of removing a large amount of sodium ions present in the SDS-protein sample with the Ziptip.

When SDS is replaced with HDS, better mass resolution is obtained. This can be seen from Figure 3.2C. Protonated peaks are the dominant signals in the molecular ion region for all four proteins. However, for cytochrome *c* and myoglobin, HDS suppresses the MALDI signals to some extent. The reason for signal reduction is likely related to the strong acidity of HDS that may affect the protein incorporation and co-crystallization with the HCCA molecules. For example, the addition of 9 mM sulfuric acid to the cytochrome *c* sample, which results in the same acidity as that of 0.5% HDS, gave the similar signal reduction effect as HDS did (spectra not shown). It should be noted that on-probe washing did improve signals for protein samples containing 0.5% dodecyl sulfate salts. However, washing must be limited for several seconds because a prolonged washing step would dissolve the matrix/analyte layer, resulting in poor signals. However, for samples containing sulfuric acid, a long washing period can be applied without dissolving the sample layer. In the case of cytochrome *c* containing sulfuric acid, if the sample spot was washed over 20 sec, the protein signal can be recovered to the same level as that obtained from a sample with no strong acid added. But, for myoglobin, the additional washing step did not result in signal recovery. These results suggest that strong acids such as HDS and sulfuric acid can affect the MALDI sensitivity with the two-layer sample preparation method and on-probe washing may not always be effective to reduce this interference.

Figure 3.2D shows the MALDI spectra of the four protein samples containing 0.5% ADS. The spectral quality is comparable to that obtained from the protein sample with no surfactant added. Each spectrum displays a strong protonated molecular ion peak, from which accurate protein molecular mass can be determined. In the case of

cytochrome *c*, ammonium adduct peaks are observed. The improvement of the MALDI spectral quality from the use of ADS over SDS is quite dramatic for the analysis of trypsinogen. With the presence of 0.5% SDS in the sample, the MALDI spectrum (Figure 3.2B) displays a broad peak which is clearly shifted to a higher mass. But, for the ADS containing sample, a sharp protonated molecular ion peak is obtained.

The improved MALDI results from the use of ADS in place of SDS are not totally surprising. The use of ammonium to suppress the sodium adduct ions in MALDI has been practiced by many researchers in the area of oligonucleotide analysis [26, 30]. In electrospray ionization MS, the addition of ammonium salts to the protein or peptide solution can also enhance the protonated peaks over the sodiated peaks. During the ionization and ion transport processes, the protein-ammonium adduct ions can be dissociated to form the protonated ions with the loss of ammonia. Note that merely adding a large amount of ammonium salts such as $NH_4Cl$ to a protein sample containing 0.5% SDS does not result in the reduction of sodium adduct ions. Sodium ions appear to interact with the protein more strongly than the ammonium ions do. This observation is similar to that found in oligonucleotide analysis by MALDI where desalting of the DNA sample prior to mixing it with the matrix and ammonium salts is critical to generate a high resolution mass spectrum [26, 30].

The results from the analysis of TDS-containing protein samples as shown in Figure 3.2E are somewhat surprising. In planning our experiments, we thought that the bulky structure of Tris would not readily form adduct ions with proteins. With the two-layer sample preparation method, MALDI signal intensity, mass resolution and mass measurement accuracy of proteins are unaffected by the presence of neutralized Tris base

at levels up to 1% (spectra not shown). In analyzing protein samples containing Tris base, an acid was added to the second-layer solution to neutralize the Tris base. Otherwise, the first crystal layer would be dissolved by the basic second-layer solution. In the cases of analyzing TDS-containing samples, for insulin and cytochrome *c*, the protonated peak is dominant, as shown in Figure 3.2E. But extensive adduct ions are observed for myoglobin and trypsinogen. Interestingly, the patterns of the adduct ion peaks are similar to those from the SDS-containing protein samples. It is possible that the adduct ion peaks shown in Figure 3.2E are also from the sodiated species. The source of the sodium ions may come from the sample, the TDS reagent, and containers used. During the matrix/analyte co-crystallization process, the bulky structure of Tris that should act as the counter ion for dodecyl sulfate and the protein-dodecyl sulfate complex has to compete with other smaller positive ions such as sodium to be incorporated into the co-crystals. In addition, the affinity of protein with sodium may be higher than the one of protein with Tris. Even though the amount of sodium ions in the solution is small compared to Tris, its readiness to be incorporated in the co-crystals can result in an accumulation of a large quantity of sodium ions in the crystals. Once the sodium ions are trapped inside the crystals, they are not readily washed away by the on-probe washing process.

Comparative analysis is also carried out for relatively high mass proteins. Figure 3.3 shows the MALDI mass spectra obtained from 5 μM carbonic anhydrase II (molecular mass 29025 Da), BSA (molecular mass 66430 Da), and lactoferrin (~ 82 kDa, a glycoprotein) with and without the dodecyl sulfate surfactants added. The spectrum shown in Figure 3.3A is from the protein samples with no surfactant added. With the

presence of 0.5% SDS in the sample (Figure 3.3B), the molecular ion peaks become very broad. With HDS (Figure 3.3C), the peaks are still broad with their centroid masses shifted to high masses. When TDS is present in the samples (Figure 3.3E), the resulting spectra are similar to those obtained from samples containing SDS. However, when ADS is used (Figure 3.3D), good resolution is obtained. For example, in the carbonic anhydrase spectrum, a reasonably sharp peak is observed at m/z 29042, which seems to correspond to the ammonium-protein adduct ion, $[M+NH_4]^+$. Peak tailing to a higher mass suggests that some other adduct ions likely from the attachment of two and three ammonium ions are also formed.

The observation of the ammonium-protein adduct ion as the dominant peak in the protein mass spectrum shown in Figure 3.3D can be explained by considering the readiness of adduct ion formation and the energetics of adduct ion dissociation. As the protein mass increases, the increase in protein chain length provides more possible binding sites for ammonium ion attachment. During the MALDI process, protein molecules are gaining internal energies through collisions, reactions and possibly direct absorption of the laser energy. However, the internal energy gained may not be sufficient to dissociate the ammonium adduct ions formed with a larger protein. As a consequence, protonated molecular ions are not extensively produced, as is the case for smaller proteins. A relatively sharp peak observed in the mass spectrum of the ADS-containing sample indicates that ions from the proteins with more than two ammonium ions attached are much less abundant than the adduct ion with one ammonium ion attached. Proteins with multiple ammonium ions attached likely dissociate into the $[M + NH_4]^+$ ions during the MALDI process. In the case of sodium adduct ions (see Figure 3.3B), since the

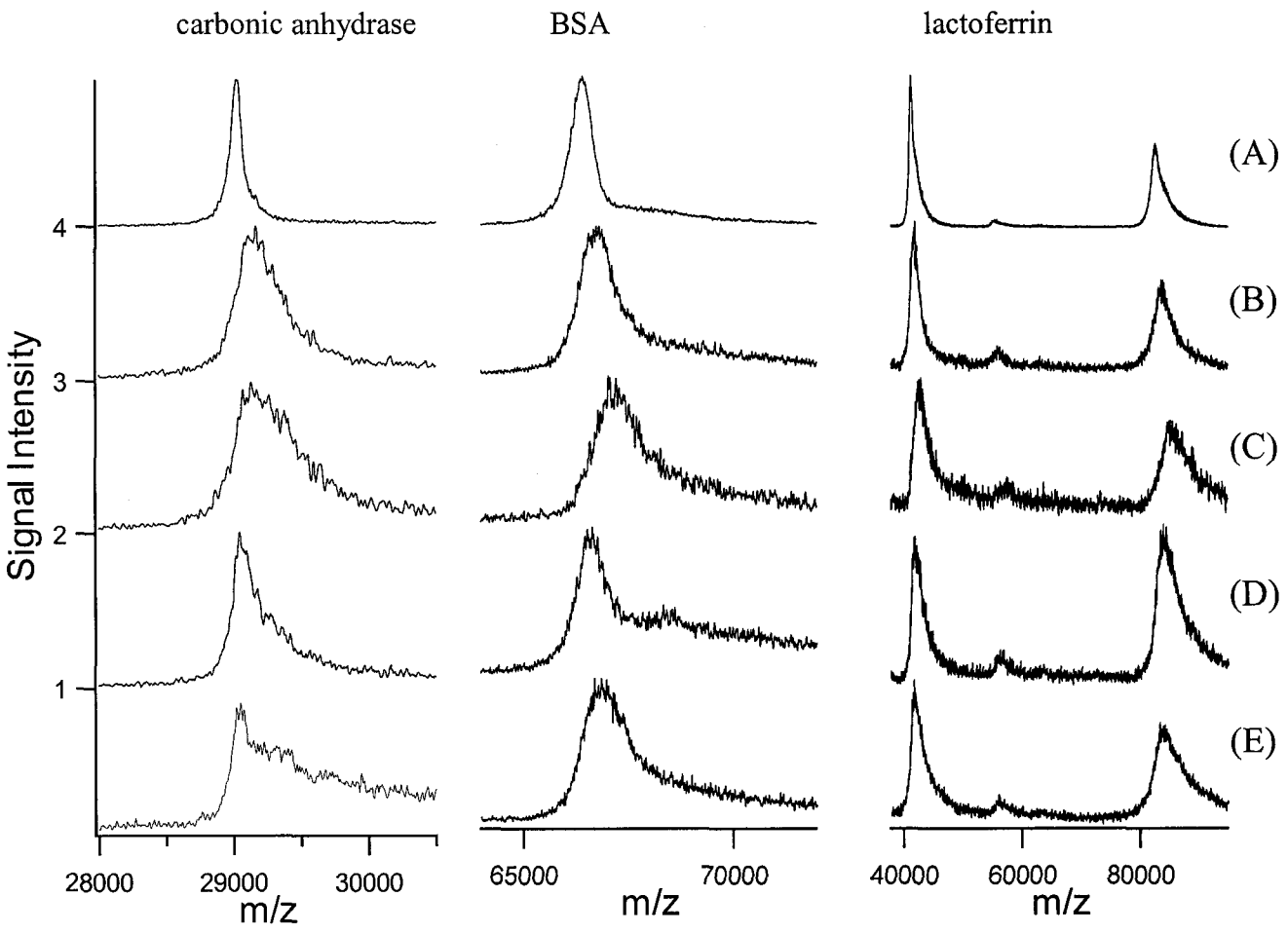



**Figure 3.3** MALDI spectra of 5 μM carbonic anhydrase, BSA and lactoferrin: (A) without surfactant, (B) with 0.5% SDS, (C) with 0.5% HDS, (D) with 0.5% ADS, and (E) with 0.5% TDS. Analyte to matrix ratio varies from 1:1 to 1:2.

binding between the sodium ion and the protein molecule is stronger, more energy is required to dissociate the multiple sodium adduct ions. An insufficient amount of internal energy placed on the protein-sodium adduct ions during MALDI would prevent them from dissociation, resulting in the observation of similar intensities of protein adduct ions with several sodium ions attached.

### 3.3.3 Negative Ion Mode Operation

The same protein samples were run in negative mode, to see if these four dodecyl sulfate compounds have similar effects on mass resolution as in positive mode. However, the sensitivity of negative ion detection of proteins is found to be much lower than that of positive ion detection. For example, for the 2 μM cytochrome c solution, the molecular ion peak is barely detectable in the MALDI spectrum. For the 2 μM insulin samples with and without the presence of any dodecyl sulfate surfactants (at the 0.5% level), similar mass resolution is observed (data not shown). However, for cytochrome c and other higher mass proteins, the signal intensity is weak even for a 5 μM protein solution. Considering that the positive ion detection gives much better detection sensitivity and that good sensitivity is critical for biomolecular analysis, we did not attempt to make any further investigation on the effects of surfactants on mass resolution for the large proteins detected under the negative ion mode operation.

### 3.3.4 Effect of ADS on Gel Electrophoresis

ADS is a surfactant that possesses similar properties as SDS [31]. Thus it can be used to replace SDS for cell extraction and protein dissolution. However, for gel electrophoresis, the use of the ammonium ion instead of the sodium ion may cause

interference. The effect of ADS on the performance of gel electrophoresis is investigated.

We have examined the use of ADS in place of all SDS for gel casting and running procedure, and also including the SDS in sample buffer (i.e., ADS-PAGE). Figure 3.4 shows a series of gel images of separation of protein mixture in 1% ADS obtained from different condition. The protein mixture contains cytochrome *c*, myoglobin, trypsinogen, carbonic anhydrase and BSA. Image A was obtained from a 12% ADS-PAGE, while image B was obtained from a normal 12% SDS-PAGE. Lane 1 and 2 show protein bands from protein mixture in 1% ADS separated in ADS-PAGE, by mixing sample with ADS-sample buffer. Lane 3 and 4 were obtained from the same sample mixed with normal SDS-sample buffer separated in SDS-PAGE. As a control, protein mixture in 1% SDS and mixed with normal SDS-sample buffer was also separated by the same gel in lane 5 and 6. The odd number lanes have loading of 0.5μg and even number lanes have 1μg. Since the ammonium ion has the potential to dissociate to form ammonia and a proton during the electrophoretic process, the gels are run under a constant current. The voltage needed for running the ADS gel is found to be slightly lower than that used for running the SDS gel. The separation pattern of the ADS-PAGE and SDS-PAGE is different, and it seems that one protein is missing in ADS-PAGE, probably because the degree of polymerization of ADS-PAGE is different from the SDS-PAGE. But this problem could be solved by changing the percentage of the gel. However, the major problem for running the ADS gel is associated with the staining process. When the ADS gel is stained by the Coomassie blue staining reagent, the background can be quite high, reducing the contrast between the protein bands and the gel background. Thus, the

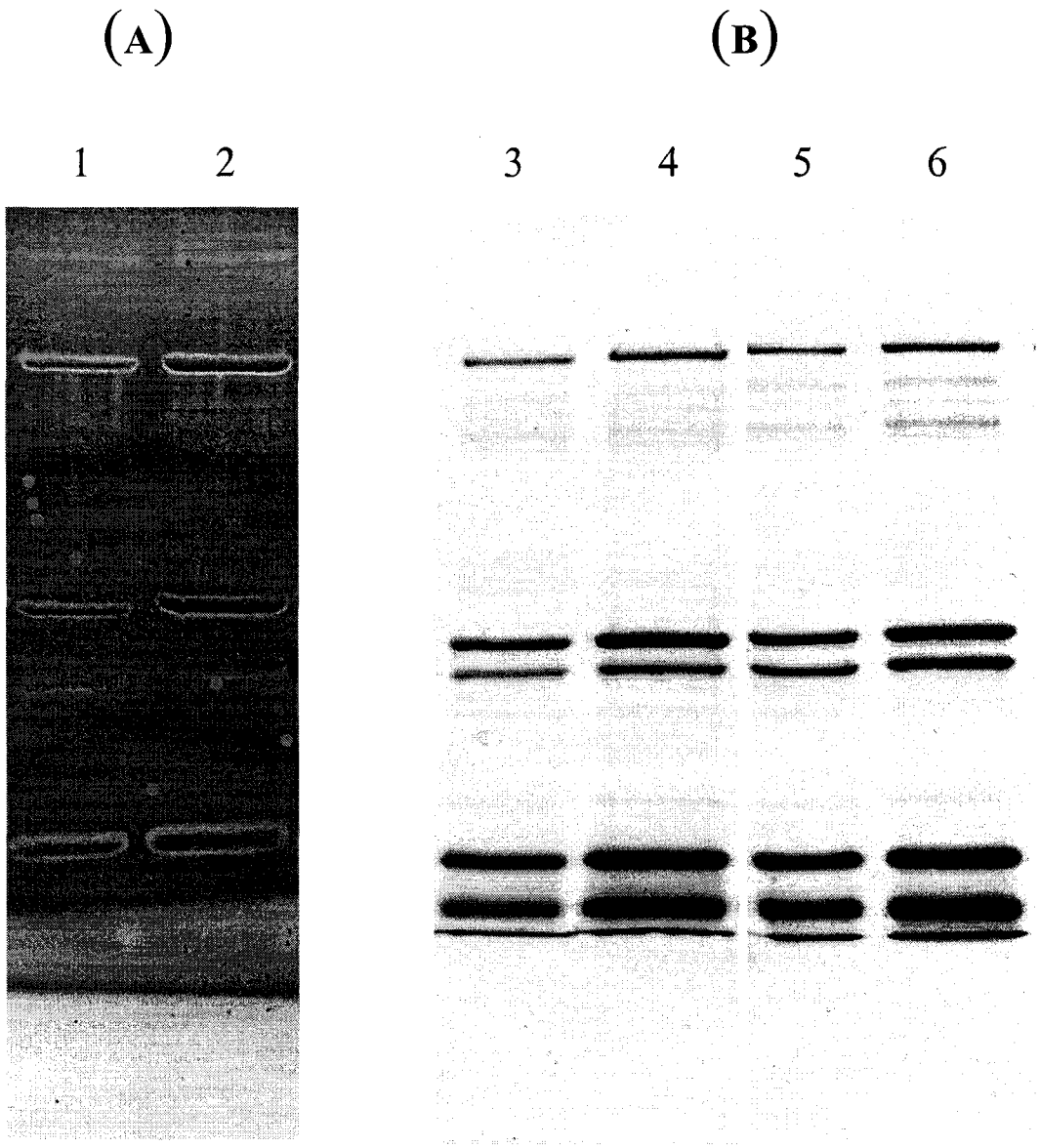



**Figure 3.4** Gel image of cytochrome *c*, myoglobin, trypsinogen, carbonic anhydrase and BSA on 12% (A) ADS-PAGE and (B) SDS-PAGE. Lane arrangement: proteins in (1-4) 1% ADS and (5,6) 1% SDS. Gel loading: (1,3,5) 0.5μg and (2,4,6) 1μg for each protein.

protein display sensitivity with Coomassie stain is lower in running ADS-PAGE. High background or reduced sensitivity is also observed to be true for silver stained ADS gels. Therefore, in terms of detection sensitivity, SDS-PAGE is preferred to ADS-PAGE.

From the image B in Figure 3.4 which shows the separation of the proteins solubilized by ADS in SDS-PAGE, we can see that no difference was observed among the bands displayed from proteins dissolved in ADS or SDS (lane 3-6). This indicated that proteins solubilized by ADS could be separated by SDS-PAGE.

### 3.3.5 Application to Hydrophobic Protein Analysis

The detergent insoluble fraction (DIF) pellet from lymphoblastoid cells or Raji/CD9 B-cells was dissolved in 1% ADS or 1% SDS. This pellet should mainly consist of very hydrophobic membrane proteins. Without the addition of ADS or SDS, no MALDI signals could be obtained by simply mixing the DIF sample with the HCCA second-layer solution in the two-layer sample preparation. Figure 3.5 shows the MALDI spectra of DIF obtained with (A) SDS or (B) ADS. These spectra can be readily reproduced. Thus the spectral differences observed in Panels A and B of Figure 3.5 reflect the different effects of SDS and ADS on MALDI signals. From the spectra shown in Figure 3.5, we can see that low mass proteins with masses up to about 20 kDa are detected in both cases. However, the spectrum obtained from ADS (Figure 3.5B) gives more peak information because the peaks are better resolved. For example, the broad peak labeled at m/z 7879 in Figure 3.5A actually consists of two major peaks as shown in Figure 3.5B at m/z 7816 and 7902. The peak shown in Figure 3.5B labeled as 12799 is not observed in Figure 3.5A. In addition, with ADS in the sample, a protein with molecular mass about 33 kDa was detected while this protein was barely seen in the case

of SDS (Figure 3.5A). As a control experiment, a portion of the DIF sample was dissolved in 1% SDS, mixed with gel loading sample buffer and separated on a 17% SDS-PAGE gel (Figure 3.6). The spectrum obtained from ADS is in agreement with the gel control experiment (Figure 3.6), which indicates that there is a large amount of protein in the low mass range and another dark band in the mass range 30-35 kDa.

It should be noted that both spectra in Figure 3.5 were obtained by adding a small amount [about 10% (v/v)] of 0.2 M diammonium hydrogen citrate (DAHC) to the mixture of analyte and matrix (HCCA saturated in 40 % MeOH/$H_2O$) [26, 30, 32]. DAHC is a reagent that is sometimes used to assist analyte ionization by enhancing protonated molecular ion signals and suppressing the sodium adduct ions in DNA and protein samples. In the analysis of DIF, without the addition of DAHC, the signals were low and only some of the low mass proteins in the mass range up to 12 kDa could be obtained.

Although the gel image shown in Figure 3.6 detects a number of high mass proteins, many of them are not detected in MALDI. This observation is similar to that obtained in analyzing water-soluble proteins [33]. It appears that for both hydrophilic and hydrophobic proteins the low mass proteins are more readily ionized and the high mass protein signals are suppressed. However, MALDI analysis of very hydrophobic proteins is not as sensitive as for less hydrophobic proteins. For the analysis of DIF, which consists mainly of very hydrophobic proteins, the laser power required to produce protein signals was higher than that used for less hydrophobic proteins. The ionization efficiency for a hydrophobic protein via protonation is expected to be low because of the lack of many protonation sites.

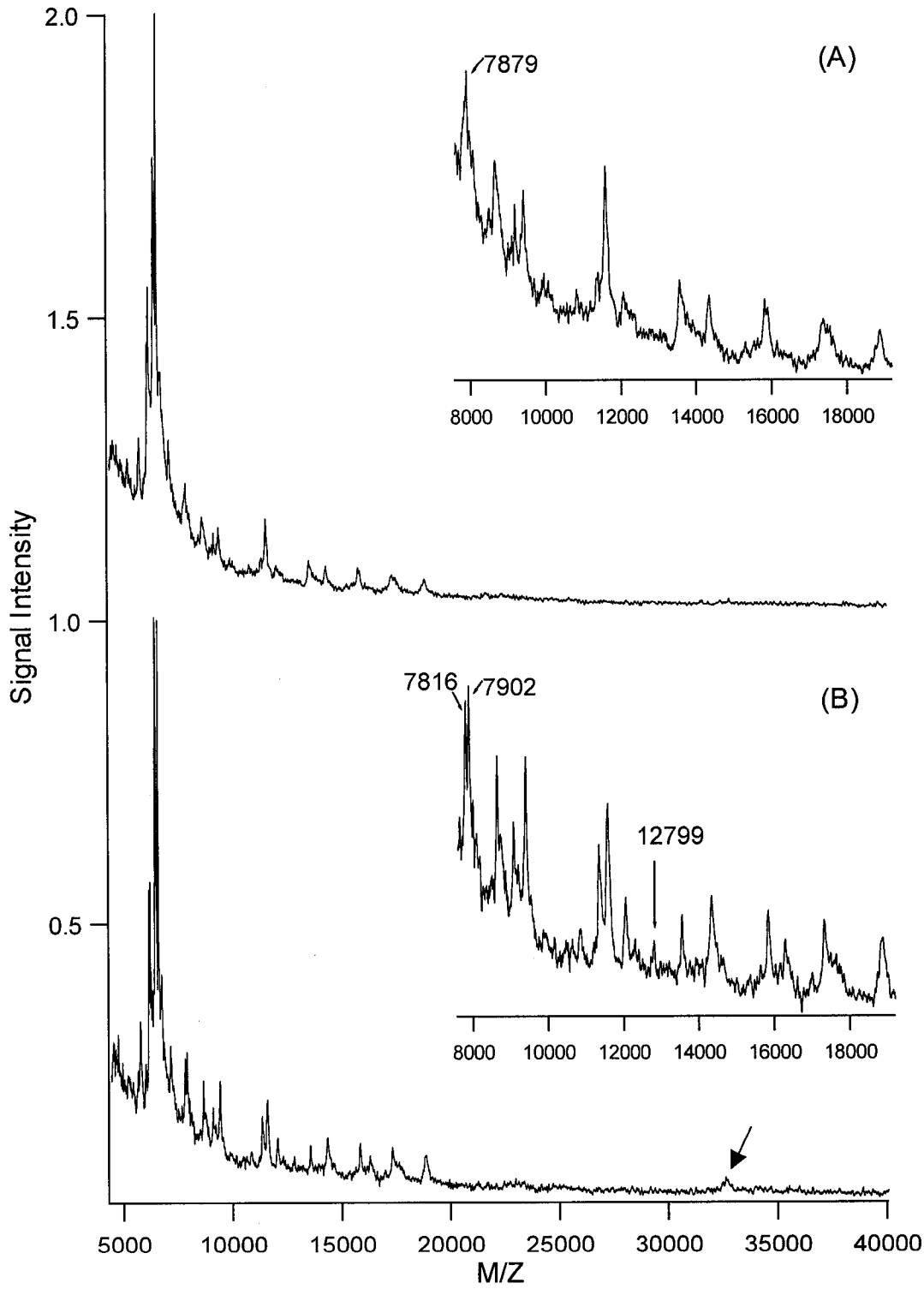



**Figure 3.5** MALDI spectra of the CD9 cell lysate Detergent Insoluble Fraction pellet. The pellet was mixed with (A) 1% SDS and (B) 1% ADS.

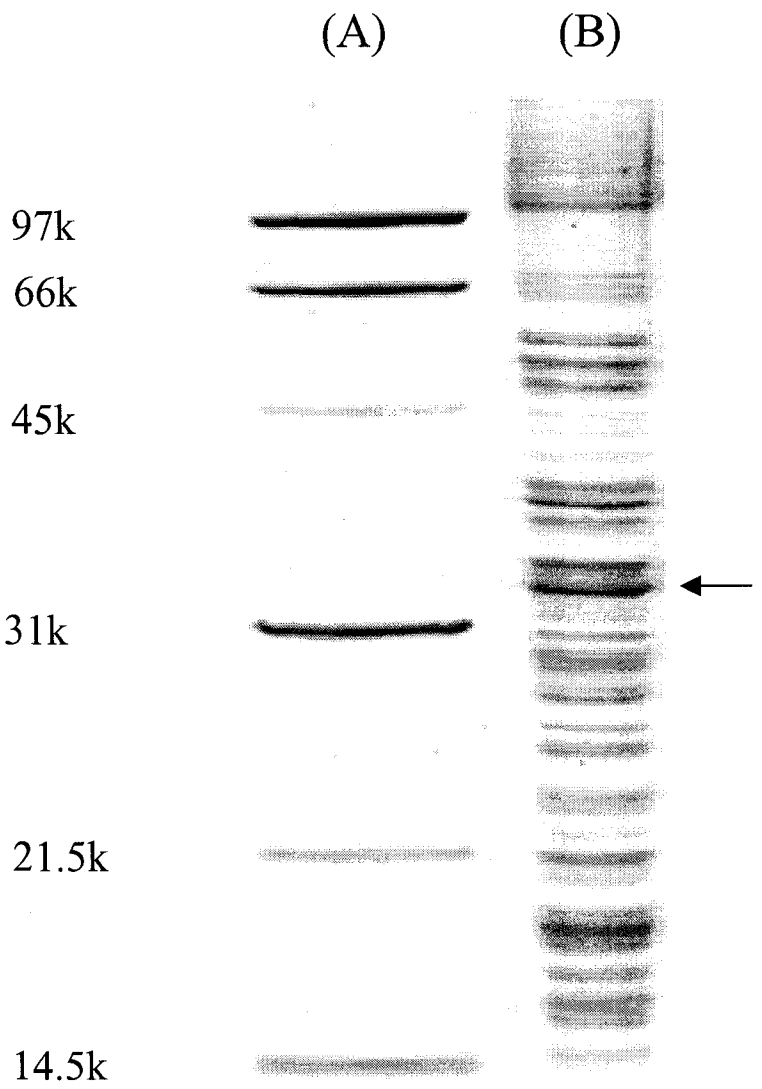



**Figure 3.6** Gel image of the CD9 cell lysate Detergent Insoluble Fraction pellet. Lane A is from molecular weight markers and Lane B is from the DIF pellet.

It is clear that the MALDI technique is compatible with the strong dodecyl sulfate surfactant, which should provide a new tool to tackle hydrophobic proteome analysis problem. For example, we are currently in the process of developing a protein digestion protocol to generate peptides directly from the DIF pellet of the lymphoblastoid cells. MALDI MS/MS will be used to generate peptide sequence information to identify the proteins present in the mixture. With the knowledge of accurate molecular masses obtained by MALDI-TOF, we are hoping to determine the origins of these low-mass proteins. These proteins could be the fragments of large membrane proteins or small gene products with or without post-translational modifications. Our future work will also be directed to the improvement of ionization efficiency of MALDI for the analysis of very hydrophobic proteins that can only be soluble in dodecyl sulfate surfactant. To this end, the use of a transition metal ion such as silver salt will be attempted to examine whether cationization can be a better means of ionization for the very hydrophobic proteins, as one would expect from the polymer MALDI work. In that case, the analysis of less polar species such as polybutadiene is usually done by using cationization with a transition metal ion [34].

## 3.4 Conclusions

While protein samples containing sodium dodecyl sulfate (SDS) up to ~1% can be analyzed by the two-layer sample/matrix preparation method, the MALDI signal quality was found to be degraded because of the formation of protein sodium adduct ions in the MALDI process. If the instrument resolving power is inadequate to resolve these adducts, peak broadening is observed and the broaden peak centroid is shifted to higher

mass. Ammonium dodecyl sulfate (ADS), hydrogen dodecyl sulfate (HDS), and tris (hydroxymethyl) aminomethane dodecyl sulfate (TDS), were explored as alternatives to SDS. We found that the ADS was least detrimental to MALDI performance. The presence of ADS in the protein samples did not significantly degrade the signal resolution and accuracy, as was the case with SDS. With ADS in protein samples, the protonated peak was dominant in MALDI analysis from moderate protein mass up to ~25kDa. As the protein molecular weight increases, the ammonium adducts become dominant in the MALDI spectrum; but ADS still yields the best results of the four surfactants. The performance of ADS-solubilized proteins for gel electrophoresis was also examined. We demonstrated that ADS solubilized cell extracts could be separated by SDS-polyacrylamide gel electrophoresis by mixing the samples with normal SDS buffer. The improved MALDI performance of an ADS-solubilized detergent insoluble fraction from Raji/CD9 B-cell extract was also demonstrated in this work. It can be concluded that ADS is a viable alternative to SDS in protein extraction and purification and should be used in MALDI analysis of very hydrophobic proteins.

## 3.5 Literature Cited


(1) Patterson, S. D.; Aebersold, R. *Electrophoresis* **1995,** *16*, 1791-1814.

(2) Rabilloud, T. *Electrophoresis* **1996,** *17*, 813-829.

(3) Beavis, R. C.; Chait, B. T. *Proc. Natl. Acad. Sci. USA* **1990,** *87*, 6873-6877.

(4) Fountoulakis, M.; Takacs, B. *Electrophoresis* **2001,** *22*, 1593-1602.

(5) Cordwell, S. J.; Nouwens, A. S.; Verrills, N. M.; Basseal, D. J.; Walsh, B. J. *Electrophoresis* **2000,** *21*, 1094-1103.

(6) Molloy, M. P.; Herbert, B. R.; Walsh, B., J.; Tyler, M. I.; Traini, M.; Sanchez, J.-C.; Hochstrasser, D. F.; Willams, K. L.; Gooley, A. A. *Electrophoresis* **1998,** *19*, 837-844.

(7) Jeannot, M. A.; Zheng, J.; Li, L. *J. Am. Soc. Mass Spectrom.* **1999,** *10*, 512-520.

(8) Rosinke, B.; Strupat, K.; Hillenkamp, F.; Rosenbusch, J.; Dencher, N.; Krueger, U.; Galla, H.-J. *J. Mass Spectrom.* **1995,** *30*, 1462-1468.

(9) Schnaible, V.; Michels, J.; Zeth, K.; Freigang, J.; Welte, W.; Buehler, S.; Glocker, M. O.; Przybylski, M. *Int. J. Mass Spectrom. Ion Process* **1997,** *169/170*, 165-177.

(10) Puchades, M.; Westman, A.; Blennow, K.; Davidsson, P. *Rapid Commun. Mass Spectrom.* **1999,** *13*, 344-349.

(11) Schaegger, H. In *A practical guide to membrane protein purification*; von Jagow, G., Schaegger, H., Eds.; Academic Press: San Diego, CA, 1994, pp 31-34.

(12) Immler, D.; Mehl, E.; Lottspeich, F.; Meyer, H. E. In *Microcharacterization of proteins*; Kellner, R., Lottspeich, F., Meyer, H. E., Eds.; Wiley-VCH: New York, 1999, pp 87-93.

(13) Feick, R. G.; Shiozawa, J. A. *Anal. Biochem.* **1990,** *187*, 205-211.

(14) Schuhmacher, M.; Glocker, M. O.; Wunderlin, M.; Przybylski, M. *Electrophoresis* **1996,** *17*, 848-854.

(15) Vestling, M. M.; Fenselau, C. *Anal. Chem.* **1994,** *66*, 471-177.

(16) Galvani, M.; Hamdan, M. *Rapid Commun. Mass Spectrom.* **2000,** *14*, 721-723.

(17) Eckerskorn, C.; Strupat, K.; Schleuder, D.; Dennis Hochstrasser, D.; Sanchez, J.-C.; Lottspeich, F.; Hillenkamp, F. *Anal. Chem.* **1997,** *69*, 2888-2892.

(18) Compton, B. J.; Lee, J.-J.; Brown, E. K.; Herbert, R. G.; Ding, J.; Livingstone, J.; Bouvier, E. S. P. *Proceedings of the 47th Conference of the American Society for Mass Spectrometry and Allied Topics*, Dallas, TX, June 13-17, 1999; p1012-1013.

(19) Ross, A. R. S.; Smith, D.; Langridge, J. I.; Gaskell, S. J. *Proceedings of the 48th Conference of the American Society for Mass Spectrometry and Allied Topics*, Long Beach, CA, June11-15, 2000; p450-451.

(20) Jensen, C.; Haebel, S.; Anderson, S. O.; Roepstorff, P. *Int. J. Mass Spectrom. Ion Processes* **1997,** *160*, 339-356.

(21) Tummala, R.; Ballard, L. M.; Breaux, G. A.; Green-Church, K. B.; Limbach, P. A. *Proceedings of the 48th Conference of the American Society for Mass Spectrometry and Allied Topics*, Long Beach, CA, June 11-15, 2000; p481-483.

(22) Saundres, C. J.; Frankel, L. K.; Green-Church, K. B.; Tummala, R.; Bricker, T.; Limbach, P. A. *Proceedings of the 48th Conference of the American Society for Mass Spectrometry and Allied Topics*, Long Beach, CA, June 11-15, 2000; p489-490.

(23) Breaux, G. A.; Green-Church, K. B.; France, A.; Limbach, P. A. *Anal. Chem.* **2000,** *72*, 1169-1174.

(24) Amado, F. M. L.; Santana-Marques, M. G.; Ferrer-Correia, A. J.; Tomer, K. B. *Anal. Chem.* **1997,** *69*, 1102-1106.

(25) Zhang, N.; Doucette, A.; Li, L. *Anal. Chem.* **2001,** *73*, 2968-2975.

(26) Nordhoff, E.; Cramer, R.; Karas, M.; Hillenkamp, F.; Kirpekar, F.; Kristiansen, K.; Roepstorff, P. *Nucleic Acids Reserch* **1993,** *21*, 3347-3357.

(27) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1996,** *68*, 2494-2500.

(28) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999,** *71*, 1087-1091.

(29) Neugebauer, J. M. In *Guide to Protein Purification*; Deutscher, M. P., Ed.; Academic Press: San Diego, CA, 1990, pp 239-253.

(30) Pieles, U.; Zuercher, W.; Schaer, M.; Moser, H. E. *Nucleic Acids Reserch* **1993,** *21*, 3191-3196.

