INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

University of Alberta

Coupling Gel Electrophoresis with Mass Spectrometry for Protein Characterization and Identification

by

Jing Zheng

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

Department of Chemistry

Edmonton, Alberta

Fall 2000

University of Alberta

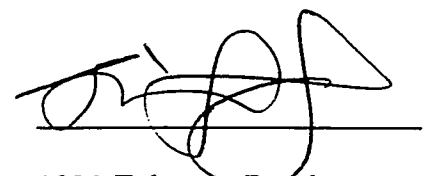
Library Release Form

Name of Author: Jing Zheng

Title of Thesis: Coupling Gel Electrophoresis with Mass Spectrometry for Protein Characterization and Identification

Degree: Master of Science

Year this Degree Granted: 2000



1031 Falconer Road,
Edmonton, Alberta
Canada
T6R 2C6

September 27, 2000

Abstract

Coupling gel electrophoresis with mass spectrometry is a powerful tool for protein mixture analysis. In this study, a passive protein elution method from gels was developed. In conjunction with the time lag focusing MALDI TOF mass spectrometer, this method enables accurate molecular weight determination for some proteins. However, the accuracy may be greatly reduced for proteins containing cysteine residues due to gel-induced modification. A crushing-gel/in-gel digestion method is also described. Compared with traditional approaches, this method has better extraction efficiency and is less time consuming. Several applications of this method are demonstrated. Along with appropriate mass spectrometric techniques including MALDI and NanoESI, seventeen proteins from biological samples separated by gel electrophoresis were identified. Some technical issues brought by these applications, such as the choice of suitable tandem mass spectrometric methods, and combining 2D gel databases with sequence databases to aide identification, are discussed. The merits and limitations of the mass spectrometric techniques are especially addressed.

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 **Coupling Gel Electrophoresis with Mass Spectrometry for Protein Characterization and Identification** submitted by Jing Zheng in partial fulfillment of the requirements for the degree of Master of Science.

Liang Li

Dr. Liang Li, Professor of Chemistry

Monica Palcic

Dr. Monica M. Palcic, Professor of Chemistry

David Wishart

Dr. David Wishart, Assistant Professor of Pharmacy and Pharmaceutical Sciences

Sept. 26, 2000

To my parents,
my husband and my children

Acknowledgements

I would like to thank my supervisor Professor Liang Li for his continuous encouragement, inspiration and enthusiasm during the last three years. I appreciate having the opportunity to work in his research group where I have obtained training with modern analytical instruments. I am impressed by his vast knowledge and keen thinking and have benefited from his constructive advice. I also thank him for hiring me as a research assistant during last two years so that I could concentrate on my research project.

I would also like to thank my oral examination committee members Professor Monica M. Palcic and Professor David Wishart for their review and comments regarding this thesis.

I would like to acknowledge the following people for providing help to my research: Professor Paul G. Scout and Dr. Kalal Derhami for the preparation of human skin fibroblasts 2D gel samples and the support of some gel reagents; Professor David Wishart and Ms. Ping Li for the preparation of human rhinovirus-3C protease gel samples and Professor Monica M. Palcic for the loan of her mini gel system for research of accurate molecular weight detection of gel-separated protein. Thanks also to Dr. Michael A. Jeannot with whom I worked at beginning of my research. I thank him for his guidance and advice.

Special thanks to the other group members of Dr. Li's group: Mr. Rui Chen, Mr. Alan Doucette, Ms. Andrea Kiceniuk, Ms. Nan Zhang, Ms. Zhengping Wang, Mr. Dave Craft, Mr. Chris MacDonald, Mr. Bernd Keller, Dr. Randy Murray Whittal and Dr. Ken

Yeung for their time and discussion in the process of writing this thesis. I especially thank Ms. Zhengping Wang for her preparation of *E. coli* sample.

I thank Department of Chemistry and University of Alberta for providing the facilities; the Natural Sciences and Engineering Research Council of Canada and Genomic Solutions Inc. for providing funding for my research.

Special thanks to my family for making my graduate studies possible and successful. I thank my husband Qizhu Ding for his support and understanding in last three years. I thank my son Chuck Ding for his help in some of the graphics formatting. I thank my daughter Lina Ding for her patience while waiting for my arrival at the door almost every day and never complaining about my lateness.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ATCC	American Type Culture Classification)
CID	collision induced association
DE	delayed extraction
2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
DTT	dithiothreitol
E. coli	*Escherichia coli*
ESI	electrospray ionization
FAB	fast-atom bombardment
HCCA	α-cyano-4-hydroxy-*trans*-cinnamic acid
HRV-3C	human rhinovirus-3C
IEF	isoelectric focusing
MALDI	matrix assisted laser desorption ionization
MS^n	tandem mass spectrometry
m/z	mass-to-charge ratio
NanoES	nano electrospray
PD	plasma desorption ionization
PSD	post source decay
rf	radio frequency
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SWIFT	stored waveform inverse Fourier transformation
TEMED	tetramethylenediamine

TLF	Time-Lag Focusing
TOF	time-of-flight
TFA	trifluoroacetic acid
v/v	volume-to-volume ratio

Chapter 1

Coupling Gel Electrophoresis with Mass Spectrometry for Protein Identification

1.1 Gel Electrophoresis for Protein Separation and Purification

1.1.1 Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

Electrophoresis is one of the most widely used laboratory techniques for biological sample analysis. This technique employs the characteristic charges of biomolecules to fulfill the separation and purification task.

Basically, a molecule which has charge, q, in an electric field, E, experiences an electric force, F, ($F = qE$) and an opposing force which, in free solution, can be calculated as fv, in which v is the velocity of the molecule and f is the frictional coefficient. Molecules in the electric field will rapidly reach steady state, terminal velocities, with the balance of these two forces: $qE = fv$. The electrophoretic mobility of the molecule, μ, defined as the steady state velocity per unit field, or $\mu = v/E = q/f$, describes mobility of the molecules in the electric field. In other words, electrophoretic mobility of the molecule is proportional to the ratio of the net charge of the molecule to its frictional coefficient, which related to the size and shape of the molecule.

But charge and shape of biomolecules are affected by several external factors including the pH of the environment, the types and amounts of ions in the solution, and denaturants such as detergents, reducing reagents and urea. As a result, molecules with the same size and charge may not have the same electrophoretic mobility unless the other factors are kept constant. Finally, in order to maximize the separation between molecules, the difference in ion mobility between the molecules must be maximized.

Polyacrylamide gel electrophoresis was introduced by Orstein [1] for protein separation. Polyacrylamide gels are formed by copolymerization of acrylamide (CH_2=CH-CO-NH_2) with a cross-linking comonomer, usually N,N'-methylenebisacrylamide (bisacrylamide) (CH_2=CH-CO-NH-CH_2-NH-CO-CH=CH_2). A free radical initiator, commonly ammonium persulfate, along with the accelerator tetramethylenediamine (TEMED), generally initiates polymerization. Cross-linking co-monomers are bifunctional acrylic reagents that covalently link to adjacent polyacrylamide chains thus form three-dimensional network (small pores). Pore sizes of the gel decrease when monomer percentage increases in a fixed, low cross-linker concentration (smaller than 5%). The gel pores act as a molecular sieve to differentiate mobility of proteins with different sizes. Different gel systems were proposed for different purposes. Among various gel systems, SDS-PAGE (sodium dodecyl sulfate polyacrylamide gel electrophoresis) [2] is able to separate proteins by their molecular weights.

In the SDS-PAGE system, proteins are denatured by SDS, a strong anionic detergent and reduced by a thiol-reducing reagent, such as 2-mercaptoethanol, to break any disulfide bonds that formed between cysteine residues. At saturation, approximately 1.4 grams of SDS are bound to per gram of protein. The amount of bound SDS is almost proportional to the molecular weight of the protein and the proteins coated with SDS have similar extended, rod-like shape. Therefore dimensions of the proteins are proportional to their molecular weight. Thus, separation based on molecular weight becomes possible.

Figure 1.1 simply illustrates the SDS-PAGE process. A simplified procedure of SDS-PAGE for protein separation can be described as follows:

1. Casting the gel: Gels are cast into flat slabs. Usually, a stacking gel is also cast on top of the resolving gel in order to achieve better resolution and

separation. A comb is placed on the top of the gels before polymerization in order to create wells for sample loading. The process of polymerization is complete in approximately thirty minutes.

2. Reduction and denaturation of proteins: Proteins are treated in a sample buffer, which contains SDS and a reducing reagent for protein denaturation and reduction. Heating the sample at 95°C for five minutes aids in this process.
3. Sample Loading: Protein samples are loaded into gel wells by gel-loading pipette tips.
4. Electrophoresis: An electric field is applied to the gels with a setting of either constant current or constant voltage. Under the electric field, the anionic constituents of the system become aligned in order of their electrophoretic mobilities. The SDS-polypeptide complex in the sample is swept between the electrodes. After migrating through a stacking gel of high porosity, the protein complexes are deposited on the surface of the resolving gel, stacking into a small volume. After stacking, sample proteins enter the smaller pore resolving gel, where they migrate by sieving and separate from each other on the basis of molecular weight.
5. Protein detection: Protein detection is usually accomplished by staining the gels with dyes or metals [3]. One of the most sensitive detection methods is silver staining which allows for the visualization of down to 1-10 ng of protein per band. Coomassie brilliant blue, the most commonly used staining reagent, gives sensitivity of 0.1-1 μg of protein per band. Negative staining methods, such as copper staining, only stain the gel background, leaving the protein bands unstained. Sensitivity similar to that of coomassie blue stain is achieved with this method.

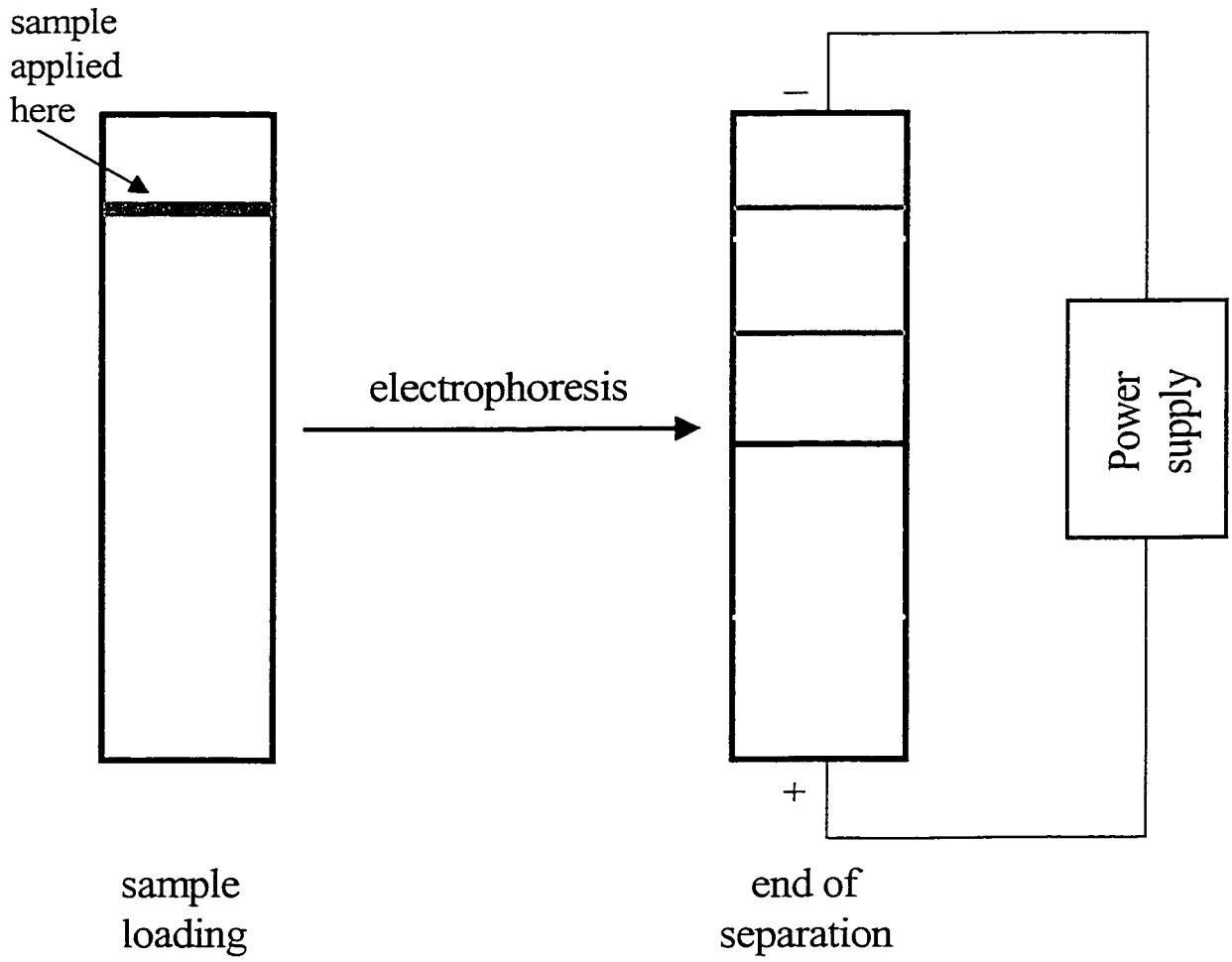


Figure1.1 A diagram illustrating SDS-PAGE process.

1.1.2 Isoelectric Focusing and Two Dimensional PAGE for Protein Separation and Purification

In reality, samples from biological sources often contain a large number of proteins, which can not be resolved using molecular weight separation exclusively. Therefore, two-dimensional separation is necessary. In two-dimensional PAGE (2D-PAGE), proteins are separated in the first dimension by isoelectric focusing (IEF) and, in the second dimension, by SDS-PAGE.