(31) Mukerjee, P.; Mysels, K. J. *Critical Micelle Concentration of Aqueous Surfactant Systems*; U. S. Department of Commerce: Washington, D. C., 1971.

(32) Tang, K.; Fu, D.; Koetter, S.; Cotter, R. J.; Cantor, C. R.; Koester, H. *Nucleic Acids Reserch* **1995,** *23*, 3126-3131.

(33) Zheng, J.; Wang, Z.; Li, L. *Proceedings of the 49$^{th}$ Conference of the American Society for Mass Spectrometry and Allied Topics*, Chicage, IL, 2001.

(34) Yalcin, T.; Schriemer, D. C.; Li, L. *J. Am. Soc. Mass Spectrom.* **1997,** *8*, 1220-1229.

# Chapter 4

# Effects of Surfactants on Protein Digestion and Peptide Mass Spectrometric Analysis

## 4.1 Introduction

Surfactants play a significant role in protein chemistry with applications from protein solubilization and stabilization to protein dis-aggregation, and protein denaturation [1-3]. Surfactants are usually employed in gel and capillary electrophoresis for protein separation and purification. In sample preparation steps, membrane proteins and membrane-bound proteins, in particular, require the presence of surfactants for solubilization.

MALDI provides a fast and accurate means of protein characterization and has become a much-valued tool for proteomics studies. However, earlier studies on proteins showed that surfactants, such as SDS, Triton-X or CHAPS, interfered with the mass spectrometric analysis, e.g., deteriorating ion formation during measurement of intact protein and reducing the resolution of mass measurement. [4-7]. Because of these interferences, surfactants should be avoided whenever possible and should be removed prior to mass spectrometric analysis [7-11]. However, it is also believed that biochemical problems cannot be totally solved without the involvement of surfactants, even in the mass spectrometric step.

More recent studies showed that surfactants might hold some benefit for hydrophobic protein and peptide analysis. Some non-ionic surfactants such as n-octyl glucoside (OG) have been employed for mass spectrometric analysis without severe

interference [6, 12]. However, non-ionic surfactants usually have mild solubility power, only solubilizing a small portion of proteins. Strong surfactants, e.g. SDS, must be used for protein solubilization in most of cases. Limbach and coworkers successfully used SDS to eliminate the discrimination on ionization of the hydrophobic peptides in a mixture containing both hydrophilic and hydrophobic peptide [13]. Furthermore, an alternative to SDS, an acid-labile surfactant (ALS) used in gel electrophoresis has been reported by Waters Corp. [14, 15], though this surfactant's performance needs to be examined further.

Fast protein identification accomplished by peptide mass mapping usually involves protein digestion and measurement of the resultant peptides by a mass spectrometer, usually MALDI MS. Positive identification requires measurement of a large number of peptides and high sequence coverage for a confident match to the sequence information in protein or genome databases [16]. In other words, an effective enzymatic digestion is essential for obtaining a definitive database match. In addition, a high sequence coverage is important for mapping the posttranslational modifications of proteins. Because surfactants are often involved in protein solubilization, it is necessary to understand how the surfactants affect the enzyme activity and subsequent MS analysis.

In this chapter, frequently used surfactants for biological membrane protein sample preparation are examined to study their effects on enzyme activity, protein sequence coverage, and their compatibility to MALDI MS analysis. The goal of this work is to explore the possibility of using a suitable surfactant for future membrane protein identification.

## 4.2 Experimental

### 4.2.1 Chemicals and Reagents

Triton X100, n-octyl glucoside (OG), sodium dodecyl sulfate (SDS), 3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate (CHAPS), dithiothreitol (DTT), α-Cyano-4-hydroxycinnamic acid (HCCA), iodoacetamide, trifluoroacetic acid (TFA), bovine trypsin, bovine serum album (BSA), Bacteriorhodopsin (BRO), and other protein standards were from Aldrich-Sigma Canada (Markham, ON, Canada). $C_{18}$ and HPL ZipTip were purchased from Millipore Corporation (Bedford, MA, USA). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA, USA). HCCA was recrystallized from ethanol (95%) before use.

### 4.2.2 Protein Standards In-solution Digest

Samples were prepared in 0%, 0.1% or 1% surfactant solutions (v/v for Triton X 100, and w/v for CHAPS, OG and SDS). Fourteen-microliter aliquots of the 1 μg/μL protein solutions were mixed with 1.5 μL of 1.0 mM $NH_4HCO_3$ and 1.4 μL of 1 μg/μL trypsin in siliconized vials. The samples were incubated at 37°C for 2 h. Following incubation, a 1-μL portion of the digested sample was mixed with 9 μL of saturated HCCA solution for MALDI detection. A control experiment was also performed by digesting protein samples prepared in water, and then adding surfactant to the digests to produce samples containing 0.1% or 1% surfactant as indicated.

### 4.2.3 Bacteriorhodopsin (BRO) Tryptic Digest

About 190 μg of BRO was weighted and added to 3.8 μL of 5% SDS (making 50 μg/μL). As well, about 170 μg BRO was added to 3.4 μL of water (also 50 μg/μL).

Samples were vortexed for 20 minutes, and heated to 95°C for 5 min, then cooled to room temperature. Tryptic digest were performed in 100 mM $NH_4HCO_3$ buffer, with a trypsin to protein ratio of 1 to 10 (by weight), and a final concentration of 2 mM $CaCl_2$ was added to the protein solution to prevent trypsin autolysis. The final protein concentration was diluted to 1 μg/μL, with SDS consequently diluted to 0.1%. The digestion was performed overnight at 37°C. To complete the digestion, a ten-microgram aliquot of trypsin was added to each sample, and the samples were left in the incubator for another two hours at 37°C.

### 4.2.4 Sample Preparation

The two-layer method was employed for MALDI sample deposition [18, 19]. Briefly, 1 μL of a 12-mg/mL solution of HCCA in 80% acetone/methanol was deposited onto the target and allowed to dry, forming a very thin layer of fine crystals. A 1-μL aliquot of the protein sample was mixed with 10 μL HCCA second-layer solution (saturated in 40% methanol/water) and a 0.5 μL aliquot of this mixture was deposited on top of the first layer and allowed to air dry. On-probe washing is always performed. The washing step involves adding 2 μL of water to the sample spot and then blowing away the water with air. This procedure was performed twice on each sample spot.

### 4.2.5 Instrumentation

MALDI MS experiments were carried out on a Bruker Reflex III time-of-flight mass spectrometer (Bremen/Leipzig, German) using the reflectron mode of operation. Ionization was performed with a 337-nm pulsed nitrogen laser. All data were processed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

## 4.3 Results and Discussion

### 4.3.1 Effects of Mild Surfactants on Digestion and MS

In biological sample preparation, surfactants are commonly employed for protein extraction and solubilization of relatively hydrophobic proteins. For example, Triton X100 or n-octyl glucoside (OG) is usually employed to fractionate the membrane proteins from the cytosolic plasma [20-22], and CHAPS is commonly used for increasing protein solubilization in the isoelectric focusing (IEF) experiment. SDS is the most commonly used surfactant that is almost the perfect match for membrane protein and hydrophobic protein treatment. Table 4.1 lists the structure and critical micelle concentration (CMC) of the surfactants used in this study.

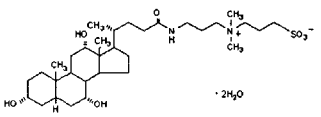
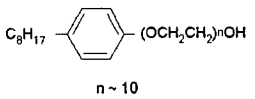
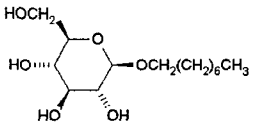
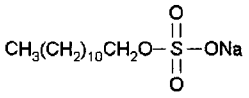
The first surfactant examined was CHAPS. Figure 4.1 shows the MALDI MS spectra of 0.5 μg/μl BSA digested without and with CHAPS. The peptide masses from a standard tryptic BSA digestion were labelled in Figure 4.1A. Examining all the four spectra (CHAPS concentrations are 0.1%, 1% and 2%), the typical BSA tryptic peptides 1479.85 and 1567.77 Da can be observed in the spectra even with 2% CHAPS present during the digestion process. Diluting the peptide mixture by 10 times, peptide signals were higher than the ones showed in Figure 4.1B, which was obtained in 0.1% CHAPS. This indicates that tryptic digestion at a typical protein concentration used in proteomics applications is still possible in the presence of CHAPS up to 2%.

However, the peptide peak intensities in the MALDI spectra are decreased with an increase in CHAPS concentration. In order to confirm that the intensity decrease is from an unsuccessful digestion or mass spectrometric response alone, a control experiment adding 0.1% CHAPS after the digestion process was performed and the

corresponding MALDI MS spectrum is shown as Figure 4.2B, while Figure 4.2A is from a standard BSA digestion. The significant difference in these two spectra is that a strong peak was observed at m/z 1229.76 Da in Figure 4.2B corresponding to the dimer of the surfactant CHAPS molecular peak (CHAPS mono-isotopic molecular weight is 614.4 Da). Comparing these two spectra, most of the tryptic peptide peaks observed in Figure 4.2A are also observed in Figure 4.2B. However, the relative intensities of these peptide peaks are changed. For example, the peak intensity of 1163.41 Da was dramatically decreased in the spectrum obtained from the sample with CHAPS added compared to the standard one. Comparing Figure 4.1B which obtained from digestion with CHAPS present and Figure 4.2B obtained from adding CHAPS after digestion, peptides detected in Figure 4.2B are also detected in Figure 4.1B, indicating the digestion itself was successful. Because the digestion was successful in both of these cases, these peak intensity variations were mainly caused by the effect of CHAPS to MALDI MS response. As CHAPS is a zwitterionic surfactant, it can easily pick up a proton and be ionized, and in the sample mixture of peptide and CHAPS, the competition of obtaining proton is favourable to CHAPS. Therefore, the presence of CHAPS will hinder the peptide ionization and as a result, decrease the MS response.

Similar experiments were also performed with surfactants TX100 and OG. TX100 has a similar effect as CHAPS on MALDI MS. Triton surfactants are modified polyethylene glycol, an easily ionized polymer, and have a molecular weight distribution range from 400 to 1200 Da, where peptides are normally observed. Therefore, the presence of Triton severely suppresses the peptide ionization and will obscure the observation of peptides (Figure 4.3A). On the other hand, OG is a "MS friendly"

**Table 4.1 Critical Micelle Concentration (CMC) of Some Surfactant in Aqueous Solvent.**

| Surfactants | MW (isotope) | Structure | CMC concentration | |
|---|---|---|---|---|
| CHAPS | 614.40 |  | 4 mM | 0.25% (v/v) |
| Triton X 100 | 576.62 (average) |  | ~0.23 mM | ~0.013% (v/v) |
| OG | 292.19 |  | 25 mM | 0.73% (w/v) |
| SDS | 288.13 |  | 8 mM | 0.23% (w/v) |

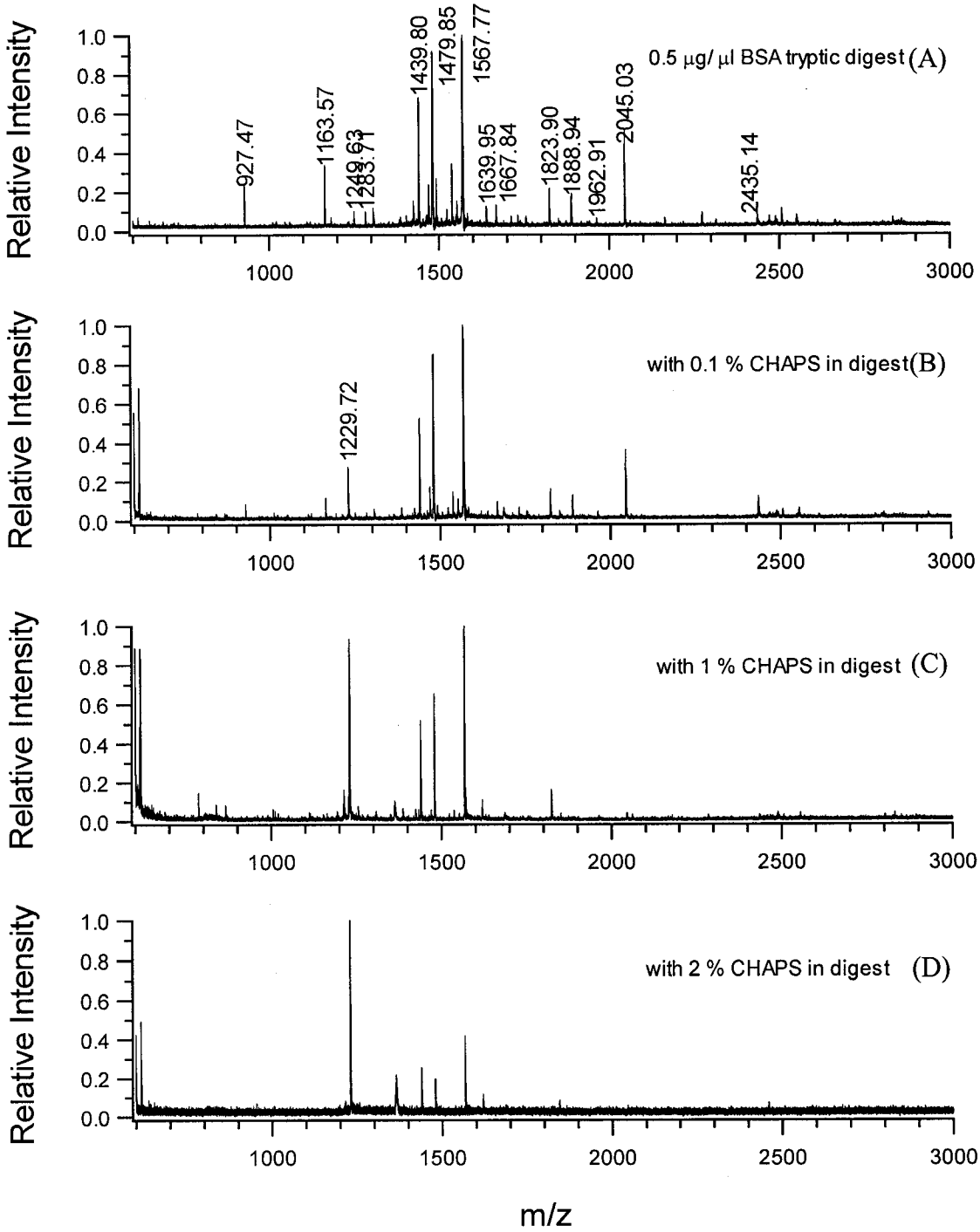



**Figure 4.1** MALDI MS spectra of 0.5 μg/μl BSA digested with presence of (A) no surfactant, (B) 0.1%, (C) 1%, and (D) 2% CHAPS.

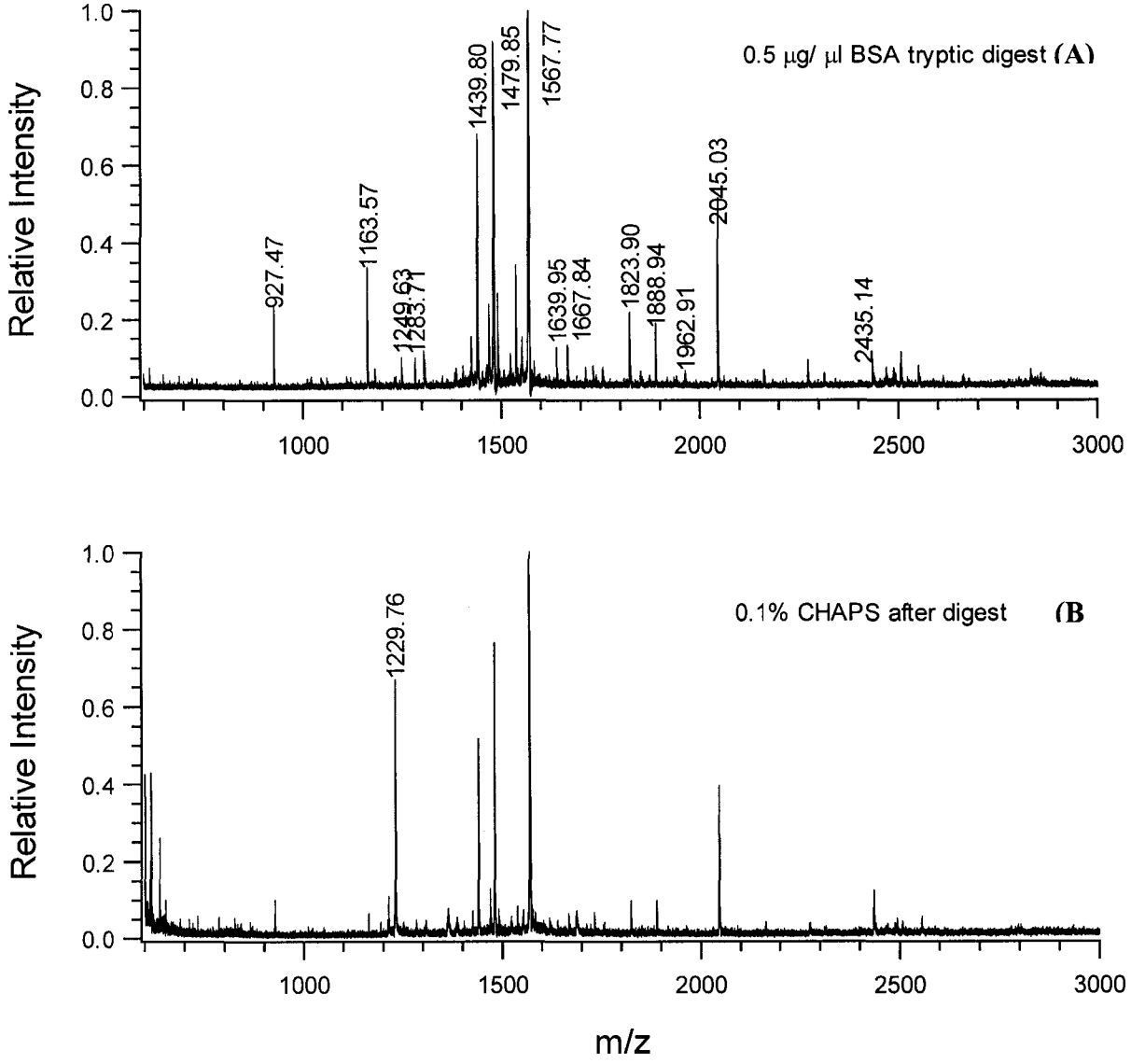



**Figure 4.2** MALDI spectra obtained from 0.5 μg/μL BSA tryptic digest in (A) water and (B) add 0.1% CHAPS after digestion.

surfactant [25], and no significant effect was observed for digestion and peptide detection with OG concentrations up 1% (Figure 4.4).

If CHAPS and TX100 have to be present for tryptic digestion in some proteomic situations, e.g., membrane protein solubilization, they definitely must be removed prior to MS analysis in order to obtain good peptide MS spectra for protein identification. A $C_{18}$ ZipTip contains the same material as a $C_{18}$ reversed-phase column, and is widely used for the removal of salts and surfactants in protein and peptide samples. Experiments show that CHAPS could be easily removed by HPLC $C_{18}$ columns [26], and therefore it can also be removed by $C_{18}$ ZipTip (data not shown). However, the removal of Triton X-100 with $C_{18}$ ZipTip turns out to be problematic. Figure 4.3A shows MALDI MS spectrum of BSA tryptic digest with 1% TX-100 present during digestion process. A $C_{18}$ ZipTip was employed to clean the surfactant TX-100 and the peptides were eluted with 50% ACN/$H_2O$ in 0.1% TFA (Figure 4.3B). However, a significant portion of TX-100 still exists in the peptide samples. This result is also in agreement with Fenselau's results where they used reversed phase $C_{18}$ column to purify protein isomers and observed strong and broad TX-100 UV absorbance peaks during the separation process [27]. Alternatively, HPL ZipTip, a hydrophilic interaction micro-column, is newly designated to cleanup hydrophilic contaminants. The effect of HPL ZipTip for the removal of TX-100 is very significant, as shown in Figure 4.3C, where BSA tryptic peptides can be recovered and clearly observed in the MS spectrum. From the above results we can conclude that if peptide samples contain TX-100, they are not suitable for separation by a reversed phase $C_{18}$ or $C_8$ column, because the Triton isomer can be observed over a broad range elutes.

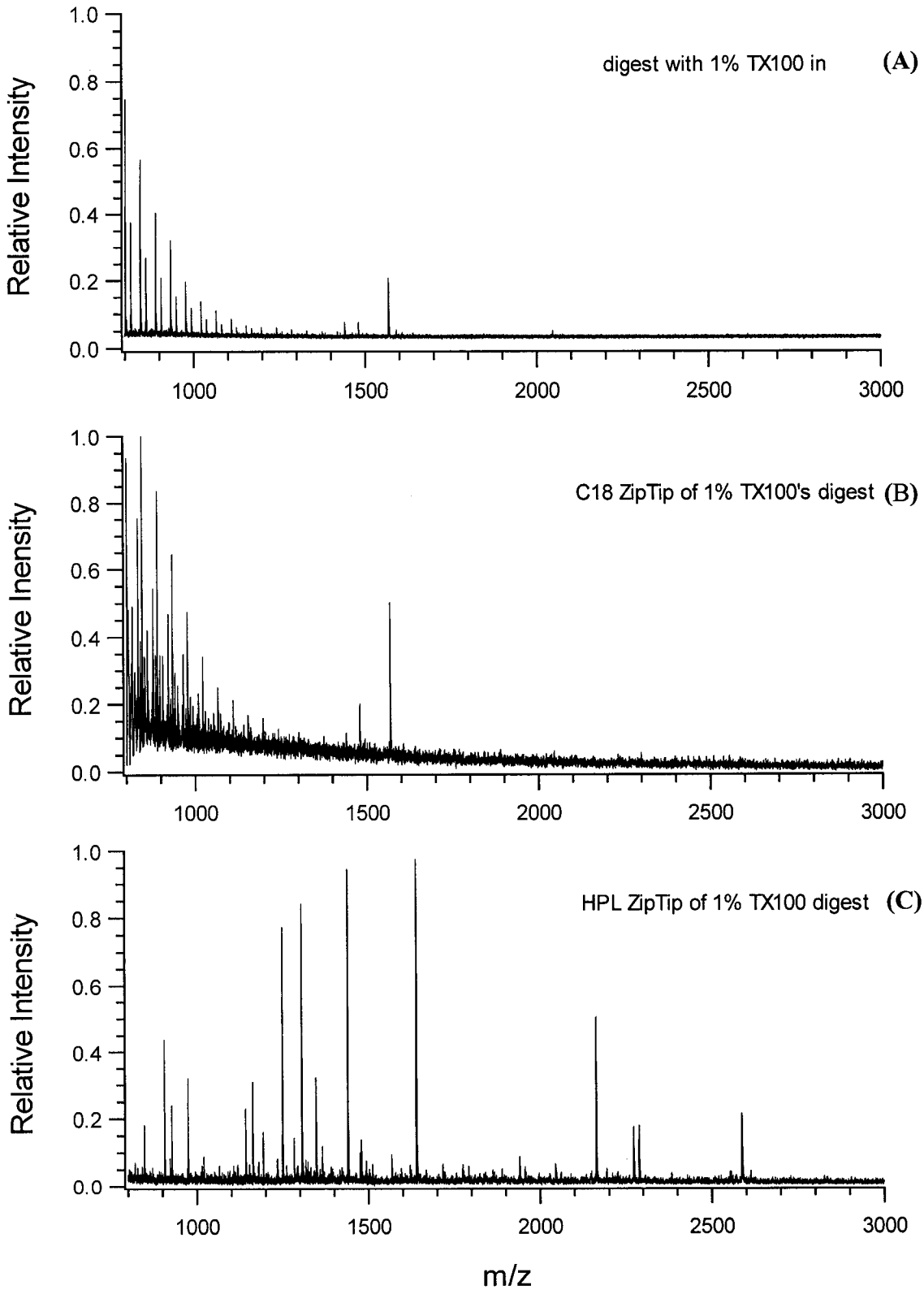



**Figure 4.3** MALDI spectra of 0.1 μg/μL BSA tryptic digestion with 1% Triton X 100 (A), (B) after $C_{18}$ ZipTip cleanup, and (C) after HPL ZipTip cleanup.

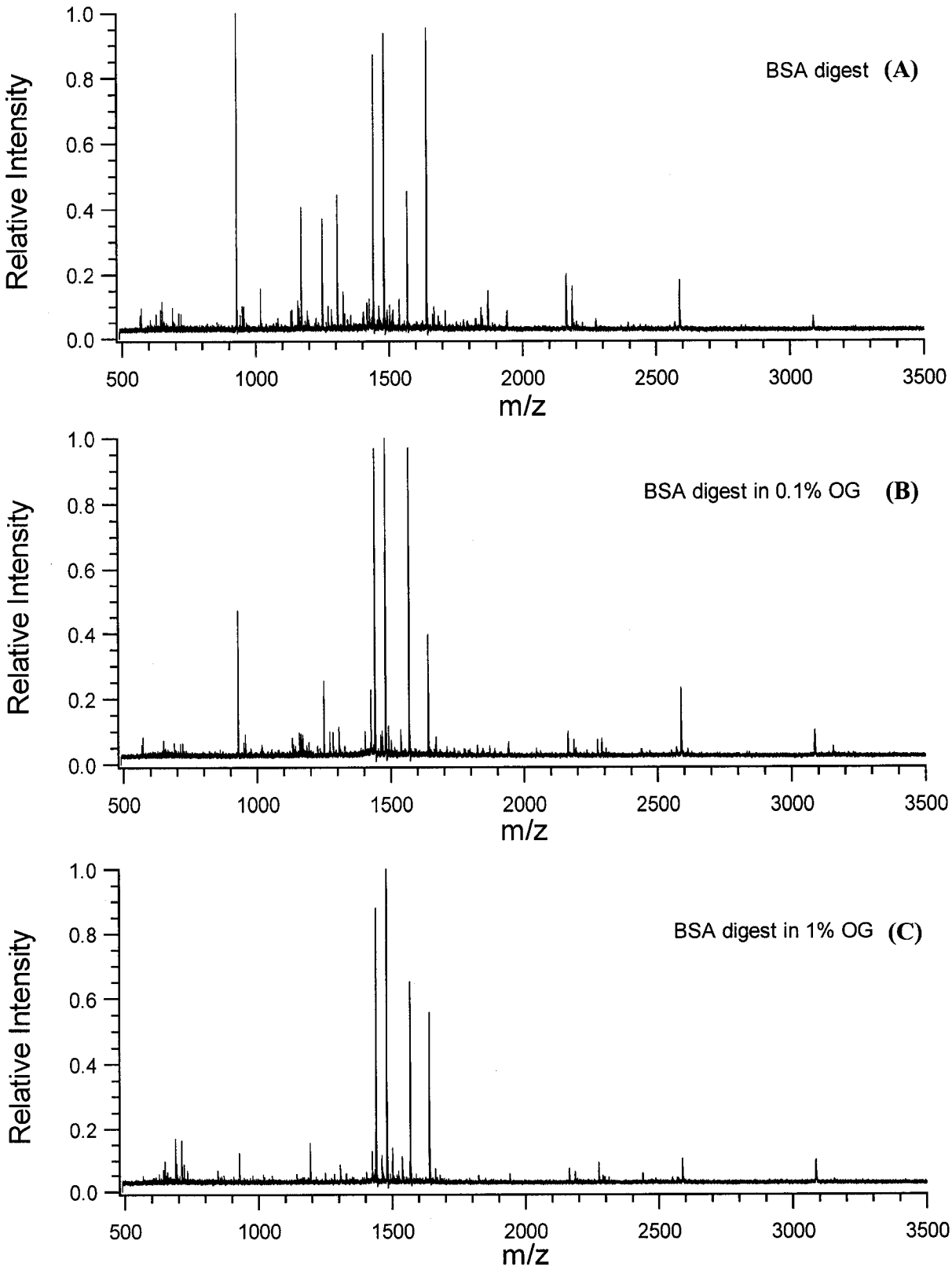



**Figure 4.4** MALDI MS spectra of 0.5 μg/μl BSA digested with the presence of (A) no surfactant, (B) 0.1%, (C) 1% OG.

### 4.3.2 Protein Digestion in the Presence of SDS

SDS is the most commonly used surfactant for biological sample preparation. As it was shown in Chapter 2, two-layer sample preparation method can be employed for the analysis of protein samples containing SDS. From Chapter 3, the resolution degradation from protein samples containing SDS is mainly because the sodium adducts to the samples. In the mass range for peptide analysis, most mass spectrometers have the ability to resolve these adducts; therefore, resolution would not be a problem for peptide detection. However, it is necessary to explore whether SDS has any interference peaks if SDS is used during enzymatic digestion and peptide analysis.

In the blank experiment, SDS and HCCA second-layer solutions were mixed in a 1:1 (v/v) ratio, with SDS concentrations of 0.01% 0.02%, 0.1%, 0.2%, 1% and 2%. No peaks was observed in the low mass range for SDS solutions up to 0.2%, and cluster peaks were observed starting in a 1% SDS even with on-probe washing step. Figure 4.5A displays these clusters observed from 1% SDS solution. In this figure, not only are SDS clusters observed, but strong matrix clusters and matrix-SDS clusters are also observed because of the presence of sodium, a cation that aids the ionization process. These peak assignments are further confirmed by MALDI MS/MS experiments. Figure 4.5B shows a MS/MS spectrum of a cluster at 810.2 Da clearly showing the peak interval of 288, corresponding to the loss of SDS. Therefore, for samples containing greater than 1% SDS, SDS-matrix cluster can be observed in the mass range below 900 Da.

Furthermore, the effect of SDS on the digestion itself was examined. This work is motivated by the fact that detection of a greater number of peptides and/or different types (hydrophobic or hydrophilic) of peptides from a single protein enhances the confidence

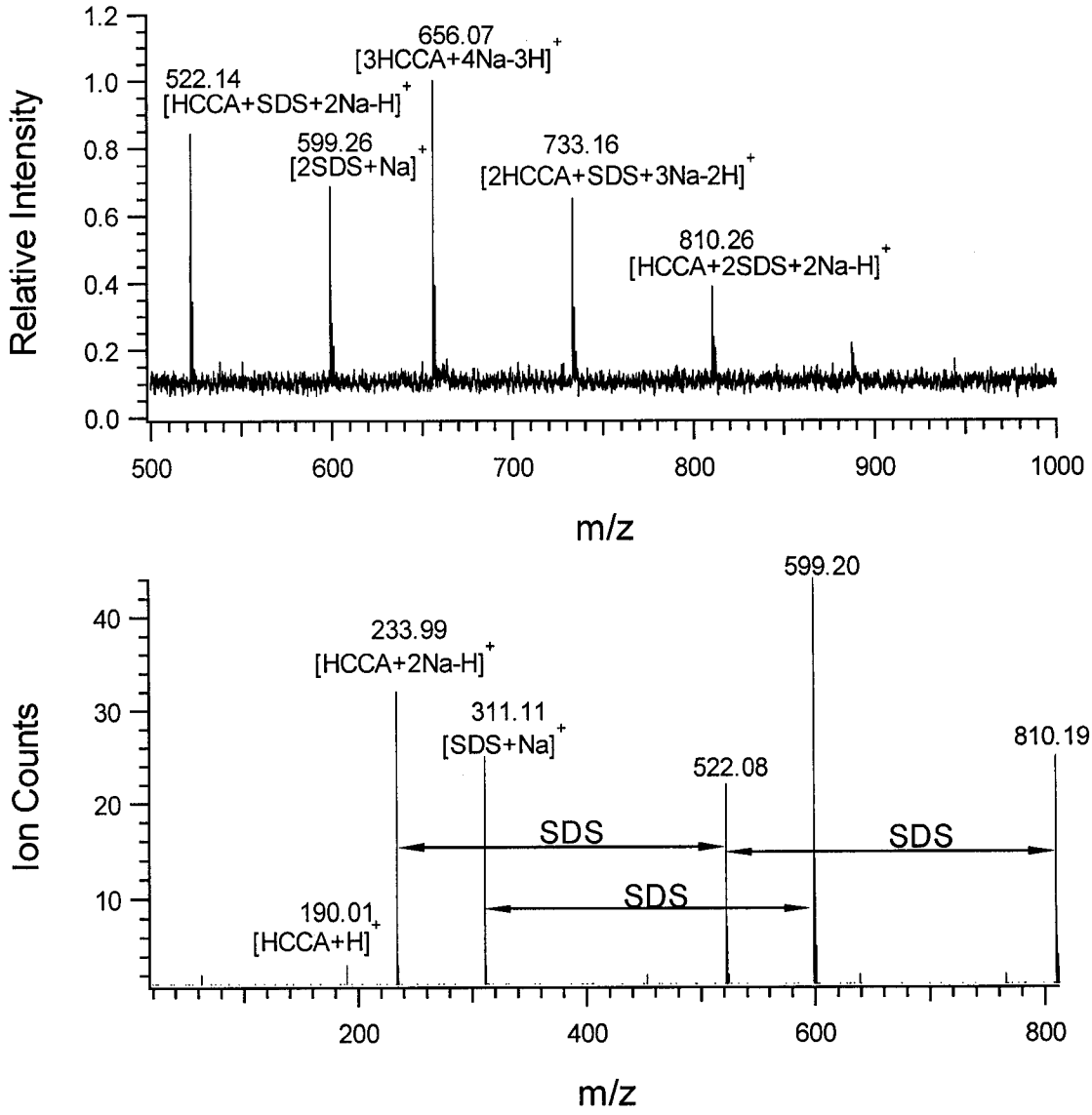



**Figure 4.5** (A) MALDI MS spectra of 1% SDS with HCCA as matrix. (B) MALDI MS/MS spectra of matrix-SDS cluster 810.2 Da.

of protein identification by using either peptide mass fingerprinting or MS/MS in proteomics.

To study the effects of SDS on the digestion of proteins by trypsin, 7 different proteins (ubiquitin, cytochrome *c*, myoglobin, BSA, actin, ovalbumin, and lactoferrin) were prepared in either 0.1% or 1% SDS and digested with trypsin (see Experimental section for details). As a control, the same protein samples were digested in the absence of SDS, and then SDS was added to the sample at concentration of 0.1% or 1%. The digests were then analyzed by MALDI MS.

The MALDI spectra obtained from the digestion of 1 μg/μL ubiquitin are shown in Figures 4.6 and 4.7. Figure 4.6 corresponds to the control in which SDS was added to the sample after the digestion was complete. At all levels of SDS, a MALDI spectrum could be obtained with detectable ubiquitin peptide fragments. However, there are some notable differences between the spectra. The addition of SDS to the sample following digestion had an effect on the relative intensities of the peaks. This is particularly noticeable for the peaks corresponding to sequence coverage 34-42, 64-72, and 55-63. However, the addition of SDS resulted in few changes to the overall number of detected peaks in each spectrum. From Figure 4.6A, it can be seen that 10 peaks were identified as peptide fragments of ubiquitin. When 0.1% SDS was added to the sample following digestion, the same 10 peaks were detected (Figure 4.6B) with an addition of a new peptide fragment corresponding to sequence 7-27. The addition of SDS to the digested sample also produced several peaks resulting from sodium attachment. The spectrum obtained from the sample containing 1% SDS is shown in Figure 4.6C. Eight of the 10 peaks originally detected in the sample without SDS were still visible, with no additional

peptide fragments being observed. However, several low mass ions corresponding to SDS-matrix ion clusters were also observed in addition to sodium attachment peaks, thus deteriorating the overall quality of the MALDI spectrum.

Ubiquitin was then digested in the presence of SDS, and the resulting spectra are shown in Figure 4.7. With 0.1% SDS added to the sample prior to digestion (Figure 4.7B), seven additional peptide fragments were observed in the MALDI spectrum that were not detected in the sample digested without SDS or in the sample with 0.1% SDS added after digestion. This result indicates that the presence of SDS in the sample has an effect on the digestion process itself, rather than exclusively on MS detection. Although additional peaks were detected, no new cleavage sites of ubiquitin were observed. The additional peaks all contained missed cleavage sites, indicating that the addition of SDS perhaps altered the trypsin specificity. This is not surprising considering that SDS is a strong denaturing reagent and can cause the change of trypsin conformations. At the 1% level of SDS (Figure 4.7C), we see that no peptide fragments of ubiquitin were observed in the low mass region of the spectrum. However, the strong intact molecular ion was still observed, indicating that the sample remained undigested and this is in agreement with the conventional brief trypsin protocol "that trypsin remains active in 0.1% SDS". Since peptides were observed in the control spectrum with 1% SDS added following the digestion, the results indicate that 1% SDS completely disrupts the activity of the trypsin, thereby preventing digestion. This is likely due to denaturation of the enzyme in the SDS containing solution.

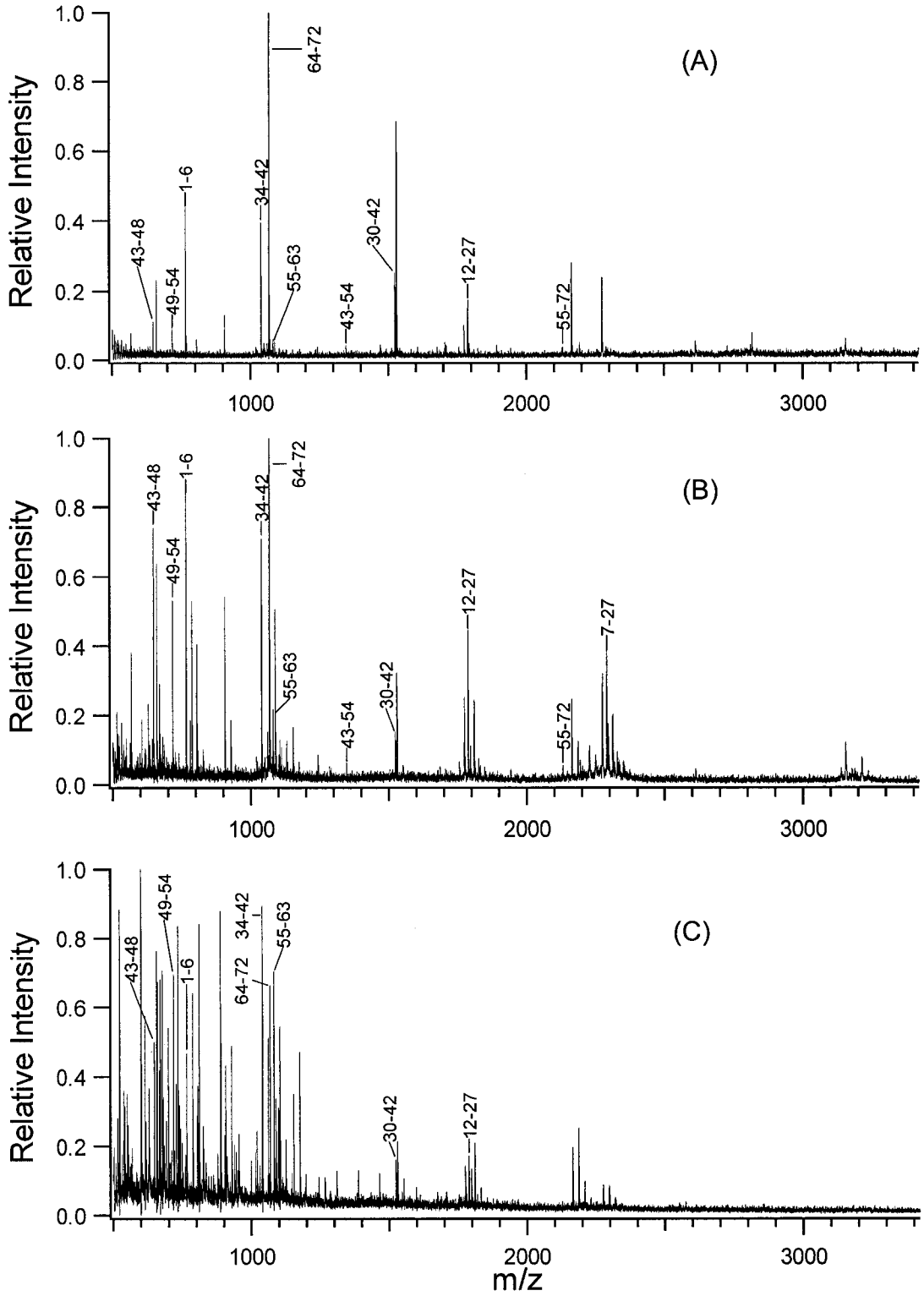



**Figure 4.6** MALDI spectra obtained from digestion of 1 μg/μL ubiquitin (A) without adding SDS, (B) with 0.1% SDS added after digestion and (C) with 1% SDS added after digestion. The peptide fragments are labelled according to the amino acid sequence coverage of ubiquitin.

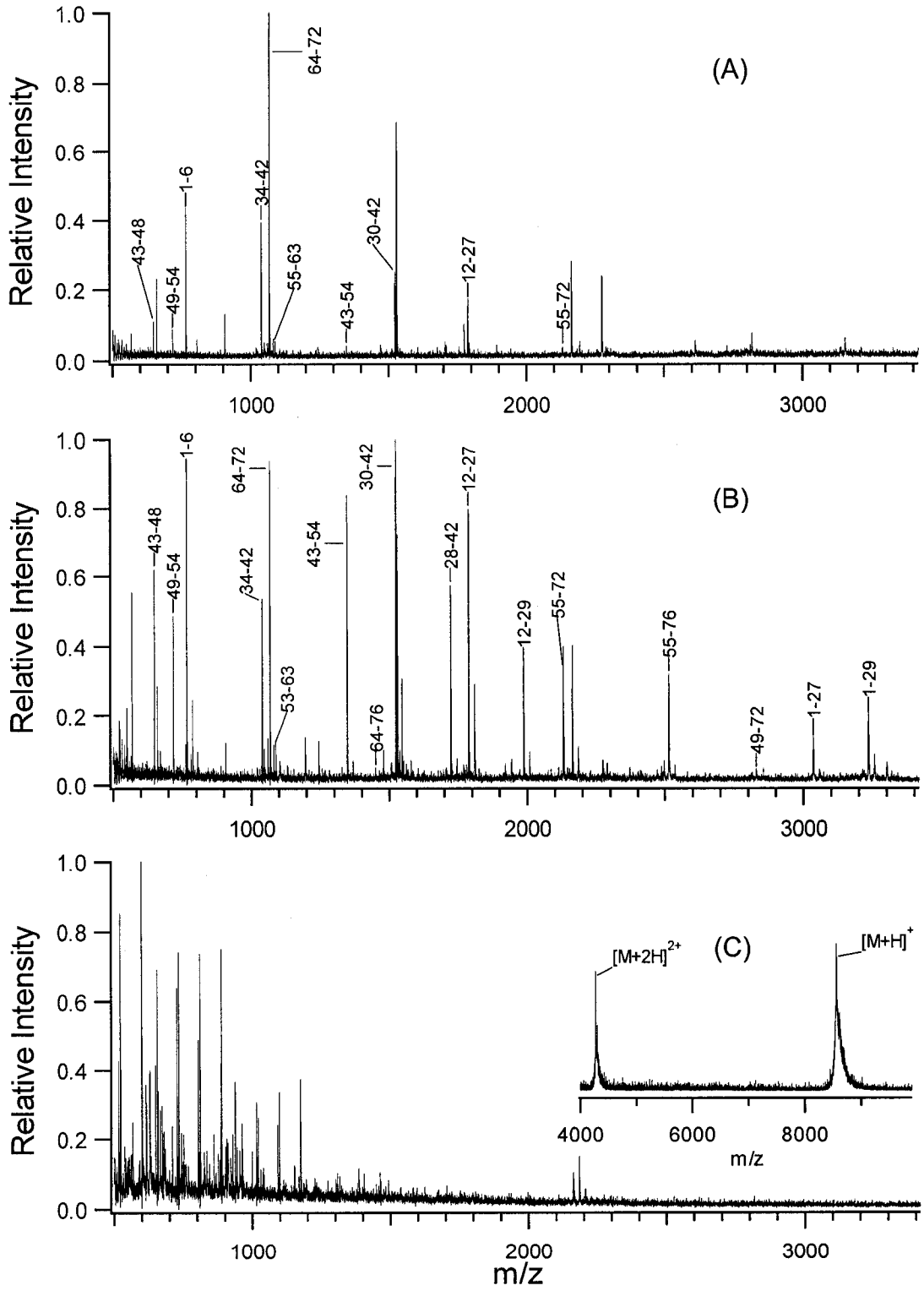



**Figure 4.7** MALDI spectra obtained from digestion of 1 μg/μL ubiquitin (A) without adding SDS, (B) with 0.1% SDS added prior to digestion and (C) with 1% SDS added prior to digestion. The peptide fragments are labelled according to the amino acid sequence coverage of ubiquitin.

Table 4.2 lists peptide sequence coverage of tryptic digests from 7 proteins with SDS added before or after digestion. When the digests of ubiquitin, cytochrome *c*, and ovalbumin were analyzed, more peaks were observed from the samples digested in the presence of SDS than from the control digests, and the peptide sequence coverage increased with SDS in digestion. However, when the digests of myoglobin, actin and BSA were analyzed, only minor differences were observed between the control digests and the samples from digestion of proteins in the presence of SDS. When the digests of cytochrome *c*, actin and ovalbumin were analyzed, peptides with new cleavage sites were observed in the digestion of proteins in the presence of SDS. By combination of adding SDS in and after digestion, 6 out 7 protein samples give increased peptide mass mapping coverage. Therefore it can be concluded that although the presence of SDS in the sample will change the digestion process, the extent of this change and the observed effect on the MALDI spectrum is dependent on the type of protein. Nevertheless, the results shown here demonstrate that proteins can be digested in the presence of low amounts of SDS, and that the resulting digests can be analyzed by MALDI MS. The additional peptides produced are useful for protein identification by peptide mass mapping in proteome analysis. The additional peptides may also increase the chances of generating useful sequence information for a protein based on MS/MS.

Table 4.3 lists the summary of the effects of these four surfactants for protein tryptic digestion and mass spectrometric analysis. OG is a MS compatible surfactant and has little effect on enzyme activity, even though it has limited solubility power; CHAPS and Triton surfactants do not affect enzyme activity but severely interfere with MS

**Table 4.2 Summary of Peptide Sequence Coverage of Tryptic Protein Digestion with (A) No SDS, (B) 0.1% SDS in presence, (C) 0.1% SDS Added after Digestion and (D) Combination of Results from A, B and C.**

| Protein name | (A) 0% SDS | (B) 0.1% SDS in digestion | (C) 0.1% SDS after digestion | (D) Combination of A, B & C |
|---|---|---|---|---|
| Ubiquitin | 65/76aa (86%) | 76/76aa (100%) | 70/76aa (92%) | 76/76aa (100)% |
| Cytochrome c | 53/104aa (51%) | 92/104aa (88%) | 90/104aa (86%) | 93/104aa (89%) |
| Ovalbumin | 134/386aa (34%) | 239/386aa (61%) | 273/386aa (70%) | 305/386aa(79%) |
| Myoglobin | 147/153aa (96%) | 109/153aa (71%) | 145/153aa (95%) | 147/153aa (96%) |
| BSA | 265/607aa (43%) | 227/607aa (37%) | 249/607aa (41%) | 363/607aa (59%) |
| Actin | 230/237aa (61%) | 211/377aa (55%) | 233/377aa (59%) | 253/377aa (67%) |
| Lactoferrin | 297/708aa (41%) | 163/708aa (23%) | 407/708aa (57%) | 431/708aa (60%) |

**Table 4.3 Summary of Some Surfactants' Effects on Protein Tryptic Digestion.**

| Surfactant name | Concentration | Effect on enzyme activity | Effect on MS |
|---|---|---|---|
| CHAPS | 0.1% | – | ++ |
| | 1% | – | ++ |
| OG | 0.1% | – | – |
| | 1% | – | – |
| TX 100 | 0.1% | – | ++ |
| | 1% | – | ++ |
| SDS | 0.1% | – | – |
| | 1% | + | + |

Note: In effect on MS, "++" represents surfactants that have interference peaks and significantly suppress the peptide signals. "+" represents surfactants that have interference peaks.

analysis; trypsin remains active in 0.1% SDS, however, SDS generates matrix-SDS cluster in mass range of 500-900 Da with concentration of 1%, and as the laser beam energy is increased, the cluster may be extended to higher mass.

### 4.3.3 Application to Hydrophobic Membrane Protein

Integral membrane proteins usually have a hydrophobic region in the membrane, and hydrophilic peptides extending to extracellular or cytoplasmic regions. This structure is also called bilayer structure as described in Figure 1.12. Bacteriorhodopsin (BRO) is commonly employed as a model for membrane protein studies. Because of its hydrophobicity, BRO cannot dissolve in aqueous solvent without the addition of surfactant. SDS is employed as one of the surfactants for the purpose of solubilization and denaturation during the tryptic digestion (see Experimental section). Figure 4.8 displays MALDI MS spectra obtained from tryptic digestion with or without SDS. Without SDS, the strong peaks in the MALDI MS spectrum (Figure 4.8A) are all from trypsin autolysis or commonly observed peaks during tryptic digestion, labelled as "T/C". Only two low intensity peaks labelled with stars (*) are BRO peptides corresponding to sequence coverage of 217-225 and 160-172. However, with 0.1% SDS added after digestion, the signal intensities of BRO peptides and trypsin autolysis peaks changed (Figure 4.8B). This is because SDS is an ionization reagent eliminating the discrimination between hydrophobic peptides from BRO and hydrophilic peptides from trypsin autolysis by increasing the ionization efficiency of the hydrophobic peptides [13]. However, the peptide pattern looks significantly different when SDS is present during the digestion (Figure 4.8C). The peptide signal intensity from sequence 160-172 significantly increases compared to trypsin autolysis peaks, and peptide 41-82 is detected

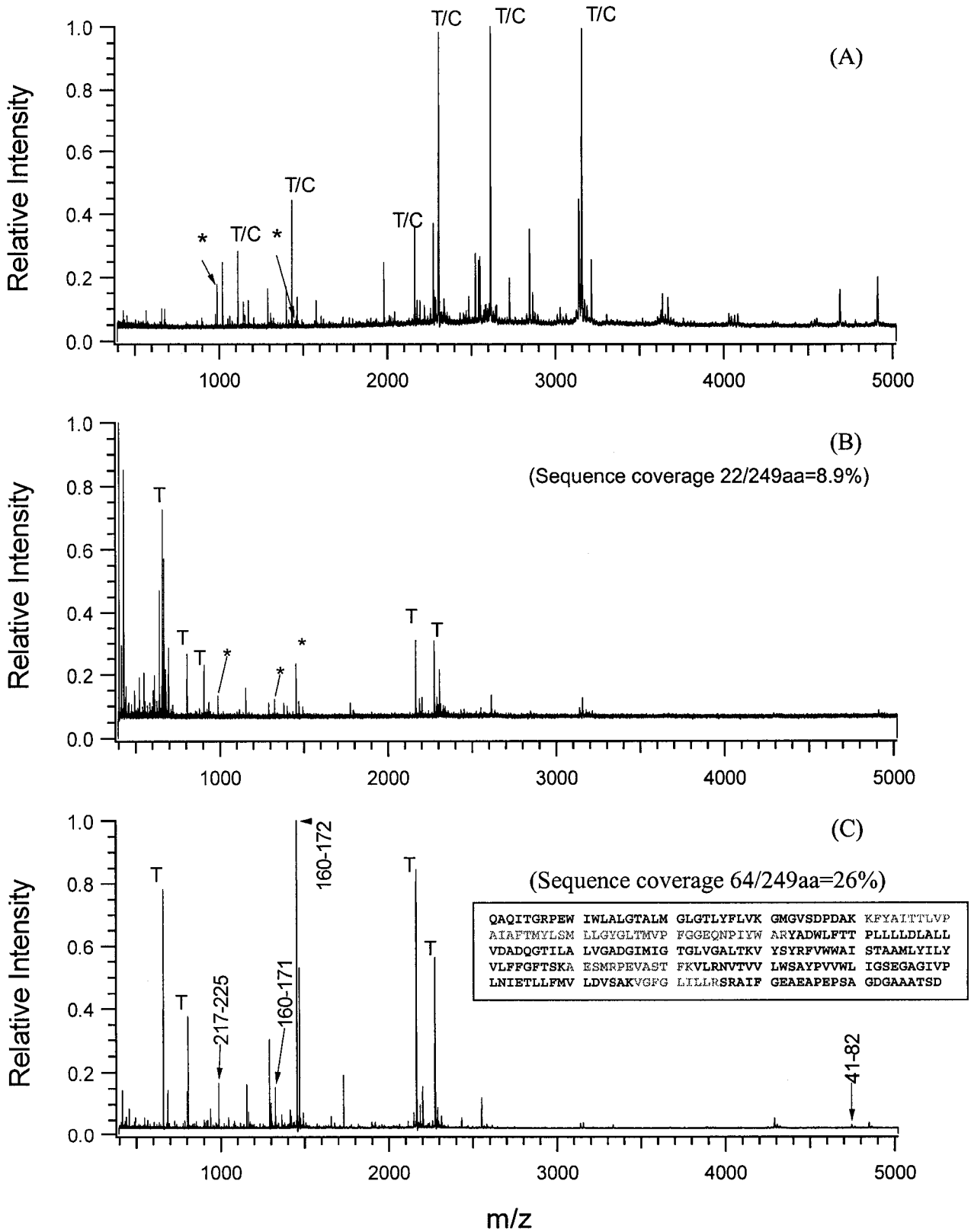



**Figure 4.8** MALDI MS spectra of BRO tryptic digestion with (A) no SDS, (B) 0.1% SDS added after digestion and (C) 0.1% SDS present during digestion. (T/C indicates trypsin autolysis or commonly observed peaks in gel analysis.)

as well. Unlike the peptides 160-172 and 217-225 which arise from the cytoplasmic segment of BRO, the peptide 41-82 is cleaved from the middle of the BRO, which includes one entire α-helix. This result is not surprising because BRO has a bilayer structure, which likely makes it difficult for trypsin to access the cleavage site effectively if the protein is not fully denatured. The observation of stronger BRO peptides in the digestion with SDS present is likely because the denaturation effect of SDS to the protein helps to unfold the protein bilayer structure and make enzyme access to the cleavage site easier, and therefore the digestion is more complete than the one without SDS.

It is clear that 0.1% SDS can be employed in protein identification from the enzyme digestion step to MALDI MS analysis without much interference. In other words, water-insoluble proteins in cells or organelles can be extracted and solubilized by SDS first and then directly subjected to proteolytic digestion, and analyzed via peptide mass mapping or tandem mass spectrometry (MS/MS). For membrane protein digestion, the denaturation of the protein holds some benefit for enzyme to access the cleavage sites, and therefore facilitate the digestion process moving forwards. A better protein sequence coverage can be achieved by dual enzyme digestion, such as CNBr/trypsin [28], because the chemical digestion is not disrupted by SDS and enzymatic digestion is workable under low amount of SDS. For a mixture of hydrophobic and hydrophilic peptides, an additional separation step (such as LC) could be employed to avoid the possible ionization suppression between hydrophobic and hydrophilic peptides [26]. Future work, to be described in the following chapters, will be directed towards the analysis of cell extracts, where surfactants are required to extract proteins out of the cells,

and protein identification is carried out by enzymatic digestion and tandem mass spectrometry.

## 4.4 Conclusions

The effects of four frequently used surfactants for biological sample preparation on enzymatic digestion and peptide MS analysis are examined in detail. Non-ionic surfactants have no effect on the enzymatic digestion, but CHAPS and Triton surfactants severely affect the MALDI MS analysis and therefore should be removed before MALDI MS analysis. If the protein can be dissolved by OG, enzymatic digestion and MS analysis can be performed with OG up to 1% without any interference. SDS can be used in tryptic digestion with concentration up to 0.1% without any interference in the MALDI MS analysis. The performance of SDS-solubilized hydrophobic membrane protein tryptic digestion was demonstrated in presence of 0.1% SDS, and it leads to an improvement in the sequence coverage. We envision that SDS extracted protein mixture can be directly analyzed without purification or SDS removal steps by direct enzyme digestion and identification with tandem mass spectrometry.

## 4.5 Literature Cited

(1) Almgren M. *Biochim. Biophys. Acta* **2000**, *1508*, 146-163.

(2) Maire, M. L.; Champeil, P.; Moller, J. V. *Biochim. Biophys. Acta* **2000**, *1508*, 86-111.

(3) Neugebauer, J. M. In *Guide to Protein Purification*; Deutscher, M. P., Ed.; Academic Press: San Diego, CA, 1990, pp 239-253.

(4) Karas, M.; Bahr, A.; Ingendoh, E.; Nordhoof, B.; Stahl, B.; Strupat, K.; Hillenkamp, F. *Anal. Chim. Acta* **1990**, *241*, 175-185.

(5) Rosinke, B.; Strupat, K.; Hillenkamp, F.; Rosenbusch, J.; Dencher, N.; Krueger, U.; Galla, H.-J. *J. Mass Spectrom.* **1995**, *30*, 1462-1468.

(6) Gharahdaghi, F.; Kirchner, M.; Fernandez, J.; Mische, S. M. *Anal. Biochem.* **1996**, *233*, 94-99.

(7) Schnaible, V.; Michels, J.; Zeth, K.; Freigang, J.; Welte, W.; Buehler, S.; Glocker, M. O.; Przybylski, M. *Int. J. Mass Spectrom. Ion Proces.* **1997**, *169/170*, 165-177.

(8) Schaegger, H. In *A Practical Guide to Membrane Protein Purification*; von Jagow, G., Schaegger, H., Eds.; Academic Press: San Diego, CA, 1994, pp 31-34.

(9) Immler, D.; Mehl, E.; Lottspeich, F.; Meyer, H. E. In *Microcharacterization of Proteins*; Kellner, R., Lottspeich, F., Meyer, H. E., Eds.; Wiley-VCH: New York, 1999, pp 87-93.

(10) Puchades, M.; Westman, A.; Blennow, K.; Davidsson, P. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 344-349.

(11) Galvani, M.; Hamdan, M. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 721-723.

(12) Cadene, M.; Chait, B. T. *Anal. Chem.* **2000**, *72*, 5655-5658.

(13) Breaux, G. A.; Green-Church, K. B.; France, A.; Limbach, P. A. *Anal. Chem.* **2000**, *72*, 1169-1174.

(14) Ross, A. R. S.; Lee, P. J.; Smith, D. L.; Langridge, J. I.; Whetton, A. D.; Gaskell, S. J. *Proteomics* **2002**, *2*, 928-936.

(15) Compton, B. J.; Lee, J.-J.; Brown, E. K.; Herbert, R. G.; Ding, J.; Livingstone, J.; Bouvier, E. S. P. *Proceedings of the 47th Conference of the American Society for Mass Spectrometry and Allied Topics*, Dallas, TX, June 13-17 1999; p1012-1013.

(16) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999**, *20*, 3551-3567.

(17) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. *Anal. Chem.* **1996**, *68*, 850-858.

(18) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999**, *71*, 1087-1091.

(19) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1996**, *68*, 2494-2500.

(20) Kenny, B.; Devinney, R.; Stein, M.; Reinscheid, D. J.; Frey, E. A.; Finlay, B. B. *Cell* **1997**, *91*, 511-520.

(21) Duchesne, L.; Deschamps, S.; Pellerin, I.; Lagree, V.; Froger, A.; Thomas, D.; Bron, P.; Delamarche, C.; Hubert, J.-F. *Kindey Int.* **2001**, *60*, 422-426.

(22) Duchesne, L.; Pellerin, I.; Delamarche, C.; Deschamps, S.; Lagree, V.; Froger, A.; Bonnec, G.; Thomas, D.; Hubert, J.-F. *J. Biol. Chem.* **2002**, *277*, 20598-20604.

(23) Jensen, C.; Haebel, S.; Anderson, S. O.; Roepstorff, P. *Int. J. Mass Spectrom. Ion Proces.* **1997**, *160*, 339-356.

(24) Jeannot, M. A.; Zheng, J.; Li, L. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 512-520.

(25) Barnidge, D. R.; Dratz, E. A.; Sunner, J.; Jesaitis, A. J. *Protein Sci.* **1997**, *6*, 816-824.

(26) Hixson, L. K.; Rodriguez, N.; Camp II, D. G.; Strittmatter, E. F.; Lipton, M. S.; Smith, R. D. *Electrophoresis* **2002**, *23*, 3224-3232.

(27) Shore, L. J.; Odell, G. B.; Fenselau, C. *Biochem. Pharm.* **1995**, *49*, 181-186.

(28) Quach, T. T.; Li, N.; Richards, D. P.; Zheng, J.; Keller, B. O.; Li, L. *J. of Proteome Research* **2003**, in press.

# Chapter 5

# Analysis of In-Solution Digest of SDS-Solubilized Proteins for Membrane Protein Identification by MS/MS Approach I: Method Development [a]

## 5.1 Introduction

In the field of proteomics, people strive to globally analyze the complete protein complement of a given organelle, cell or tissue. Up to now, 2D gel electrophoresis separation of proteins coupled with trypsin digestion and peptide mass mapping is still the method of choice to achieve this goal. More recently, with the development of 2D HPLC [1-3], ICAT technology [4] and high throughput automation, solution-based protein identifications have gained wider attention. Generally, in this solution-based approach, an aqueous protein mixture is enzymatically digested and the resultant peptide mixture is separated by LC and sequenced by ESI-MS/MS. Because LC-ESI is easily automated, this approach has been applied to many systems with great success and has already raised excitement of whether it will replace gel-based techniques soon. However, it should be noted that the solution-based method is more for "regular" proteins that are easily extracted/dissolved into aqueous solution, or in other words, hydrophilic. In the case of hydrophobic membrane protein systems, application of solution-based approach might become complicated and problematic. On the other hand, although the analysis of membrane protein by 2D gel electrophoresis is also difficult, it can be simply

[a] Mr. P. Semchuk of the Institute for Biomolecular Design of the University of Alberta generated part of the ESI data.

accomplished by 1D gel electrophoresis where a strong surfactant sodium dodecyl sulfate (SDS) is added to denature the membrane proteins [5-7].

It is difficult to believe that biochemical problems involved solution-based techniques could be solved without the use of surfactant. The idea of using surfactants to aid protein solubilization has been introduced to solution-based approaches. However, SDS, a strong ionic surfactant, is known to suppress analyte ionization in ESI MS unless extensive clean up procedures are applied prior to detection [8-10]. MALDI MS, on the other hand, can tolerant contaminants including SDS to a much higher extent compared to ESI [11]. Our previous studies have shown that the effect of SDS on protein analysis by MALDI MS can be minimized by choosing the correct matrix, solvent, and deposition method [12]. The effect of SDS on peptide analysis by MALDI is much less than that for protein samples. In fact, Limbach and co-workers have shown that the presence of SDS in a sample can change the relative intensity ratio of peptides with varied hydrophobicity and even increase the number of detectable peaks without interference with MALDI MS [13].

Here we demonstrate a solution-based digestion, separation and detection method for membrane protein identification. In this technique, 1% SDS was first added to solubilize membrane hydrophobic proteins. Trypsin digestion was then carried out in the presence of 0.1% SDS. A home built offline HPLC-MALDI instrument was used for MS and MS/MS peptide detection and sequencing. All steps, including digestion, separation and peptide sequencing, were successful and efficient without the use of any purification steps to eliminate SDS. This method was validated with a membrane protein-enriched sample from a cultured human cancer cell line HT29 and was proven very useful for the

identification of membrane proteins. For the purpose of identification and method comparison, the same sample system was also analyzed by an ESI approach and conventional in-gel digestion/peptide mass mapping approach. Comparison of the three results reveals that each method has its own merits and should be combined to yield a better and more complete proteome map.

## 5.2 Experimental

### 5.2.1 Chemicals and Reagents

α-Cyano-4-hydroxycinnamic acid (HCCA), bovine serum albumin (BSA), bovine trypsin, dithiothreitol (DTT), iodoacetamide, Triton X-100, trifluoroacetic acid (TFA) and sodium dodecyl sulfate (SDS) were purchased from Sigma-Aldrich Canada (Markham, ON, Canada). HPLC-grade acetonitrile and acetone were from Fisher Scientific Canada (Edmonton, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA, USA). HCCA was recrystallized from ethanol (95%) before use.

### 5.2.2 Subcellular Fractionation of Adherent Cells

Adherent human HT 29 cells were used as a model system in this experiment. The HT29 cells were fractionated into cytosolic and membrane fractions by subsequent extractions with saponin and Triton X-100 [14]. Briefly, the cytosolic proteins were first released by 0.2% saponin lysis buffer containing 50 mM Tris-Cl pH 7.5 and 1 mM phenylmethyl sulfonyl fluoride (PMSF) at 4 °C. Membrane proteins were then extracted by buffer containing 1% TX100, 50 mM Tris-Cl pH 7.5 and 1 mM PMSF at 4 °C.

### 5.2.3 Acetone Precipitation and Protein in Solution Digestion

Triton X-100 can be removed from the membrane subcellular fraction with acetone precipitation [15]. Acetone was pre-cooled at –80°C and added to the protein extracts to a final concentration of 80% (v/v). After vortexing for 2 min, the mixture was kept cool at -20 °C for 4 hours and then centrifuged for 10 min at 4°C at 16,000 g to obtain the precipitate. The process was repeated with a small volume of acetone before the final pellet was collected and lyophilized for further experiments.

For the in solution digestion, one milligram lyophilized protein powder was dissolved in 17 μL 1% SDS (w/v) in 20 mM Tris. After the standard reduction and alkylation procedures [16], the mixture was diluted to a final SDS concentration of 0.1% and pH was adjusted with 100 mM $NH_4HCO_3$ to ~8.5. Finally, 50 μL of 1 μg/μL bovine trypsin and 7 μL of 20 mM $CaCl_2$ were added to the mixture, and the digestion was performed overnight at 30°C. Additional SDS was added if precipitation was observed during any dilution steps.

### 5.2.4 HPLC-MALDI Analysis

Peptide separation was performed on an Agilent (Palo Alto, CA, USA) 1100 series capillary HPLC equipped with an auto sampler. Chromatographic analysis was performed with a reversed-phase 1.0 × 150 mm Vydac $C_8$ column equipped with a 1.0 × 10 mm Nest Group SDS guard cartridge (Southborough, MA, USA). A flow-rate of 40 μL/min was used for separation. Gradient elution was performed with solvent A (0.1%, v/v, aqueous TFA) and B (0.1%, v/v, TFA in acetonitrile). About 30 μg (equivalent to 8 × $10^5$ cells) of peptide mixture was injected and UV wavelength was set at 210 nm.

HPLC fractions were directly collected onto a 100-well MALDI target by our home built device (semi-offline HPLC-MALDI device, patent pending) [17]. After the separation was completed, 1 μL of 50% acetonitrile/water (v/v) saturated by HCCA was added on top of each spot on the sample plate, and allowed to air dry. The sample spots were ready for both MALDI MS and MALDI MS/MS analysis.

### 5.2.5 LC-ESI MS/MS Analysis

1D LC-ESI MS/MS was carried out on a ThermoFinnigan LCQ Deca ion trap instrument equipped with a Surveyer LC/MS system (San Jose, CA, USA) and sample was loaded onto a 300 μm × 150 mm $C_{18}$ capillary column with the same 1.0 × 10 mm SDS guard cartridge used in HPLC-MALDI experiment. For offline 2D LC-ESI MS/MS analysis, the sample was first separated by a 1.0 × 250 mm cation exchange column (8 μm particle size, 900Å pore size) (solvent A: 20% Acetonitrile 0.1% TFA (v/v), solvent B: the same as A with 0.5 M NaCl) and then the individual fractions (1 fraction/minute) were injected on a 75 μm × 150 mm $C_{18}$ column followed by analysis by a LCQ Deca XP MS/MS (San Jose, CA, USA).

### 5.2.6 SDS-PAGE and In-Gel Digestion

SDS-PAGE was carried out in a Bio-Rad mini-Protein III system using 4%/12% stacking/separating polyacrylamide mini-gels. The SDS sample buffer contained 2% mercaptoethanol (v/v), 1% SDS, 12% glycerol, 50 mM Tris-HCl and trace amount of bromophenol blue. Prior to electrophoresis, the protein sample was mixed in the sample buffer and heated at 95°C for 4 min. Visualization of protein bands was done using Bio-Rad's Biosafe Coomassie blue stain reagent.

Protein bands of interest were excised from the gel and digested with trypsin according Mann's procedure [16]. Briefly, the gel pieces were cut into small segments and washed with fresh water for 20 min. After reduction and alkylation, gel pieces were dehydrated and covered with 10 ng/μL trypsin in 0.1 mM $NH_4HCO_3$ and 2 mM $CaCl_2$ for overnight digestion at 30 °C. Peptides were extracted 3 times in 60% acetonitrile in 0.25% TFA (volume: enough to cover the gel pieces) and once in 100% acetonitrile by a 20 min shaking each time. The pooled extracts were evaporated to ~ 1 μL.

### 5.2.7 Analysis of Tryptic Peptides from In-Gel Digestion

The two-layer sample preparation method was used for MALDI MS and MS/MS analysis of gel samples [18, 19]. Peptide mass mapping was performed using the Mascot search engine. The proteins identified from peptide mass mapping were confirmed by MALDI MS/MS. CID experiments were performed on the unmatched peaks after peptide mass mapping, starting from the high signal/noise peak to the lower ones, as with the HPLC-MALDI MS/MS experiment, until the sample is consumed. Usually about 20 MS/MS spectra could be collect from one sample spot.

### 5.2.8 Protein Identification from CID Data

Fragmentation information from CID spectra was first submitted to the sequence query program in Matrix Science (http://www.matrixscience.com) for possible identities with a precision tolerance of 0.3 Da for both the parent peptide and MS/MS fragments. These potential protein matches were then put into the MS-product (http://prospector.ucsf.edu) program and the theoretical fragments were compared with

experimental results, which may lead to definite identification of proteins. All the peptide identities were manually analyzed after database search.

### 5.2.9 Instrumentation

MALDI MS experiments were carried out on a Bruker Reflex III time-of-flight mass spectrometer (Bremen/Leipzig, German) using the reflectron mode of operation. Ionization was performed with a 337-nm pulsed nitrogen laser. MALDI MS/MS experiments were carried out on a MDS Sciex QSTAR Pulsar QqTOF mass spectrometer equipped with an orthogonal UV-MALDI source (Concord, ON, Canada). All data were processed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR, USA).

## 5.3 Results and Discussion

It is known that some surfactants will deteriorate the MALDI signals for protein and peptide analysis [20-22]. Because Triton X-100 is commonly used to extract the membrane proteins out of cells, the effect of Triton X-100 on protein digestion and peptide detection by MALDI MS was examined (Chapter 4). It was concluded that Triton X-100 had no effect on the digestion process itself, since it is a polyethylene glycol (PEG) derivative and thus a mild surfactant. However, as a PEG, Triton X-100 is easily ionized in mass spectrometry. In addition, Triton X-100 has a wide molecular weight distribution ranging from 400 to 1200 Da, a mass range where tryptic peptides are commonly observed. Therefore, Triton X-100 severely suppresses and obscures the peptide signals in MALDI. The effect of HPLC clean-up was also examined and although Triton X-100 does not interfere with HPLC separation, its monomers are present in a long HPLC fraction [23], making it difficult to obtain pure peptide fractions,

which in turn will affect the mass spectrometric detection. Therefore, Triton X-100 must be removed by acetone precipitation before digestion.