Proteins are amphoteric molecules whose net charges are determined by the pH of their local environment. When proteins move through a medium with variable pH, their net charges change according to the pH they encounter. Under an electric field, proteins in a pH gradient will migrate until they focus at the position where the pH of the environment is equal to its isoelectric point (pI), and thus its net charge is zero. As a result, proteins with different pI volues are separated. There are three types of IEF gels, classified according to the way the gel is cast: tubes, strips or slabs. To combine with SDS-PAGE for two-dimensional separation, tube or strip IEF gels must be selected. Synthetic carrier ampholytes (which usually are mixtures of polyamino-polycarboxylate that have closely spaced pI and high conductivity) are either added to the monomer solution, such that they are included in the tube gel after polymerization, or immobilized into the IEF strips, such that an immobilized pH gradient is obtained. In the former case, when a voltage is applied across the gel, the ampholytes partition into a smooth pH gradient, increasing monotonically from anode to cathode. In the later case, the gradient was formed when the strips were made.

In the experiments described throughout this thesis, tube gels were exclusively used for IEF-PAGE.

In conclusion, 2D-PAGE is currently the most powerful technique for separating proteins of very complex biological samples, such as the set of expressed proteins of an organelle, a tissue, or an organism [4-6].

1.2 Mass Spectrometry for Protein Identification

Mass spectrometers measure the mass to charge ratio (m/z) of gaseous ions of the analyte, where m is the ion mass in atomic mass units of Daltons (Da) and z is the number of elemental charge units. Generally, a mass spectrometer has two main components: an ion source for the production of ions and introduction of ions into gas-phase, and a mass analyzer for examining masses of the ions. It does not always measure masses, sometimes it measures some other properties related to mass, like momentum or kinetic energy. For the high mass molecules, generation of gas-phase ions is particularly challenging, especially for fragile biomolecules such as proteins, peptides, and DNA. Traditional ionization methods, such as electron impact, will only generate fragments of large biomolecules. Earlier ionization methods include fast-atom bombardment (FAB) and plasma desorption ionization (PD) are much softer processes which can generate intact biomolecular ions instead of fragments [7]. But the low ionization efficiency, the mass range of ionizable proteins and other practical issues restricted the application of these two ionization methods.

The breakthrough in mass spectrometry of proteins and peptides came in 1988 with the introduction of two soft ionization methods that introduce large, intact molecular ions into the gas phase. These are electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) [8-10] methods.

1.2.1 MALDI-TOF Mass Spectrometry for Protein Analysis

In the process of MALD ionization [11], analyte proteins and peptides are added to a solution that contains the matrix molecule. A drop of this mixture is deposited onto a sample probe and allowed to dry. Polypeptides co-crystallize with matrix and form a thin layer on the probe tip. Matrices are weak aromatic acids that are soluble in protein solvents and enhance polypeptide solubility. The most important aspect is that the matrix should have the ability to strongly absorb energy from the laser used for ionization (often a pulsed nitrogen laser at 337 nm). Irradiated by a short-pulsed laser, matrix molecules are ionized and energized so that they are able to lift analyte molecules up into gas-phase in an intact ion form. The process of ionization of biomolecules in MALDI is still not completely understood.

MALD ionization methods tend to generate low charge state ions. In the peptide mass ranges (usually m/z less than 5000), singly charged protonated molecular ions dominate the spectra. This characteristic makes MALDI spectra easily interpreted, especially when dealing with complex mixtures. Ion suppression is a common problem especially when dirty samples (samples which contain salts and/or surfactants) or mixture analytes are beening anlyzes by mass spectrometry. On-probe washing can efficiently remove most of the salt on the MALDI probe tip, therefore eliminating ion suppression effect in the MALD ionization process.

MALDI is a discontinuous ion production method and should be coupled with a pulsed mass analyzer. In it's most common configuration, MALDI is coupled with a time-of-flight (TOF) mass analyzer (see Figure 1.2). In a TOF mass analyzer, a packet of ions given the same kinetic energy fly down a field-free tube and arrive at an ion detector in order of their mass to charge values. All ions generated by a single shot are recorded. So the sensitivity is very high compared with a scanning analyzer such as a quadrupole.

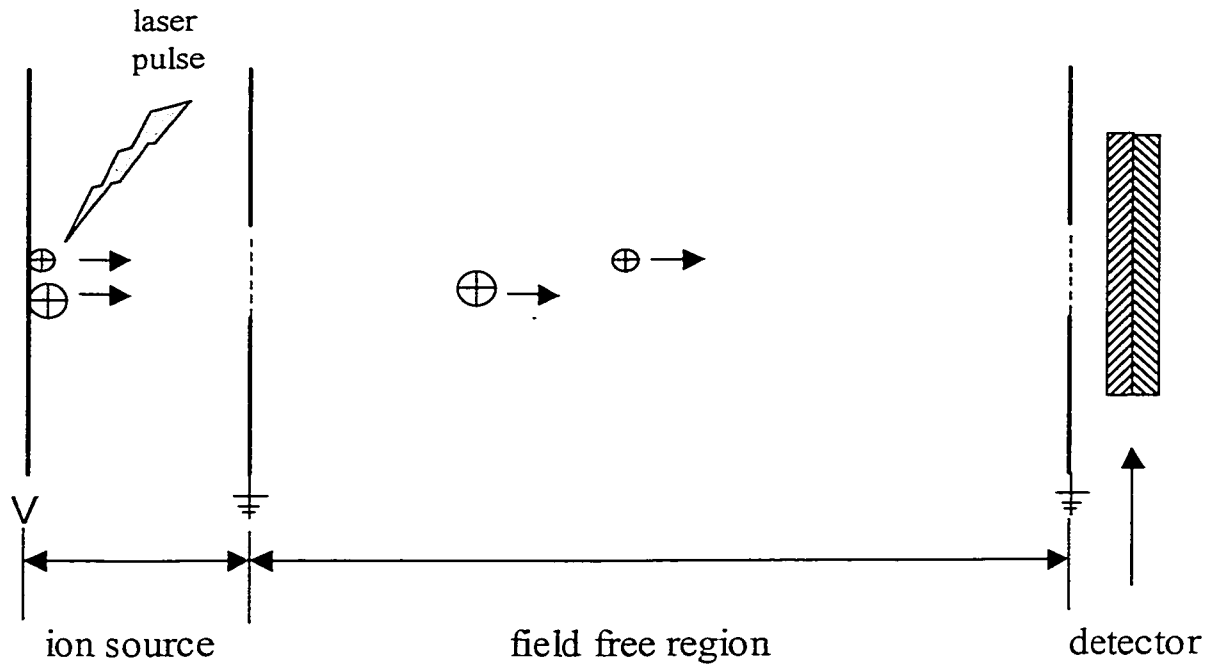


Figure 1.2 Schematic of a MALDI-TOF mass spectrometer.

When a post acceleration device (i.e., a device that accelerates ions at the end of the flight tube) is included, sensitive detection of high mass molecules becomes possible by TOF. Another advantage of a TOF analyzer is the enormous mass range. Theoretically, the mass range of a TOF mass spectrometer is unlimited. Practically, TOF gives the highest direct mass detection over all mass analyzers.

In conclusion, MALDI-TOF is the first choice for polypeptide analysis when dealing with (A) high mass protein detection, (B) mixtures of peptides such as in tryptic digests, and (C) dirty samples such as those extract from gels or biosamples which contain buffer ions and surfactants.

1.2.2 ES and NanoES Ionization Processes

In the electrospray (ES) ionization process [12], a solution containing the sample proteins or peptides is passed through a fine needle held at high potential, creating a fine spray of highly charged droplets that is directed across a small inlet orifice towards the analyzer. Between the spray needle and the orifice plate, the droplets are desolvated by drying gas and/or heating, resulting in the formation of gaseous-phase ions. This method of generation of gas phase ions is considered to be a soft ionization process. In the most common configuration, ES ion sources are interfaced with quadrupole or ion trap mass analyzers. Because the ionization process occurs in solution, ESI is easy to couple with liquid chromatography (LC-MS), and plays an important role in automation of complex biosample separation and protein identification.

ESI is characterized by the formation of multiply charged species. Ions from a pure polypeptide result in several, sometimes tens of different charge states, creating several, sometimes tens of multiple charged peaks in the mass spectrum. This is very

helpful in detecting proteins whose molecular weight are higher than the mass range of a quadrupole or ion trap analyzer, which is typically less than m/z 4000. For example, a protein with mass up to 200,000 Da can be analyzed, by the multiple charged peaks, using a quadrupole analyzer with 2400 Da mass limit [13]. Multiple charged peaks give better mass accuracy, because the mass information of a single molecule is held in several charged states, thereby providing the measurement of independent m/z. But the drawback of this characteristic is the complexity and difficulty of interpreting the ESI mass spectra, especially when analyzing polypeptide mixtures. Because of the surface ionization process associated with ESI, the ion suppression effect is always an important issue. Impurity tolerance of ESI-MS is also very poor. Even small amounts of detergents will cause uninterpretable complexity of mass spectra. High concentrations of salt can lead to signal suppression and clustering effects. Therefore sample cleanup before operation of ESI-MS is necessary for analyzing real world samples.

When dealing with a small amount sample and multicomponent mixtures, a nano electrospray (NanoES) source is particularly useful. NanoES source was developed and introduced by Mann's group in 1995 for protein analysis [14,15]. The NanoES source is essentially a nano scale ES source. A finely drawn gold-coated glass capillary placed 1-2 mm in front of the plate, which has an orifice leading to the mass spectrometer. The capillary tip has an orifice of 1-3 μm in diameter (compared to > 20 μm in conventional ES), which results in flow rate of 20-30 nL/min. This generates a reduced droplets diameter (calculated at 180 nm whereas conventional ESI generate droplets of 1-2 μm in diameter) [15] and a higher charge to volume ratio, which results in greater accessibility of ionization for analyte. In other words, the ionization process of NanoESI has less ion suppression effects compare with conventional ESI processes.

1.2.3 Peptide Mass Mapping

Peptide mass mapping is one of the most common ways to identify proteins by mass spectrometry [16]. In this process, proteins are digested by a protease such as trypsin. The resulted peptides are analyzed by mass spectrometry and the detected peptide masses are entered into a protein database search program in order to match the characteristic peptides to a theoretical digestion of the proteins. A number of programs of protein identification by sequence database and mass spectrometry are available on the world wide web sites. Examples are **MS-Fit** developed by the UCSF mass spectrometry group, **PeptIdent** by Swiss Institute of Bioinformatics, and **PeptideSearch** by the EMBL Protein & Peptide Group. These programs offer convenient tools for protein identification by peptide mass information (called peptide mass mapping). Their related Web sites are http://prospector.ucsf.edu/ucsfhtml3.2/msdi-gest.html, http://expasy.cbr.nrc.ca/tools/peptident.html, and http://www.mann.emblhei-delberg.de/Services/PeptideSearch/PeptideSearchIntro.html. The following information is usually entered for a protein identification: protein mass range, cleavage reagent, peptide mass accuracy, peptide masses, modification during the sample preparation process, and the number of missed cleavages. The retrieved proteins are ranked according to the numbers of matching peptide masses/or least unmatched masses and some statistic calculation. A scoring program named mouse score or a simple ranking system can usually be chosen. Note that these programs can only provide the statistically most probable identification. The evaluation needed by the investigator based on species of origin, apparent mass, sequence coverage and relative intensity of matched peptide peaks in the spectrum may help to determine the true identity of the protein.

1.3 Tandem Mass Spectrometry for Protein Analysis

In some cases, proteins can be identified by peptide mass mapping alone. But ambiguities arise in many situations caused by incomplete digestion, lack of adequate amount of protein or complexity of the sample. The results of this experiment may need to be confirmed by the information from another dimension such as primary structural information obtained from tandem mass spectrometry.

Mass spectrometers, which have tandem MS ability, include triple quadrupole, multiple sector, hybrid sector/quadrupole or time-of-flight, hybrid quadrupole/time-of-flight, ion trap and reflectron-TOF. Only the last two types of instruments will be described here, as they have been exclusively used in this research.

1.3.1 PSD TOF Mass Spectrometry

Reflectron-TOF was initially introduced to obtain increased resolving power by compensating for the initial energy distribution of the ions. However, it was found that by stepwise reducing the voltage applied to the ion mirror from its nominal value down to zero, one is able to detect metastable ions fragmented in the region between the ion source and the ion mirror (see Figure 1.3). Metastable ions are produced in following process: intact molecular ions leaving the ion source having acquired a sufficient internal energy during the desorption processes by photo activation or in-source collision will fragment while traveling in the field-free region of the instrument. This fragmentation process is called post source decay (PSD) [17,18]. These fragment ions have the same velocity as their precursor ions but a different kinetic energy. So, in a linear TOF mass spectrometer, the metastable ions will arrive at the detector at the same time as their precursor ions. In a reflectron-TOF, however, fragment ions can be

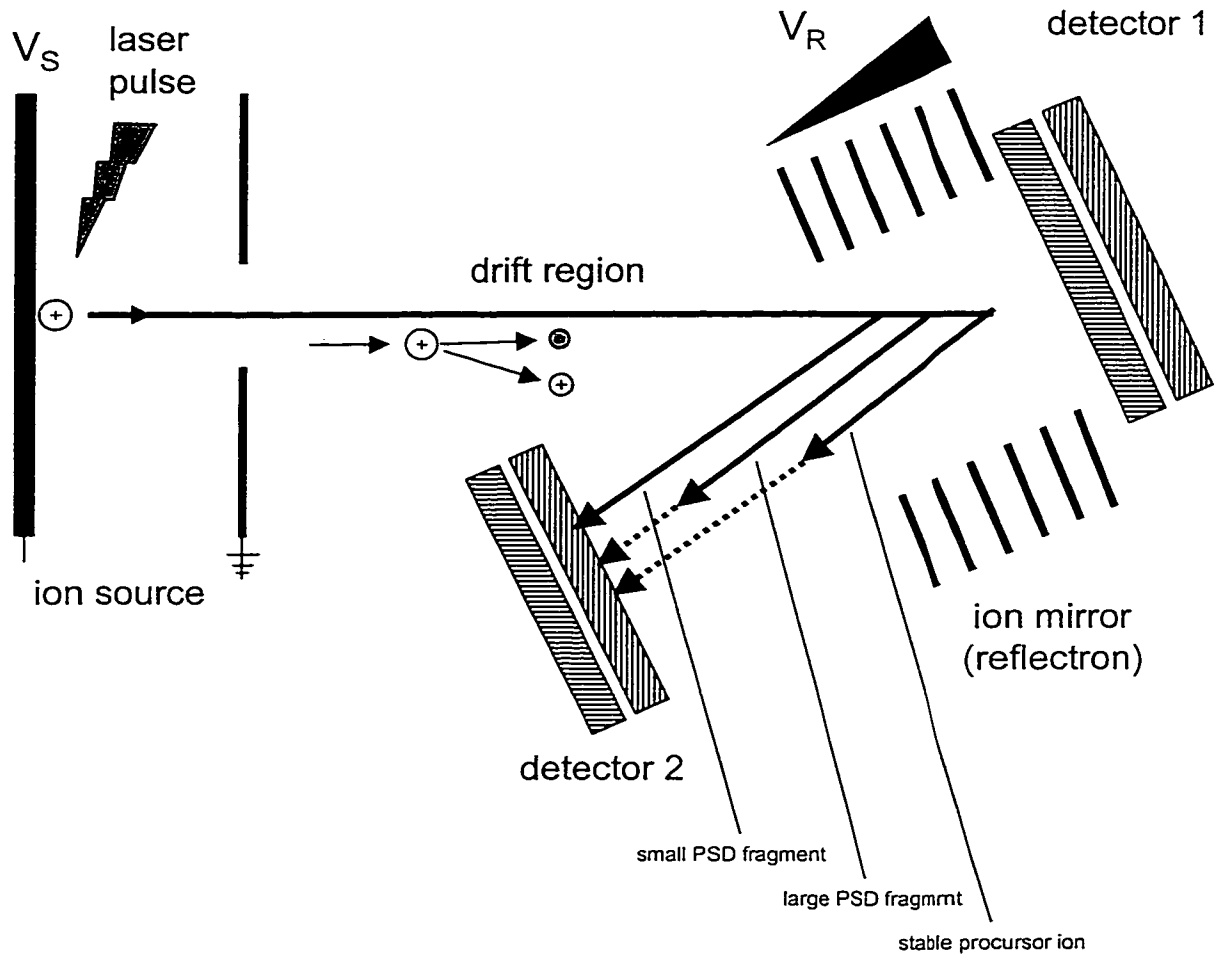


Figure1.3 Principle of PSD analysis in a Reflectron-TOF mass spectrometer.

differentiated by their kinetic energy by the time dispersions induced by the electrostatic reflector or ion mirror. Larger fragment ions (with higher kinetic energy) will penetrate deeper into the reflectron and will reach the detector later. By using an ion gate, the precursor ions such as a peptide ion from a peptide mixture can be selected while excluding other ions and the PSD spectrum of the peptide ions can be obtained. In operation, the potential applied to the reflectron is gradually reduced in order to acquire a complete spectrum by concatenation of several PSD spectra.

Sequencing with PSD is easy to perform because it can be achieved by the same sample used for peptide mass mapping. MALDI PSD TOF mass spectrometry retains the advantages of both MALD ionization and the TOF mass analyzer (see Section 1.2.1), but it should be noted that no ion accumulation or ion storage is possible whereas some tandem mass spectrometers like those with an ion trap are able to accumulate the ions of interest before the fragmentation step. So the source of fragment ions is limited. Generally, only relatively intense peaks can be selected for a PSD experiment. The resolution of a PSD fragment ion is not very good due to the broad kinetic energy distribution of the fragment ions during the fragmentation process.

1.3.2 Ion Trap Tandem Mass Spectrometry

Ion traps have been widely used in trapping and mass analysis [19,20]. Basically, the device contains a chamber formed by a ring electrode and two end-cap electrodes (see Figure 1.4). Radio frequency (rf) voltages are applied between the ring electrode and the two end-cap electrodes establishing a hyperbolic electric field. This field may create a stable trajectory so that certain mass range ions are trapped. Mass spectra are obtained by increasing the rf over time. The ions of successively greater mass-to-charge ratio drop into an unstable trajectory and are ejected and detected.

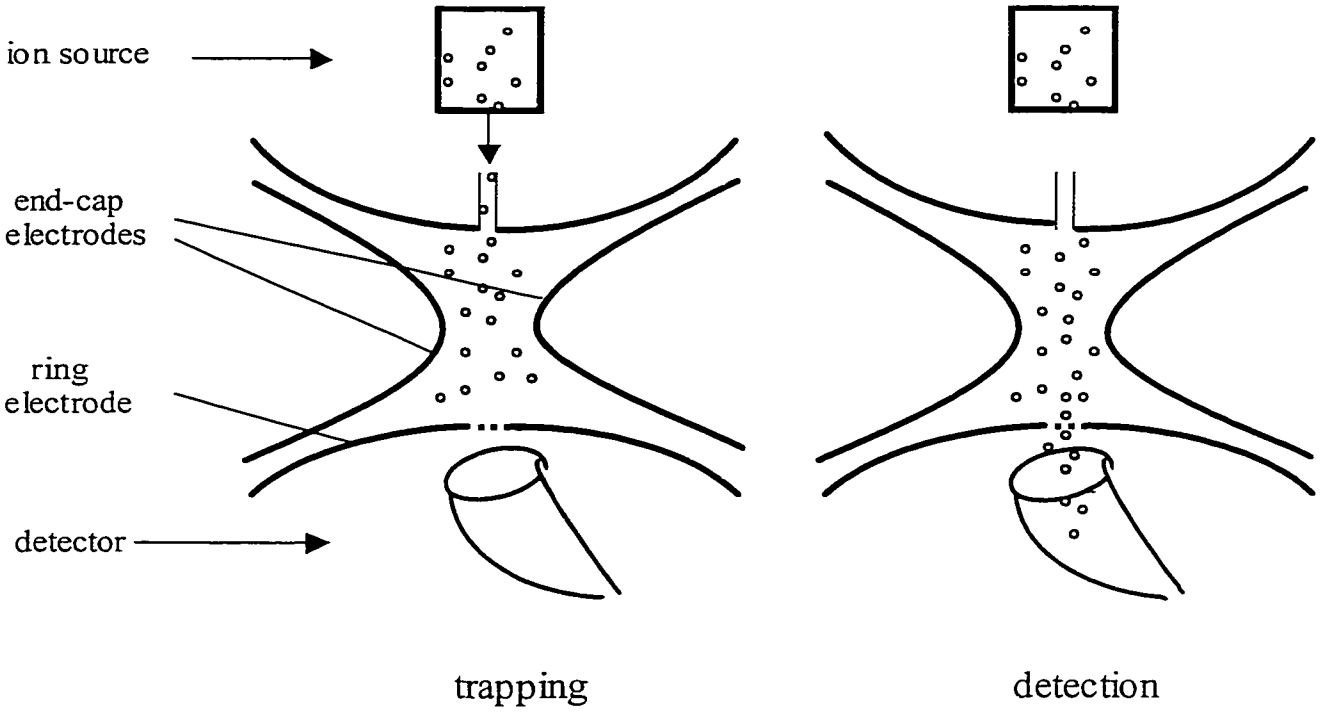


Figure1.4 Schematic of an ion trap mass spectrometer.

When a precursor ion of interest is selected, a stored waveform inverse Fourier transformation (SWIFT) pulse is applied to one of the end-caps, allowing isolation and accumulation. Thus, even low abundance ions have possibilities for tandem MS.

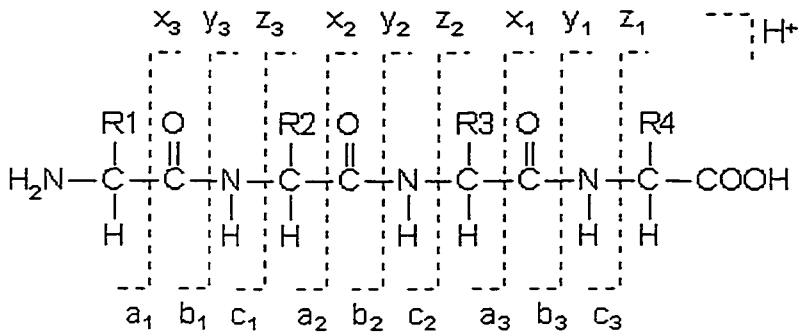
The ion trap initially was interfaced to an ESI source [20]. This configuration is perfect when dealing with clean and abundant samples. Shortcomings emerge when a small amount or a mixture of analytes needs to be analyzed (see Section 1.2.2 about the limitation of ESI).

Interfacing NanoESI source with ion trap MS preserves the advantages of both ion trap and NanoESI source. These include micro-scale sample consumption, longer analyzing time, lower ion suppression effect and high resolution. Therefore, NanoES ion trap is especially suited when dealing with a small amount of analyte mixture in tandem mass spectrometry.

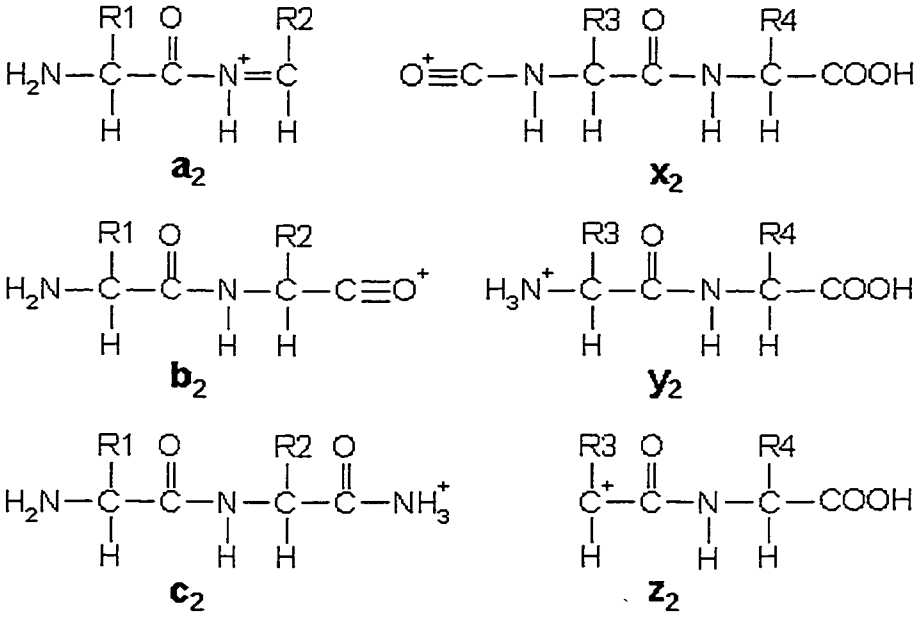
1.3.3 Interpretation of Tandem Mass Spectra

The types of fragment ions observed in a tandem spectrum depend on severe factors including the primary sequence and the charge state of the precursor ions, the amount of internal energy and how the energy was introduced. The accepted nomenclature for fragment ions was first proposed by Roepstorff and Fohlman [21]. Biemann first reviewed the fragment pattern and nomenclature for fragment ions in the 1980's [22]. Interpretation of peptide CID (collision induced association) tandem mass spectra was comprehensively reviewed by Papayannopoulos [23].

Fragments will only be detected if they carry at least one charge. Considering singly charged peptide ions, if the charge is retained on the N-terminal fragment, the ions are classified as either **a**, **b** or **c**. If the charge is retained on the C-terminal, the ions are named either **x**, **y** or **z**. The fragment on the other side is neutral. A subscript indicates the names of fragment ions:

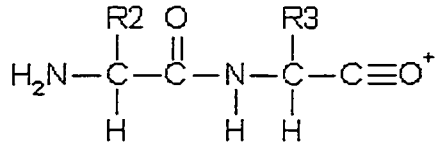


In addition to the proton(s) carrying the charge, **c** ions and **y** ions abstract an additional proton from the precursor peptide. Thus, the structures of the six singly charged sequence ions are:

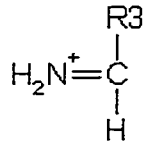


Other than sequence ions, double backbone cleavage gives rise to internal fragments. Usually, these are formed by a combination of **b** type and **y** type cleavage to produce the illustrated structure, an amino-acylium ion. Sometimes, internal cleavage ions can be formed by a combination of **a** type and **y** type cleavage, an amino-

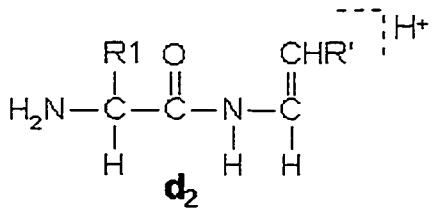
immonium ion. Internal fragments are labeled with their 1 letter amino acid code (e.g. KLDN):



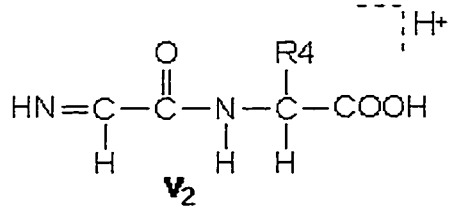
An internal fragment with just a single side chain formed by a combination of **a** type and **y** type cleavage is called an immonium ion. These ions are labeled with the 1 letter code for the corresponding amino acid (e.g. F).