After the removal of the surfactant Triton X-100, another surfactant, which has similar or stronger solubilization properties, is required to solubilize those proteins. n-Octylglucoside is one possible choice as the alternative surfactant. It is a commonly used MS-friendly surfactant and has similar solubilization properties to Triton X-100. However, most of the 1 mg acetone-precipitated proteins would not dissolve in 250 μL of 2% n-octylglucoside after several hours of vortexing. Although more n-octylglucoside might help the solubilization process, the final protein solutions would be too diluted for digestion. Conversely, when the strong surfactant SDS was used, the same amount of lyophilized protein powder was easily dissolved with 17 μL 1% SDS. In addition, SDS is widely used for biological system for membrane or hydrophobic protein solubilization and thus would be a better choice. Therefore SDS was chosen in the experiments to solubilize the precipitated protein powder.

### 5.3.1 Protein Digestion in the Presence of SDS

Because of the interference of SDS on mass spectrometry, more effort has been devoted to the removal of SDS. However, removal of SDS often results in the loss of hydrophobic or low abundant protein samples [24, 25]. It would therefore be preferred and more practical to develop a MS method with SDS present. In Chapter 4, it was demonstrated that tryptic digestion was still successful in the presence of 0.1% SDS and the presence of SDS does not interfere with MALDI MS analysis of peptides. For several proteins, the addition of SDS enabled the observation of larger peptides and higher sequence coverage. For membrane proteins, the presence of SDS aided the digestion

process as surfactants aid the solubilization and denaturation process of those hydrophobic proteins [26-28].

The SDS-dissolved HT29 membrane protein mixture was tryptically digested with 0.1% SDS. Figure 5.1 shows the MALDI MS spectrum of the digested protein mixture. From the MS spectrum, we can quickly conclude that the tryptic digestion was effective. Because there may be tens of relatively high abundance proteins in the protein mixture and therefore hundreds of peptides presence in this mixture, it is not surprising to see a spectrum like this where peptide signal suppression was very severe. A separation of the complex peptides digest mixture is required and essential prior to MS detection to obtain signals for further MS and MS/MS experiments.

### 5.3.2 HPLC Separation

HPLC separation of the SDS-containing tryptic digest can be carried out with the use of a SDS guard column. The performance of this guard column was pre-examined with BSA standard tryptic digest with 0.1% SDS (data not shown) and the use of a SDS guard column showed positive performance improvements on the HPLC separation. Figure 5.2 shows a portion of the HPLC chromatogram of the SDS-containing tryptic peptide mixture, and Figure 5.3 shows MALDI MS spectra of three HPLC fractions at elution times 23 min, 42 min and 54 min (from the fractions labeled with asterisks). In contrast to the whole digest spectrum showed in Figure 5.1, Figures 5.3A, B and C clearly proved the improvements of quality on the MALDI MS detection. This in turn also demonstrates the efficiency of the HPLC separation. In addition, as the HPLC fractions are directly concentrated and spotted onto the MALDI target in a semi-automated fashion and no vials or further drying steps are needed, sample loss is

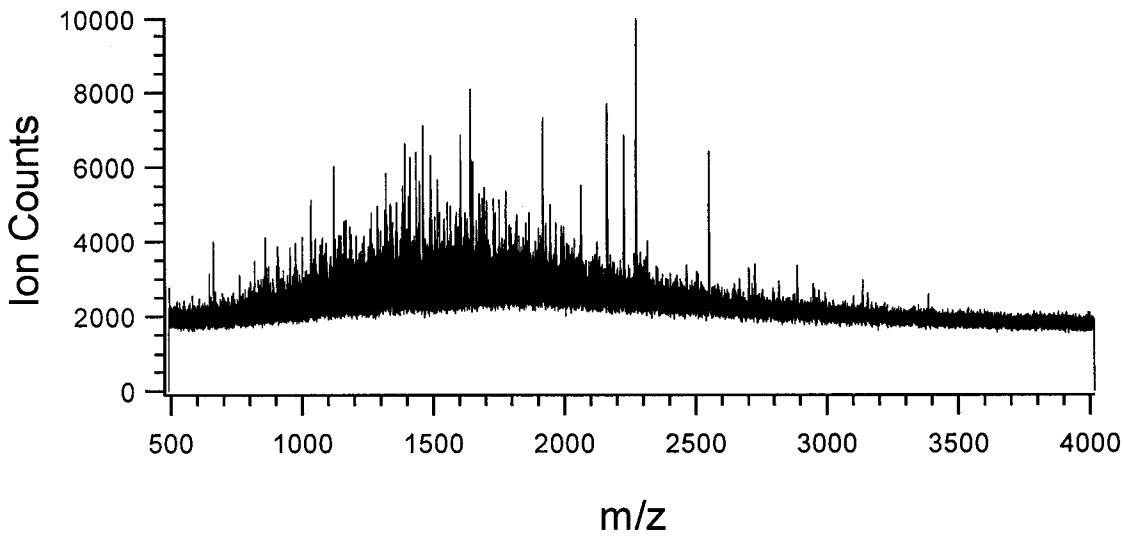



**Figure 5.1** MALDI MS spectrum of tryptic digest of membrane protein mixture with 0.1% SDS present.

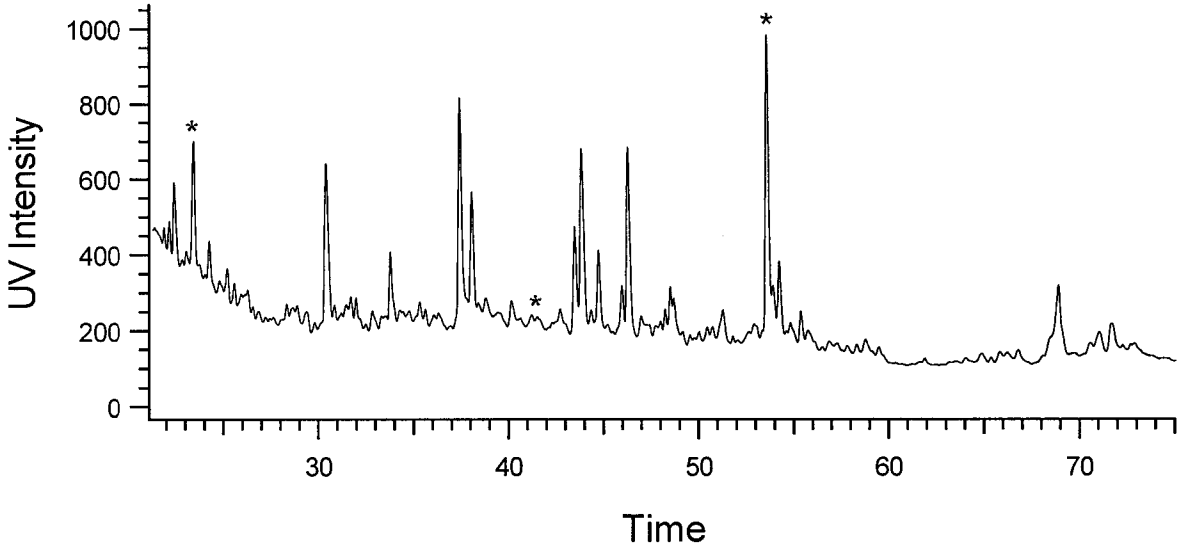



**Figure 5.2** Reversed-phase HPLC chromatogram of tryptic digest of SDS-solubilized membrane protein mixture. (The MALDI MS spectra of the fractions labeled with * are shown in Figure 5.3)

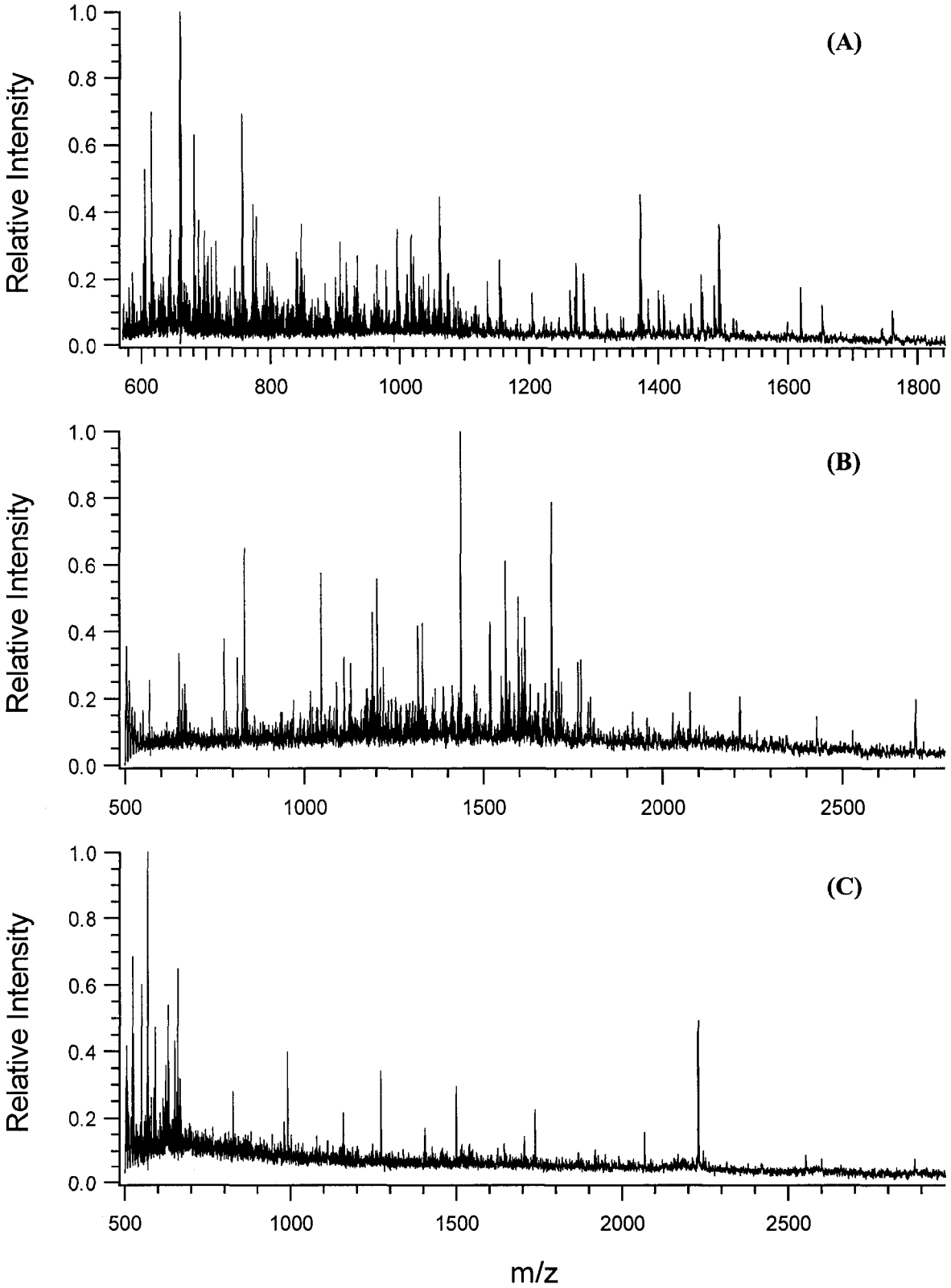



**Figure 5.3** MALDI MS spectra obtained from HPLC fractions of (A) 23 min, (B) 42 min and (C) 54 min.

minimized. This is very helpful in the analysis of hydrophobic peptides where adsorption on the vials during dying step is very severe.

### 5.3.3 MALDI MS and MS/MS

After the HPLC experiment, saturated HCCA in 50% acetonitrile/water was added directly to the fractions on the MALDI target and allowed to air dry. These sample spots were ready for peptide analysis by MALDI MS and MS/MS. Because trypsin digestion of proteins and HPLC fractionation were performed and the resultant peptides are separated according to their hydrophobicity, peptides generated from one protein are impossible to collect in a single fraction. This means peptides in one HPLC fraction could belong to several different proteins. As a result, protein identification through peptide mass mapping is not achievable. Peptide sequence tags obtained from MALDI MS/MS data have to be used for protein identification in this technique [29-31]. This has become more routine with the development of the MALDI QqTOF instrument. (Our group has been using Sciex MALDI QqTOF extensively for peptide sequencing with good success.) However, MALDI MS experiments with a conventional MALDI TOF instrument (Bruker Reflex III) must be carried out to obtain the correct tryptic peptide mass information. This is because in-source fragmentation is often observed in the MALDI QqTOF instrument in MS mode when HCCA is used as the matrix [32-34]. These fragmented peptides are not tryptic peptides and will fail a database search if treated as such.

In the HPLC-MALDI MS/MS experiment, peptide sequencing was performed sequentially from the most to least intense peaks and from high to low mass until the

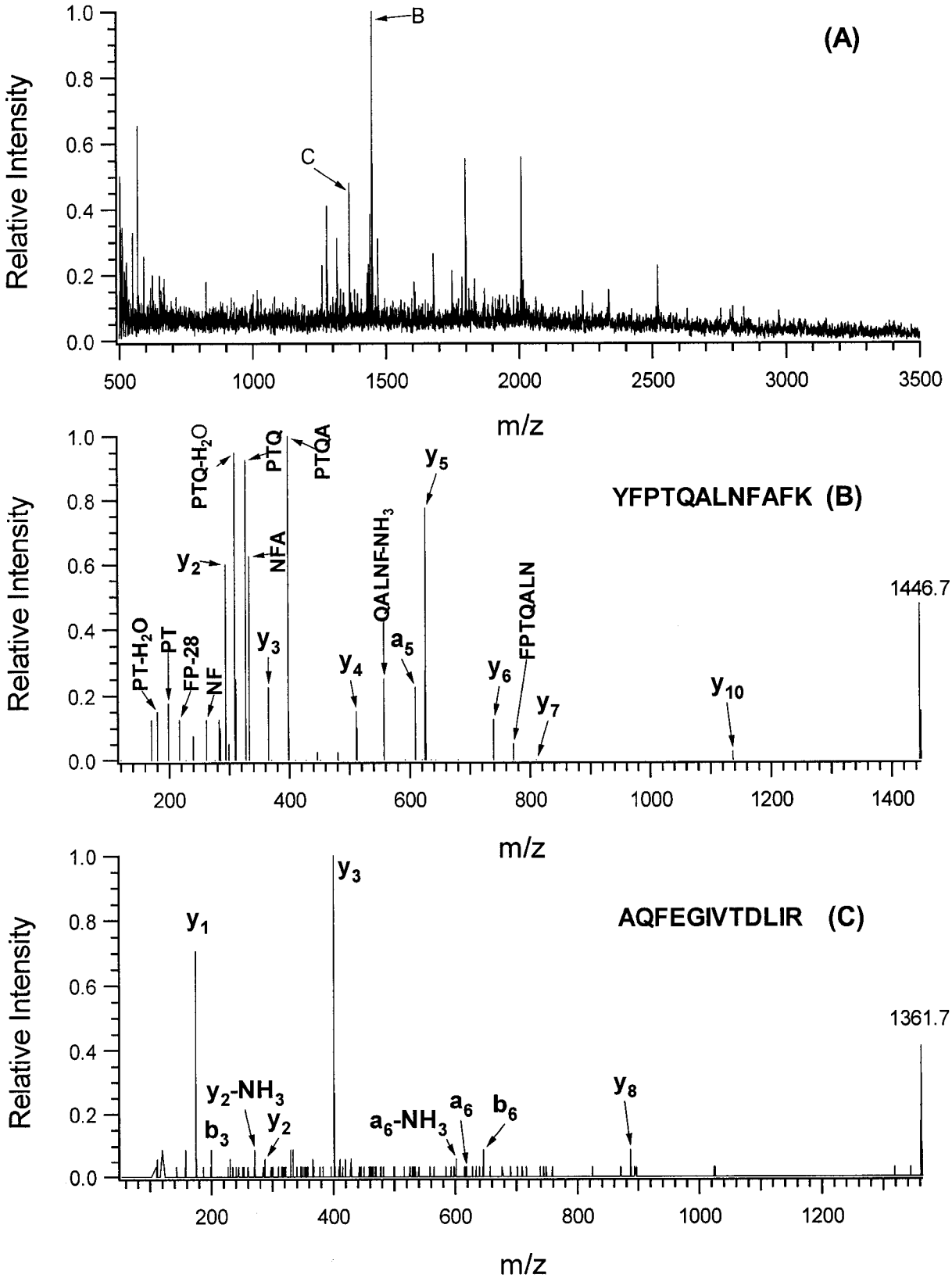



**Figure 5.4** MALDI MS spectrum of HPLC 50 min fraction (A), and MALDI MS/MS spectra of peptides 1446.6 Da (B) and 1361.7 Da (C).

sample spot was consumed or no further peptides could be selected. Figure 5.4 shows a MALDI MS spectrum obtained from one HPLC fraction (A) and MS/MS spectra obtained from two peptides in that fraction, one with high intensity (B) and one with low intensity (C). Two peptide sequences AQFEGIVTDLIR and YFPTQALNFAFK were identified based on these two MS/MS spectra and led to the identification of stress 70 protein (P38646) and ADP, ATP carrier protein, either from P05141 or P12235, which was further identified as P05141 by other peptides. Generally speaking, an average of 3-5 peptides were identified in each fraction, and a total of 157 peptides from 73 proteins were identified, as shown in protein Table 5.1. This demonstrates that the developed technique of solubilizing membrane proteins with SDS, digesting tryptically in SDS followed by semi-automated HPLC-MALDI MS and MS/MS is efficient and successful. Fifty-three of the seventy-three identified proteins were membrane or membrane associated proteins, indicating this approach is very useful in membrane protein identifications.

### 5.3.4 LC-ESI MS/MS

In ESI-MS, the nonvolatile surfactant SDS severely destroys the analyte signals [10, 35-38]. Hydrophobic proteins or peptides are usually solubilized in organic solvents for ESI analysis [39-42]. Ogorzalek Loo and coworkers examined the effect of SDS on ESI by flow injection and showed that with 0.01% and 0.1% SDS in 6 pmol/μL myoglobin, 10% of the original myoglobin signal was retrieved [9, 38]. Kirby and coworkers obtained ESI signals from 150 mM peptides with 1.25 mM (0.04%) SDS using 1.25 mM Genapol C-100 to shield the surfactant's effect [8]. Vissers, Salzmann, and coworkers stated that they successfully obtained LC-MS/MS data from 8 pmol

cytochrome *c* tryptic digest in the presence of 0.1% SDS using a SDS removal cartridge coupled with a LC column using a triple quadruple ESI instrument [43]. More recently, Hixson and coworker analyzed bacteriorhodopsin peptides by LC-ESI-MS and MS/MS after digestion in 0.015% SDS and overnight dialysis [44]. Han in the Aebersold group successfully analyzed a tryptically digested protein mixture containing less than 0.05% SDS by a 2D LC-ESI MS/MS method [45].

However, when the Thermo Finnigan LCQ Deca with Surveyer LC system was used to analyze our SDS-containing digest and the same SDS removal guard column as that used for HPLC-MALDI was used, only 10 peptides were detected by this LC ESI method. This is because the SDS-removal guard column is not 100% efficient and the remaining SDS still severely ruins the electrospray signal [8, 10]. Alternatively, we also tried using Millipore HPL hydrophobic or $C_{18}$ ZipTip to remove the SDS, or using polyethylene glycol lauryl ether to reduce the SDS suppression in ESI [8], or loading various sample amounts into the capillary column. However, all attempts failed to give as good a result as the one obtained from HPLC-MALDI experiment.

To determine the reason for the low number of proteins identified by ESI MS, another control experiment was performed. Figure 5.5 shows the ESI analysis on 8 and 0.8 pmol cytochrome c tryptic peptides with or without SDS. In the case of the analysis of 8-pmol cytochrome c tryptic peptide containing no SDS, twelve peptides were identified and the base peak chromatogram is shown in Figure 5.5A. With 0.1% SDS added into the digested sample and SDS removal guard cartridge coupled onto the LC column, the intensity of base peak chromatogram, as shown in Figure 5.5B, decreased, especially for the hydrophobic peptides which were eluted at high acetonitrile

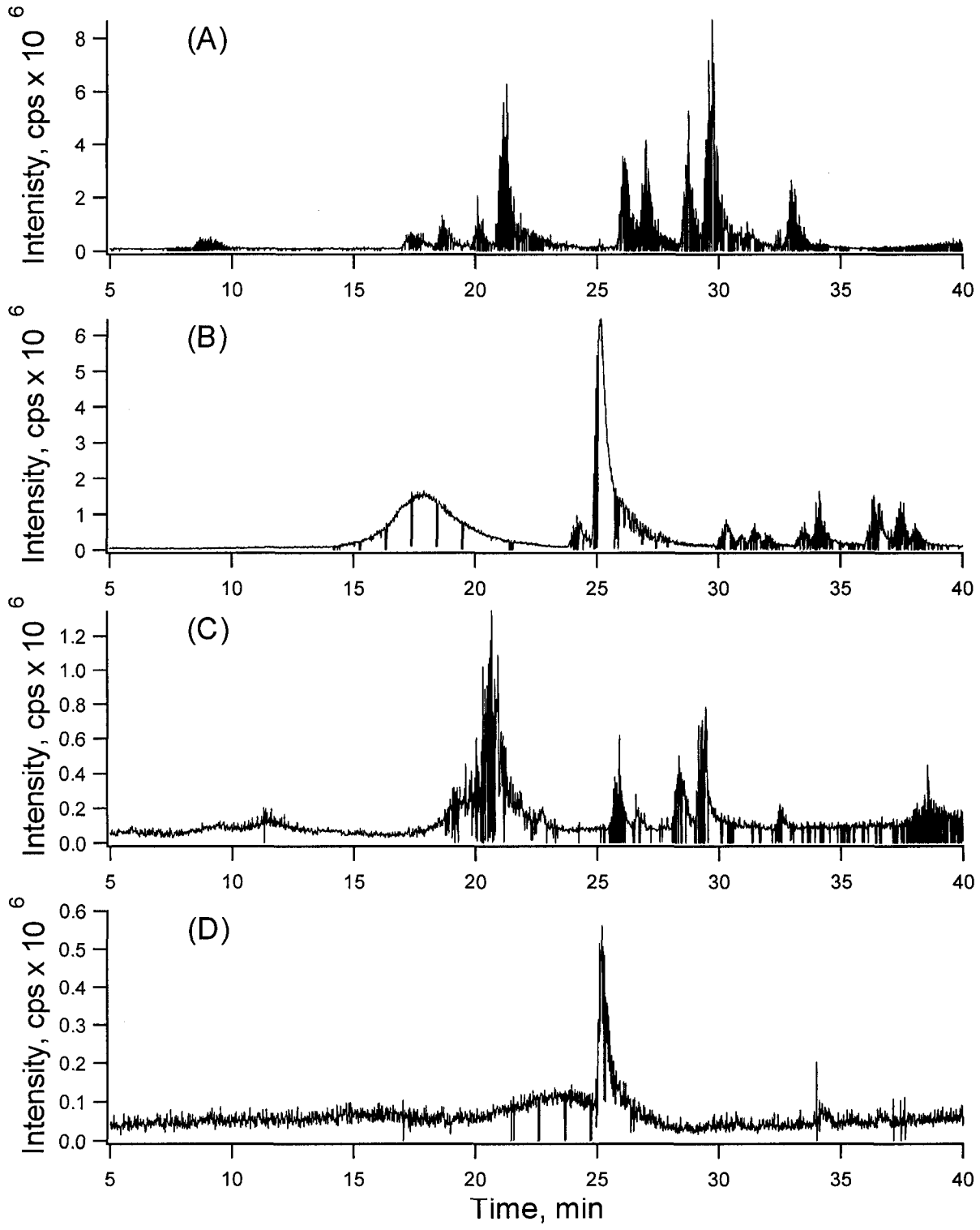



**Figure 5.5** LCQ Deca chromatograms of cytochrome c tryptic digest. (A) 8-pmol sample, (B) 8-pmol sample in 0.1% SDS with SDS removal guard column, (C) 800-fmol sample and (D) 800-fmol sample in 0.1% SDS with SDS removal guard column.

composition and only five peptides were identified. The constant elution time delays in SDS-containing samples are caused by the addition of the guard cartridge. In the case where only 0.8-pmol of the digest is used, in the absence of SDS only 3 peptides are identified (Figure 5.5C); with 0.1% SDS in presence, the SDS suppression effect becomes even more dominant and severe and no peptide is identifiable (Figure 5.5D). In other words, as the protein (or peptide) concentration decreases, the suppression effect from SDS becomes dominant and deteriorates ESI MS and MS/MS analysis. For a HT 29 cell extract which may contain hundreds of proteins, even though the total protein concentration is high, the individual protein concentration in the mixture can be very low. As a result, the presence of SDS may have a strong negative effect on ESI analysis and only a few peptides from the high quantity protein were identified by 1D ESI MS and MS/MS approach.

To further study this sample by ESI, a 2D LC separation method coupled with ESI MS and MS/MS was also applied [45]. The same amount of sample as in the HPLC-MALDI experiment was used for 2D LC-ESI experiment. The 2D separation was achieved by a cation exchange separation followed by reversed-phase fractionation. In cation exchange, cations are adsorbed onto the resin until eluted by a salt gradient. Most peptides are eluted with a salt strength of 0.1 – 0.2 M NaCl. Therefore, for a peptide mixture containing 0.1% SDS (3 mM), the sodium amount in this sample itself (3 mM) can be ignored compared to the sodium concentration in the mobile phase. Figure 5.6A shows the UV chromatogram of the cation exchange (IE) separation of 30 μg of the tryptic digest from the membrane protein mixture. Eighteen fractions (2 to 20 min) were collected based on 1 min interval, and each fraction was individually injected onto a

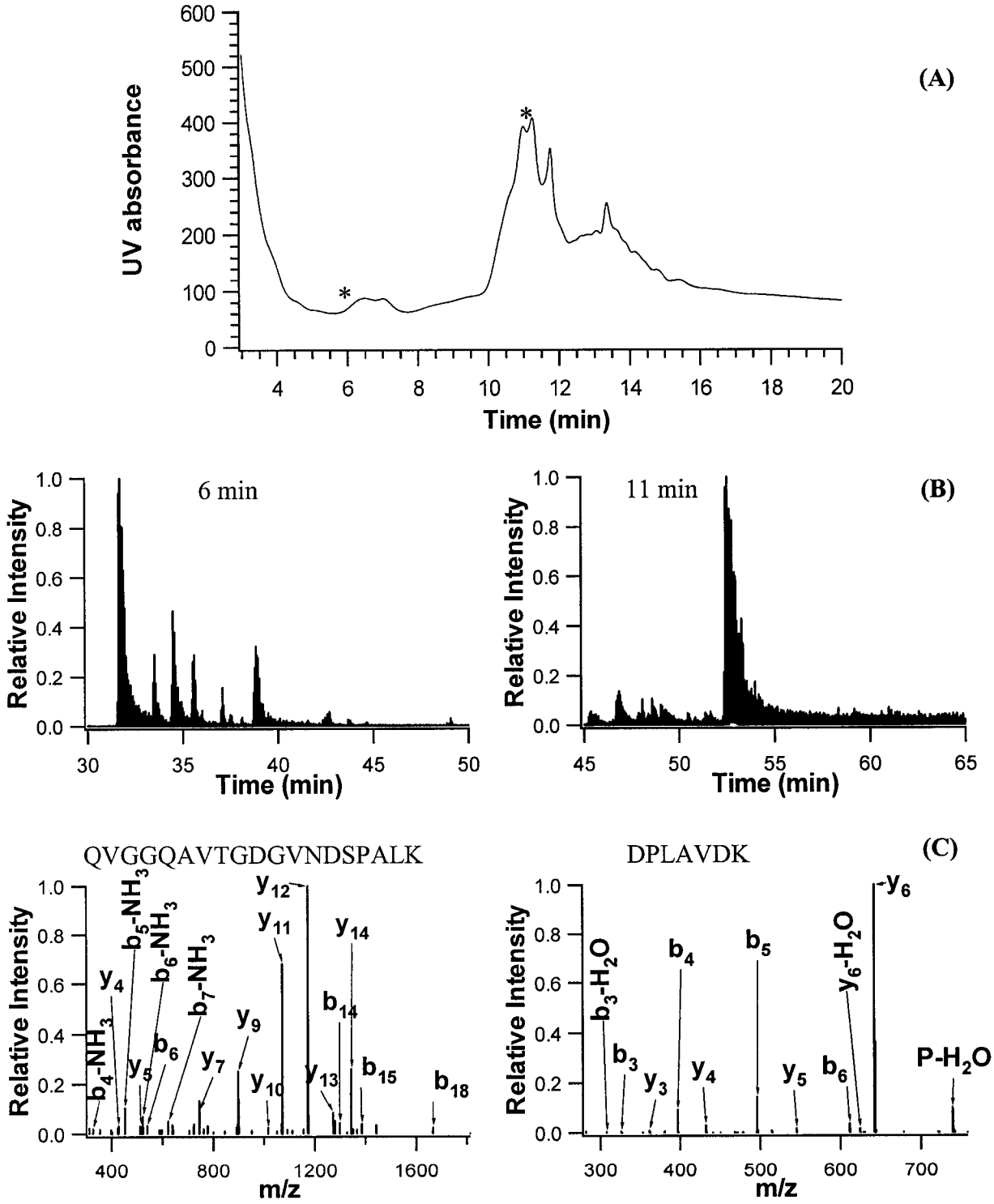



**Figure 5.6** (A) UV chromatogram of the cation exchange separation of the tryptic digest from the membrane protein mixture. (B) LC-ESI base peaks and (C) selected peptide MS/MS spectra from fractions 6 min and 11 min respectively.

reversed-phase column followed by ESI MS/MS analysis. The reversed-phase separation base peaks of IE fractions at 6 min and 11 min were shown in Figure 5.6B. The corresponding MS/MS spectra were recorded and searched for protein identities with the Mascot search engine (Figure 5.6C). In this approach, a total of 51 peptides were analyzed from 37 unique proteins. Among the proteins identified, 20 had already been identified by the HPLC-MALDI MS and MS/MS analysis and 17 were newly identified with this 2D LC-ESI approach. The proteins identified by the 2D LC-ESI MS/MS approach were also listed in Table 5.1.

Generally speaking, less proteins are identified by the ESI method. This is not surprising as offline 2D LC involves several steps, such as fraction collection, sample solvent evaporation and re-injection onto a reversed-phase separation column. Conversely, semi-offline HPLC-MALDI separation was achieved as a one-step process — the eluate from LC separation was directly deposited onto a MALDI target, which significantly reduces the sample loss during any offline sample handling processes. Furthermore, several researchers have reported that complementary peptides can be identified from the same protein sample by ESI and MALDI MS [46-48].

In general, combining the proteins identified by the MALDI and ESI approaches, a total of 90 proteins were identified from the HT 29 membrane fraction. This indicates that the developed solution-based SDS-aided solubilization and digestion of membrane protein method is effective and practical.

### 5.3.5 Protein Identification by SDS-PAGE, In-Gel Digestion, Peptide Mass Mapping and Sequencing

To further validate the use of the developed in-solution membrane protein identification approach, a comparison experiment using the same membrane-enriched protein sample was performed using the in-gel approach. In this experiment, one-dimensional SDS-PAGE is employed to separate the membrane protein sample to avoid the possible solubilization problems often encountered in the rehydration step in 2D gel separation [5-7]. About 30 μg of protein extracts, the same amount injected for the HPLC-MALDI and LC-ESI experiments, was mixed with SDS sample buffer and loaded onto a 1D gel. The commassie blue stained gel is shown in Figure 5.7. All visible gel bands were excised, in-gel tryptically digested, and identified with MALDI MS and MALDI MS/MS (see experimental section). For example, Figure 5.8 shows the MALDI MS spectrum of the tryptic peptides digested from the gel band indicated with the arrow in Figure 5.7. All strong peptide peaks were loaded into the database for protein identification by peptide mass mapping. The unmatched peptide peaks from peptide mass mapping were sequenced for potential protein identification. This is possible as several proteins with similar molecular weight can co-migrate into the same band position. In this particular gel band, protein ATP synthase B chain, P24539, was first tentatively identified by peptide mass mapping and all peptides that may belong to this protein were marked as "1" in Figure 5.8A. This identification was confirmed by the MALDI MS/MS spectrum of a peptide HYLFDVQR (m/z 1077.52) specific to protein ATP synthase B chain as depicted in Figure 5.8B. For the unmatched peaks, MS sequence query was used for protein identification and the same criteria as HPLC-MALDI MS/MS experiment were also used here. Figure 5.8C is a MALDI MS/MS spectrum from an unmatched peptide from peptide mass mapping results, which led to

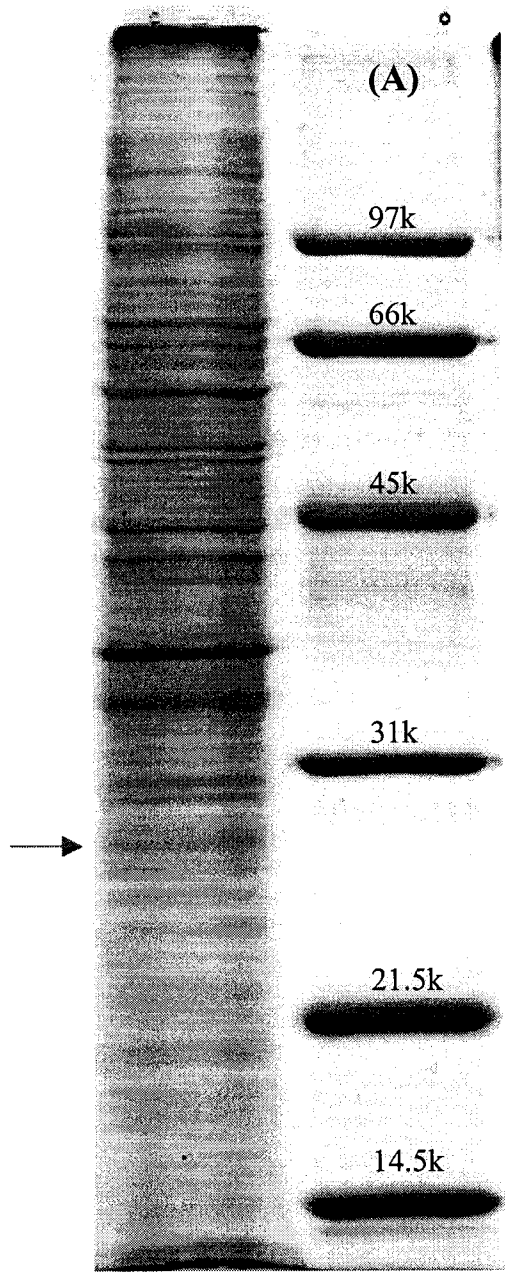



**Figure 5.7** SDS-PAGE Gel image of separated membrane proteins. Arrow indicated band was analyzed in Figure 5.8.

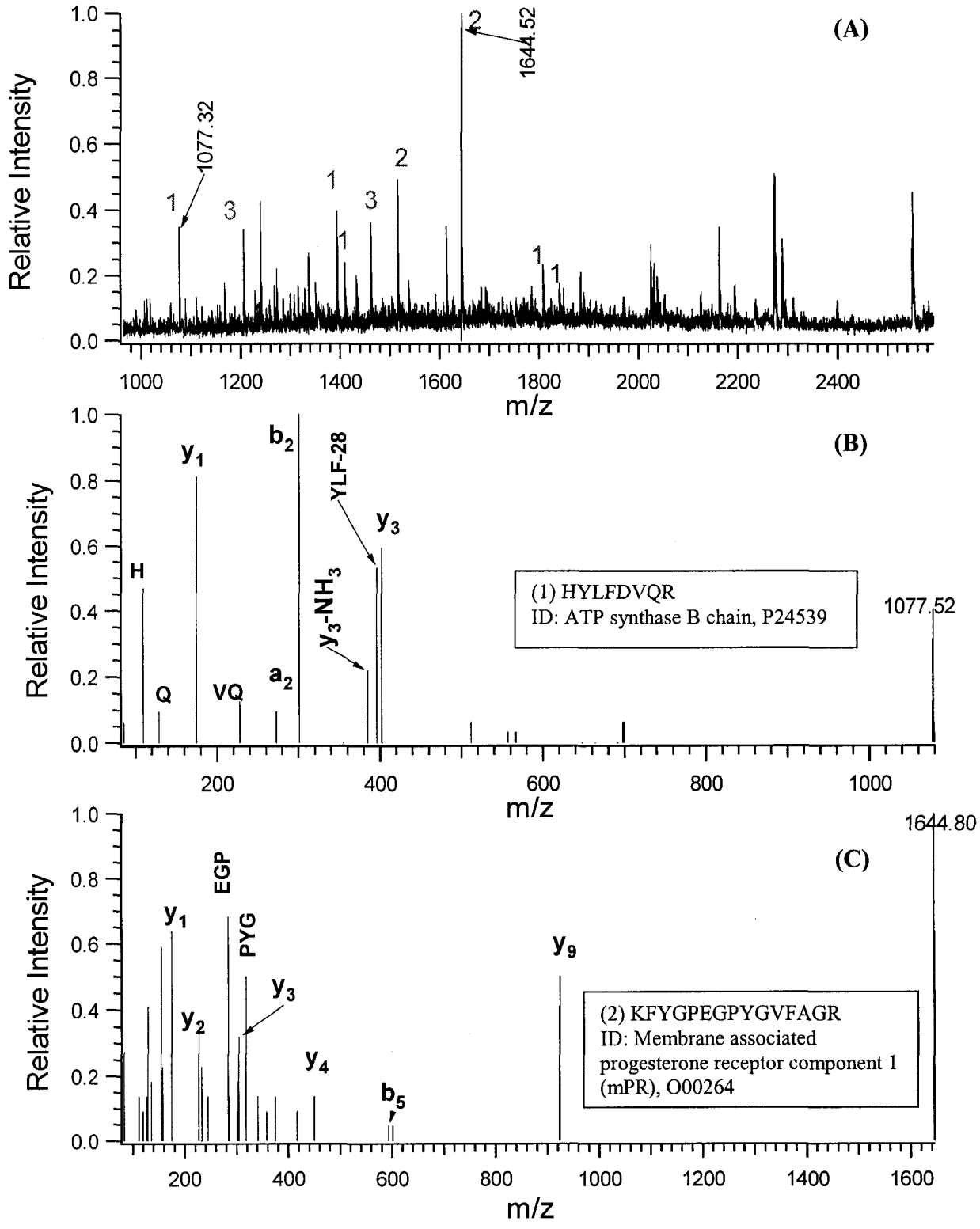



**Figure 5.8** (A) MALDI MS spectrum of tryptic peptides from gel band marked with arrow in Figure 5.7. MALDI MS/MS spectra of peptides 1077.5 Da (B) and 1644.8 Da (C) along with their sequences and the protein IDs.

the identification of the protein "membrane associated progesterone receptor component 1 (mPR)", O00264. A theoretical trypsin digest was performed on this newly identified protein and one of the unmatched peptides that matches the newly identified protein (from peptide 1644.8 Da) was subjected to MALDI MS/MS sequencing and identification, followed by labelling all peptides that belong to the second protein as "2". By this method, fewer samples are consumed to identify more proteins present in this band. This process was repeated several times until sample is consumed or no more protein can be identified.