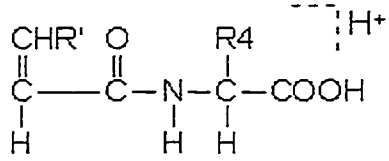
High energy CID (usually at keV) can generate additional fragment ion types due to side chain cleavages:



d_2



v_2



w_2

In these structures, R' is the substituent at the beta carbon following the loss of the gamma carbon by side chain cleavage. For example, side carbon structure at cleavage site of **d_2** undergoes the following cleavage: C-CH-R' ⟶ C=CH-R'

R''

Energy for fragmentation in both ion trap and PSD without a collision cell is considered low energy CID. In this situation, a peptide carrying a positive charge mainly fragment along its backbone, generating **a, b** and **y** ions and ions that have lost ammonia (-17 Da) or water (-18 Da). Beside these, internal cleavages are often observed. The reason for this is the low energy collision, which usually contains multistep collisions, so there are more chances for the backbone to be cleaved at two sites.

In a tandem ion search, confidence that a protein has been identified correctly comes largely from getting multiple matches to peptides from the same protein. An MS/MS ion search on database from an isolated peptide may actually be less specific than a peptide mass mapping. But it reveals the sequence-specific information of the peptide, so tandem mass spectrometry is a powerful tool for protein confirmation. By combining the information obtained from peptide mass mapping with tandem mass spectrometry, a high confidence level in the identification can be achieved. In practice, if the expected sequence of the precursor peptide is available, one can easily calculate the theoretical fragment masses from world wide web programs such as **Mass-Product** by UCSF via web site http://prospector.ucsf.edu/ucsfhtml3.2/msprod.html. Interpretation of the tandem mass spectra can be achieved by comparing the PSD mass spectral results with the theoretical fragment masses of the candidate peptides. Another program called **Mass-Tag** under the same web site was designed for the searching of peptide sequence and possible proteins by the fragment ion data. Mass information of the precursor ion and the fragment ions is needed for **Mass-Tag** searching. The search results can be saved and combined with those of peptide mass mapping. Combination of these two programs can give a much higher confidence level than using only one program alone.

1.4 Identification of Proteins by Mass Spectrometry from Polyacrylamide Gel Electrophoresis

Strategies of coupling gel electrophoresis with mass spectrometry have been widely studied and reviewed [24,25]. When dealing with real world samples, proteins are most often separated by polyacrylamide gels. Mass spectrometry of gel-separated samples is a challenging assignment. Proteins separated by gel electrophoresis have the possibility of modification in the separation process [24]. Possible modifications are the formation of cysteine-acrylamide adducts, cysteine-β-mercaptoethanol adducts, N-terminal-acrylamide adducts, methionine oxidation and dodecyl sulfate adducts from detergent SDS. These modifications not only cause misinterpretation of the mass spectra, but also affect detection sensitivity and resolution in mass spectrometry. The observation and study of gel-induced modifications are frequently reported [25-37].

1.4.1 Molecular Weight Detection of Gel Induced Proteins by Mass Spectrometry

Conventionally, the molecular weight of gel-induced proteins can only be estimated by comparing the electrophoretic mobility of the proteins of interest with that of standard proteins. This method generates an accuracy of 10% and therefore is not sufficient to deduce any structural information. Accurate mass detection of gel-induced proteins will be necessary and several approaches were reported recently using both ESI MS and MALDI MS [30,32]. ESI MS has better mass accuracy than MALDI MS while MALDI MS is able to obtain better sensitivity (see Section 1.2.1). Proteins with masses over 200000 Da are considered unsuitable for ESI MS [30]. MALDI MS has higher mass detection limit and better impurity tolerance. But the existence of SDS will affect both ionization processes and result in ion suppression.

One of the challenges is to effectively remove large protein molecules from the gel matrix for sensitive detection [38]. Various strategies, involving electro elution

[39,40], electro blotting [41,42], direct desorption from ultra thin gels [43,44] and passive (i.e., diffusive) elution [38,45], have been employed. While all these methods have been used successfully, the passive elution technique has the advantage that no special equipment is required for elution, and standard gel thickness can be used.

Another important issue is related to mass resolution and accuracy in analyzing the gel-separated proteins. In many reported studies, the observed molecular ion peak is quite broad, and this has often been attributed to residual SDS interference in the MALDI analysis. Free acrylamide molecules in the gels are always considered dangerous for the proteins separated by the gel, especially when the protein is rich in cysteine residues (because of the possibilities of protein modifications by free acrylamide) [31]. The mass accuracy of gel-separated proteins is relatively poor (typically no better than 0.1%), compared with the analysis of free solution proteins [44,45].

In Chapter 2, a passive elution method for the mass spectrometric analysis of proteins from copper-stained/destained gels using a minimal number of steps is reported. The effect of SDS, acrylamide as well as some other gel reagents on mass detection is discussed. In addition, some experimental results relevant to the reduction of gel-induced covalent modifications as well as detection sensitivity are discussed.

1.4.2 Peptide Mass Mapping and Tandem Mass Spectrometry of Gel Induced Proteins

Gel separated proteins can be digested either in-gel (or on-membrane) or in-solution after being extracted from a gel. Extraction of peptides should be much easier than proteins because of their smaller size and fewer possibilities of covalent linkage with gel components. Among the in-gel digestion protocols, the protocol modified by Mann's group [46] has been frequently used. In this protocol, a gel band is excised and

dehydrated by acetonitrile. Reduction, alkylation and digestion of the protein is performed by hydration of the gel piece by the reaction solution. The digestion step takes up to 12 hours.

In Chapter 3, a simple in-gel digestion method is introduced. In this method, gel pieces are crushed before digestion, making better contact of the reagent and enzyme with protein molecules in the gel. The digestion step takes only 2-3 hours. Application of this digestion method is demonstrated by identifying the gel-separated proteins from human rhinovirus 3-C (HRV-3C) protease culture and an *E. coli* protein fraction. Techniques used for protein identification include peptide mass mapping and tandem mass spectrometry. MALDI TOF, MALDI PSD TOF and NanoES ion trap were employed for the detection process. Three proteins were successfully identified.

In Chapter 4, 2-D gel electrophoresis, in-gel digestion and mass spectrometric techniques are combined to identify proteins adsorbed to different implant materials in order to examine the changes in the entire proteome of cells in response to different growth substrates. This investigation of biocompatibility may be valuable in the search for specific proteins involved in biological response to a material.

1.5 Literature Cited

1. Ornstein, L. Ann. *N.Y. Acad. Sci.* **1964**, 121, 321.
2. Laemmli, U. K. *Nature* **1970**, 227, 680.
3. Wirth, J. P., Romano, A. *J. Chromatogra. A* 1995, 698, 123.
4. Gersten, D. **1996**, John Wiley and sons Ltd., Chichester, W. Sussex, UK
5. Allen, R. C., Budowle, B. **1994**, Water De Gruyter and Co., Berlin, Germany
6. Dunn, M. J. BIOS Scientific Publishers, Oxford, UK
7. MacCloskey, J. A., *Methods Enzymol.* **1990**, *193, 960.*

8. Covey, T. R., Bonner, R. F., Shushan, B. I. Henion, J., *Rapid Commun. Mass Spectrom.* **1988**, 2, 249.
9. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M., Science **1989**, 246, 64.
10. Karas, M., Hillenkamp, F. *Anal. Chem.* **1988**, 60, 2299.
11. Zenobi, R., Knochenmuss, R. *Mass Spectrom. Rev.* **1998**, 17, 337.
12. Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M. *Mass Spectrom. Rev.* **1990**, 9, 37.
13. Feng, R., Konish, Y., *Anal. Chem.* **1992**, 64, 2090.
14. Körner, R., Wilm, M., Morand, K., Schubert, M., Mann, M. *J. Am. Soc. Mass Spectrom.* **1996**, 7, 150.
15. Wilm, M., Mann, M. *Anal. Chem.* **1996**, 68, 1.
16. Gevaert , K., Vandekerckhove, J. *Electrophoresis* **2000**, 21, 1145.
17. Spengler, B. *J. Mass Spectrom.* **1997**, 32, 1019.
18. Chaurand, P, Luetzenkirchen, F., Spengler, B. *J. Am. Soc. Mass Spectrom.* **1999**, 10, 91.
19. Cooks, R. G., Kaiser, R. E. *Acc. Chem. Res.* **1990**, 23, 213.
20. Cooks, R. G.; Glish, G. L., Mcluckey, S. A., Kaiser, R. E. *Chem. Eng. News* **1991**, March 25, 26.
21. Roepstorff, P, Fohlman, J. *Biomed. Mass Spectrom.* **1984**, 11, 601.
22. Biemann, K., Martin, A. S., *Mass Spectrom. Rev.* **1987**, 6, 1.
23. Papayaannopoulos, I. A., *Mass Spectrom. Rev.* **1995**, 14, 49.
24. Patterson, D. S. *Electrophoresis* **1995**, 16, 1791.
25. Chiari, M., Righetti, P. G., Negri, A., Ceciliani, F., Ronchi, S. *Electrophoresis* **1992**, 13, 882.
26. Patterson, D. S. *Anal. Biochem.* **1994**, 221, 1.

27. Jungblut, P., Thiede, B. *Mass Spectrom. Rev.* **1997**, 16, 145.

28. Ploug, M., Jensen, A. L., Barkholt, V. *Anal. Biochem.* **1989**, 181, 33.

29. Brunce, D. C. *Anal. Biochem.* **1992**, 207, 285.

30. Le Maire, M., Deschamps, S., Moller, J. V., Le Caer, J. P., Rossier, J. *Anal. Biochem.* **1993**, 214, 50.

31. Clauser, K. R., Hall, S. C., Smith, D. M., Webb, J. W., Andrews, L. E., Tran, H. M., Epstein, L. B., Burlingame, A. L. *Proc. Natl. Acad. Sci. USA* **1995**, 92, 5072.

32. Haebel, S., Jensen, C., Andersen, S. O., Roepstorff, P. *Protein Sci.* **1995**, 4, 394.

33. Mortz, E., Vorm, O., Mann, M., Roepstorff, P. *Biol. Mass Spectrom.* **1994**, 23, 249.

34. Sutton, C. W., Pemberton, K. S., Cottrell, J. S., Corbett, J. M., Weeler, C. H., Dunn, M. J., Pappin, D. J., *Electrophoresis* **1995**, 16, 308.

35. Hall, S. H., Smith, D. M., Masiarz, F.R., Soo, V. W., Tran, H. M., Epstein, L. B., Burlingame, A. L. *Proc. Natl. Acad. Sci. USA* **1993**, 90, 1927.

36. Klarscov, K., Roecklin, D., Bouchon, B., Sabatie, J., Van Dorssalaer, A., Bischoff, R. *Anal. Biochem.* **1994**, 216, 127.

37. Van Dorssalaer, A., Bitsch, F., Green, B., Jarvis, S. Lepage, P., Bischoff, R., Kolbe, H. V. J., Roisch, C. *Biomed. Environ. Mass Spectrom.* **1990**, 19, 692.

38. Cohen, S. L., Chait, B. T. *Anal. Biochem.* **1997**, 247, 257.

39. Schuhmacher, M., Glocker, M.O., Wunderlin, M., Przybylski, M. *Electrophoresis* 1996, 17, 848.

40. Clarke, N.J., Li, F., Tomlinson, A.J., Naylor, S. *J. Am. Soc. Mass Spectrom.* **1998,** 9, 88.

41. Vestling, M.M., Fenselau, C. *Anal. Chem.* **1994,** 66, 471.

42. Liang, X.; Bai, J., Liu, Y. -H., Lubman, D.M. *Anal. Chem.* **1996,** 68, 1012.

43. Ogorzalek Loo, R.R., Stevenson, T.I., Mitchell, C., Loo, J.A., Andrews, P.C. *Anal. Chem.* **1996,** 68, 1910.

44. Ogorzalek Loo, R.R., Mitchell, C., Stevenson, T.I., Martin, S.A., Hines, W.M., Juhasz, P., Patterson, D.H., Peltier, J.M., Loo, J.A., Andrews, P.C. *Electrophoresis* **1997,** 18, 382.

45. Ehring, H., Stromberg, S., Tjernberg, A., Noren, B. *Rapid Commun. Mass Spectrom.* **1997,** 11, 1867.

46. Shevchenko, A., Wilm M., Vorm O., Mann M. *Anal. Chem.* **1996,** 68, 850.

Chapter 2

Observation of SDS-PAGE Gel-Induced Protein Modifications and Its Implications for Accurate Molecular Weight Determination of Gel-Separated Proteins by MALDI-TOF MS[a]

2.1 Introduction

Sodium dodecylsulfate polyacrylamide gel electrophoresis (SDS-PAGE) and related techniques such as 2-D gel electrophoresis remain in the forefront as the method of choice for protein separation in many applications. The combination of electrophoretic separation with mass spectrometric analysis provides a very powerful tool for protein identification [1, 2]. Knowledge of the protein molecular weight can be very useful for unambiguous protein identification. This is particularly true for the analysis of proteins with single or a few amino acid substitutions and with post-translational modifications. In peptide mapping, the digest fragments often do not cover the entire sequence of a protein and the missing fragments may include the modified

---

[a] A form of this chapter has been published as: Michael A. Jeannot, Jing Zheng and Liang Li "Observation of SDS-PAGE Gel-Induced Protein Modifications and Its Implications for Accurate Molecular Weight Determination of Gel-Separated Proteins by MALDI-TOF MS" *J. Am. Soc. Mass Spectrom.* **1999**, 10, 512-520. Dr. Jeannot did experiments on the effects of major gel components on MALDI signals, gel-induced protein modifications and MALDI-MS of gel-separated proteins (Sections 2.3.1, 2.3.2. and 2.3.3.).

peptides. This would result in miss-identification of the protein against a library of known proteins. In this case, the molecular weight information can be useful in confirming whether the protein in question is a modified form or mutant of a known protein in the database. In addition, for preparative gel-separation methods, accurate molecular weight information can be used to confirm the identity of proteins.

Matrix-assisted laser desorption ionization (MALDI) mass spectrometry has been developed to analyze proteins separated from polyacrylamide gels. One of the technical challenges in combining MALDI MS with SDS-PAGE is to effectively remove the large proteins from the gel matrix for sensitive detection [3]. Various strategies, involving electroelution [4, 5], electroblotting [6, 7], direct desorption from ultrathin gels [8,9] and passive (i.e., diffusive) elution [3, 10] have been employed. While all these methods have been used successfully, the passive elution technique has the advantage that no special equipment is required for elution, and standard gel thickness can be used.

Another important issue is related to mass resolution and mass measurement accuracy in analyzing the gel-separated proteins. In many reported studies, the observed molecular ion peak is quite broad, and this has often been attributed to residual SDS interference in the MALDI analysis. Mass accuracy is also relatively poor (typically no better than 0.1% accuracy), compared to the analysis of proteins from free solutions. This was reported to be the case for both continuous extraction and time-lag focusing or delayed extraction MALDI systems [9, 10].

In this study, we report a passive elution method for the analysis of proteins from copper-stained/destained gels using a minimum number of steps. This method minimizes the SDS interference in MALDI analysis, thus providing the opportunity of discerning other factors of peak broadening from the SDS effect. With the use of a time-lag focusing MALDI time-of-flight (TOF) mass spectrometer, we demonstrate

unequivocally that gel-induced covalent protein modifications by acrylamide play a major role in governing mass resolution and accuracy in direct analysis of gel-separated proteins. The gel-induced modifications could broaden the molecular ion peak if the instrument is not capable of resolving these covalent adducts. Furthermore, covalent modifications during gel separation can have a consequence in accurate molecular weight determination of intact proteins. It is further demonstrated that for proteins without covalent modifications it is entirely possible to obtain high resolution MALDI spectra and accurate molecular weight information by using the passive elution method in a time-lag focusing instrument. In addition, some experimental results relevant to the reduction of the extent of gel-induced covalent modifications as well as detection sensitivity are discussed.

2.2 Experimental

2.2.1 Materials

Tris-Tricine Ready Gels (16.5%), copper stain/destain kit, and Tris buffer were obtained from Bio-Rad (Hercules, CA). Bovine ubiquitin, equine cytochrome c, chicken lysozyme, equine myoglobin, bovine β-lactoglobulin B, bovine trypsinogen, bovine carbonic anhydrase II, bovine serum albumin (BSA), sodium dodecylsulfate (SDS), tricine, and 2-mercaptoethanol were obtained from Sigma (St. Louis, MO). Other reagents included glycerol (Caledon, Georgetown, ON, Canada), sinapinic acid (Aldrich, Milwaukee, WI), bromphenol blue (BDH, Toronto, ON, Canada), and acetonitrile (Fisher, Fair Lawn, NJ).

2.2.2 SDS-PAGE

SDS-PAGE was carried out according to Schagger and von Jagow [11] using 16.5% tris-tricine gels, substituting bromphenol blue for Serva blue G as the tracking

dye. Prior to electrophoresis, protein samples were treated at ~40°C for 30 min in sample buffer containing 2% mercaptoethanol (v/v), 4% SDS, 12% glycerol, 50 mM Tris, and 0.01% bromphenol blue (the pH of the buffer was adjusted to 6.8 by HCl). Ten- or fifteen-μL samples were loaded in the sample wells. The amount of proteins loaded on the gel for each experiment is specified in the Results and Discussion. The proteins were normally stacked at 30.0 V for one hour, followed by separation on the resolving gel at 90.0 V for 2.5 hours. Localization of the protein bands was carried out using standard Bio-Rad copper staining protocols (Instruction Manual 161-0470). For the silver stain experiment, Bio-Rad's standard protocols were used (Instruction Manual 161-0449).

2.2.3 Gel Excision and Protein Elution

For copper-stained gels, the protein band of interest was cut out using a scalpel. The dimensions of the excised gel piece were about 1.5 mm x 6 mm. The gel pieces were destained as described by Qin et al [12]. Each gel piece was cut into four 1.5 mm x 1.5 mm segments for extraction. For the detection limit study, the gel piece was cut into two equal segments for extraction. Protein extraction was carried out by addition of 5 μL of a saturated solution of sinapinic acid in 75% acetonitrile / 1 M HCl (aq) to a single 1.5 mm x 1.5 mm or 10 μL to a 1.5 mm x 3 mm gel piece. The HCl was necessary to react with all the tris buffer and lower the pH such that the MALDI matrix remains in the acid form. The extraction time was about 30 s and the extraction was performed at room temperature. The gel was crushed along with the eluting solvent in a 0.25 mL microcentrifuge tube. The supernatant was then used for MALDI analysis.

2.2.4 MALDI MS

MALDI MS analysis was performed on either a Model G2025A linear time-of-flight system with continuous extraction (Hewlett-Packard, Reno, NV) or a high-

resolution home-built linear time-of-flight system equipped with time-lag focusing (TLF) [13]. Both instruments are equipped with a pulsed nitrogen laser (337 nm). Spectra were acquired and processed with Hewlett-Packard supporting software. They were further processed for presentation with the Igor Pro software package (WaveMetrics Inc., Lake Oswego, OR). All mass spectra shown in the figures are the smoothed spectra using 5- or 15-point Savitzky-Golay smoothing. No baseline correction was performed. External calibration was performed in all cases, except those specifically mentioned in the Results and Discussion.

Samples were typically prepared by using the two-layer method [14]. The first layer was formed by depositing ~1 μL of 30 mg/mL sinapinic acid in acetone on the probe tip [15], followed by addition of 0.4 to 1 μL of standard or sample mixed with a matrix. For external calibration purposes, 5 μM protein standards in 0.1% TFA were mixed 1:4 with a saturated solution of sinapinic acid in 75% acetonitrile / 1 M HCl (aq) (i.e., the same solution used to extract the proteins from the gel). In the case of internal calibration, 1 μL of a 1-5 μM mixture of two proteins whose molecular weights bracketed that of the analyte was added to the extract. After drying, the samples were washed on-probe with water. Multiple shots were averaged to obtain a spectrum. For the TLF instrument, the pulse delay was set to 1.0 μs. The pulse voltage was optimized for each protein using a 1:4 mixture of the protein in 0.1% TFA plus a saturated solution of sinapinic acid in 25% acetonitrile / 0.1% TFA loaded on top of the first layer.

For the SDS studies, solutions of 1 μM ubiquitin, cytochrome c, lysozyme, and myoglobin were prepared containing 0, 0.005, 0.01, 0.02, 0.03, 0.05, 0.07, and 0.1% SDS, as follows: Ten μL of a 10 μM protein solution in 0.1% TFA was mixed with 10 μL of a 10-fold concentrated SDS solution in 0.1% TFA, followed by addition of 80 μL of a saturated solution of sinapinic acid in 25% acetonitrile / 0.1% TFA (aq). One μL of this solution was placed on the probe tip on top of the first layer (sinapinic acid from

acetone solution). Samples were washed on-probe, and MALDI MS was performed on the high-resolution TLF instrument focused on myoglobin. The system was calibrated with ubiquitin and myoglobin containing no SDS.

For the glycerol and Tris studies, appropriate amounts of glycerol or Tris-HCl were mixed with 20 μL of a 5 μM protein solution in 0.1% TFA and 80 μL of saturated sinapinic acid in 25% acetonitrile / 0.1% TFA(aq). Final glycerol concentrations of 0, 0.5, 1, 2, 3, 5, 7, and 10% (w/v) and final tris-HCl concentrations of 0, 0.05, 0.1, 0.2, 0.3, 0.5, 0.7, and 1.0 M were obtained. MALDI MS analyses were carried out as in the SDS studies.

For the studies involving direct reaction of acrylamide with proteins, a protein mixture containing 5 μM of ubiquitin, cytochrome c, lysozyme, and myoglobin in tris-HCl buffer was mixed with 2-mercaptoethanol (0.1% v/v) and incubated at ~40°C for 30 min. After the solution was cooled to room temperature, acrylamide was added to 50 mM. At various time intervals, the protein/acrylamide solution was taken for MALDI analysis.

2.3 Results and Discussion

2.3.1 Effects of Major Gel Components on MALDI Signals

Although it is generally assumed that SDS degrades both signal-to-noise ratio and resolution in MALDI, conflicting reports have appeared regarding the effect of SDS at various concentrations using continuous extraction MALDI-TOF [16-18]. The gels used herein contain 0.1% SDS, and the effects of SDS at and below this concentration have been studied for ubiquitin, cytochrome c, lysozyme, and myoglobin. Figure 2.1 shows mass spectra of these proteins in the presence of 0 to 0.1% SDS, obtained by using the time-lag focusing MALDI-TOF instrument. Clearly, with only 0.005% SDS, there is already a reduction in signal-to-noise ratio. Above 0.02% SDS, there is

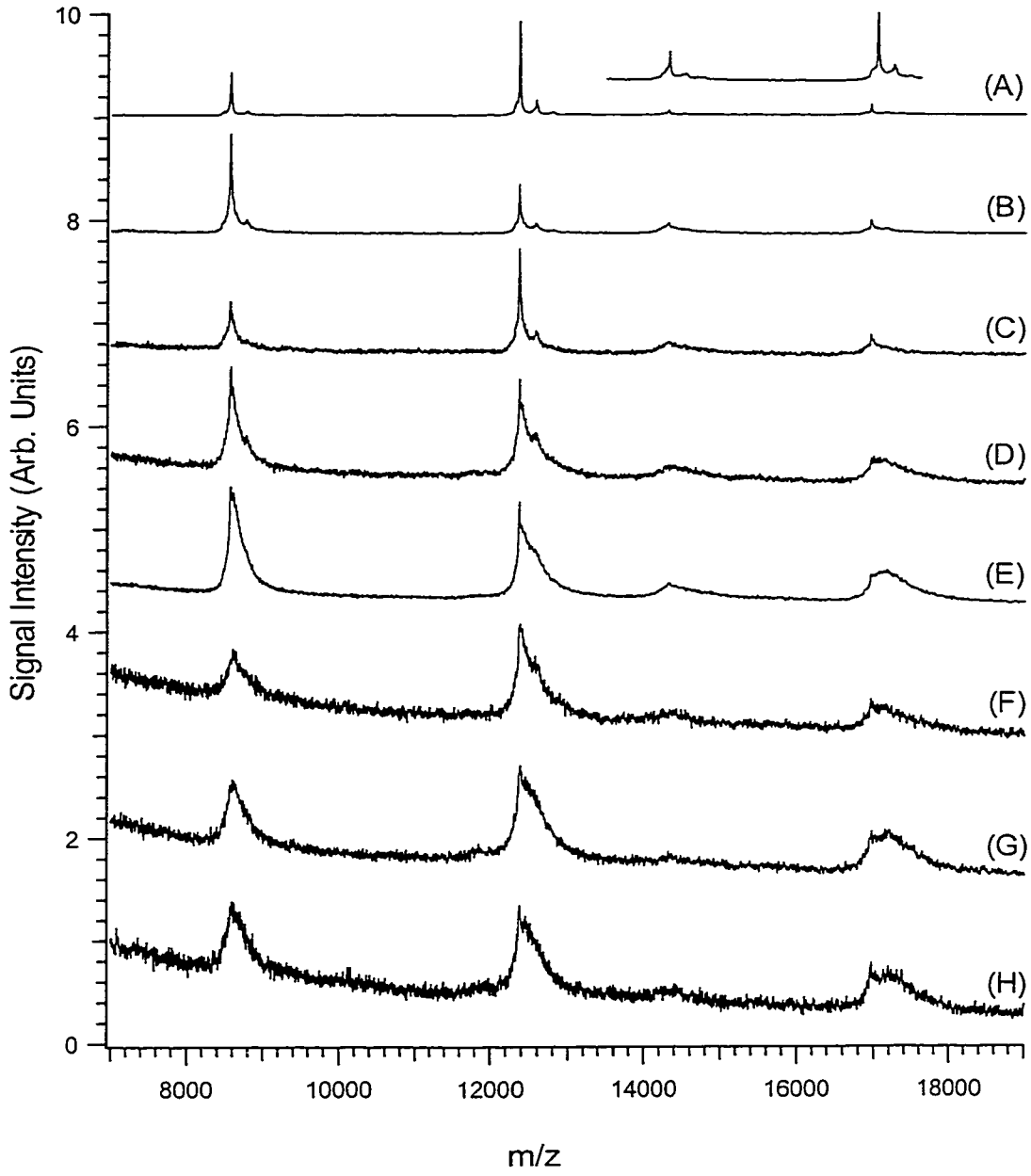


Figure 2.1 MALDI mass spectra of a mixture of 1μM each of bovine ubiquitin (MH^+=8565.87), equine cytochrome c (MH^+=12361.18), chicken lysozyme (MH^+=14307.17), and equine myoglobin (MH^+=16952.50) obtained by time-lag focusing TOFMS with the following concentrations of SDS in the mixture: (A) 0%, (B) 0.005%, (C) 0.01%, (D) 0.02%, (E) 0.03%, (F) 0.05%, (G) 0.07%, and (H) 0.1%.

significant broadening to the high mass end of the true molecular weight. Above 0.05% SDS, lysozyme cannot be readily detected. While mass spectra can be obtained for the other three proteins even at 0.1% SDS, the loss in mass resolution in these cases is considered to be very significant by current MALDI standards. The SDS-PAGE gels used in this work contain 0.1% SDS; however the copper staining/destaining process likely removes a significant amount of SDS [3], and the proteins are eluted into a solution which contains no SDS. Thus, the final concentration of SDS in the MALDI extract is expected to be much less than 0.1% (see below).