By this gel approach, 45 proteins are identified and among them 37 proteins are membrane or membrane-associated proteins and eight are cytoplasmic proteins.

### 5.3.6 Discussion

Table 5.1 lists all the proteins identified from this membrane protein fraction of HT29 cells using in-gel separation/peptide mass mapping/MALDI MS/MS, in-solution based 1D HPLC-MALDI MS/MS and in-solution based 2D LC-ESI MS/MS. Figure 5.9 summarizes the results obtained from the three methods. As Figure 5.9A shows, a total of 112 proteins were identified. LC-MALDI identified 73 proteins, compared to 37 from 2D LC-ESI and 45 from the gel-based method. There were proteins identified by more than one method, as well as a number of proteins identified by a single method. This finding is not surprising considering the significant differences in experimental conditions used among the three methods. It is clear that for comprehensive proteome profiling, these methods complement each other.

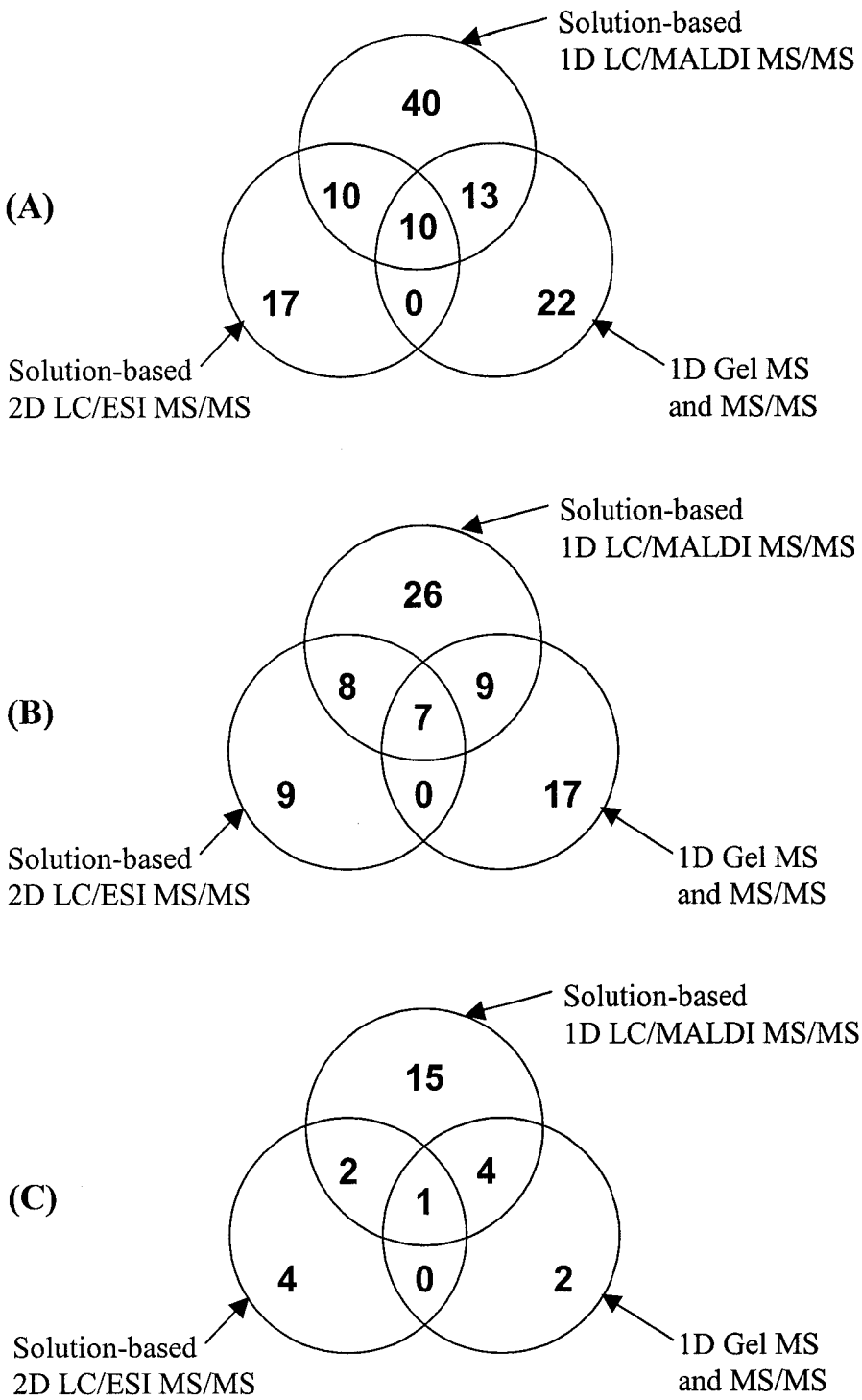



**Figure 5.9** Schematic graph shows the overlap of proteins identified by solution-based HPLC-MALDI MS/MS method, solution-based LC-ESI MS/MS and gel-peptide mass mapping-MALDI MS/MS. (A) A total of 112 proteins were identified. Among them, (B) 76 are membrane and membrane-associated proteins, which include (C) 28 integral membrane proteins.

One significant observation is that with 1D LC-MALDI, a greater number of proteins were identified compared to the other methods. This difference is even more striking when only the membrane proteins are considered. Among the 112 proteins identified, 76 proteins are membrane and membrane-associated proteins including 28 integral membrane proteins. Figure 5.9B illustrates that 50 of the 76 membrane proteins were identified by LC-MALDI (66%) vs. 24 (32%) from 2D LC-ESI and 33 (43%) from the gel method. For the 28 integral membrane proteins (Figure 5.9C), 22 were identified by LC/MALDI (79%), 7 were found by 2D LC-ESI (25%), and 7 were detected by the gel method (25%). The challenge of detecting membrane proteins, particularly integral membrane proteins, is well documented [6, 49] and thus the identification results from the gel method are not unexpected. But it is surprising to discover that there is a great difference in membrane protein detection ability between the ESI and MALDI proteomic methods. This difference may be attributed to several factors including the possibility of higher efficiency of MALDI in ionizing more hydrophobic peptides and the concentration effect during LC fractionation in LC-MALDI that enriches the peptides cleaved from low abundance membrane proteins and thus increases the probability of their detection.

Ninety proteins were identified with in-solution approach while only 45 proteins identified with in-gel approach. For the 28 integral membrane proteins, 26 were covered in the in-solution approach. The discussion below lists several possible reasons behind this observation.

Proteins only identified by the in-gel approach might be a result of digestion bias by in-solution digest. Proteins in solution can have very different concentrations and

some may be much easier to digest than others. Those relatively low abundance proteins or difficult-to-digest proteins may not be able to compete for trypsin effectively with those proteins with high concentrations or easier-to-digest ones. As a result, digestion is not efficient and complete and some proteins will be missing from the identification lists. On the contrary, in the in-gel approach, proteins are first separated and localized in different positions of the gel. Each protein, to a certain degree, is "enriched" in its own molecular weight region for effective digestion.

Several reasons can be proposed to explain why so many more proteins were identified by the in-solution digest approach. First of all, hydrophobic proteins might not be fully denatured by gel sample buffer or the conformation might be restored during the sample loading process. As a result, these proteins may not migrate into the gel or only a small portion of the total amount of these proteins is actually being separated. This might be the reason why a dark band was observed at the top of the separating gel in the sample lane, while there is nothing in the same region in the protein standard marker lane (Figure 5.7). If the proteins are not present in the gel bands, they will not be identifiable by the in-gel approach. Second, in the in-solution identification approach, proteins are not only solubilized by SDS but also digested in the presence of SDS. SDS will maintain proteins in the denatured condition during the entire digestion process which in turn aids the enzyme to work on the proteins, especially hydrophobic proteins. This is opposite to the in-gel digestion process where extensive washes are applied to remove SDS and other contaminants. Third, extraction of peptides from gel pieces after in-gel digestion, especially those hydrophobic peptides, may not be very efficient and adsorption on vial walls of these peptides can also be a severe problem, while these factors are not present

in this developed in-solution technique. Last but not least, co-migrated hydrophilic proteins in a gel band can suppress the digestion, extraction and detection of hydrophobic peptides [13, 50], leading to the failure to identify these hydrophobic proteins, even though additional LC-MS/MS experiment after in-gel digestion can solve the detection bias problem.

One protein identified in this in-solution method, ADP, ATP carrier protein (SwissProt P05141) will be used as an example for the above discussion. This protein is an integral membrane protein, has a molecular weight of 32.9 kDa, and contains 6 trans-membrane regions. Six peptides including three in the transmembrane regions were detected by the in-solution method which conclusively identified this protein. Although this protein was not identified in the gel experiment, we found two small peptide peaks that match the masses of ADP, ATP carrier protein's theoretic tryptic digest after careful examination of the MALDI MS spectra from the band with molecular weight around 33 kDa. These two peptides may belong to this membrane protein, but they were too weak to obtain good MS/MS spectra for a positive identification. Again, as a hydrophobic membrane protein, it might partially migrate into the gel and the amount is not enough for a successful digestion. Trypsin may not successfully target this membrane protein and the extraction and detection of hydrophobic peptides are problematic as well. More interestingly, the band at the 33 kDa region which gave the two small peaks that may belong to this membrane protein was identified to be Prohibitin (Swiss Prot P35232), a very hydrophilic protein. This strongly supports the above hypotheses.

It is clear that the developed method of in-solution digestion in SDS and MS/MS identification of membrane proteins is successful and very useful in membrane proteomic

research. However, neither this method nor the routine in-gel approach is perfect. They each have their own advantages and merits and supply complementary information to each other.

The developed in-solution approach could be further improved. It is not difficult to imagine that a rough separation of protein samples to reduce the sample complexity will increase the coverage range of the proteome. More efficient peptide separation with, for example, 2D-HPLC can also decrease the suppression effect and thus increase the total number of detectable peptides. Another way is to increase the HPLC sample loading, and this could be easily achieved by utilizing a larger HPLC column (2.1 mm ID column is still compatible with our semi-offline HPLC-MALDI technique). By increase the sample loading amount, the low abundance proteins may be identified as the increasing of absolute sample amount. The in-solution approach can also combine with ICAT technique [4] for quantification of membrane proteins.

## 5.4 Conclusions

In this work, we have demonstrated a new solution-based digestion and identification method for membrane proteomic research in the presence of SDS. Membrane protein analysis usually requires the use of surfactants for extraction and solubilization, which increases the difficulty of analyzing them by solution-based LC-ESI MS/MS and thus the identification is usually achieved by a gel-based approach. In this solution approach, we use SDS to solubilize membrane proteins without worrying about the subsequent tryptic digestion, HPLC separation, and peptide detection and sequencing experiments. Both HPLC-MALDI and 2D LC-ESI were applied for peptide sequencing

and protein identification. However, the in-solution approach combined with LC-MALDI MS/MS offers several unique advantages over the LC-ESI MS/MS. The use of multidimensional LC-MALDI can further increase the number of proteins detected by this method. This work clearly illustrates that, for comprehensive analysis of proteomes of interest, a preferred proteomics approach should incorporate all the three technologies.

## 5.5 Literature Cited


(1) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates III, J. R. *Nat. Biotech.* **1999,** *17*, 676.

(2) Wagner, K.; Miliotis, T.; Marko-Varga, G.; Bischoff, R.; Unger, K. K. *Anal. Chem.* **2002,** *74*, 809-820.

(3) Wolters, D. A.; Washburn, M. P.; Yates III, J. R. *Anal. Chem.* **2001,** *73*, 5683-5690.

(4) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. *Nat. Biotechnol.* **1999,** *17*, 994-999.

(5) Wilkins, M. R.; Gasteriger, E.; Sanchez, J.-C.; Bairoch, A.; Hochstrasser, D. F. *Electrophoresis* **1998,** *19*, 1501-1505.

(6) Santoni, V.; Molloy, M.; Rabilloud, T. *Electrophoresis* **2000,** *21*, 1054-1070.

(7) Adessi, C.; Miege, C.; Albrieux, C.; Rabilloud, T. *Electrophoresis* **1997,** *18*, 127-135.

(8) Kirby, D.; Greve, K. F.; Foret, F.; Vouros, P.; Karger, B. L. *Proceedings of the 42nd Conference of the American Society for Mass Spectrometry and Allied Topics*, Chicago, IL, May 29-June 3 1994; 1014.

(9) Ogorzalek Loo, R. R.; Dales, N.; Andrews, P. C. *Protein Sci.* **1994,** *3*, 1975-1983.

(10) Rundlett, K. L.; Armstrong, D. W. *Anal. Chem.* **1996,** *68*, 3493-3497.

(11) Kinter, M.; Sheerman, N. E. In *Protein Sequencing and Identification Using Tandem Mass Spectrometry*; Wiley-Interscience: New York, 2000. pp38-39.

(12) Zhang, N.; Doucette, A.; Li, L. *Anal. Chem.* **2001,** *73*, 2968-2975.

(13) Breaux, G. A.; Green-Church, K. B.; France, A.; Limbach, P. A. *Anal. Chem.* **2000,** *72*, 1169-1174.

(14) Scherer, C. A.; Cooper, E.; Miller, S. I. *Mol. Microbiol.* **2000,** *37*, 1133-1145.

(15) Immler, D.; Mehl, E.; Lottspeich, F.; Meyer, H. E. In *Microcharacterization of Proteins*; Kellner, R., Lottspeich, F., Meyer, H. E., Eds.; Wiley-VCH: New York, 1999, pp 87-93.

(16) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. *Anal. Chem.* **1996,** *68*, 850-858.

(17) Zhang, B.; McDonald, C. M.; Li, L. manuscripts submitted to *Anal. Chem.* **2003**.

(18) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1996,** *68*, 2494-2500.

(19) Dai, Y.; Whittal, R. M.; Li, L. *Anal. Chem.* **1999,** *71*, 1087-1091.

(20) Schnaible, V.; Michels, J.; Zeth, K.; Freigang, J.; Welte, W.; Buehler, S.; Glocker, M. O.; Przybylski, M. *Int. J. Mass Spectrom. Ion Processes* **1997,** *169/170*, 165-177.

(21) Xiang, F.; Beavis, R. C. *Organic Mass Spectrom.* **1993,** *28*, 1424-2429.

(22) Amado, F. M. L.; Santana-Marques, M. G.; Ferrer-Correia, A. J.; Tomer, K. B. *Anal. Chem.* **1997,** *69*, 1102-1106.

(23) Shore, L. J.; Odell, G. B.; Fenselau, C. *Biochem. Pharmacol.* **1995,** *49*, 181-186.

(24) Eckerskorn, C.; Strupat, K.; Schleuder, D.; Dennis Hochstrasser, D.; Sanchez, J.-C.; Lottspeich, F.; Hillenkamp, F. *Anal. Chem.* **1997,** *69*, 2888-2892.

(25) Galvani, M.; Hamdan, M. *Rapid Commun. in Mass Spectrom.* **2000,** *14*, 721-723.

(26) Katayama, H.; Nagasu, T.; Oda, Y. *Rapid Commun. in Mass Spectrom.* **2001,** *15*, 1416-1421.

(27) Barnidge, D. R.; Dratz, E. A.; Sunner, J.; Jesaitis, A. J. *Protein Sci.* **1997,** *6*, 816-824.

(28) Krungkrai, J.; Kanchanarithisak, R.; Krungkrai, S. R.; Rochanakij, S. *Experimental Parasitology* **2002,** *100*, 54-61.

(29) Mann, M.; Wilm, M. *Anal. Chem.* **1994,** *66*, 4390-4399.

(30) Bonner, R.; Shushan, B. *Rapid Commun. in Mass Spectrom.* **1995,** *9*, 1077-1080.

(31) Gooley, A. A.; Ou, K.; Russell, J.; Wilkins, M. R.; Sanchez, J.-C.; Hochstrasser, D. F.; Williams, K. L. *Electrophoresis* **1997,** *18*, 1068-1072.

(32) Loboda, A. V.; Krutchinsky, A. N.; Bromirski, M.; Ens, W.; Standing, K. G. *Rapid Commun. in Mass Spectrom.* **2000,** *14*, 1047-1057.

(33) Zhang, N.; Li, N.; Quach, T.; Keller, B. O.; Li, L. *Proceedings of the 50th Conference of the American Society for Mass Spectrometry and Allied Topics*, Orlando, Florida, June 2-6 2002.

(34) Raska, C. S.; Parker, C. E.; Huang, C.; Han, J.; Glish, G. L.; Pope, M.; Borchers, C. H. *J. Am. Soc. Mass Spectrom.* **2002,** *13*, 1034-1041.

(35) Tomer, K. B.; Deterding, L. J.; Parker, C. E. In *Advances in Chromatography*; Brown, P. R., Grushka, E., Eds.; Marcel Dekker: New York, 1995; Vol. 35, pp Charpter 2.

(36) Smith, R. D.; Wahl, J. H.; Goodlett, D., R.; Hofstadler, S. A. *Anal. Chem.* **1993,** *65*, 574A-584A.

(37) Cai, J.; Henion, J. *J. Chromatogr. A* **1995,** *703*, 667-692.

(38) Ogorzalek Loo, R. R.; Dales, N.; Andrews, P. C. *Methods Mol. Biol.* **1996,** *61*, 141-160.

(39) Hufnagel, P.; U, S.; Eckerskorn, C.; Oesterhelt, D. *Anal. Biochem.* **1996,** *243*, 46-54.

(40) Maire, M. L.; Deschamps, S.; Moller, J. V.; Caer, J. P. l.; Rossier, J. *Anal. Biochem.* **1993,** *214*, 50-57.

(41) Schindler, P. A.; Dorsselaer, A. V.; Falick, A. M. *Anal. Biochem.* **1993,** *213*, 256-263.

(42) Barnidge, D. R.; Dratz, E. A.; Jesaitis, A. J.; Sunner, J. *Anal. Biochem.* **1999,** *269*, 1-9.

(43) Vissers, J., P. C.; Chervet, J.-P.; Salzmann, J.-P. *J. Mass Spectrom.* **1996,** *31*, 1021-1027.

(44) Hixson, L. K.; Rodriguez, N.; Camp II, D. G.; Strittmatter, E. F.; Lipton, M. S.; Smith, R. D. *Electrophoresis* **2002,** *23*, 3224-3232.

(45) Han, D. K.; Eng, J.; Zhou, H.; Aebersold, R. *Nat. Biotechnol.* **2001,** *19*, 946-951.

(46) Saundres, C. J.; Frankel, L. K.; Green-Church, K. B.; Tummala, R.; Bricker, T.; Limbach, P. A. *Proceedings of the 48th Conference of the American Society for Mass Spectrometry and Allied Topics*, Long Beach, CA, June 11-15 2000; 489-490.

(47) Ball, L. E.; Oatis, J. E., Jr.; Dharmasiri, K.; Busman, M.; Wang, J.; Cowden, L. B.; Galijatovic, A.; Chen, N.; Crouch, R. K.; Knapp, D. R. *Protein Sci.* **1998,** *7*, 758-764.

(48) Hansen, K. C.; Schmitt-Ulms, G.; Chalkley, R. J.; Hirsch, J.; Baldwin, M. A.; Burlingame, A. L. *Mol. Cell Proteomics* **2003,** *2*, 299-314.

(49) Wu, C. C.; Yates III, J. R. *Nat. Biotechnol.* **2003,** *21*, 262-267.

(50) Keller, B.; Li, L. *Anal. Chem.* **2001,** *73*, 2929-2936.

**Table 5.1 Total Proteins Identified by Three Approaches Evaluated in This Chapter.**

| # | Protein Description | Swiss-prot | MW (kDa) | HPLC-MALDI | 2D LC-ESI | Gel | Subcellular location |
|---|---|---|---|---|---|---|---|
| 1 | Ubiquitin | P02248 | 8.6 | N | Y | N | Nuclear and cytoplasmic |
| 2 | 10 kDa heat shock protein, mitochondrial (Hsp10) | Q04984 | 10.8 | Y | Y | N | |
| 3 | ATP synthase G chain, mitochondrial | O75964 | 11.4 | Y | Y | N | |
| 4 | Calpactin I light chain | P08206 | 11.7 | N | Y | N | Annexin ligand |
| 5 | CD59 glycoprotein (Membrane attack complex inhibition factor) | P13987 | 14.2 | Y | N | N | Attached to the membrane by a GPI-anchor |
| 6 | Cytochrome b5 STANDARD VARSPLIC | P00167 | 15.2 | N | N | Y | Microsomal membrane |
| 7 | 40S ribosomal protein S19 | P39019 | 15.9 | Y | N | N | |
| 8 | Cytochrome b5 outer mitochondrial membrane isoform | O43169 | 16.3 | Y | N | N | Mitochondrial outer membrane |
| 9 | Calmodulin STANDARD VARSPLIC | P02593 | 16.7 | N | N | Y | Calcium binding protein |
| 10 | ATP synthase delta chain | P30049 | 17.4 | N | Y | N | Mithchondrial |
| 11 | Peptidyl-prolyl cis-trans isomerase A | P05092 | 17.9 | N | N | Y | Cytoplasmic |
| 12 | ATP synthase D chain | O75947 | 18.4 | N | N | Y | |
| 13 | Thioredoxin | Q99757 | 18.4 | Y | N | N | Mitochondrial |
| 14 | Translocon-associated protein, delta subunit | P51571 | 19.0 | N | N | Y | Type I membrane protein |
| 15 | Splicing factor, arginine/serine-rich 3 | P23152 | 19.3 | Y | Y | Y | Nuclear |
| 16 | Cytochrome c oxidase subunit IV isoform 1 | P13073 | 19.6 | N | N | Y | Mitochondrial inner membrane |
| 17 | Membrane associated progesterone receptor component 1 | O00264 | 21.5 | Y | N | Y | Membrane-bound |
| 18 | Transmembrane 4 superfamily, member 1 | P30408 | 21.6 | Y | N | N | Integral membrane protein |
| 19 | NADH-ubiquinone oxidoreductase B22 subunit | Q9Y6M9 | 21.7 | N | N | Y | Mitochondrial inner membrane |
| 20 | Programmed cell death protein 6 | O75340 | 21.9 | N | N | Y | Calcium binding domains |
| 21 | Claudin-9/O15551/O14493/P56747 | O95484 | 22.8 | Y | N | N | Integral membrane protein |
| 22 | Glutathione S-transferase P | P09211 | 23.2 | N | N | Y | |
| 23 | 60S ribosomal protein L14 | P50914 | 23.2 | Y | N | N | |
| 24 | ATP synthase oligomycin sensitivity conferral protein | P48047 | 23.3 | N | N | Y | Mitochondrial matrix |
| 25 | 60S ribosomal protein L13A | P40429 | 23.4 | N | Y | N | |
| 26 | Ras-related protein Rab-7 | P51149 | 23.5 | N | N | Y | Found on late endosomes |
| 27 | Ras-related protein Rab-5C | P51148 | 23.6 | Y | N | N | Lipid Binding protein |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 28 | ER lumen protein retaining receptor 1 | P24390/ P33947 | 24.5 | Y | N | N | Integral membrane protein |
| 29 | CD9 antigen | P21926 | 25.3 | Y | N | N | Integral membrane protein |
| 30 | CD63 antigen (Melanoma-associated antigen ME491) | P08962 | 25.5 | Y | N | N | Integral membrane protein |
| 31 | Splicing factor, arginine/serine-rich 1 | Q07955 | 27.6 | Y | N | N | Nuclear |
| 32 | Thioredoxin-dependent peroxide reductase, mitochondrial | P30048 | 27.7 | Y | N | Y | Mitochondrial |
| 33 | Electron transfer flavoprotein beta-subunit | P38117 | 27.8 | Y | N | N | Mitochondrial matrix |
| 34 | B-cell receptor-associated protein 31 | P51572 | 28.0 | Y | Y | N | Integral membrane protein |
| 35 | ATP synthase B chain, mitochondrial | P24539 | 28.9 | N | N | Y | Mitochondrial |
| 36 | 40S ribosomal protein S4, X isoform | P12750 | 29.5 | N | N | Y | |
| 37 | Prohibitin | P35232 | 29.8 | Y | Y | Y | Cytoplasmic |
| 38 | NADH-ubiquinone oxidoreductase 30 kDa subunit | O75489 | 30.2 | Y | N | Y | Matrix and cytoplasmic side of the mitochondrial inner membrane |
| 39 | Surfeit locus protein 4 STANDARD VARSPLIC | O15260 | 30.4 | Y | N | N | Integral membrane protein, ER |
| 40 | Enoyl-CoA hydratase, mitochondrial | P30084 | 31.4 | N | N | Y | Mitochondrial matrix |
| 41 | Nucleophosmin | P06748 | 32.6 | Y | N | N | Nuclear |
| 42 | Syndecan-1 | P18827 | 32.6 | N | Y | N | Type I membrane protein |
| 43 | ADP,ATP carrier protein, fibroblast isoform | P05141 | 32.9 | Y | N | N | Mitochondrial inner membrane |
| 44 | Thiosulfate sulfurtransferase | Q16762 | 33.3 | N | N | Y | Mitochondrial matrix |
| 45 | Heterogeneous nuclear ribonucleoproteins C1/C2 (hnRNP C1 / hnRNP C2) | P07910 | 33.7 | Y | N | Y | Nuclear |
| 46 | Malate dehydrogenase, mitochondrial | P40926 | 35.5 | Y | N | N | Mitochondrial matrix |
| 47 | Annexin A4 (Annexin IV) | P09525 | 35.8 | Y | N | Y | binding site for calcium and phospholipid. |
| 48 | Annexin V | P08758 | 35.8 | Y | Y | Y | Mitochondrial matrix |
| 49 | Galectin-4 (Lactose-binding lectin 4) | P56470 | 35.9 | Y | N | Y | |
| 50 | Delta3,5-delta2,4-dienoyl-CoA isomerase | Q13011 | 36.0 | Y | Y | N | mitochondrial and peroxisomal |
| 51 | Annexin III (Lipocortin III) | P12429 | 36.4 | Y | N | N | binding site for calcium and phospholipid. |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 52 | Heterogeneous nuclear ribonucleoproteins A2/B1 | P22626 | 37.4 | Y | N | N | Nuclear |
| 53 | Probable mitochondrial import receptor subunit TOM40 homolog | O96008 | 37.9 | Y | N | Y | Integral membrane protein, Mitochondrial outer membrane |
| 54 | Voltage-dependent anion-selective channel protein 2 | P45880 | 38.1 | Y | Y | N | Outer mitochondrial membrane |
| 55 | Annexin I (Lipocortin I) | P04083 | 38.6 | Y | N | N | |
| 56 | Heterogeneous nuclear ribonucleoprotein A1 | P09651 | 38.7 | Y | N | N | Nuclear |
| 57 | Phosphate carrier protein | Q00325 | 40.1 | Y | Y | N | Mitochondrial inner membrane |
| 58 | Actin (alpha or beta) | P02570 | 41.7 | Y | N | Y | Cytoplasmic |
| 59 | 3-ketoacyl-CoA thiolase, mitochondrial | P42765 | 42.0 | Y | N | N | Mitochondrial |
| 60 | Splicing factor, arginine/serine-rich 7 | P16629 | 42.3 | N | N | Y | |
| 61 | Annexin II (Lipocortin II) | P07355 | 42.7 | Y | Y | Y | Lipid Binding protein |
| 62 | Flotillin-1 | O75955 | 47.3 | Y | N | N | |
| 63 | Aspartate aminotransferase | P00505 | 47.5 | N | N | Y | Mitochondrial matrix |
| 64 | Ubiquinol-cytochrome C reductase complex core protein 2 | P22695 | 48.4 | Y | Y | N | Mitochondrial inner protein |
| 65 | Heterogeneous nuclear ribonucleoprotein H | P31943 | 49.2 | Y | N | N | Nuclear |
| 66 | Elongation factor Tu | P49411 | 49.5 | Y | N | N | Mitochondrial |
| 67 | Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit | P39656 | 50.7 | Y | N | Y | Type I membrane protein |
| 68 | Heterogeneous nuclear ribonucleoprotein K | Q07244 | 51.0 | Y | Y | N | Nuclear |
| 69 | Citrate synthase, mitochondrial | O75390 | 51.7 | N | N | Y | Mitochondrial |
| 70 | Ubiquinol-cytochrome C reductase complex core protein I | P31930 | 52.6 | N | N | Y | Mitochondrial inner membrane |
| 71 | Keratin, type II cytoskeletal 8 | P05787 | 53.5 | Y | N | N | Cytoplasmic |
| 72 | Large neutral amino acids transporter small subunit 1 | Q01650 | 55.7 | N | Y | N | Integral membrane protein |
| 73 | ATP synthase beta chain, mitochondrial | P06576 | 56.6 | Y | Y | Y | Mitochondrial |
| 74 | AHNAK-related protein | Q13727 | 57.7 | N | Y | N | |
| 75 | 4F2 cell-surface antigen heavy chain (4F2hc) | P08195 | 57.9 | Y | N | N | Type II membrane protein |
| 76 | T-complex protein 1, delta subunit | P50991 | 58.3 | N | Y | N | Cytoplasmic |
| 77 | ATP synthase alpha chain | P25705 | 59.8 | Y | Y | Y | Mitochondrial inner membrane |
| 78 | 60 kDa heat shock protein | P10809 | 61.1 | Y | Y | Y | Mitochondrial matrix |
| 79 | Lamin B1 | P20700 | 66.3 | Y | N | N | Lipid Binding protein |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 80 | Calnexin | P27824 | 67.6 | Y | N | Y | Type I membrane protein, ER |
| 81 | Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 67 kDa subunit | P04843 | 68.6 | Y | Y | Y | Type I membrane protein |
| 82 | Probable RNA-dependent helicase p72/p68 | Q92841/ P17844 | 69.1 | Y | N | N | |
| 83 | Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 63 kDa subunit | P04844 | 69.3 | Y | N | N | Type I membrane protein |
| 84 | Matrix metalloproteinase-16 | P51512 | 70 | N | Y | N | Type I membrane protein |
| 85 | Acyl-CoA dehydrogenase | P49748 | 70.7 | N | Y | N | Mitochondrial inner membrane. |
| 86 | Heat shock cognate 71 kDa protein | P11142 | 70.9 | Y | Y(2/3) | Y | |
| 87 | Heterogeneous nuclear ribonucleoprotein R | O43390 | 70.9 | Y | N | N | Nuclear |
| 88 | Succinate dehydrogenase [ubiquinone] flavoprotein subunit | P31040 | 72.7 | Y | N | N | Mitochondrial inner membrane |
| 89 | Stress-70 protein | P38646 | 73.7 | Y | Y(1/3) | Y | Mitochondrial |
| 90 | Nucleolin | P19338 | 76.2 | Y | Y | N | Nuclear |
| 91 | NADPH-cytochrome P450 reductase | P16435 | 76.6 | N | N | Y | Anchored to the ER membrnane by its N-terminal hydrophobic region |
| 92 | NADH-ubiquinone oxidoreductase 75 kDa subunit, | P28331 | 79.6 | N | N | Y | Matrix and cytoplasmic side of the mitochondrial inner membrane |
| 93 | Heat shock protein 75 kDa | Q12931 | 80.0 | Y | N | Y | mitochondrial |
| 94 | Trifunctional enzyme alpha subunit | P40939 | 83.7 | N | Y | N | Mitochondrial matrix |
| 95 | Delta 1-pyrroline-5-carboxylate synthetase | P54886 | 87.3 | Y | N | Y | Mitochondrial inner membrane |
| 96 | Integrin beta-1 (Fibronectin receptor beta subunit) (CD29 antigen) | P05556 | 88.5 | Y | N | N | type I membrane protein |
| 97 | Heterogenous nuclear ribonucleoprotein U (hnRNP U) | Q00839 | 90.5 | Y | N | N | Nuclear |
| 98 | RasGAP-activating-like protein 1 | O95294 | 90.9 | N | Y | N | |
| 99 | Alpha-1 catenin | P35221 | 100.7 | N | Y | N | Found at cell-cell boudary |
| 100 | Hexokinase, type I | P19367 | 102.5 | Y | N | N | membrane bound |
| 101 | Catenin delta-1 (p120 catenin) | O60716 | 108.2 | Y | N | N | Nuclear |
| 102 | Sodium/potassium-transporting ATPase alpha-1 chain (Commom one in their family) (P13637/P50993) | P05023 | 112.9 | Y | Y | N | Integral membrane protein |
| 103 | Sarcoplasmic/endoplasmic reticulum calcium ATPase 2 | P16615 | 114.7 | Y | N | N | integral membrane protein |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 104 | Potassium-transporting ATPase alpha chain 2 | P54707 | 116.7 | N | Y | N | integral membrane protein |
| 105 | Integrin alpha-6 precursor (VLA-6) (CD49f) | P23229 | 126.6 | N | N | Y | Type I membrane protein |
| 106 | Reticulon 4 (Neurite outgrowth inhibitor) | Q9NQC3 | 129.9 | Y | N | N | Integral membrane protein. ER |
| 107 | ATP-dependent RNA helicase A | Q08211 | 142.1 | N | Y | N | Nuclear |
| 108 | Potential phospholipid-transporting ATPase IC | O43520 | 143.7 | Y | N | N | Integral membrane protein |
| 109 | 130 kDa leucine-rich protein (LRP 130) | P42704 | 145.2 | Y | N | Y | Nuclear |
| 110 | Nuclear pore complex protein Nup155 | O75694 | 156.7 | N | Y | N | Nuclear pore |
| 111 | Integrin beta-4 (GP150) (CD104 antigen) | P16144 | 202.2 | Y | N | N | Type I membrane protein |
| 112 | Neuroblast differentiation associated protein AHNAK | Q09666 | 312.6 | N | Y | N | Nuclear |

# Chapter 6

# Analysis of In-Solution Digest of SDS-Solubilized Proteins for Membrane Protein Identification by HPLC-MALDI MS/MS Approach II: Application to the Analysis of THP1 Lipid Raft Proteome [a]

## 6.1 Introduction

In the previous chapters, HPLC-MALDI MS/MS has been developed as a useful technique for analyzing membrane proteome. In this chapter, the application of this technique for proteome analysis of lipid rafts is described. Lipid rafts are cell-type specific glycolipid and cholesterol-enriched membrane microdomains implicated in receptor signaling, and protein trafficking. To investigate the function of lipid rafts in monocytes, a gel-based proteomic approach to the systematic identification of proteins within low buoyant-density gradient fractions of Triton X-100-solubilized THP-1 cells was performed [1]. In that study, 52 proteins were identified, 28 of which had not previously been reported in lipid rafts. However, gel-based approaches may not detect all of the proteins present within a population. Reasons why proteins escape detection include incomplete digestion and recovery of very hydrophobic and low abundance proteins from polyacrylamide gels.