In addition to SDS, the polyacrylamide gel contains other components, such as glycerol and Tris buffer. Even very high concentrations of glycerol (up to 10% w/v) have negligible effect on signal-to-noise, resolution, and mass accuracy. The glycerol residue on the MALDI probe tip is effectively removed by the on-probe washing step. In the case of Tris buffer, it is important to use the hydrochloride (or alternately to add HCl to the Tris solution) to keep the pH low for the MALDI matrix. Here too, there is negligible loss in resolution or mass accuracy, even with high concentrations of Tris (up to 1 M). However, the relative sensitivity changes with increasing amounts of Tris. As the concentration of tris increases, ubiquitin and lysozyme signals become more intense, but myoglobin intensity decreases. This effect could be attributed either to the Tris itself, or to differences in pH with different concentrations of Tris-HCl. Nevertheless, it is quite clear that both glycerol and Tris do not seriously interfere with the MALDI analysis.

2.3.2 MALDI MS of Gel-Separated Proteins

MALDI analysis of ubiquitin, cytochrome c, lysozyme, myoglobin, β-lactoglobulin B, trypsinogen, and carbonic anhydrase II isolated from SDS-PAGE was performed on both continuous extraction (DC) and time-lag focusing instruments. Comparative spectra are shown in Figures 2.2 and 2.3, and the results of measured

molecular weights and mass resolution are summarized in Table 2.1. The mass spectral patterns of the molecular ion regions shown in Figures 2.2 and 2.3 are quite different.

Figure 2.2 illustrates that MALDI mass spectra with high resolution and good signal-to-noise ratio can be obtained for gel-separated ubiquitin, myoglobin, and carbonic anhydrase II. These results suggest that the concentration of SDS in the final solution of gel-separated proteins must be quite low (0.01% or less, comparing with Figure 2.1). As Table 2.1 shows, good mass measurement accuracy can also be obtained for these three proteins. The mass accuracy can be improved by employing internal calibration. For example, the average masses of MH^+ for myoglobin using external calibration (Table 2.1) is 16954.5 ± 0.3, which is statistically higher than the calculated value of 16952.50, representing an error of +118 ppm. By using internal calibration, a value of 16952.1 ± 0.8 is obtained. The somewhat worse mass accuracy with external calibration is likely due to small differences in the thickness of the crystal layer between calibration and sample measurements.

It is interesting to compare the MALDI results of these three proteins with those reported by others using time-lag focusing TOFMS for the analysis of gel-separated proteins [10]. For proteins extracted from Coomassie Brilliant Blue stained SDS-PAGE by using a solvent mixture of formic acid, acetonitrile, isopropanol and water [10], the mass resolution obtained was 550 for ubiquitin (50 pmol gel loading), 400 for myoglobin (60 pmol gel loading), and 80 for carbonic anhydrase II (200 pmol gel loading). The mass accuracy with external calibration was 0.2-0.3%. It is clear that the experimental conditions used in the reported work [10] were quite different from ours. Our results demonstrate that it is entirely possible to obtain MALDI data from these three gel-separated proteins with high resolution and accuracy. One advantage of using

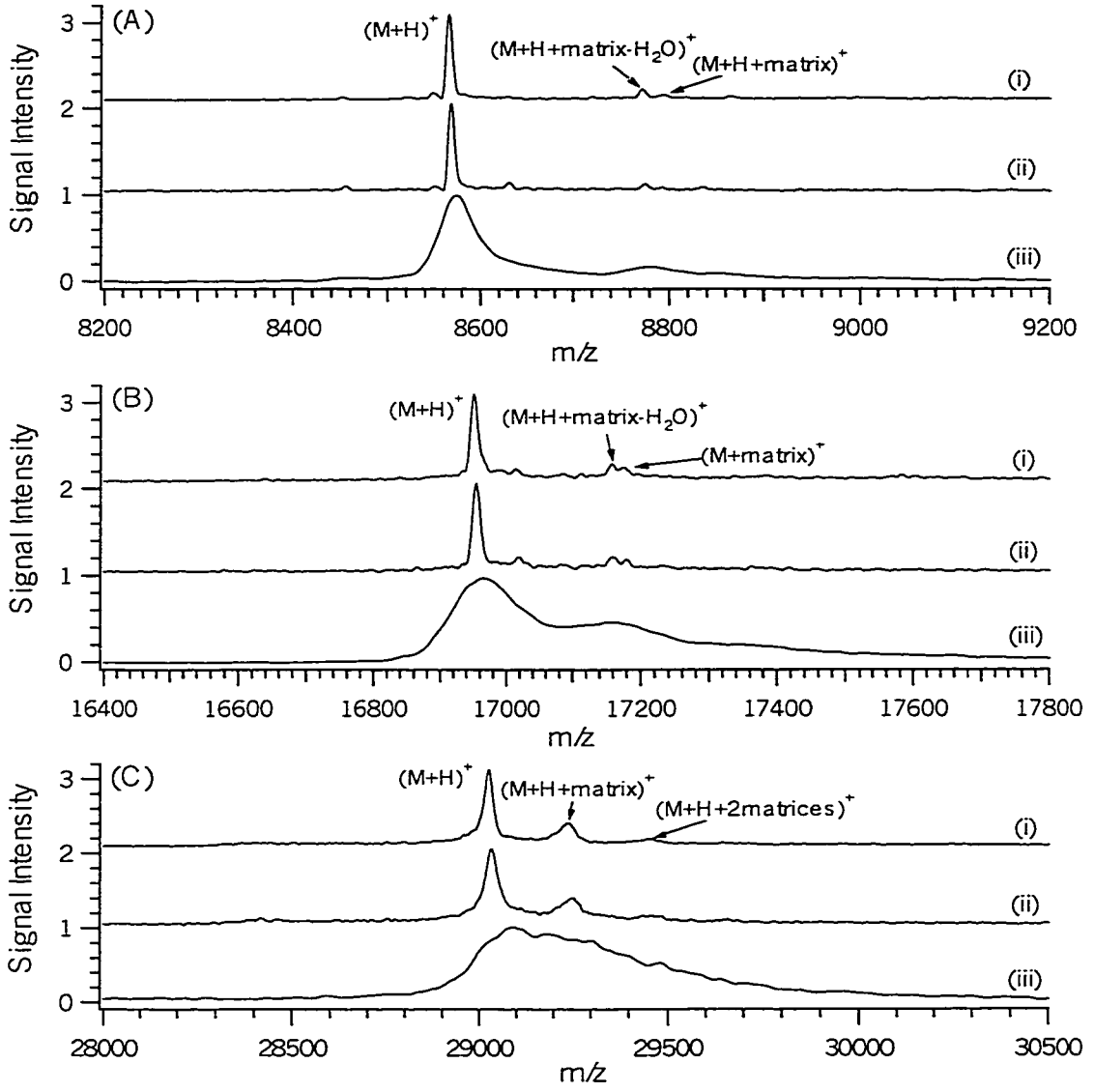


Figure 2.2 MALDI mass spectra of (A) ubiquitin, (B) myoglobin, and (C) carbonic anhydrase II obtained under different experimental conditions: (i) standard preparation with TLF instrument; (ii) gel extract with TLF instrument; (iii) gel extract with DC instrument. For TLF experiments, the delay time was 1 μs and the pulse voltage was 2.72 kV for ubiquitin, 3.35 kV for myoglobin, and 3.90 kV for carbonic anhydrase II. The amount of each protein loaded on the gel was 50 pmol.

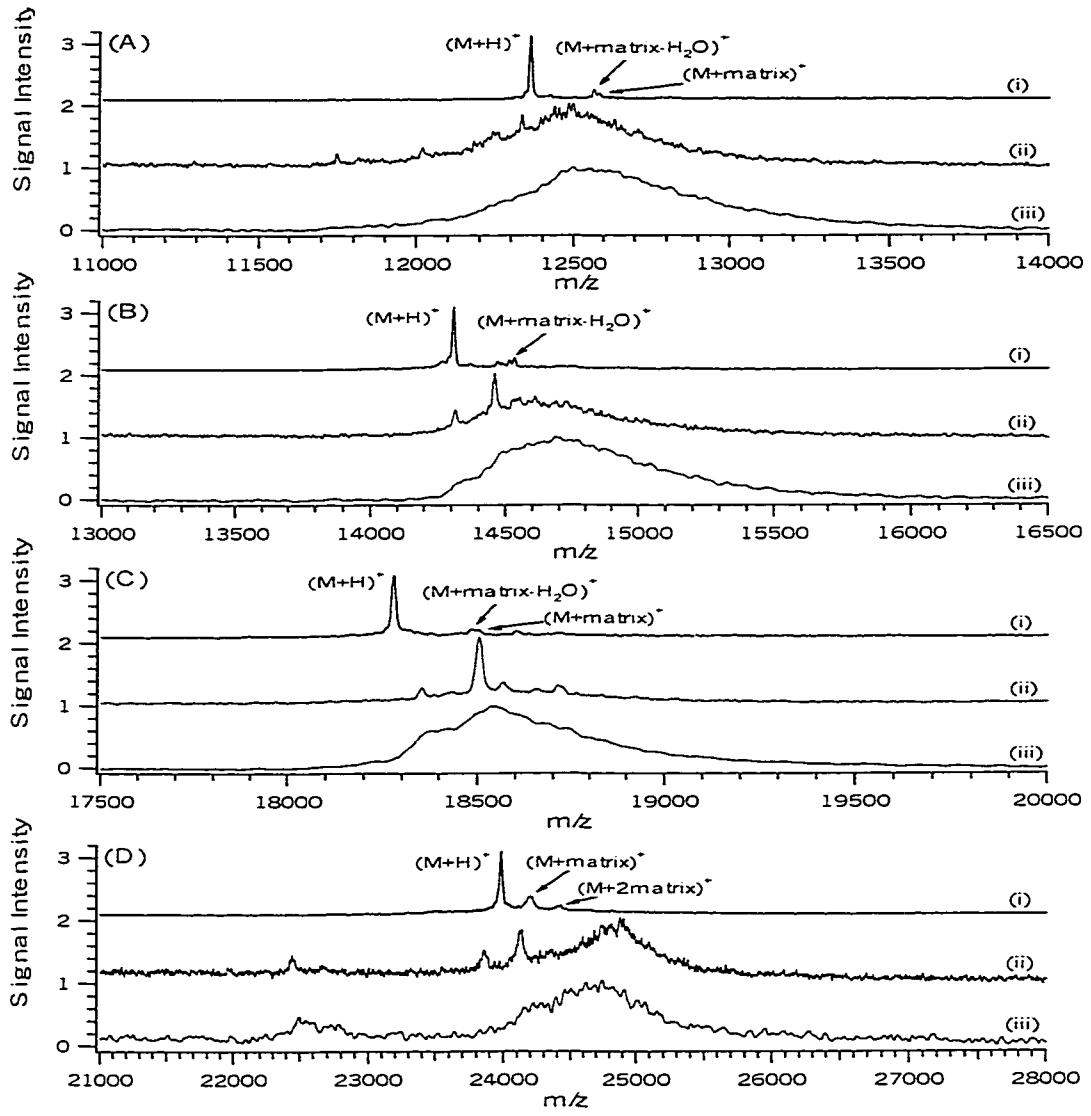


Figure 2.3 MALDI mass spectra of (A) cytochrome c, (B) lysozyme, (C) β-lactoglobulin B, and (D) trypsinogen obtained under different experimental conditions: (i) standard preparation with TLF instrument; (ii) gel extract with TLF instrument; (iii) gel extract with DC instrument. For TLF experiments, the delay time was 1 μs and the pulse voltage was 3.00 kV for cytochrome c, 3.15 kV for lysozyme, 3.45 kV for β-lactoglobulin B, and 3.70 kV for trypsinogen. The amount of each protein loaded on the gel was 50 pmol.

Table 2.1 Mass Accuracy and Resolution of Proteins Isolated from SDS-PAGE and Analyzed by Time-Lag Focusing (TLF) MALDI-MS

Protein	Calc. MH^+	Obs. MH^+ (from gel)[a, b]	Resolution (standard)[b]	Resolution (from gel)[b]
Ubiquitin	8565.87	8568.3 ± 0.4	1066 ± 22	1039 ± 22
Cytochrome C	12361.18	c	1163 ± 9	c
Lysozyme	14307.17	c	1066 ± 76	c
Myoglobin (apo)	16952.50	16954.5 ± 0.3	1454 ± 132	1398 ± 266
β-lactoglobulin B	18278.26	c	1421 ± 81	c
Trypsinogen	23981.97	c	1051 ± 107	c
Carbonic Anhydrase	29024.72	29030.5 ± 1.5	985 ± 49	898 ± 169

[a]Using external calibration

[b]Average and standard deviation from 3 replicate experiments

[c]Peak is broadened by covalent protein modifications

copper staining instead of Coomassie Brilliant Blue is the absence of adduct ions from the staining reagent in MALDI spectra (the adduct ions from the matrix are labeled in Figures 2.2 and 2.3). In copper staining, the proteins are not fixed in the gels and the gels can be destained. The negatively stained images can be visualized with intermediate sensitivity between Coomassie blue and silver staining. We attempted to use the same extraction procedure to extract proteins from silver stained gels; but no spectra could be obtained with gel loading of 500 ng of ubiquitin, carbonic anhydrase II, or BSA. With this amount of loading, we can readily obtain mass spectra of ubiquitin extracted from copper-stained gels (see below).

However, for cytochrome c, lysozyme, β-lactoglobulin B, and trypsinogen, the molecular ion region consists of a broad band with several peaks, as shown in Figure 2.3. Since SDS does not appear to cause a significant problem, it is likely that some other phenomenon is responsible for the appearance of the mass spectral patterns shown in Figure 2.3. It is noted that ubiquitin, myoglobin, and carbonic anhydrase II contain no cysteines, while cytochrome c, lysozyme, β-lactoglobulin B, and trypsinogen contain various numbers of cysteine residues.