To obtain a more comprehensive map of the lipid raft proteome, the novel technique of using HPLC-MALDI MS/MS combined with SDS-aided protein solubilization and digestion as described in previous chapter was employed in this work.

[a] Dr. N. Li collected the MS/MS data; Mr. A. Mak from Professor Andy Shaw's group performed the protein function assay. Cell culture and lipid raft proteins isolation were done in Prof. Andy Show's group.

This technique identified more than 70 proteins, of which 45 were not detected using gel-based analysis [1]. 26 of the proteins, including the lipid raft markers flotillin 1 and 2 [2, 3], and the raft-associated GPI-linked proteins aminopeptidase M [4] and carboxypeptidase M [5], were detected in both the solution-based and gel-based digests confirming that both techniques identify known raft proteins. However, the identification of 45 new proteins by in-solution digestion indicates that conventional approaches to sample preparation fail to detect a substantial component of the protein population. The new proteome results provide a route map for further functional studies of lipid rafts.

## 6.2 Experimental

### 6.2.1 Chemicals and Reagents

α-Cyano-4-hydroxycinnamic acid (HCCA), bovine serum albumin (BSA), cytochrome c, bovine trypsin, dithiothreitol (DTT), iodoacetamide, Triton X-100, trifluoroacetic acid (TFA) and sodium dodecyl sulfate (SDS) were purchased from Sigma-Aldrich Canada (Markham, ON, Canada). HPLC-grade acetonitrile was from Fisher Scientific Canada (Edmonton, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA, USA). HCCA was recrystallized from ethanol (95%) before use.

### 6.2.2 Cell Culture

The monocytic cell line THP-1 was obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA). Cells were maintained between $2 \times 10^5$ and $6 \times 10^5$ cells per mL in a humidified incubator at 5% carbon dioxide in RPMI 1640

medium supplemented with antibiotics, L-glutamine (Sigma, Markham, ON Canada) and 10% fetal calf serum (ATCC).

### 6.2.3 Isolation of Lipid Raft Proteins

Raft proteins were obtained by buoyant-density fractionation over a discontinuous sucrose-density gradient. $1.5 \times 10^8$ cells were washed in cold PBS and lysed in MES buffer [25 mM 2-(4-Morpholino)ethane sulfonic acid containing 150 mM NaCl, 5 mM $MgCl_2$, 5 mM sodium pyrophosphate, 1 mM $Na_3VO_4$, 2 mM NaF pH 6.5, 1% Triton X-100, 1 mM EGTA, 1 mM iodoacetamide, 1 mM PMSF, and 2 mM of a proteinase inhibitor cocktail containing aprotinin, leupeptin, and pepstatin] for 90 minutes at 4°C. The lysate was subjected to Dounce homogenization using a loose-fitting pestle, and clarified by centrifugation at 250G for 5 minutes. The post-nuclear supernatant was then adjusted to 45% sucrose and overlaid with aliquots of 30%, and 5% sucrose. Buoyant density centrifugation was performed at 165,000 RCF for 18 hours at 4°C in a Beckman SW60Ti rotor. Eight fractions were collected from the gradient and concentrated by diluting the sucrose with MES buffer, followed by centrifugation of insoluble proteins and separation by 1D SDS-PAGE electrophoresis. Triton-insoluble pelleted proteins were dissolved in a non-reducing loading buffer containing 1% SDS, 25 mM Tris, pH 6.8 and applied to 12% polyacrylamide gels (180 × 160 × 1.5 mm) and run for 5 hours with cooling at a constant current of 65 mA.

### 6.2.4 In-Solution Enzymatic Digestion

Six milligram lyophilized raft powder equivalent to $3.5 \times 10^8$ cells was dissolved in 90 μL 1% SDS (w/v) in 20 mM Tris. After sufficient mixing, samples were

centrifuged and the supernatant containing the dissolved raft proteins was removed for subsequent experiments. First, standard reduction and alkylation procedures were applied [6]. Then the mixture pH was adjusted with 100 mM $NH_4HCO_3$ to 8.5. Finally, 100 μL of 1 μg/μL bovine trypsin and 1 μL of 1 M $CaCl_2$ were added to the mixture. The final SDS concentration in the digestion mixture was 0.1% in a total volume of 700 μL. Digestion was performed overnight at 30°C.

### 6.2.5 HPLC-MALDI Analysis

Peptide separation was performed on an Agilent (Palo Alto, CA, USA) 1100 series capillary HPLC with a reversed-phase 1.0 × 150 mm Vydac $C_8$ column connected to a 1.0 × 10 mm Nest Group SDS guard cartridge (Southborough, MA, USA). Gradient elution was performed with solvent A (0.1%, v/v, aqueous TFA) and B (0.1%, v/v, TFA in acetonitrile). 50 μL of peptide mixture (equivalent to a starting amount of $2.5\times10^7$ cells) was injected and UV wavelength was set at 210 nm. HPLC fractions were directly collected onto a 100-well MALDI target by a home-built device (patent pending) [7]. After the separation had finished, 1 μL of 50% acetonitrile/water (v/v) saturated by HCCA was added on top of each spot on the sample plate. After air-drying, the sample spots were ready for MALDI MS and MALDI MS/MS analysis.

### 6.2.6 Protein Identification from CID Data

Fragmentation information from CID spectra was first submitted to sequence query program in Matrix Science (http://www.matrixscience.com) for possible identifications with precision tolerance of 0.3 Da for both the parent peptide and MS/MS fragments. These potential protein matches were then put into MS-product

(http://prospector.ucsf.edu) and the theoretical fragments were compared with experimental results, which may lead to definite identification of proteins. All the peptide identities were manually analyzed after database search.

### 6.2.7 Instrumentation

MALDI MS experiments were carried out on a Bruker Reflex III time-of-flight mass spectrometer (Bremen/Leipzig, German) using the reflectron mode of operation. Ionization was performed with a 337-nm pulsed nitrogen laser. MALDI MS/MS experiments were carried out on a MDS Sciex QSTAR Pulsar QqTOF mass spectrometer equipped with an orthogonal UV-MALDI source. All data were processed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

## 6.3 Results and Discussion

### 6.3.1 Proteome Identification by HPLC-MALDI MS/MS

Due to the high lipid content and existence of hydrophobic membrane proteins in the lipid raft insoluble material, SDS was added to aid the solubilization process. Under optimized conditions, the presence of SDS did not affect the tryptic digestion process as demonstrated by successful MALDI MS analysis of the digestion product, as shown in Figure 6.1. However, only a limited number of peptide peaks from the raft proteins were detected. Many of the dominant peaks were from either trypsin autolysis or keratin contamination indicated by "T/K". It was believed that this is mainly due to the suppression during the ionization and detection processes, and can be solved by a separation step prior to the MS experiment.

Typically, an online LC/MS system is used for complex peptide mixture analysis where an ESI MS is connected to the HPLC, serving as an online detector. However, the SDS in the raft digest was found to deteriorate the signal from the 1D LC separation-ESI MS system. MALDI, on the other hand, has a much greater ability to handle contaminants including surfactants. The employment of this semi-automated HPLC off-line MALDI system significantly enhanced the peptide recovery and minimized the sample loss. HPLC eluate was directly deposited onto a MALDI plate that was subjected to MALDI MS analysis after addition of matrix. Figure 6.2 shows the UV chromatogram of the HPLC separation on this SDS-containing raft digest (A) and the MALDI spectra of three consecutive fractions at 34, 35 and 36 min (B, C and D). Several conclusions can be drawn based on Figure 6.2. First, the separation was efficient. This was demonstrated by the high quality of the HPLC chromatogram. Second, the interface did not introduce any carryover from one fraction to the next. Table 1 lists all the peptide masses measured from the three consecutive fractions at 34, 35 and 36 minutes. Only very few major peaks were detected in more than one fractions. This suggests that there is not significant carryover during fractionation. Third, the whole setup was sensitive enough to detect peptides present in very low abundance such as those in fraction 36 min.

Based on results from MS profiling, peptide peaks were selected for MS/MS sequencing experiments. This was done on the same plate that was used for the MALDI MS experiment. Fraction 36 min is presented as an example. Peptides detected in MALDI MS (Figure 6.3A) were selected one by one, from the most intense to the least, for MALDI MS/MS sequencing until sample on the spot was consumed. Figures 6.3B and C show the MS/MS sequencing experiments on two peptides with medium intensity

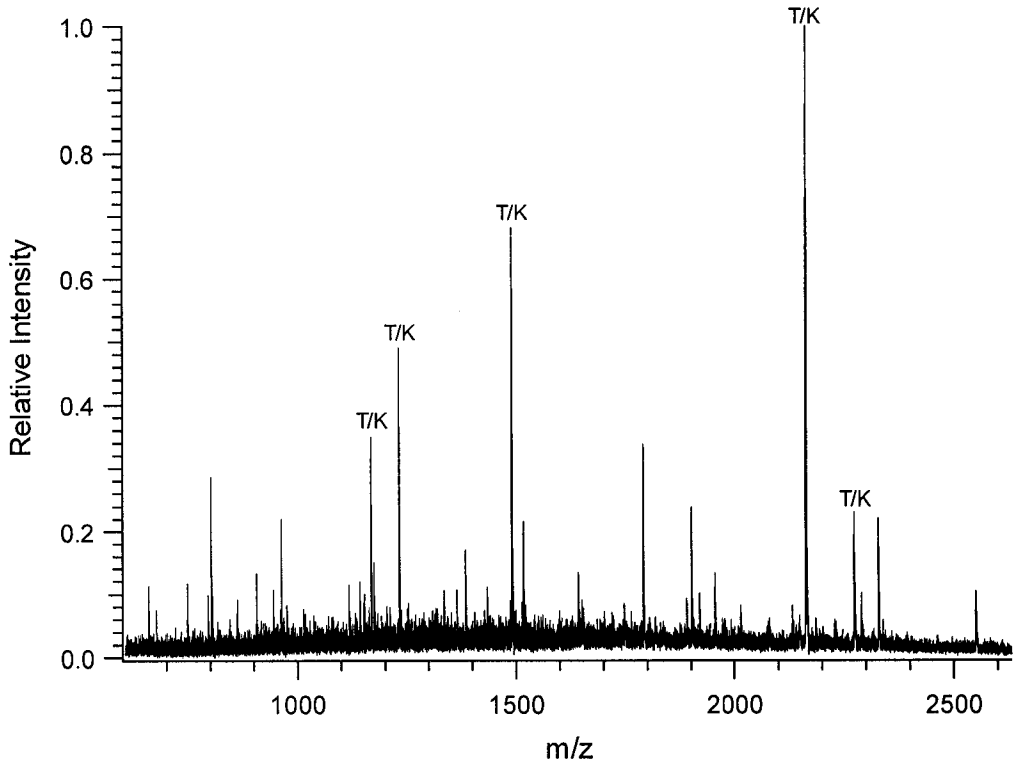



**Figure 6.1** Direct MALDI MS profiling on the SDS-containing raft digest. Digestion was efficient. However, most of the dominant peaks were from Tryspin autolysis or Keratin contamination, indicated by T/K. A limited of peptides from raft proteins were detected.

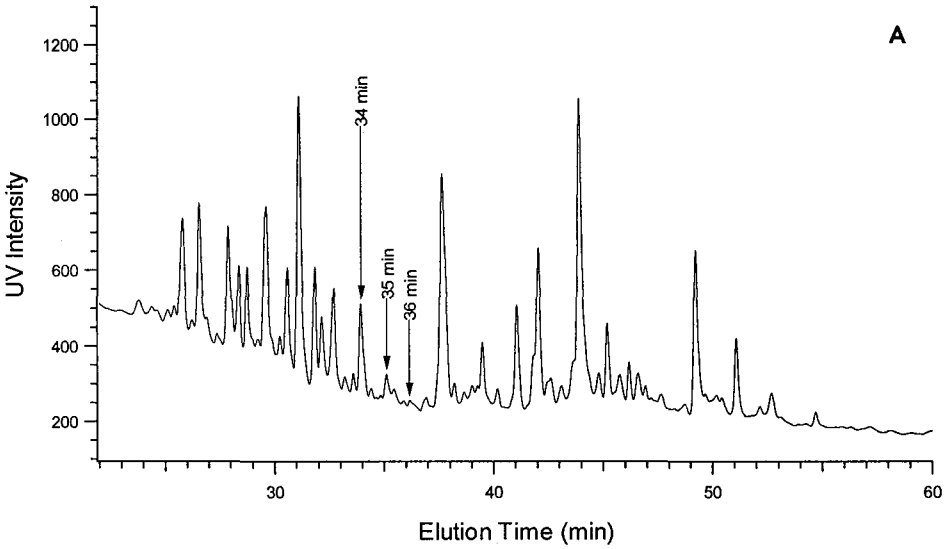



**Figure 6.2A** HPLC separation of SDS-containing raft tryptic digest.

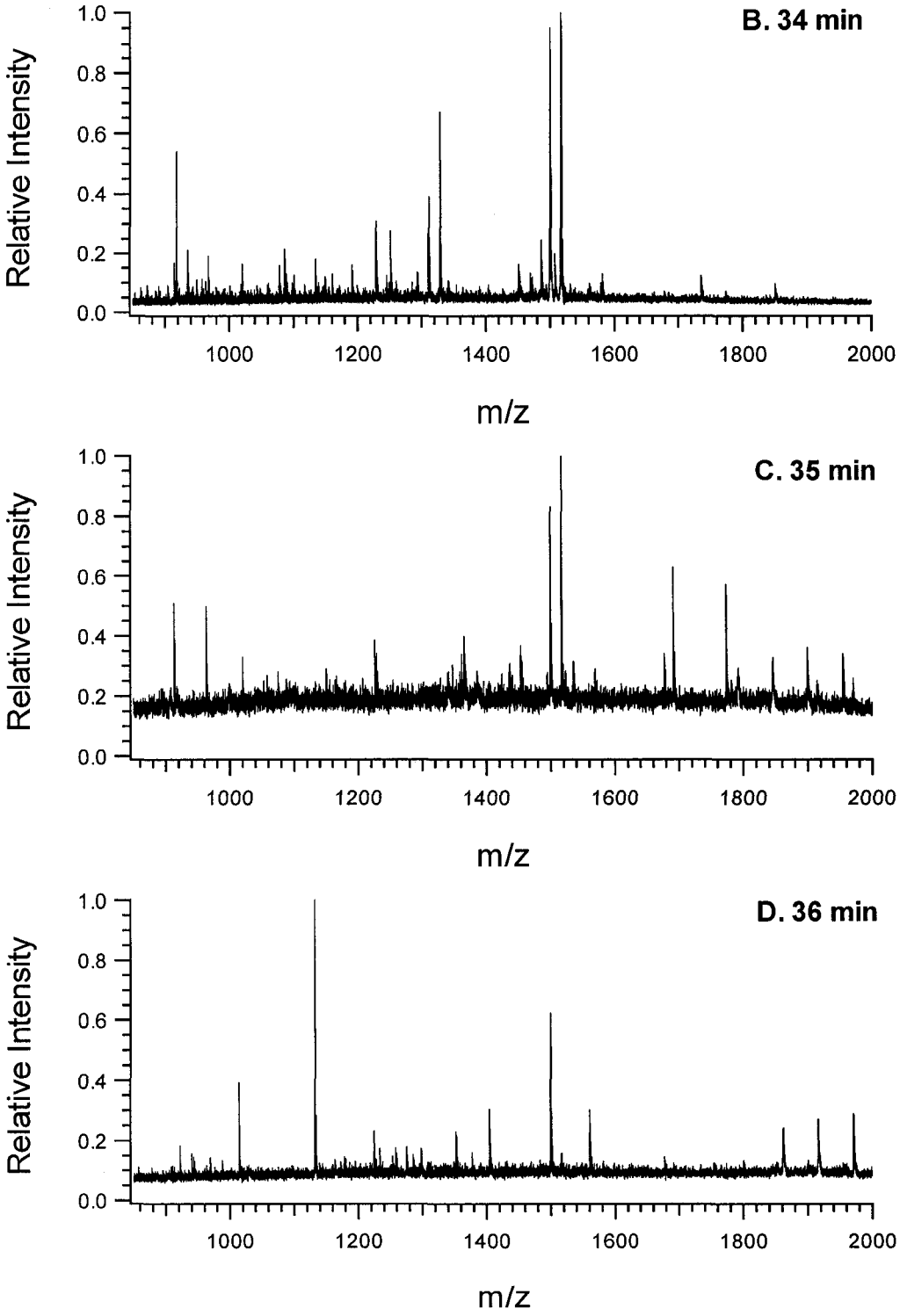



**Figure 6.2** MALDI MS spectra of fractions (B) 34, (C) 35 and (D) 36 min.

**Table 6.1 Lists of Peptides Detected from HPLC Fractions 34, 35 and 36 Minutes.**

| 34min | 35min | 36min | 34min | 35min | 36min |
|---|---|---|---|---|---|
| | 913.4 | | | | 1352.4 |
| 914.5 | | | | 1361.4 | |
| 917.6 | | | | | 1377.5 |
| | | 922.4 | | | 1404.6 |
| 935.5 | | | | 1435.4 | |
| | | 940.4 | 1451.7 | | |
| | | 943.4 | | 1453.4 | |
| | 963.3 | | 1486.7 | | |
| 967.6 | | | **1499.6** | **1499.4** | **1499.5** |
| | | 969.3 | 1507.6 | | |
| | | 1014.4 | **1516.7** | **1516.4** | **1516.5** |
| 1020.6 | | | | 1523.4 | |
| 1078.6 | | | | 1535.4 | |
| 1086.6 | | | | | 1560.5 |
| | | 1132.4 | | 1568.4 | |
| 1134.6 | | | 1578.7 | | |
| 1191.7 | | | 1581.7 | | |
| | | 1224.3 | | **1677.4** | **1677.6** |
| | 1226.4 | | | 1691.4 | |
| 1228.6 | | | 1735.7 | | |
| | 1229.3 | | **1773.7** | **1773.5** | |
| | | 1233.5 | 1850.8 | | |
| 1251.6 | | | | | 1861.5 |
| | | 1253.4 | | 1899.7 | |
| | | 1258.4 | | | 1915.8 |
| | | 1275.5 | | 1954.6 | |
| | | 1285.5 | | 1969.4 | |
| | | 1297.8 | | | 1970.7 |
| 1310.6 | | | | 2103.4 | |
| 1328.6 | | | | | 2137.9 |
| | 1347.3 | | | | 2762.9 |

Note: Only very few peaks were detected in more than one fraction, highlighted in bold in this table.

in fraction 36 min. The fragmentation patterns of these two peptides led to the identification of proteins monocarboxylate transporter 4 (O15427), and urokinase plasminogen activator surface receptor, GPI-anchored form (Q03405) respectively. In total, five proteins were identified based on sequencing data of the six peptides in this fraction alone. Theoretically, MS/MS experiments can be applied to every peptide detected in the fraction, provided that there is enough sample for all the experiments. A total of 220 peptides were analyzed and sequences were assigned to 103 of them, allowing the identification of 71 proteins. It is worth mentioning that only 50 μL of the digest product was injected to HPLC which is only equivalent to $2.5 \times 10^7$ cells. On the other hand, the previous in-gel analysis used approximately 14 times more cell extract for the separation and protein identification [1]. This suggests that the sensitivity of this identification method with this solution-based semi-offline HPLC-MALDI MS/MS design is much better than that of the typical gel separation and in-gel digestion identification method. And if multiple experiments were performed, more proteins would be likely identified.

Among these 71 identified protein, 23 were consistent with the previous report obtained using the gel-based MS approach [1]. More interestingly and importantly, 45 proteins were revealed only by the solution-based approach, providing significant complementary information on raft function. Table 6.2 lists all the proteins identified by the in-solution HPLC-MS/MS method.

### 6.3.2 Protein Function Analysis

The α– and β-tubulin subunits, tubulin specific chaperone A (a protein required

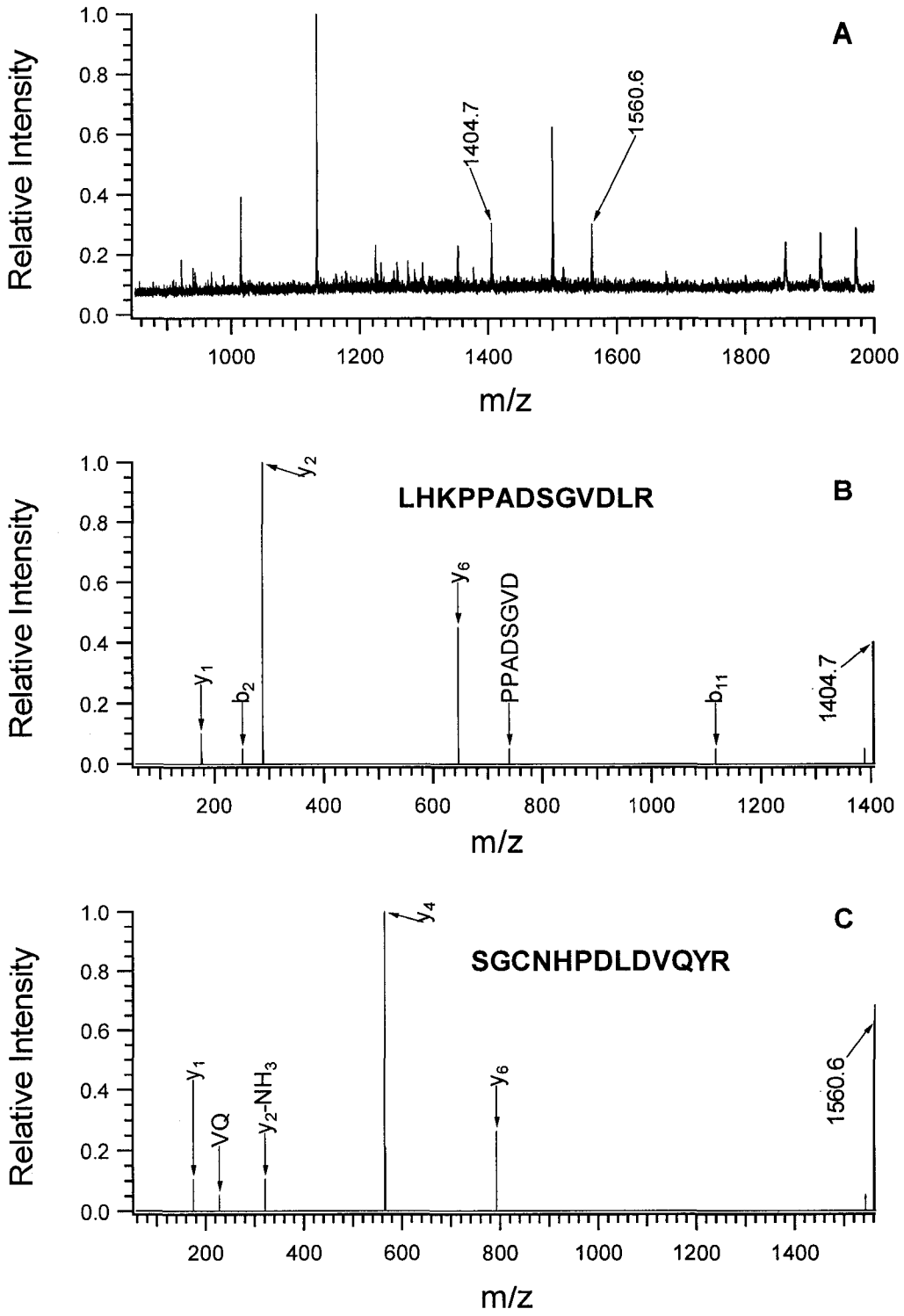



**Figure 6.3** MALDI MS spectrum of HPLC fraction 36 min (A) and sequencing patterns of peptides 1404.7 and 1560.6 (B and C). Two sequences were assigned as shown, leading to the identifications of protein monocarboxylate transporter 4, O15427, and urokinase plasminogen activator surface receptor, GPI-anchored form, Q03405, respectively.

for the assembly of αβ tubulin heterodimers), and the microtubule associated plus-end motor protein kinesin-like protein 13, are all proteins associated the assembly and maintenance of microtubules. All four proteins were identified within the low buoyant density fractions of proteins separated on sucrose density gradients followed by in-solution digestion of lipid raft proteins. As these microtubule-associated proteins were not identified by mass spectrometric analysis following conventional in-gel-digestion of THP-1 lipid raft proteins [1], it seems likely that the microtubule-associated proteins are either excluded from the gel, or enter the gel but resist digestion or peptide recovery.

To differentiate between these possibilities, an experiment was carried out to investigate whether α-tubulin was detectable within the fractions of a buoyant density gradient before and after separation on polyacrylamide gels. As a control for protein loading, flotillin-2, the lipid raft marker and caveolin-like protein, was chosen, since it was detected following either in-solution or in-gel digestion and peptide analysis. α-tubulin and flotillin-2 were readily identified by immunoblotting Triton X-100 cell lysates applied to nitrocellulose in a dot blot apparatus prior to gel separation. Both proteins were detected in high (7 and 8) and low (3 and 4) density fractions, but poorly represented in between, commensurate with dynamic equilibrium between detergent soluble (high density) and detergent insoluble (low density) pools (Figure 6.4A). To corroborate the fractions containing lipid raft proteins, the sample immunoblotted with an antibody against the lipid raft marker and Src family kinase Lyn. Lyn was present exclusively within fractions 3 and 4, identifying the raft-like nature of those fractions. The absence of Lyn from the high density fractions contrasts with the presence of tubulin and flotillin-1 suggesting that Lyn is recruited entirely into the raft.

Following separation on polyacrylamide gels, only flotillin-2 was detectable in the low-density fractions although both proteins were well represented in the high-density fractions (Figure 6.4B). The absence of α-tubulin from low, but not high-density fractions, demonstrates that raft-associated tubulin is preferentially excluded from the gel. The loss of α-tubulin following gel-separation is an indication that tubulin may be incorporated into higher-order protein assemblies of a size too large to enter the gel. This may explain why tubulin subunits and tubulin chaperone and motor proteins are notably absent from descriptions of lipid raft proteomes despite the inclusion of cytoskeletal proteins associated with the actin and intermediate filament cytoskeletons. However, tubulin has been identified within the lipid raft-like domains of neurones using photoactivatable crosslinkers attached to a ganglioside analogs [8,9]. In these experiments tubulin was specifically enriched within lipid rafts, but was unaffected by inhibitors of microtubule assembly, suggesting the tubulin was not polymerized. Similarly, the recent mass spectrometric analysis of secreted B cell vesicles or 'exosomes' demonstrated that the vesicles were enriched in cholesterol and gangliosides as well as α– and β-tubulin [10]. Exosomes are secreted and are unlikely to possess polymerized tubulin. Tubulin may therefore be a common component of lipid rafts, but is inconsistently recognized because tubulin polymerizes into assemblies that are excluded from polyacrylamide gels.

The recruitment of tubulin into lipid rafts could have important implications for membrane restructuring. The mechanisms governing the capture, docking and maintenance of microtubule plus ends at the plasma membrane are not defined. Microtubule extension is regulated by chaperone proteins that add cytosolic αβ tubulin

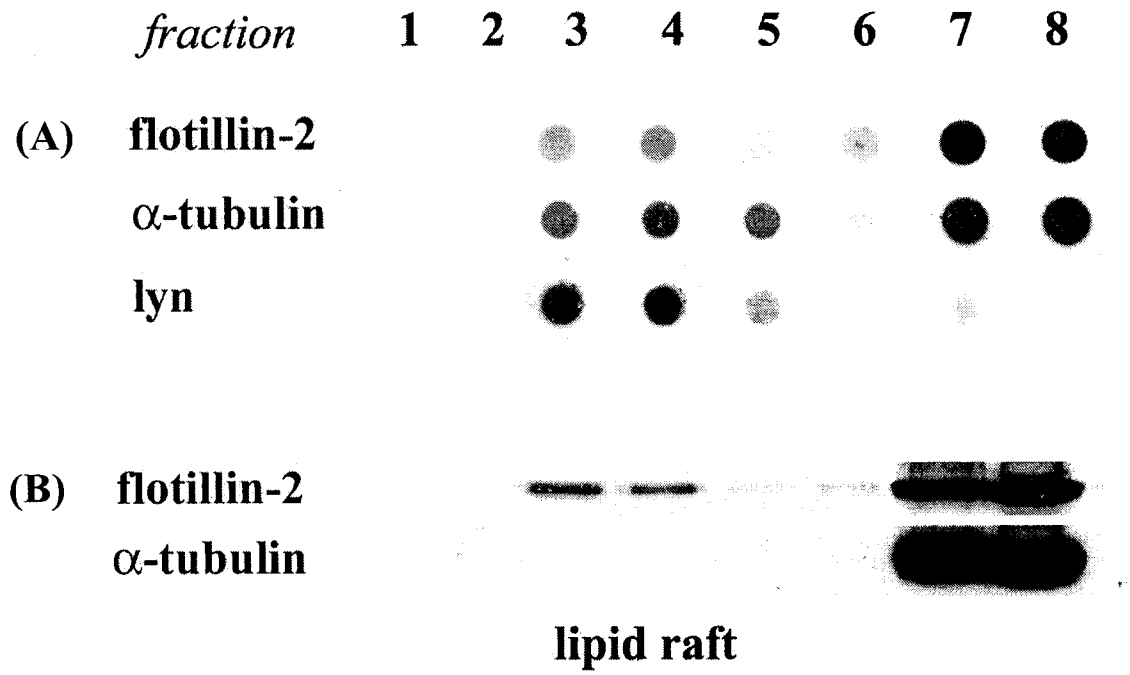



**Figure 6.4** THP-1 cells were lyzed in Triton X-100, the nuclei removed by low speed centrifugation and the remaining material separated on a 45-5 % sucrose buoyant density gradient. Ten fractions (1 low density, 10 highest density) from the buoyant density gradient were applied directly to nitrocellulose filters under vacuum using a dot-blot apparatus and immunoblotted with mAb against flotillin-2, α-tubulin, and lyn. Proteins in the fractions were concentrated by centrifugation of diluted buoyant density fractions, solubilized in SDS and separated on 12% polyacryalamide gels as described in materials and methods. The proteins were transfered to nitrocellulose and immunoblotted with mAb against flotillin-2 and α-tubulin.

heterodimers to the plus ends of the microtubules. However, addition of tubulin from the cytosol cannot continue if microtubules dock with the plasma membrane. It is of considerable interest therefore that ganglioside-associated tubulin contains palmitic acid, a modification that targets many proteins to lipid rafts. The addition of palmitoylated tubulin to the plus ends of microtubules within lipid rafts could maintain the length of microtubules associating with the plasma membrane. Such a mechanism could confer the additional benefit of targeting microtubule-associated vesicles and their protein cargos to regions of receptor engagement thereby contributing to membrane restructuring. These possibilities are currently under investigation.

Polyacrylamide gel electrophoresis is a highly successful and convenient separating medium for sizing and analysis of proteins. However, SDS-PAGE imposes limitations on the size and solubility of protein assemblies that can be analyzed. Subsequently, a very significant component of proteins within lipid rafts have been overlooked because they are part of a macromolecular assembly that fails to enter a 12% polyacrylamide gel. In-solution digestion provides important additional information about the composition of protein complexes that are vital to the investigation of macromolecular protein assemblies that resist SDS-PAGE. In-solution digestion is therefore particularly applicable to the interrogation of complex protein assemblies involving the cytoskeleton and the extracellular matrix which involve covalent modification and crosslinking of proteins. Why in-solution digestion detected only half the proteins detected by in-gel digestion is not yet clear, but may reflect the relative abundance of the proteins in the sample since in-solution digestion HPLC-MALDI technique requires much smaller amounts of protein. Alternatively, in-solution digestion

may selectively identify those proteins that are bound with highest affinity to macromolecular assemblies. If the latter should be the case the two methods of analysis are not only complementary, but provide a tool to dissect functionally important protein-protein interactions involving macromolecular assemblies.

## 6.4 Conclusions

In this work we have demonstrated that the SDS-aided in-solution based protein identification approach for Triton X-100 insoluble fraction THP-1 raft is very effective. From the results of Chapter 5 and this one, the developed SDS-aided solution-based protein identification method is practically useful for membrane protein identifications. Combined with semi-offline HPLC-MALDI interface, this method can be employed for SDS-solubilized membrane protein identification, whereas not suitable for common ESI method because of the presence of SDS. In ESI, removal of SDS by an additional experimental step is required. The LC/MALDI method is found to provide complementary results to conventional gel electrophoresis-MALDI approach. A combination of these two methods would provide a better proteome map in terms of protein coverage. In the future, THP-1 raft protein identification will be attempted with a 2D-HPLC MALDI MS/MS system where the same interface design can be used. Since only a very small portion of the raft digest was used for the 1D HPLC MALDI MS/MS study, we expect that more unknown raft-associated proteins can be revealed after 2D-HPLC separation and MALDI MS/MS sequencing.

## 6.5 Literature Cited

(1) Li, N., Mak; A., Richards, D. R.; Naber, C., Keller; B. O., Li, L.; Shaw, A. R. *Proteomics,* **2003**, *3*, 536-548.

(2) Salzer, U.; Prohaska R. *Blood*, **2001,** *97*, 1141-1143.

(3) Bickel, P. E.; Scherer, P. E.; Schnitzer, J. E.; Oh, P.; Lisanti, M. P.; Lodish, H. F. *J. Biol. Chem.* **1997**, *272,*13793-13802.

(4) Santos, A. N.; Langner, J.; Herrmann, M.; Riemann, D. *Cell. Immunol.* **2000,** *201,* 22-32.

(5) Tan, F.; Chan, S. J.; Steiner, D. F.; Schilling, J. W.; Skidgel R. A. *J. Biol. Chem.* **1989**. *264,* 13165-13170.

(6) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. *Anal. Chem.* **1996,** 68, 850-858.

(7) Zhang, B.; McDonald, C. J.; Li, L. manuscripts submitted to *Anal. Chem.*

(8) Palestini, P.; Pitto, M.; Tedeschi, G.; Ferraretto, A.; Parenti, M.; Brunner, J.; Masserini, M. *J. Biol. Chem.* **2000,** *275,* 9978-9985.

(9) Loberto, N.; Prioni, S.; Prinetti, A.; Ottico, E.; Chigorno, V.; Karagogeos, D, Sonnino, S. *J. Neurochem.* **2003,** *85*, 224-33.

(10) Wubbolts, R.; Leckie, R. S.; Veenhuizen, P. T.; Schwarzmann, G.; Mobius, W.; Hoernschemeyer, J.; Slot, J. W.; Geuze, H. J.; Stoorvogel, W. *J. Biol. Chem.* **2003** *278,* 10963-10972.

**Table 6.2 Lists of Protein Identified from In-Solution Digestion - HPLC-MALDI-MS/MS Experiment.**

| # | Identified proteins | Mass [kDa] | Ref. |
|---|---|---|---|
| 1 | Myosin Vc | 204 | Q9NQX4 |
| 2 | Kinesin-like protein KIF 13A | 203.8 | Q9H1H9 |
| | Kinesin-like protein KIF 13B | 203.9 | Q9NQT8 |
| 3 | Adenylate cyclase, type VIII | 141.3 | P40145 |
| 4 | Sodium/potassium-transporting ATPase alpha-1 chain | 114.1 | P05023 |
| 5 | Aminopeptidase N | 110 | P15144 |
| 6 | Vacuolar proton translocating ATPase 116 kDa subunit A isoform 3 | 93.7 | Q13488 |
| 7 | Gelsolin precursor, plasma | 86 | P06396 |
| 8 | Heat shock protein HSP 90 - beta | 83.5 | P08238 |
| 9 | Platelet endothelial cell adhesion molecule precursor | 83.4 | P16284 |
| 10 | Heat shock protein HSP 90 - alpha | 84.9 | P07900 |
| | Heat shock protein HSP 90 - beta | 83.5 | P08238 |
| | Heat shock protein 75 kDa | 80.3 | Q12931 |
| 11 | Heat shock protein HSP 90 - alpha | 84.9 | P07900 |
| | Heat shock protein HSP 90 - beta | 83.5 | P08238 |
| 12 | Stress-70 protein, mitochondria precursor | 73.9 | P38646 |
| 13 | Heat shock cognate 71 kDa proein | 71.1 | P11142 |
| | Heat shock-related 70 kDa protein 2 | 70.3 | P54652 |
| 14 | Heat shock 70 kDa protein 6 | 71.2 | P17066 |
| | Heat shock 70 kDa protein 7 | 27 | P48741 |
| 15 | Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 63 kDa subunit precursor | 69.4 | P04844 |
| 16 | Vacuolar ATP synthase catalytic subunit A, ubiquitous isoform | 68.7 | P38606 |
| 17 | Calnexin precursor | 68 | P27824 |
| 18 | Vacuolar protein sorting-associated protein 45 | 65.4 | Q9NRW7 |
| 19 | ATP synthase alpha chain, mitochondria precursor | 59.8 | P25705 |
| 20 | Vacuolar ATP synthase subunit B, brain isoform | 56.8 | P21281 |
| 21 | ATP synthase beta chain, mitochondria precursor | 56.5 | P06576 |
| 22 | Vitronectin precursor | 55.1 | P04004 |
| 23 | Monocarboxylate transporter 1 | 54.6 | P53985 |
| 24 | Solute carrier family 2, facilitated glucose transporter member 1 | 54.4 | P11166 |
| 25 | Carboxyppeptidase M precursor | 50.9 | P14384 |
| 26 | Tubulin beta-1 chain | 50.2 | P07437 |
| | Tubulin beta-5 chain | 50.1 | P05218 |
| | Tubulin beta-5 chain | 50.9 | P04350 |
| | Tubulin beta-4q chain | 48.9 | Q99867 |
| | Tubulin beta-4q chain | 50.9 | Q13509 |
| | Tubulin beta-2 | 50.3 | P05217 |
| 27 | Tubulin alpha-6 chain | 50.5 | Q9BQE3 |
| | Tubulin alpha1 chain | 50.8 | P05209 |
| 28 | Tubulin beta-5 chain | 50.1 | P05218 |
| | Tubulin beta-1 chain | 50.2 | P07437 |

| | | | |
|---|---|---|---|
| | Tubulin beta-2 chain | 50.3 | P05217 |
| | Tubulin beta-4 chain | 50.9 | Q13509 |
| **29** | Elongation factor 1 - alpha 1 | 50.5 | P04720 |
| **30** | Monocarboxylate transporter 4 | 50.1 | O15427 |
| | Alpha enolase | 47.4 | P06733 |
| **31** | Beta enolase | 47.2 | P13929 |
| | Gamma enolase | 47.5 | P09104 |
| **32** | Beta-1, 4-galactosyltransferase | 42.3 | O60909 |
| **33** | Actin, beta *or* gamma | 42.1 | P02570<br>P02571 |
| **34** | Actin, beta | 42.1 | P02570 |
| **35** | Flotillin -2 | 41.8 | Q14254 |
| **36** | HLA class I histocompatibility antigen (*and many different chains could be possible, as a sequence of* **FIAVGYVDDTQFVR** was identified, so just list *one* Swiss Prot number and *its* MW info) | 41.1 | P30443 |
| | G(T), alpha-2 subunit | 40.6 | P19087 |
| | G(T), alpha-1 subunit | 40.3 | P11488 |
| **37** | G(K), alpha subunit | 40.9 | P08754 |
| | G(i) alpha-1 subunit | 40.8 | P04898 |
| | G(i) alpha-2 subunit | 40.9 | P04899 |
| **38** | G(i), alpha-2 subunit | 40.9 | P04899 |
| **39** | G(K), alpha subunit | 40.9 | P08754 |
| **40** | Vacuolar ATP synthase catalytic subunit d | 40.8 | P12953 |
| **41** | Galectin-8 | 40.4 | O00214 |
| **42** | Annexin II | 38.7 | P07355 |
| **43** | Urokinase plasminogen activator surface receptor, GPI-anchored form precursor | 38.6 | Q03405 |
| **44** | G(I)/G(S)/G(T), beta subunit 2 | 38 | P11016 |
| | G(I)/G(S)/G(T), beta subunit 1 | 38.2 | P04904 |
| | G(I)/G(S)/G(T) beta subunit 1 | 38.2 | P04904 |
| **45** | G(I)/G(S)/G(T) beta subunit 1 | 38 | P11016 |
| | G protein, beta subunit 4 | 38.3 | Q9HAV0 |
| **46** | Chromosome 8 open reading frame 2 | 38 | gi/6005721 |
| **47** | ADP-ribosyl cyclase 2 precursor | 36.3 | Q10588 |
| **48** | Glyceraldehyde 3-phosphate dehydrogenase, liver | 36.1 | P04406 |
| **49** | Sodium/potassium-transporting ATPase beta-3 chain | 31.8 | P54709 |
| **50** | Basigin precursor | 29.4 | P35613 |
| **51** | 14-3-3 protein beta/alpha *or* eta *or* zeta/delta *or* gamma *or* tau *or* epsilon -- only MW and Swiss Prot number of the *first one* was listed here as an example -- Identified sequence: DSTLIMQLLR | 28 | P31946 |
| **52** | Proteasome activator complex subunit 2 | 27.5 | Q9UL46 |
| **53** | Chloride intracellular channel protein 1 | 27.2 | O00299 |
| **54** | Cytochrome C oxidase polypeptide II | 25.7 | P00403 |
| **55** | Synaptosomal-associated protein 23 | 23.7 | O00161 |
| **56** | Heat shock 27 kDa protein | 22.8 | P04792 |
| **57** | Brain acid soluble protein 1 | 22.5 | P80723 |

| | | | |
|---|---|---|---|
| **58** | Ras-related protein RAP-1A | 21.3 | P10113 |
| | Ras-related protein RAP-1B | 21 | P09526 |
| **59** | Ras-related protein RAP-2b | 20.7 | P17964 |
| **60** | Peptidyl-prolyl cis-trans isomerase A | 18.1 | P05092 |
| **61** | Chromosome 6 open reading frame 9; G18.2 protein | 18 | gi/11545817 |
| **62** | Cytochrome C oxidase subunit Va | 16.9 | gi/18999392 |
| **63** | Galectin -1 | 14.9 | P09382 |
| **64** | Mitogen-activated protein kinase kinase 1 interacting protein 1 | 13.7 | Q9UHA4 |
| **65** | Vacuolar ATP synthase subunit F | 13.4 | Q16864 |
| **66** | Tubulin-specific chaperone A | 12.8 | O75347 |
| **67** | Uniquinol-cytochrome C reductase complex 11 kDa protein, mitochondrial precursor | 11 | P07919 |
| **68** | Cytochrome C oxidase polypeptide Vb (Bos taurus) | 10.9 | P00428 |
| **69** | Cytochromoe C oxidase polypeptide VIb | 10.3 | P14854 |
| **70** | Nonhistoe chormosomal protein HMG-17-like 3 | 9.2 | O00479 |
| | Nonhistone chromosomal protein HMG-17 | 9.2 | P05204 |
| **71** | Ubiquitin | 8.6 | P02248 |

**Note**: 71 proteins were identified (some of them may be overlapped due to several possible options led by identified peptide sequences). Among these proteins, 24 of them were also identified by 1D gel-trypsin digestion-MALDI MS and MALDI MS/MS experiment. The other 45 of them were only identified using this approach.

# Chapter 7

# Comprehensive Lipid Raft Proteome Analysis by Offline 2D-HPLC Separation and Mass Spectrometry [a]

## 7.1 Introduction

Lipid rafts are cell-type specific dynamic assemblies of glycolipid and cholesterol which form membrane microdomains for receptor signaling and protein trafficking. One of the most important properties of lipid rafts is that they can include or exclude proteins to variable extents, e.g. they act as molecular sieves to allow certain proteins to transfer with important consequences for the functional organization of the plasma membrane or intracellular membranes [1-3].

To fully recognize the raft-dependent functions, it is imperative to have an extensive and unbiased understanding of raft protein composition, or the raft proteome. Until recently, study of lipid rafts was preformed by an immunoblotting technique that identifies certain proteins without the full composition.

Recent developments in mass spectrometry make it the powerful technique for proteomics research and can provide definitive information of protein content. Analysis of lipid raft proteins by a MS technique is challenging because the proteins are insoluble in aqueous solution and in nonionic detergents. Therefore, the analysis of lipid raft proteome is only achieved by a combination of MS technique and gel electrophoresis where the strong detergent SDS is employed for protein solubilization [4, 5].

---

[a] Mr. L. Burke of the Institute for Biomolecular Design of the University of Alberta generated the ESI data. Cell culture and lipid raft proteins isolation were done in Dr. A.Shaw's group.

Recently, the shotgun proteomics approach has emerged as a powerful technique for proteome identification and quantitation [6-8]. It is based on the digestion of proteins from extracts of whole cells, organelles, or specific fractions thereof, followed by liquid chromatography (LC) tandem mass spectrometry (MS/MS) analysis of the resulting peptides. For complex peptide mixtures, multidimensional liquid chromatography technique (usually two-dimensional LC) can be employed to separate the peptides prior to analysis by mass spectrometry [6, 9-11]. Two-dimensional LC can effectively resolve peptides from a complex mixture based on peptides' charge and hydrophobicity. Ion exchange chromatography is used as the primary separation technique because it possesses high loading capacity, and reversed-phase chromatography is used as the secondary separation because of its compatibility with ESI or MALDI mass spectrometry.

In Chapter 5, an in-solution membrane protein identification method by using SDS as a solubilization reagent to "dissolve" the proteins prior to enzymatic digestion has been developed, and this method was applied to the identification of lipid raft protein 1D HPLC/MALDI MS/MS (see Chapter 6), where 72 proteins are identified and 45 of them were not identified by previous gel/MS approaches. If sample amount is not a problem for proteome identification, optimization of the separation technique, such as multidimensional separation, and an increase in the sample loading amount of the LC-MALDI MS/MS approach will surely achieve better and more complete proteome identification. Here a method is presented to analyze more complete lipid raft proteome using the in-solution digestion method combined with multidimensional separation technique and newly developed LC-MALDI interface that combines microbore LC with

matrix-assisted laser desorption ionization (MALDI) MS/MS. Using the 2D LC-ESI and 2D LC MALDI techniques, it is shown that comprehensive lipid raft proteome analysis can be achieved.

## 7.2 Experimental

### 7.2.1 Chemicals and Reagents

α-Cyano-4-hydroxycinnamic acid (HCCA), bovine serum albumin (BSA), cytochrome c, bovine trypsin, dithiothreitol (DTT), iodoacetamide, Triton X-100, trifluoroacetic acid (TFA) and sodium dodecyl sulfate (SDS) were purchased from Sigma-Aldrich Canada (Markham, ON, Canada). HPLC-grade acetonitrile and acetone were purchased from Fisher Scientific Canada (Edmonton, Canada). Water was obtained from a Milli-Q Plus purification system (Millipore, Bedford, MA, USA). HCCA was recrystallized from ethanol (95%) before use.

### 7.2.2 Cell Culture

The monocytic cell line THP-1 was obtained from the American Type Culture Collection (ATCC, Manassas, VA, USA). Cells were maintained between $2 \times 10^5$ and $6 \times 10^5$ cells per mL in a humidified incubator at 5% carbon dioxide in RPMI 1640 medium supplemented with antibiotics, L-glutamine (Sigma, Markham, ON Canada) and 10% fetal calf serum (ATCC).

### 7.2.3 Isolation of Lipid Raft Proteins

Raft proteins were obtained by buoyant-density fractionation over a discontinuous sucrose-density gradient. $1.5 \times 10^8$ cells were washed in cold PBS and lysed in MES

buffer [25 mM 2-(4-Morpholino)ethane sulfonic acid containing 150 mM NaCl, 25 mM MES, 5 mM $MgCl_2$, 5 mM sodium pyrophosphate, 1 mM $Na_3VO_4$, 2 mM NaF pH 6.5, 1% Triton X-100, 1 mM EGTA ([Ethylenebis(oxyethylenenitrilo)] tetraacetic acid), 1mM iodoacetamide, 1 mM PMSF, and 2 mM of a proteinase inhibitor cocktail containing aprotinin, leupeptin, and pepstatin] for 90 minutes at 4°C. The lysate was subjected to Dounce homogenization using a loose-fitting pestle, and clarified by centrifugation at 250G for 5 minutes. The post-nuclear supernatant was then adjusted to 45% sucrose and overlaid with aliquots of 30%, and 5% sucrose. Buoyant density centrifugation was performed at 165,000 RCF (relative centrifugal force) for 18 hours at 4°C in a Beckman SW60Ti rotor. 8 fractions were collected from the gradient and concentrated by diluting the sucrose with MES buffer, followed by centrifugation of insoluble proteins.

### 7.2.4 In-Solution Protein Digestion

Ten milligram lyophilized raft powder equivalent to $6 \times 10^8$ cells was dissolved in 120 μL 1% SDS (w/v) in 20 mM Tris. After sufficient mixing, samples were centrifuged and the supernatant containing the dissolved raft proteins was removed for following experiments. First, standard reduction and alkylation procedures were applied [12]. Then the mixture pH was adjusted with 100 mM $NH_4HCO_3$ to 8.5. Finally, 200 μL of 1 μg/μL bovine trypsin and 1.2 μL of 1 M $CaCl_2$ were added to the mixture. The final SDS concentration in the digestion mixture was 0.1% in a total volume of 1200 μL. Digestion was performed overnight at 30°C.

### 7.2.5 Peptide Mixture Dialysis

The digested protein mixture was dialyzed using Spectra/Pro tube cellulose ester molecular weight cutoff of 500 Da membrane (Houston, TX, USA). The dialysis was performed at 4°C with de-ionized distilled water for 29 hours.

### 7.2.6 Liquid Chromatography

Peptide separation was performed on an Agilent (Palo Alto, CA, USA) 1100 series HPLC equipped with an auto sampler. The first dimension, strong cation exchange separation, was performed on a Vydac VHP400 series 1 × 25 mm strong cation exchange column with a flow rate of 100 μL/min. Solvent A is 0.1% TFA in 20% acetonitrile, and solvent B is the same as A with 0.5 M NaCl. The pH of digested sample was adjusted to pH 2. The ion exchange fractions were collected by a fraction collector at 1 min intervals. The fractions were dried down to 50 μL to get rid of all acetonitrile. Then one-eighth of each fraction was used for LC-ESI MS/MS experiment by a QTOF instrument equipped with a 75 μm $C_{18}$ column. The rest of the fraction was loaded onto a Vydac 1 × 150 mm $C_{18}$ reversed-phase column, to perform the HPLC-MALDI MS/MS experiment. The HPLC-MALDI experiment is the same as described in Chapter 5.

### 7.2.7 Protein Identification from CID Data

Fragmentation information from CID spectra was first submitted to the sequence query program in Matrix Science (http://www.matrixscience.com) for possible identifications with a precision tolerance of 0.3 Da for both the parent peptide and MS/MS fragments. These potential protein matches (Score greater than homology) were then put into MS-product (http://prospector.ucsf.edu) and the theoretical fragments were

compared with experimental results, which may lead to definite identification of proteins. All peptide identities were manually analyzed after the database search.

### 7.2.8 Instrumentation

MALDI MS experiments were carried out on a Bruker Reflex III time-of-flight mass spectrometer (Bremen/Leipzig, German) using the reflectron mode of operation. Ionization was performed with a 337-nm pulsed nitrogen laser. MALDI MS/MS experiments were carried out on a MDS Sciex QSTAR Pulsar QqTOF mass spectrometer equipped with an orthogonal UV-MALDI source. LC-ESI MS/MS experiments were carried out on a Micromass Q-Tof 2 mass spectrometer equipped with a LC source (Manchester, UK). All data were processed using the Igor Pro Software package (WaveMetrics, Lake Oswego, OR).

## 7.3 Results and Discussion

### 7.3.1 Buffer Selection for Cation Exchange Column of Peptide Separation

Ion exchange chromatography separates proteins and peptides by charge primarily through electrostatic interactions between charged amino acid side chains and the surface charge of the ion-exchange resin. Strong cation-exchange chromatography retains biomolecules by the interaction of sulfonic acid groups on the surface of the ion-exchange resin with histidine (pK=6.5), lysine (pK=10) and arginine (pK=12). The mobile phase is buffered to maintain its pH at 6 or 7 in order to keep the basic side chain protonated. At higher pH, the basic side chains begin to deprotonate and the retention decreases. Below pH 6, retention is dependent on the numbers of basic amino acids present in the peptide and protein.

Proteins or peptides with the same number of basic amino acids can often be separated by adjusting the mobile phase pH between 3 and 5 where aspartic acid (pK=3.9) and glutamic acid (pK=4.3) are partially protonated. A suitable buffer in this pH range is acetate buffer (HAc, pKa 4.2). Figure 7.1 shows the cation exchange chromatograms of the separation of 40 μg standard protein tryptic digest under four different mobile buffers, namely acetate buffer (pH 4), phosphate buffer (pH 4), 0.1% TFA and 0.1% TFA in 20% acetonitrile, and with a gradient of 0.5 M NaCl. Figure 7.1A shows the chromatogram obtained using 25 mM acetate buffer in mobile phase. However, the peptide separation efficiency is not good under this buffer condition. Besides the separation efficiency, acetate has strong UV absorbance at wavelength 210 nm that is commonly used for peptide and protein UV detection and therefore interferes with peptide detection. Figure 7.1B shows the chromatogram of separation of the same protein standard digest with phosphate buffer at pH 4, and the separation efficiency is improved compared to acetate buffer. However, phosphate buffer of pH 4 only has a hundredth of the buffer capacity (phosphate has a $pKa_1$ of 2.1), and this capacity may not be enough for real sample separation or when sample loading is high. Other acids that could be used as buffer are formic acid and TFA. Because proteins or peptides are easily formylated with formate, formic acid is seldom used if an alternative can be found. 0.1% TFA itself is a buffer with pH 2.2 (TFA, pKa = -0.3), and because reversed-phase separation of peptides will proceed after the ion exchange separation, where TFA is commonly used as ion- pair reagent for reversed-phase, 0.1% TFA could be another choice for buffer.

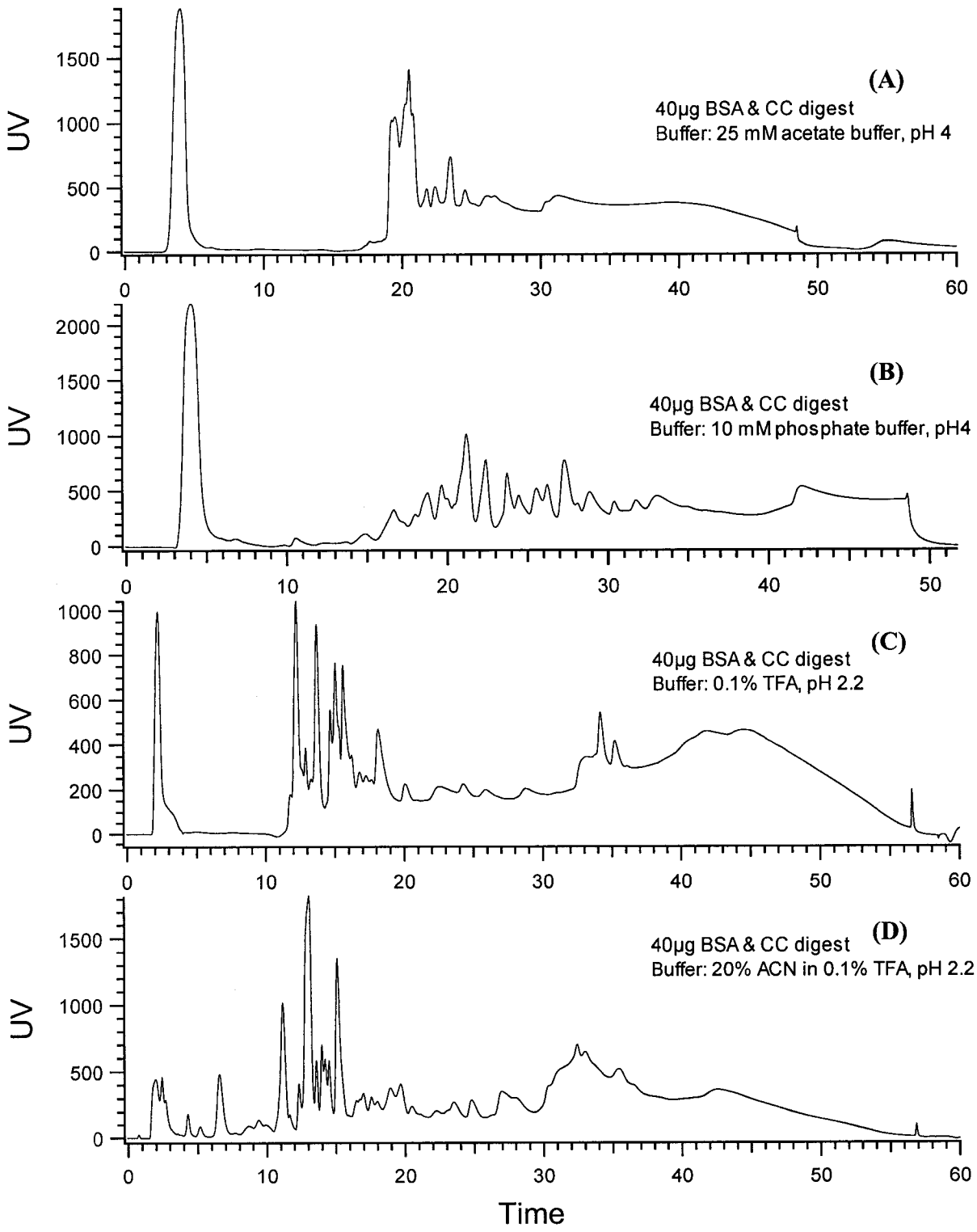



**Figure 7.1** Chromatograms of the cation exchange separation of 40 µg BSA and cytochrome c tryptic digest under different mobile phase buffers. Solvents: A is buffer indicated in the graph; B is 0.5M NaCl in A.

Figure 7.1C shows the chromatogram of standard peptide separation with 0.1% TFA as buffer under the same salt gradient. The separation performance is acceptable compared to the one with phosphate buffer. The peaks are narrower and sharper than the chromatogram obtained with phosphate buffer. The hump at 40 min is possibly from the absorbance of salts or other impurities. To further improve the ion exchange separation, addition of 20% ACN to the buffer of 0.1% TFA is employed as mobile phase buffer (Figure 7.1D) and the chromatogram peaks are sharper and more resolved. The organic solvent acetonitrile adds hydrophobic interaction effects to the ion interaction, the primary interaction on ion exchange column. Hydrophobic interaction may play a role because some digested peptides are small and very hydrophobic. Adding acetonitrile in ion-exchange mobile phase buffer can change the separation through complex effects on solubility and hydrophobic interaction. Because more and more researchers employ a certain amount acetonitrile in ion exchange separations [7, 9, 13] and the employment of acetonitrile is in agreement with our mobile phase optimization results (Figure 7.1), all subsequent experiments to be described in this chapter will use 20% ACN 0.1% TFA as a buffer system.

### 7.3.2 Salt Effect on Ion Exchange Column

Cation exchange resin has a sulfonate group as a counter ion on the resin stationary phase. SDS, an ionic surfactant should not be absorbed onto the stationary phase. The only possible interference from SDS with a cation exchange resin is from the sodium, the same cation used in mobile phase. The sodium concentration of 0.1% SDS (3 mM) is very low compared to commonly used salt gradient concentration, 0.5 M NaCl,

even at the concentration of 0.1 - 0.2 M NaCl where most peptides are eluted. Therefore, the sodium in SDS should not affect cation exchange performance.

The sample loading capacity of a 1 × 250 mm cation exchange column is 4.7 mg, according to Vydac, calculated based on separation of pure protein samples. However, real protein digest sample usually contains contaminants, such as DTT and iodoacetamide (they may get positively charged at pH 2.2). Because the strong cation exchange employs ion interaction for peptide separation, the existence of salts may have an effect on the column separation efficiency. Figure 7.2 shows the effects of salts on the peptides separation with a 1 × 250 mm strong cation exchange column. Figure 7.2A shows the chromatogram of 40 μg standard proteins (BSA and cytochrome c) tryptic digest without DTT and iodoacetamide added prior to the digestion process. With 18 mM DTT and 36 mM iodoacetamide (the common concentrations used in protein digestion) added after digestion, the cation exchange separation chromatogram is shown in Figure 7.2B. In Figure 7.2B, the injection peak is very broad, having a peak width of 5 min. The peptide peaks at 10 to 20 min are overlaid and become broader compared to that obtained from pure standard sample. This indicates that, for the same amount of peptide samples, the salt component in the sample will reduce the separation efficiency.

Another example is the ion exchange separations of the tryptic digest of *E. coli* extract with different size of columns. The sample was tryptic digested in 2.5 M urea, 15 mM DTT and 30 mM iodoacetamide. When 24 μg of the sample was loaded onto a 1 × 250 mm cation exchange column, a broad injection peak was observed and the separated peptides were eluted out of column from 10 to 16 min (Figure 7.2C). However, when 10 μg of the same sample was injected to a capillary 0.3 × 150 mm cation exchange column,

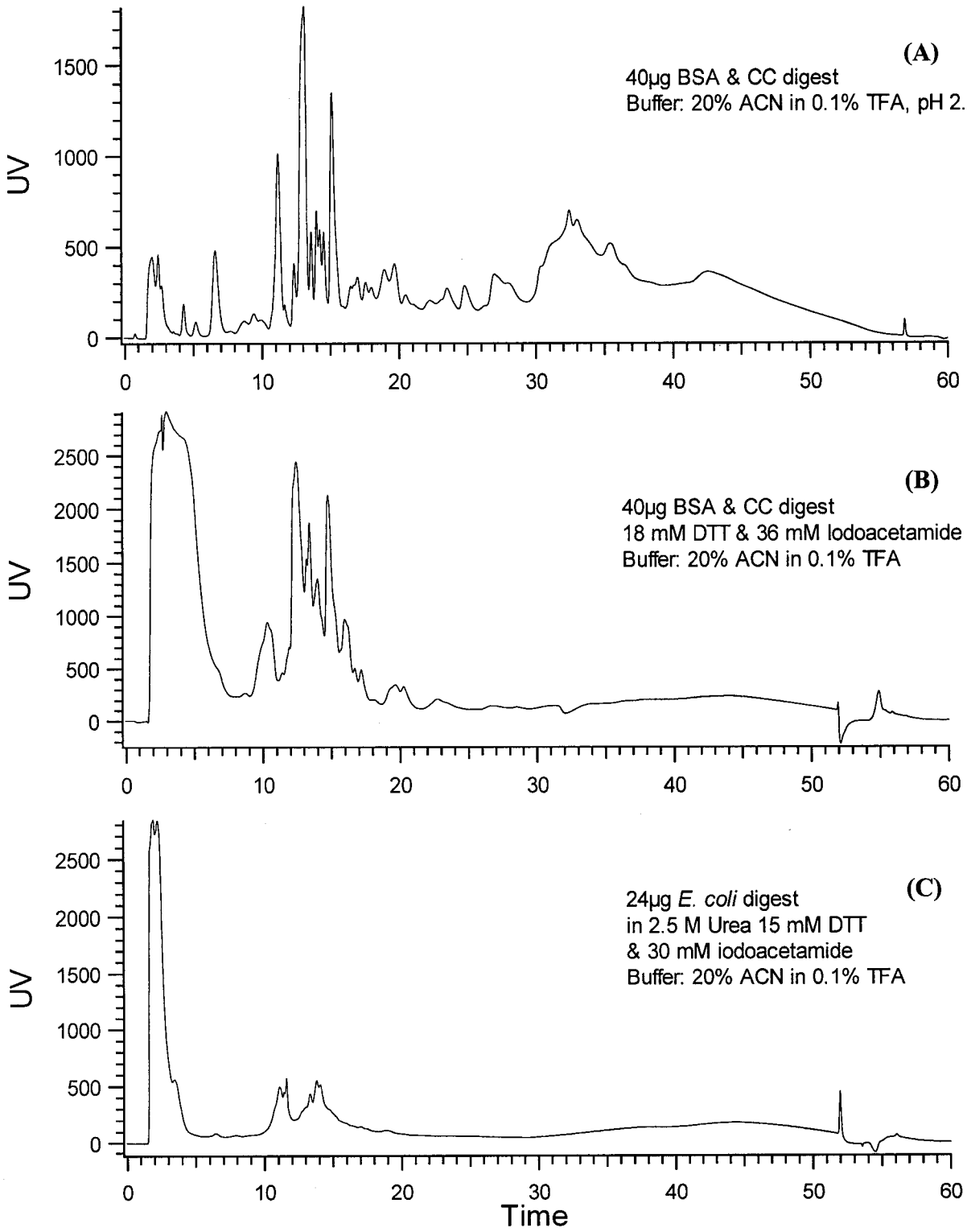



**Figure 7.2** Salts effects on ion exchange separation.

the injection peak became twice as broad as the one in Figure 7.2C. The broad saturated injection peak plateau extended up to 15 min, and almost no peptide separation was observed (data not shown). This indicates that the salt content can affect ion exchange chromatography performance significantly. On smaller columns, the ion exchange sites were totally saturated by the salt; therefore, peptides cannot be adsorbed onto the resin and co-eluted with the salt component. However, because the large column has 18 times the resin volume of the smaller one, there was still some sites left to interact with the peptides and separation peaks can be observed in the chromatogram from the large column. Thus, to obtain a better ion exchange separation, salts have to be removed as much as possible prior to sample loading onto the exchange resin. Based on this understanding, dialysis was employed to tryptic-digested samples to remove salts and SDS if possible.

### 7.3.3 2D LC-ESI MS/MS

The work in Chapter 5 clearly illustrated that, for comprehensive analysis of proteomes of interest, a preferred shotgun proteomics approach should incorporate both LC/MALDI and LC/ESI MS/MS technologies. Therefore, both ESI and MALDI are employed for analysis of the lipid raft proteome. A flow chart of the experimental design is presented in Figure 7.3. Offline two-dimensional LC separation is employed in this case not only for combination of LC-MALDI as second-dimensional, but offline 2D LC itself holds some advantages as well, such as a high organic solvent can be used during 1D ion exchange separation (only 5% ACN can be used for online 2D).