2.3.3 Gel-Induced Protein Modifications

It has been reported that residual acrylamide monomers in the gel can covalently bind with cysteine residues on proteins during electrophoresis [19-28]. Evidence for this comes from ESI analysis of whole proteins or tryptic digest peptides [19-21], MALDI MS of tryptic digest peptides [22-24], MS/MS analysis of tryptic peptides [25, 26], and amino acid analysis / sequencing data [27, 28]. Acrylamide adduction gives rise to peaks of mass M + n(71) where M is the protein molecular weight (with cysteines reduced) and n is the number of acrylamide adducts. Thus, a distribution of molecular weights may be expected for proteins containing cysteine residues.

Ubiquitin, myoglobin, and carbonic anhydrase II contain no cysteines. Consistent with this, the observed mass spectral peaks are sharp with reasonably good mass accuracy.

For the other four proteins (cytochrome c, lysozyme, β-lactoglobulin B, and trypsinogen), a large shift to the high mass end is observed. While the continuous extraction mass spectra show only a broad peak, TLF is able to partially resolve the distribution and give information regarding the extent of modification. For example, in the case of lysozyme (calculated MH^+=14307.17, fully reduced=14315.23) there is a peak at 14304.8 ± 0.3 as shown in Figure 3B (determined by internal calibration), corresponding to the unmodified protein (with all of the four disulfide bonds re-formed) and a more intense, well-resolved peak at 14453.4 ± 0.8. This is consistent with the reduction of two disulfide bonds (+4) and the addition of 2 acrylamides (+142). Further additions are apparent, although it is difficult to accurately measure the mass of higher adducts, because of weak signals. Note that, in the MALDI spectra of gel-separated proteins with internal calibrants (not shown), high resolution was observed for the calibrant peaks.

In the case of β-lactoglobulin B (Figure 2.3C), which contains 5 cysteines, the molecular ion of the unmodified protein (calculated MH^+=18278.26, reduced=18282.29) is not observed. However, acrylamide addition is evident from the peaks at 18350.5 ± 0.8 (+72) and 18500.6 ± 0.2 (+222) (determined by internal calibration), corresponding to the addition of 1 and 3 acrylamides respectively with partial reduction of disulfide bonds. Since the molecular ion peak of the unmodified protein is absent in the spectrum, the peak at m/z 18500.6 cannot be from the matrix adduct of the unmodified protein. Under the experimental conditions used, the matrix adduct peaks are generally much less intense than the protonated protein peak (see Figure 2.2 and top trace of each spectrum in Figure 2.3).

The most intense peak for trypsinogen (Figure 2.3D) (12 cysteines, calculated MH^+=23981.97, reduced=23994.07) is at ~24130, corresponding to an addition of +148 (addition of two acrylamides with partial reduction of disulfide bonds). Further adducts are apparent from the broad distribution above this, although they are not resolved.

In the case of cytochrome c, as shown in Figure 2.3A, no distinctly resolved peaks are apparent. The two cysteines in cytochrome c are covalently bound to the heme group, and therefore would not be expected to react with acrylamide. It is possible that the covalently attached heme group is modified under electrophoretic conditions. Acrylamide is also known to react with hydroxy-compounds (e.g., serine) and amino-compounds (e.g., lysine) [29]. Mass spectral peaks consistent with acrylamide addition to non-cysteine residues have been reported [21, 22]. Furthermore, for cytochrome c and other proteins, other modifications such as oxidation of methionine or cysteine-acrylamide adducts [24,30], and 2-mercaptoethanol addition [31] are possible. These modifications will further complicate the molecular ion region of the mass spectra, resulting in the difficulty of determining the actual molecular ion peak.

The difficulty in resolving and accurately measuring protein-acrylamide adducts for some proteins may be attributed to partial re-formation of disulfide bonds during and after electrophoresis when the proteins are no longer in a reducing environment. Each re-formed disulfide bond reduces the molecular weight by 2 mass units. A distribution of the number of re-formed disulfide bonds for each acrylamide adduct species will broaden the peak.

2.3.4 Attempts to Reduce Gel-Induced Modifications

Several experiments were performed to examine methods of reduction or elimination of the gel-induced modifications. First of all, the mass spectral patterns

shown in Figure 2.3 were observed only for the gel-separated proteins. If the protein solution containing ubiquitin, cytochrome c, lysozyme, and myoglobin is first treated with 2-mercaptoethanol, followed by mixing with acrylamide (see Experimental for details), the MALDI spectra obtained from this solution after incubation over 4 hours are similar to those obtained from the initial standard solution. This indicates that proteins do not readily react with acrylamide in free standing solution under the experimental conditions used and the use of a high electric field in gel separation may play a strong role in inducing the acrylamidation of proteins.

It was found that the gel-induced modification takes place at the early stage of the separation. After the protein sample (e.g., lysozyme) was loaded on the stacking gel, in about 30 min, the sample migrated to the front of the resolving gel. At this point, gel separation was stopped and the protein was extracted from the gel for MALDI analysis. The mass spectra of proteins obtained from this early stage of separation were very similar to those obtained after normal separation (i.e., Figure 2.3). Thus, decreasing separation time cannot reduce the extent of gel-induced modifications.

There were reports of using pre-electrophoresis with an acrylamide scavenger to reduce the amount of unpolymerized acrylamide in gel separation [22,28,32,33]. In this approach, an initial electrophoresis was run on the entire gel system by using 0.5 M Tris, 0.1% w/v SDS and 12 mM 3-mercaptopropanoic acid (scavenger) as cathode buffer and 0.2 M Tris as anode buffer at 15 mA/gel for about 4 hours at room temperature [32]. This gel was then used for protein separation. We have attempted to use this approach in the hope of reducing gel-induced modification by acrylamide. The mass spectra (not shown) of cytochrome c and lysozyme extracted from gels treated with pre-electrophoresis are similar to the spectra shown in Figure 2.3A,B. No significant improvement is obtained and extensive modifications are still observed. Variations of the pre-electrophoresis conditions, including the use of fresh gels, freshly

prepared buffer solutions, higher concentration of 3-mercaptopropanoic acid (up to 20 mM), and higher electric current (up to 30 mA/gel), did not result in the reduction of the extent of modifications. If we assume modification takes place between the unpolymerized acrylamide and the protein, these results would suggest that pre-electrophoresis did not remove the acrylamide effectively to a level that would not cause extensive modification. Alternatively, polyacrylamide may degrade during electrophoresis of proteins to acrylamide or oligoacrylamide that in turn reacts with the proteins. The actual reaction mechanism and, in fact, mechanisms of many intriguing reactions in gel separation cannot be clearly defined [34].

Since proteins with cysteines are modified by acrylamide in gel separation, a seemingly obvious approach to obviate the problem with acrylamide addition is to reduce and alkylate the proteins prior to gel electrophoresis. Figure 2.4 shows several mass spectra of lysozyme obtained from the samples after reduction and alkylation with iodoacetamide at different reaction times. The most abundant peak shown at m/z 14772 is from the fully alkylated lysozyme (8 cysteines). The peaks at m/z 14715 and m/z 14658 are from lysozyme with 7 and 6 cysteines alkylated, respectively. It is clear that complete alkylation is not achieved even when a very long reaction time is used. This is not surprising in light of the fact that the reactivity of cysteines at different moieties of the protein chain can be quite different. Thus, alkylation of proteins prior to electrophoresis still faces the problem of uncertainty on the extent of modification, hence, the uncertainty of defining the true molecular ion peak. It should be noted that, for an unknown protein, even if the reaction conditions are in favor of complete unknown protein, without the knowledge of the number of cysteines present in the protein, accurate molecular weight determination is still not possible.

Figure 2.4 shows several peaks superimposed on a broad hump. The peaks with m/z above 14772 cannot be readily assigned to known structures and/or matrix/salt adducts of the m/z 14715 or 14772 species. The broad hump observed in Figure 2.4 resembles those observed in the spectra of proteins with gel-induced modifications (Figure 2.3). These broad humps are likely due to multiple uncharacterized covalent adducts formed in the gel and/or their matrix/salt adducts formed in the MALDI process.

2.3.5 Detection Sensitivity

The overall detection sensitivity in MALDI analysis of gel-separated proteins is mainly governed by the separation and visualization sensitivity, extraction and matrix/sample preparation efficiency, and the detection sensitivity of the MALDI instrument. In our method, the proteins are extracted by simply macerating a portion of the destained-gel piece with 5 μL of an acetonitrile-HCl (aq) solution containing sinapinic acid for a few seconds. A two-layer method is then used for sample preparation. This sample/matrix preparation method generates very reproducible results. The commonly used dried-droplet method resulted in very poor reproducibility from sample to sample; thus it was not extensively used in this work. The merits of the two-layer method over other sample preparation methods in the direct analysis of complex peptide and protein mixtures from raw samples such as milk, bacteria, and egg white is described in a separate publication [35].

To assess the sensitivity limiting factors in our method of MALDI analysis of gel-separated proteins, several tests were performed. With copper staining, BSA and myoglobin can be visualized with a loading of >25 ng of protein and ubiquitin can be visualized with a loading of >50 ng. For MALDI analysis of standard proteins with the two-layer sample preparation performed in our instruments, ubiquitin, myoglobin, and

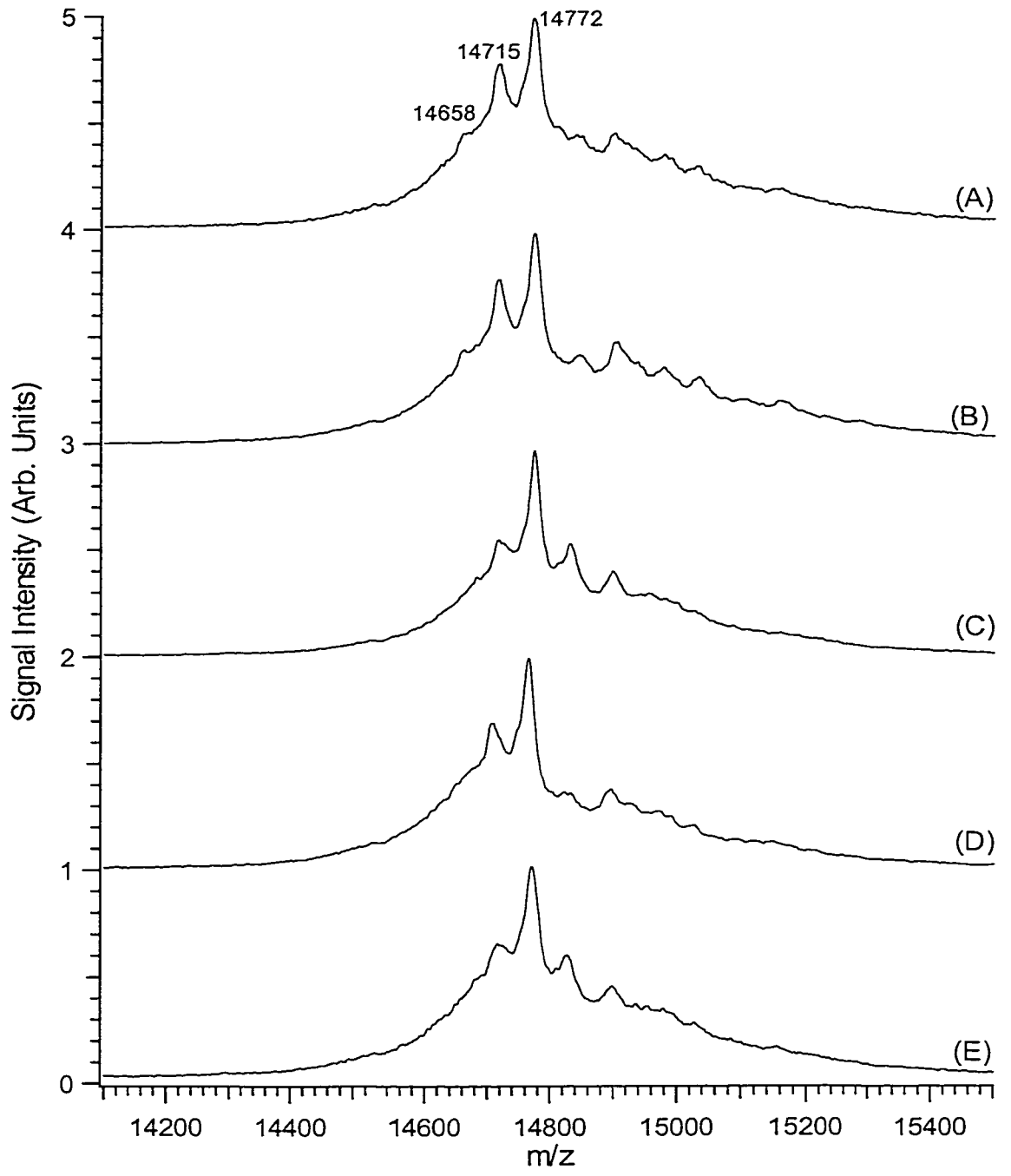


Figure 2.4 MALDI mass spectra of lysozyme obtained from the samples after reduction and alkylation with iodoacetomide at different reaction times: (A) 15 min; (B) 1 hour; (C) 2 hours; (D) 7.5 hours; and (E) 21 hours.

BSA can be detected with a sample loading of as low as 5 fmol, 20 fmol, and 60 fmol, respectively. Figure 2.5 shows the MALDI spectra obtained from gel loadings of 50 ng (6 pmol) ubiquitin, 100 ng (6 pmol) myoglobin, and 1 μg (15 pmol) BSA. For Figure 2.5, the protein gel lane was cut into 2 pieces and one piece was extracted by using 10 μl of the sinapinic acid/CH_3CN/HCl solution. 0.4 μl of the extract was deposited to the probe. Thus, the total amount of proteins loaded on the MALDI probe is less than 1/50 of the amount loaded on the gel.