Each ion exchange (IE) fraction is split to two portions, one for LC-ESI MS/MS and the other for LC-MALDI MS/MS experiments. Because the purpose of this study is

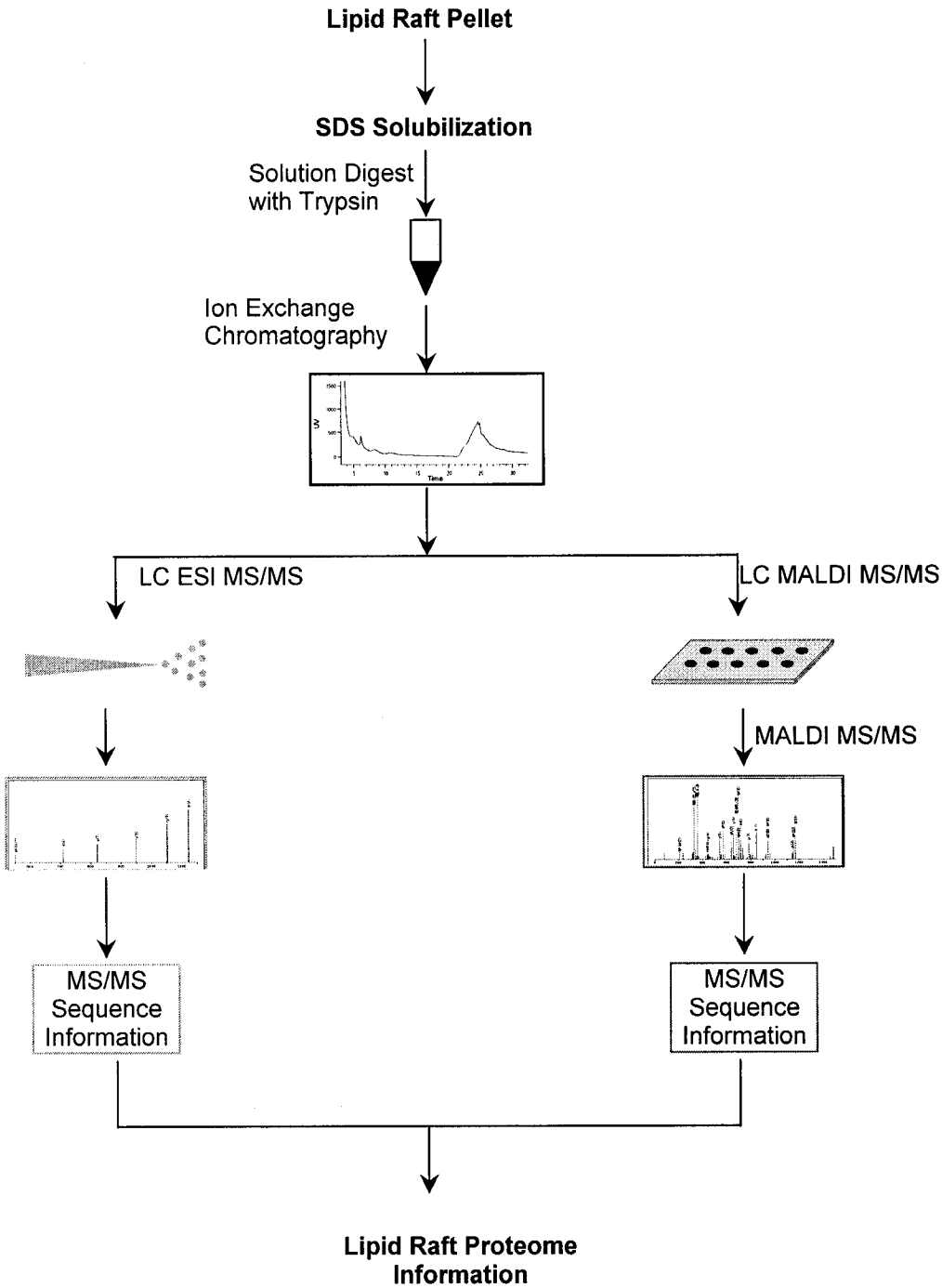



**Figure 7.3** Lipid raft proteome analysis using combination of 2-dimensinal chromatography and tandem mass spectrometry.

to identify as many proteins as possible, the samples are split in a way that the sample loading of both LC-ESI and LC-MALDI can be optimized. At the present time, ESI instruments are commonly equipped with capillary LC column to keep the sample concentration high. Because the ESI signal strength is dependent on the analyte concentration, the use of capillary LC or low flow rate LC can keep the analyte concentration high even with a small amount of sample injected onto the column. However, the capillary LC column also has a very limited sample loading capacity for separation, which can affect the detection dynamic range. It is difficult to detect the low abundant components in a small amount of proteome digest because their solution concentration can readily fall below the concentration required to generate good database-searchable spectra.

Thirty IE fractions (2 to 31min) were analyzed by $C_{18}$ on the LC-ESI QTOF instrument. Only 15 of the fractions gave positive protein identification results. From these fractions, 118 peptides were identified leading to the identification of 60 unique proteins. Among the 60 proteins, 30 were not identified in the previous LC-MALDI MS/MS experiments accomplished in Chapter 6, and 29 were not identified in the previous lipid raft studies with a gel/MALDI approach as well [4].

### 7.3.4 2D LC-MALDI MS/MS

Unlike ESI, MALDI sensitivity is not dependent on the solution concentration of the analyte, but on the absolute amount on the target or the analyte concentration in the matrix crystals. Thus a dilute solution with a large volume can be concentrated on the MALDI plate and achieve a similar sensitivity as in the case of using a concentrated solution with a small volume deposition. As a result, micro-bore column (1 mm in

diameter) can be used, which has a much higher sample loading capacity compared to a capillary column. Therefore a higher dynamic range of detection can be achieved with LC/MALDI, compared to ESI MS/MS.

The 30 IE fractions are also analyzed by LC-MALDI via our interface (see experimental section in Chapter 5). Briefly, each fraction was two dimensionally separated with a $C_{18}$ 1 × 150 mm column and the reversed-phase fractions were directly collected on the MALDI target. Figure 7.4 shows two examples of the LC-MALDI experiments from ion exchange fractions of 6 and 24 min. Figure 7.4A shows the UV chromatogram of the ion exchange separation, while Figure 7.4B shows UV chromatograms obtained from the reversed-phase (RP) separation of IE fractions 6 and 24 min. Peptides can be observed from the fraction even with low UV absorbance. Two representative MALDI MS spectra are presented in Figure 7.4C, indicating the number of peptide co-eluted in each reversed-phase fraction. Figure 7.4D illustrates representative MALDI MS/MS spectra along with the best-matched sequences from database searching. From the ion exchange fraction at 24 min, a total of 282 MS/MS spectra were collected and among them, 115 peptides were identified, leading to the identification of 76 unique proteins. Table 7.1 lists the peptides and proteins identified from IE fraction 24 by reversed-phase LC-MALDI MS/MS

From the 10 IE fractions that have been subjected to LC-MALDI MS/MS experiments, more than 1900 MALDI MS/MS spectra were collected and a total of 491 peptides have been identified leading to the identification of 185 unique proteins from these fractions.

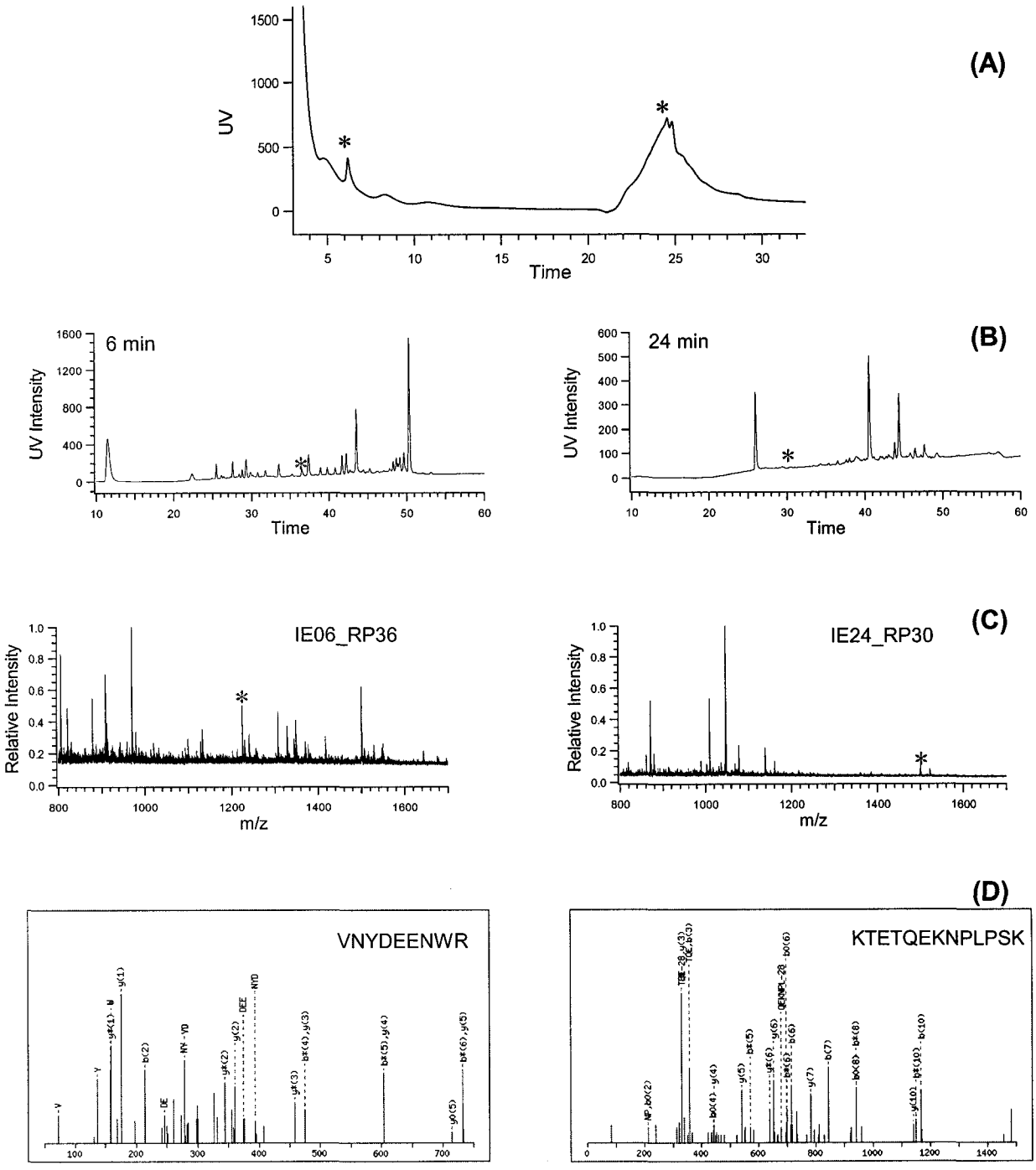



**Figure 7.4** 2D LC-MALDI MS/MS analysis of lipid raft. (A) UV chromatogram of ion exchange (IE) separation. (B) UV chromatograms of reversed-phase (RP) separation of IE fractions 06min and 24 min respectively. (C) MALDI MS spectra of RP fractions indicated in B. (D) MALDI MS/MS spectra from the peptides labeled in C with best database matches.

From these results, the advantage of 2D LC-MALDI is very significant compared to the 2D LC-ESI shotgun approach. Generally speaking, 40-100 peptides can be identified from each ion exchange fraction. At this stage, manual data collection was used for each MALDI plate with QqTOF instrument equipped with a UV MALDI source. Because two consecutive IE fractions have substantial overlaid peptides observed in their reversed-phase fractions, the advantage of offline 2D separation and manual data collection is to avoid MS/MS data collection on duplicate peptides observed in consecutive IE fractions and focus on low abundance peptides in the same fraction by increasing data collection time. Another advantage of manual data collection is fine-tuning of the instrument parameters leads to higher quality, database-searchable spectra.

Compared to LC-ESI, one major advantage of LC-MALDI is that the sample spots on the MALDI plate can be reexamined after the first pass of analysis, whereas samples are completely consumed after ESI analysis. Even though the first pass is by manual data collection, by fine-tuning the MALDI QqTOF tuning parameter (such as changing the parent ion gate window from low resolution to unit resolution), additional positive sequence assignment can be obtained [14]. It is not difficult to imagine that if the first pass is achieved in an automated manner, reexamination of the spots by fine-tuning the parameters, such collision energy, collection time and parent ion gate window, can improve the MS/MS spectra quality and lead to additional peptide sequence assignment. In the author's experience, the sample spot with matrix can be stored more than a month without observation of signal intensity difference.

Future work will mainly focus on development of an automated data collection system for the MALDI QqTOF instrument as part of the improvement of the 2D LC-

MALDI shotgun technology. With data collection fully automated, data collection for one plate of LC-MALDI sample can be completed in 24 hours. And the database searching can be finished the next day. Thus, LC-MALDI can be carried out within two days, a reasonable time for proteomic studies.

## 7.4 Conclusions

The lipid raft proteome is analyzed with incorporation of two offline shotgun approach, 2D LC-ESI MS/MS and 2D LC-MALDI MS/MS. There were 60 proteins identified with 2D LC-ESI approach, and 185 proteins identified with 2D LC-MALDI approach from 50% of the ion exchange fractions. Herein, we have demonstrated that LC-MALDI offers several unique features not seen in LC-ESI. The results suggest that offline 2D LC-MALDI is a robust technique for shotgun proteomic research.

## 7.5 Literature Cited


(1) Simons, K.; Toomre, D. *Nat. Rev. Mol. Cell Biol.* **2001,** *1*, 31-9.

(2) Zajchowski, L. D.; Robbins, S. M. *Eur. J. Biochem.* **2002,** *269*, 737-52.

(3) Shaw, A. R.; Li, L. *Curr. Opinion in Mol. Therapeutics* **2003,** *5*, 294-301.

(4) Li, N.; Mak, A.; Richards, D. P.; Naber, C.; Keller, B. O.; Li, L.; Shaw, A. R. *Proteomics* **2003,** *3*, 536-548.

(5) von Haller, P. D.; Donohoe, S.; Goodlett, D. R.; Aebersold, R.; Watts, J. D. *Proteomics* **2001,** *1*, 1010-1021.

(6) Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotech.* **2001,** *19*, 242.

(7) Han, D. K.; Eng, J.; Zhou, H.; Aebersold, R. *Nat. Biotech.* **2001,** *19*, 946-951.

(8) Blonder, J.; Goshe, M. B.; Moore, R. J.; Pasa-Tolic, L.; Masselon, C. D.; Lipton, M. S.; Smith, R. D. *J. Proteome Res.* **2002,** *1*, 351-60.

(9) Link, A. J.; Eng, J.; Schieltz, D. M.; Carmack, E.; Mize, G. J.; Morris, D. R.; Garvik, B. M.; Yates, J. R. *Nat. Biotech.* **1999,** *17*, 676.

(10) Washburn, M. P.; Ulaszek, R.; Deciu, C.; Schieltz, D. M.; Yates, J. R. *Anal. Chem.* **2002,** *74*, 1650-1657.

(11) Wolters, D. A.; Washburn, M. P.; Yates, J. R. *Anal. Chem.* **2001,** *73*, 5683-5690.

(12) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. *Anal. Chem.* **1996,** *68*, 850-858.

(13) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. *J. Proteome Res.* **2003,** *2*, 43-50.

(14) Zhao, J. Y.; Scott, G.; Keller, B. O.; Li, L. *Proceedings of the 50$^{th}$ Conference of the American Society for Mass Spectrometry and Allied Topics*, Orlando, Florida, June 2-6 2002.

**Table 7.1 Total Peptides and Proteins Identified from Ion Exchange Fraction at 24 min.**

| # | Description | Ref# | Peptide Sequence | MW (kDa) | Subcellular location |
|---|---|---|---|---|---|
| 1 | Thymosin beta-4 | P01253 | KTETQEKNPLPSK<br>TETQEKNPLPSKETIEQEK | 4.9 | Cytoplasmic |
| 2 | G(I)/G(S)/G(O) gamma-10 subunit | P50151 | DALLVGVPAGSNPFREPR | 7.4 | |
| 3 | Ubiquitin | P02248 | ESTLHLVLR | 8.6 | Nuclear and cytoplasm |
| 4 | SH3 domain-binding glutamic acid-rich-like protein | O75368 | GDYDAFFEAR | 12.8 | |
| 5 | alpha globin | gi\|28549 | VGAHAGEYGAEALER | 13.6 | |
| 6 | Vacuolar ATP synthase subunit G 1 | O75348 | mTILQTYFR | 13.9 | |
| 7 | Histone H2A.family | P02261 | HLQLAIR | 14.0 | |
| 8 | Galectin-1 | P09382 | VRGEVAPDAK<br>DSNNLCLHFNPR | 14.9 | |
| 9 | Hemoglobin beta chain | P02023 | VKVDEVGGEALGR<br>LHVDPENFR | 16.0 | |
| 10 | Calmodulin | P02593 | VFDKDGNGYISAAELR | 16.7 | Calcium binding proteins |
| 11 | synaptophysin homolog | gi\|2136213 | TVTATFGYPFR | 16.9 | |
| | synaptophysin-like protein isoform a | gi\|5803185 | | 28.9 | |
| | synaptophysin-like protein isoform b | gi\|33239443 | | 26.7 | |
| 12 | Myosin light chain alkali | P16475 | EAFQLFDR | 17.0 | |
| 13 | Peripheral-type benzodiazepine receptor | P30536 | FVHGEGLR | 18.9 | Integral membrane protein |
| 14 | Bone marrow stromal antigen 2 | Q10589 | RENQVLSVR | 20.0 | Type II membrane protein |
| 15 | Prothymosin alpha | P06454 | RAAEDDEDDDVDTKK | 21.1 | Nuclear |
| 16 | Brain acid soluble protein 1 | P80723 | EKPDQDAEGKAEEK<br>KGYNVNDEK | 22.5 | |
| 17 | Heat shock 27 kDa protein | P04792 | QDEHGYISR<br>LFDQAFGLPR | 22.8 | Cytoplasmic |
| 18 | Synaptosomal-associated protein 23 | O00161 | AHQITDESLESTR<br>RILGLAIESQDAGIK<br>DRIDIANAR | 23.7 | Plasma membrane |
| 19 | Ras-related protein Rab-7 | P51149 | FQSLGVAFYR | 23.8 | Found on late endosomes |
| 20 | Ras-related protein Rab-2B | Q8WUD1 | GAAGALLVYDITR | 24.4 | |
| | Ras-related protein Rab-2A | P08886 | | 23.7 | |
| 21 | Galectin-3 | P17931 | GNDVAFHFNPR | 26.1 | Nuclear |
| 22 | Protein Plunc | Q9NP55 | VTDPQLLELGLVQSPDGHR | 26.8 | Secreted |

| | | | | | |
|---|---|---|---|---|---|
| 23 | Chloride intracellular channel protein 1 | O00299 | GFTIPEAFR | 27.2 | Nuclear membrane |
| 24 | Acidic leucine-rich nuclear phosphoprotein 32 family member A | P39687 | LLPQLTYLDGYDR | 28.7 | Cytosol |
| 25 | Lymphocyte function-associated antigen 3 | P19256 | DKVAELENSEFR | 28.8 | Type I membrane protein |
| 26 | Carbonic anhydrase II | P00918 | YDPSLKPLSVSYDQATSLR | 29.2 | Cytoplasmic |
| 27 | F-actin capping protein beta subunit | P47756 | KLEVEANNAFDQYR | 31.6 | |
| 28 | Crk-like protein | P46109 | IHYLDTTTLIEPAPR | 33.9 | |
| 29 | Galectin-8 | O00214 | EEITYDTPFKR<br>ADVAFHFNPR<br>DIALHLNPR | 35.7 | |
| 30 | Glyceraldehyde 3-phosphate dehydrogenase | P04406 | LVINGNPITIFQER | 36.1 | Cytoplasmic |
| 31 | ADP-ribosyl cyclase 2 | Q10588 | FmPLSDVLYGR | 36.3 | Attached to membrane by a GPI-anchor. |
| 32a | Guanine nucleotide-binding protein beta subunit 4 | Q9HAV0 | AGVLAGHDNR | 37.6 | |
| | G(I)/G(S)/G(T) beta subunit 1 | P04901 | | 38.2 | |
| | G(I)/G(S)/G(T) beta subunit 2 | P11016 | | 38.0 | |
| 32b | Guanine nucleotide-binding protein beta subunit 4 | Q9HAV0 | ELPGHTGYLSCCR | 37.6 | |
| | G(I)/G(S)/G(T) beta subunit 2 | P11016 | | 38.0 | |
| 32c | G(I)/G(S)/G(T) beta subunit 1 | P04901 | IYAmHWGTDSR | 38.2 | |
| | G(I)/G(S)/G(T) beta subunit 2 | P11016 | | 38.0 | |
| 33 | chromosome 8 open reading frame 2 | gi\|6005721 | IEEGHIGVYYR<br>VTKPNIPEAIR | 38.0 | |
| 34 | Urokinase plasminogen activator surface receptor | Q03405 | SGCNHPDLDVQYR | 38.6 | Attached to membrane by a GPI-anchor |
| 35 | Annexin A2 | P07355 | DLYDAGVKR<br>GVDEVTIVNILTNR | 38.7 | Lamina beneath the plasma membrane |
| 36 | Fructose-bisphosphate aldolase A | P04075 | AAQEEYVKR<br>ADDGRPFPQVIK | 39.7 | |
| 37 | G(i), alpha-2 subunit | P04899 | FEDLNKR<br>IIHEDGYSEEECR | 40.9 | |
| 38 | G(i), alpha-1 subunit | P04898 | EIYTHFTCATDTK | 40.8 | |
| | G(i), alpha-2 subunit | P04899 | | 40.9 | |
| | (G(i) alpha-3) | P08754 | | 40.9 | |
| 38b | G(i), alpha-1 subunit | P04898 | TTGIVETHFTFK | 40.8 | |
| | G(i), alpha-2 subunit | P04899 | | 40.9 | |
| | (G(i) alpha-3) | P08754 | | 40.9 | |
| | G(O), alpha subunit 1 | P09471 | | 39.9 | |
| | G(O), alpha subunit 2 | P29777 | | 40 | |

| | | | | | |
|---|---|---|---|---|---|
| 39 | Flotillin-2 | Q14254 | DVYDKVDYLSSLGK | 41.8 | Membrane-associated protein of caveolae |
| 40 | Actin, cytoplasmic 1 | P02570 | qEYDESGPSIVHRK | 42.0 | |
| | Actin, cytoplasmic 2 | P02571 | (D)ESGPSIVHR | 42.1 | |
| | | | KDLYANTVLSGGTTmYPGIADR | | |
| | | | QEYDESGPSIVHR | | |
| 40b | Actin, alpha skeletal muscle | P02568 | SYELPDGQVITIGNER | 42.4 | |
| | Actin, cytoplasmic 1 | P02570 | HQGVmVGmGQK | 42.0 | |
| | Actin, cytoplasmic 2 | P02571 | DSYVGDEAQSKR | 42.1 | |
| | Actin, aortic smooth muscle | P03996 | AVFPSIVGRPR | 42.4 | |
| | Actin, aortic smooth muscle | P04270 | AVFPSIVGRP | 42.3 | |
| | Actin, gamma-enteric smooth muscle | P12718 | | 42.2 | |
| 41 | Actin-like protein 2 | O15142 | GYAFNHSADFETVR | 45.0 | |
| 42 | Nogo-66 receptor homolog 1 | gi\|30425563 | HLQALEELDLGDNR | 46.8 | |
| 43 | Tubulin beta-2 chain | P05217 | IREEYPDR | 50.3 | |
| | Tubulin beta-5 chain | P05218 | | 50.1 | |
| | Tubulin beta-1 chain | P07437 | | 50.2 | |
| | Tubulin beta-4q | Q99867 | | 48.9 | |
| 44 | Protein disulfide isomerase A6 | Q15084 | NRPEDYQGGR | 48.5 | Endoplasmic reticulum lumen |
| 45 | Adenylyl cyclase-associated protein 1 | Q01518 | LEAVSHTSDMHR | 51.9 | Cell membrane |
| | | | SGPKPFSAPKPQTSPSPK | | |
| 46 | Aldehyde dehydrogenase 7 | P43353 | NLATQLDSAFIR | 52.4 | |
| 47 | Solute carrier family 2, facilitated glucose transporter, member 1 | P11166 | GTADVTHDLQEmKEESR | 54.4 | Integral membrane protein |
| | | | SFEmLILGR | | |
| 48 | ATP synthase beta chain, mitochondrial | P06576 | AHGGYSVFAGVGER | 56.5 | Mitochondrial |
| 49 | Protein disulfide isomerase | P07237 | YKPESEELTAER | 57.5 | Endoplasmic reticulum lumen |
| | | | (A)DAPEEEDHVLVLR | | |
| 50 | Keratin, type I cytoskeletal 10 | P13645 | IRLENEIQTYR | 59.7 | |
| 51 | Catalase | P04040 | LFAYPDTHR | 59.8 | Peroxisomal |
| 52 | Tripeptidyl-peptidase I | O14773 | LYQQHGAGLFDVTR | 61.7 | Lysosomal |
| 53 | Glypican-1 | P35052 | GFSLSDVPQAEISGEHLR | 62.7 | Attached to the membrane by a GPI-anchor. |
| 54 | Glypican-4 | P75487 | YTEQLKPFGDVPR | 63.4 | |
| 55 | CTP synthase | P17812 | KLYGDADYLEER | 67.3 | |
| 56 | Ezrin | P15311 | APDFVFYAPR | 69.3 | Microvillar peripheral membrane protein |
| | Moesin | P26038 | | 67.8 | |
| | Radixin | P35241 | | 68.6 | |

| | | | | | |
|---|---|---|---|---|---|
| 57 | Heat shock cognate 71 kDa protein | P11142 | STAGDTHLGGEDFDNR | 71.1 | |
| | Heat shock-related 70 kDa protein 2 | P54652 | FEELNADLFR<br>IINEPTAAAIAYGLDKK | 70.3 | |
| 58 | L-plastin | P13796 | NEALIALLR | 70.8 | |
| 58b | L-plastin | P13796 | LSPEELLLR | 70.8 | |
| | T-plastin | P13797 | | 70.4 | |
| | I-plastin | Q14651 | | 44.3 | |
| 59 | unconventional myosin 1G valine form | gi\|14269502 | GHTAAEASYAR | 72.7 | |
| | FLJ00121 protein | gi\|18676448 | VIGFSPEEVESVHR | 50.4 | |
| | similar to Myosin Id | gi\|29734952 | | 107.1 | |
| 60 | Growth factor receptor-bound protein 2 | P29354 | FGNDVQHFK | 74.8 | |
| 61 | ATP-binding cassette, sub-family D, member 3 | P28288 | STHSELLEDYYQSGR | 75.9 | Integral membrane protein |
| 62 | Annexin A6 (Annexin VI) | P08133 | ILISLATGHR<br>DLEADIIGDTSGHFQK | 76.0 | Binding site for calcium and phospholipid |
| 63 | Heterogeneous nuclear ribonucleoprotein M | P52272 | FEPYANPTKR | 77.7 | Nuclear |
| 64 | Heat shock protein HSP 90-beta | P08238 | LGIHEDSTNR | 83.4 | Cytoplasmic |
| 65 | Heat shock protein HSP 90-alpha | P07900 | TKPIWTR<br>ALLFVPR | 84.5 | Cytoplasmic |
| 66 | Gelsolin, plasma (Actin-depolymerizing factor) | P06396 | HVVPNEVVVQR | 85.7 | Cytoplasmic |
| 67 | Vacuolar proton translocating ATPase 116 kDa subunit A isoform 3 | Q13488 | QGHEPQLAAAHTDGASER<br>LGELGLVEFR<br>qGHEPQLAAAHTDGASER | 93.7 | Integral membrane protein |
| 68 | Vacuolar proton translocating ATPase 116 kDa subunit A isoform 1 | Q93050 | NFLELTELK | 96.5 | |
| 69 | Ras GTPase-activating protein 3 | Q14644 | NFLDLISSSGR | 97.0 | |
| 70 | Hypothetical protein KIAA0143 | Q14156 | FANIEEDTPSYHR | 99.9 | Integral membrane protein |
| 71 | Aminopeptidase N | P15144 | LPNTLKPDSYR<br>YLSYTLNPDLIR | 109.7 | Type II membrane protein |
| 72 | signaling inositol polyphosphate 5 phosphatase SIP-110 | gi\|1245337 | KFLPSTANR | 109.8 | |
| | signaling inositol polyphosphate 5 phosphatase SIP-145 | gi\|1277082 | QQQYADLLSHDQLLTER | 132.9 | |
| 72b | signaling inositol polyphosphate 5 phosphatase SIP-145 | gi\|1277082 | AYALCVLYR | 132.9 | |
| 73 | Signal transduction protein CBL-B | Q13191 | ADAAEFWR | 110.9 | Nuclear |
| 74 | Sodium/potassium-transporting ATPase alpha-1 chain | P05023 | LSLDELHR<br>DAFQNAYLELGGLGER | 114.1 | Integral membrane protein |

| | | | | | |
|---|---|---|---|---|---|
| 75 | Clathrin heavy chain 1 | Q00610 | IVLDNSVFSEHR | 193 | Cytoplasmic face |
| 76 | Clathrin heavy chain 2 | P53675 | GQFSTDELVAEVEKR | 189.0 | Cytoplasmic face of coated pits and vesicles |

# Chapter 8

## Conclusions and Future Work

The goal of this work is to solve challenging biological problems, such as membrane protein analysis, with mass spectrometry and the use of surfactants. Because membrane proteins play an important role in biological systems, much attention and effort have been devoted to proteomics studies of this type of proteins. The physical properties of membrane proteins requires surfactants to be present in the sample workup for protein solubilization and the prevention of protein aggregation. However, the presence of surfactants in samples is a challenging issue because mass spectrometry requires relatively clean protein samples to achieve sensitive and accurate mass analysis. In this thesis, with understanding of the principle of how and why surfactants affect mass spectrometric analysis, surfactant-, especially SDS-, aided mass spectrometric proteomic analysis methods for membrane proteins have been developed.

In Chapter 2, a MALDI sample deposition method has been applied to the analysis of protein samples containing the surfactant, sodium dodecyl sulfate. SDS is commonly used surfactant in biological sample work up, such as gel electrophoresis, cell extract or hydrophobic protein solubilization. Earlier MS studies showed that SDS is detrimental to MS analysis even with low amounts present in a protein sample. The results described in Chapter 2 demonstrated that a two-layer sample preparation method can be very useful in analyzing protein and peptide samples containing SDS by MALDI MS. This method involves first the formation of a thin layer of matrix crystals, followed by deposition of a mixture of sample and matrix dissolved in a proper solvent and with an

adjustable sample to matrix ratio. Several variables related to the sample preparation are shown to affect the intensity of the MALDI signal. With this method, an on-probe washing step should be performed. The matrix to SDS ratio is not critical, but the protein to SDS ratio can have a significant effect on the MALDI signal response. Thus dilution or concentration of the sample solution may be required when this method is applied to analyze samples with unknown protein concentration. With the two-layer sample deposition method samples containing SDS up to 1% can be readily analyzed by MALDI MS without a significant effect to the signal intensity. The results also demonstrate that hydrophobic protein samples in the presence of SDS can be readily detected by MALDI. For direct analysis of cell extract by MALDI, it is now possible to incorporate solvent extraction involving SDS to examine more hydrophobic proteins.

While protein samples containing SDS up to ~1% can be analyzed by the two-layer sample/matrix preparation method, the MALDI signal quality was found to be degraded in these samples. In Chapter 3, the cause of this degradation was examined and proven to be the formation of protein sodium adduct ions in the MALDI process. If the instrument resolving power is inadequate to resolve these adducts, peak broadening is observed and the centroid mass assignment is shifted to higher mass. Ammonium dodecyl sulfate (ADS), hydrogen dodecyl sulfate (HDS), and tris (hydroxymethyl) aminomethane dodecyl sulfate (TDS) were explored as alternatives to SDS. We found that the ADS was least detrimental to MALDI performance. The presence of ADS in the protein samples did not significantly degrade the signal resolution and accuracy, as was the case with SDS. With ADS in protein samples, the protonated peak was dominant in MALDI analysis from moderate protein mass up to ~25kDa. As the protein molecular

weight increases, the ammonium adducts become dominant in the MALDI spectrum; but ADS still yields the best results of the four surfactants. The results demonstrated that ADS solubilized cell extracts could be separated by SDS-polyacrylamide gel electrophoresis by mixing the samples with normal SDS buffer. The improved MALDI performance of an ADS-solubilized detergent insoluble fraction from Raji/CD9 B-cell extract was also demonstrated in this work. It can be concluded that ADS is a viable alternative to SDS in protein extraction and purification and should be used in MALDI analysis of very hydrophobic proteins.

In Chapter 4, the performance of surfactants in proteomics analysis was further studied. The effects of four frequently used surfactants, namely CHAPS, Triton X-100, OG and SDS, for biological sample preparation on enzymatic digestion and peptide MS analysis are examined in detail. CHAPS, TX-100 and OG have no effect on the enzymatic digestion, but CHAPS and Triton surfactants severely affect the MALDI MS analysis and therefore should be removed before MALDI MS peptide analysis. If the protein samples can be dissolved by OG, enzymatic digestion and MS analysis can be performed with OG up to 1% without any interference. SDS can be used in tryptic digestion with concentration up to 0.1% without any interference in the MALDI MS analysis. The performance of SDS-solubilized hydrophobic membrane protein tryptic digestion was demonstrated in presence of 0.1% SDS, and it leads to an improvement in the sequence coverage. We envision that SDS extracted protein mixture can be directly analyzed without purification or SDS removal steps by direct enzyme digestion and identification with tandem mass spectrometry.

In Chapter 5, a new solution-based SDS-aided digestion and protein identification method for membrane proteomic research was developed with the facilitation of LC-MALDI interface. Membrane protein analysis usually requires the use of surfactants for extraction and solubilization, which increases the difficulty of analyzing them by solution-based LC-ESI MS/MS and thus the identification is usually achieved by a gel-based approach. In this solution approach, we used SDS to solubilize membrane proteins without worrying about the subsequent tryptic digestion, HPLC separation, and peptide detection and sequencing experiments. Both HPLC-MALDI MS/MS and 2D LC-ESI MS/MS were applied to achieve solution-based peptide sequencing and protein identification by using the membrane protein fraction of cancer cell line HT29 as a model. Three approaches, solution-based HPLC-MALDI MS/MS, solution-based 2D LC-ESI MS/MS and conventional gel peptide mass mapping-MALDI MS & MS/MS, were employed to equal amounts of membrane protein sample, and a total of 112 proteins were identified. However, in-solution-based HPLC-MALDI MS/MS method identified the largest number of proteins. This work clearly illustrates that, for comprehensive analysis of proteomes of interest, a preferred shotgun proteomics approach should incorporate both LC/MALDI and LC/ESI MS/MS technologies.

In Chapter 6, the membrane protein identification method developed in Chapter 5 is applied to lipid raft protein identification of Triton X-100 insoluble fraction THP-1 cell. 71 proteins are identified by solution-based HPLC-MALDI MS/MS approach, while 45 of them were not identified by previous gel/MALDI approach. The results of Chapters 5 and 6 show that the developed SDS-aided solution-based protein identification method is very useful for membrane protein identifications. Combined

with semi-offline HPLC-MALDI interface, this method can be employed for SDS-solubilized membrane protein identification. However, the method is also suitable for ESI mass spectrometry if SDS can be removed prior to analysis, such as SDS removal with a 2D-LC system. This method is found to provide complementary results to conventional gel electrophoresis-MALDI approach for lipid raft analysis. The results in Chapter 5 and 6 clearly show that, for a complete analysis of the proteome of interest, a combination of conventional gel approach and shotgun technologies would provide a better proteome map in terms of protein identification.

In Chapter 7, the solution-based SDS-aided shotgun proteomics approach is further improved by utilizing two-dimensional separation technology. If the sample amount is unlimited, the implementation of 2D LC with SDS-aided proteomics method can fulfill the mission of providing a complete proteome map. The lipid raft proteome is analyzed with incorporation of two offline shotgun approach, 2D LC-ESI MS/MS and 2D LC-MALDI MS/MS. There are 60 proteins identified with 2D LC-ESI approach and half the proteins are not identified from the previous studies. The 2D LC-MALDI MS/MS provided more exciting results and 185 proteins have already been identified. By increasing the sample loading amount, the low abundant components in the raft proteome can be analyzed with LC-MALDI shotgun approach. The results demonstrated that LC-MALDI offers several unique features that are not seen in LC-ESI. The results also indicate that offline 2D LC-MALDI is a robust technique for shotgun proteomic research.

Future work will be directed to the improvement of shotgun LC-MALDI technology by establishment of automation data collection system for the MALDI QqTOF instrument. With data collection fully automated, analysis of one plate LC-

MALDI sample can be carried out in reasonable time frame (about 2 days). Because the LC-MALDI interface is readily interfaced with narrowbore or microbore columns, large quantity of sample can be easily analyzed by LC-MALDI MS/MS technique. Compared with LC-ESI, LC-MALDI requires a fraction of the anlysis time per sample, making the latter a much more attractive method for large numbers of samples. The fully automated MALDI QqTOF instrument can make the developed SDS-aided solution-based protein identification method more useful and robust.

The methods described in this thesis can be further applied to proteomics research, such as cancer studies, in terms of protein identification. Another application of the described methods would be the incorporation of a SDS-aided solution-based technique with protein quantitation reagents or methods, e.g. ICAT, to perform membrane protein quantitation. It is hoped that further application of the methods described in this thesis with automated MALDI QqTOF will provide useful techniques for proteomics research.