It should be noted that the data shown in Figure 2.5 can be readily reproduced from experiments using the same or different batches of gel. Judging from the signal-to-noise ratio in each spectrum, one would expect that a useful spectrum could be obtained with a sample loading of less than the amount used in Figure 2.5. However, it was found that by cutting the amount by just half (e.g., loading 0.5 μg BSA), protein signals were not always detected in MALDI spectra. Generally speaking, two out of three trials failed to generate any protein signals. This finding is consistent with what was observed in the work of Cohen and Chait [3]. For example, in their work, a very good spectrum was obtained for myoglobin with a gel loading of 25 pmol protein using their fast extraction method, suggesting useful spectra could be obtained with much lower amount of gel loadings; but no useful spectrum could be obtained with a loading of 6 pmol. A plausible explanation is that the extraction efficiency decreases considerably when the concentration of the protein in the gel is decreased below a certain threshold. Based on the detection limits of our method for standard proteins (i.e., 5 fmol for ubiquitin, 20 fmol for myoglobin, and 60 fmol for BSA) and assuming the spectra shown in Figure 5 represent the detection limits for the analysis of gel-separated proteins, the extraction efficiency can be estimated to be 4%, 17%, and 20% for ubiquitin, myoglobin and BSA, respectively.

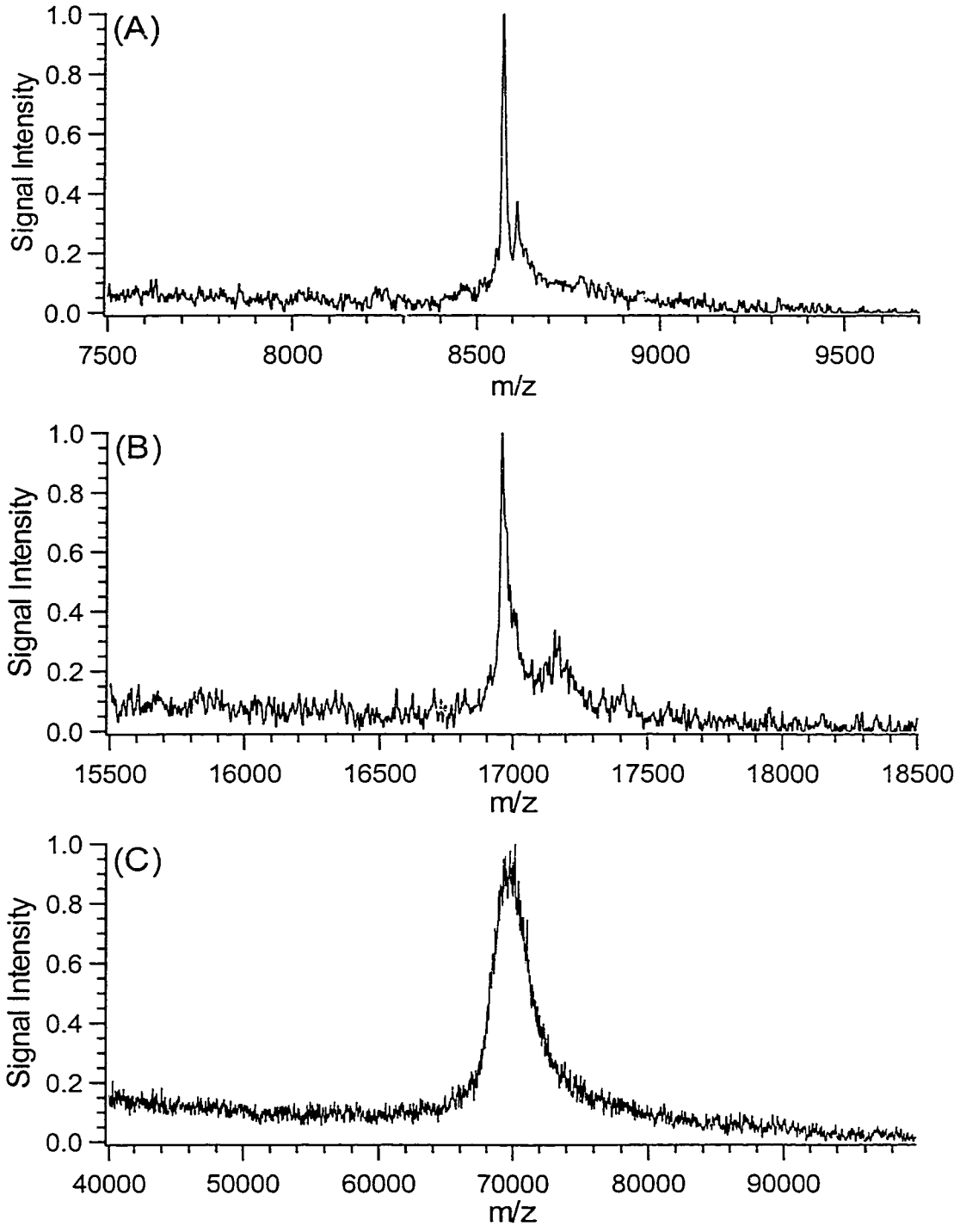


Figure 2.5 MALDI spectra obtained from gel loadings of (A) 50 ng (6 pmol) ubiquitin, (B) 100 ng (6 pmol) myoglobin, and (C) 1 μg (15 pmol) BSA.

As it can be seen from Figure 2.5, the detection limit of the current method is in the low pmol region. However, the overall detection sensitivity of the method described in this work may potentially be improved by concentrating the extract into a smaller volume prior to sample deposition to the MALDI probe. In the two-layer method, the total volume of sample needed to load onto the probe is usually less than 1 μl. Note that concentrating the extract with vacuum drying did not yield any improvement in detection sensitivity. This is understandable in light of the fact that the extraction solution comprises a saturated matrix solution. Concentrating the extracts will result in matrix precipitation, which incorporates the analyte, thus losing the analyte for the final analysis in MALDI. In addition, any impurities in the extract will also be concentrated, which may decrease the MALDI detection sensitivity. The use of a saturated matrix solution in the extraction solvent is critical in achieving the detection sensitivity illustrated in Figure 2.5. An attempt to use a 10-fold diluted matrix solution, instead of a saturated solution, for extraction, followed by an about 10-fold concentration of the extract did not yield any useful spectra for 6 pmol gel loading of myoglobin and 16 pmol of BSA. Thus, a sensible approach for improving the overall detection sensitivity of the passive extraction/MALDI analysis method is to develop concentration methods that will allow selective concentration of proteins from gel extracts. Work in this direction is currently underway by our group [36].

2.4 Conclusions

With the use of an appropriate gel-extraction method and the time-lag focusing MALDI-TOF mass spectrometer, accurate molecular weight determination for SDS-PAGE separated proteins is possible for some proteins, particularly those containing no cysteine residues. However, the accuracy may be greatly reduced as a result of gel-

induced protein modifications. In these cases, the molecular weight of the unmodified species cannot be determined with a high degree of accuracy without *apriori* knowledge of the number of modifications. However, in such cases, a reasonable estimate of the molecular weight can be made, and such a "window" may still be a useful complement to tryptic digest data for protein identification. This work clearly illustrates that covalent protein modifications, not the SDS detergent, limit the accuracy of protein molecular weight determination by MALDI TOFMS. It is further shown that the use of a short gel-separation time and pre-electrophoresis with an acrylamide scavenger did not reduce the extent of protein modifications. Finally, the detection sensitivity of the passive extraction/MALDI method is demonstrated to be in the low pmol for several proteins with molecular weights up to 67,000.

2.5 Literature Cited

1. Patterson, S.D., Aebersold, R. *Electrophoresis* **1995,** 16, 1791.
2. Jungblut, P., Thiede, B. *Mass Spectrom. Rev.* **1997,** 16, 145.
3. Cohen, S.L., Chait, B.T. *Anal. Biochem.* **1997,** 247, 257.
4. Schuhmacher, M., Glocker, M.O., Wunderlin, M., Przybylski, M. *Electrophoresis* **1996,** 17, 848.
5. Clarke, N.J., Li, F., Tomlinson, A.J., Naylor, S. *J. Am. Soc. Mass Spectrom.* **1998,** 9, 88.
6. Vestling, M.M., Fenselau, C. *Anal. Chem.* **1994,** 66, 471.
7. Liang, X., Bai, J., Liu, Y.-H., Lubman, D.M. *Anal. Chem.* **1996,** 68, 1012.
8. Ogorzalek Loo, R.R., Stevenson, T.I., Mitchell, C., Loo, J.A., Andrews, P.C. *Anal. Chem.* **1996,** 68, 1910.

9. Ogorzalek Loo, R.R., Mitchell, C., Stevenson, T.I., Martin, S.A., Hines, W.M., Juhasz, P., Patterson, D.H., Peltier, J.M., Loo, J.A., Andrews, P.C. *Electrophoresis* **1997,** 18, 382.

10. Ehring, H., Stromberg, S., Tjernberg, A., Noren, B. *Rapid Commun. Mass Spectrom.* **1997,** 11, 1867.

11. Schagger, H., von Jagow, G. *Anal. Biochem.* **1987,** 166, 368.

12. Qin, J., Fenyo, D., Zhao, Y., Hall, W.W., Chao, D.M., Wilson, C.J., Young, R.A., Chait, B.T. *Anal. Chem.* **1997,** 69, 3995.

13. Whittal, R.M., Li, L. *Anal. Chem.* **1995,** 67, 1950.

14. Dai, Y.Q., Whittal, R.M., Li, L. *Anal. Chem.* **1996**, 68, 2494.

15. Vorm, O., Roepstorff, P., Mann, M. *Anal. Chem.* **1994,** 66, 3281.

16. Amado, F.M.L., Santana-Marques, M.G., Ferrer-Correia, A.J., Tomer, K.B. *Anal. Chem.* **1997,** 69, 1102.

17. Rosinke, B., Strupat, K., Hillenkamp, F., Rosenbusch, J., Dencher, N., Kruger, U., Galla, H.-J. *J. Mass Spectrom.* **1995,** 30, 1462.

18. Jensen, C., Haebel, S., Andersen, S.O., Roepstorff, P. *Int. J. Mass Spectrom. Ion Processes* **1997,** 160, 339.

19. Le Maire, M., Deschamps, S., Moller, J.V., Le Caer, J.P., Rossier, J. *Anal. Biochem.* **1993,** 214, 50.

20. Bonaventura, C., Bonaventura, J., Stevens, R., Millington, D. *Anal. Biochem.* **1994,** 222, 44.

21. Klarskov, K., Roecklin, D., Bouchon, B., Sabatie, J., Van Dorssalaer, A., Bischoff, R. *Anal. Biochem.* **1994,** 216, 127.

22. Haebel, S., Jensen, C., Andersen, S.O., Roepstorff, P. *Protein Sci.* **1995,** 4, 394.

23. Mortz, E., Vorm, O., Mann, M., Roepstorff, P. *Biol. Mass Spectrom.* **1994,** 23, 249.

24. Sutton, C.W., Pemberton, K.S., Cottrell, J.S., Corbett, J.M., Wheeler, C.H.; Dunn, M.J.; Pappin, D.J. *Electrophoresis* **1995,** 16, 308.

25. Hall, S.C., Smith, D.M., Masiarz, F.R., Soo, V.W., Tran, H.M., Epstein, L.B., Burlingame, A.L. *Proc. Natl. Acad. Sci. USA* **1993,** 90, 1927.

26. Clauser, K.R., Hall, S.C., Smith, D.M., Webb, J.W., Andrews, L.E., Tran, H.M., Epstein, L.B., Burlingame, A.L. *Proc. Natl. Acad. Sci. USA* **1995,** 92, 5072.

27. Brune, D.C. *Anal. Biochem.* **1992,** 207, 285.

28. Plough, M., Jensen, A.L., Barkholt, V. *Anal. Biochem.* **1989,** 181, 33.

29. *Chemistry of Acrylamide, Bulletin PRC 109,* Process Chemicals Department, American Cyanamid Co., Wayne, N.J., 1969.

30. Fantes, K.H., Furminger, I.G.S. *Nature* **1967,** 215, 750.

31. Papov, V.V., Gravina, S.A., Mieyal, J.J., Biemann, K. *Protein Sci.* **1994,** 3, 428.

32. Ploug, M., Stoffer, B., Jensen, A.L. *Electrophoresis* **1992**, 13, 148.

33. Dunbar, B., Wilson, S.B. *Anal. Biochem.* **1994**, 216, 227.

34. Garfin, D.E. in *Introduction to Biophysical Methods for Protein and Nucleic Acid Research*, Glasel, J.A., Deutscher, M.P., Eds., Academic Press: New York, **1995**, Chap 2.

35. Dai, Y.Q.; Whittal, R.M.; Li, L. *Anal. Chem.* **1999**, 71, 1087.

36. Doucette, A., Craft, D., Li, L. *Anal. Chem.* **2000**, 72, 3355.

Chapter 3

Identification of SDS-PAGE Separated Proteins by Mass Spectrometry and Tandem Mass Spectrometry

3.1 Introduction

Biological samples are often complex mixtures of proteins, DNAs and impurities. Therefore, effective separation is a prerequisite for protein identification. 1D/2D SDS PAGE is a well established and widely used protein separation technique that offers incomparable resolution. Proteins can be identified by coupling gel separations and modern mass spectrometry (MS) techniques, such as matrix-assisted laser desorption ionization (MALDI), electrospray ionization (ESI) and nanoelectrospray ionization (NanoESI) MS. Peptide masses can be obtained by in-gel digestion, peptide extraction, followed by MS analysis [1-4]. Then, the mass data are used for protein identification through peptide mass mapping (see Section 1.2.3). In an ideal case, the mass data from a single MS analysis is sufficient for a positive identification. However, ambiguity may arise if the sequence coverage is low. This is particularly true when accurate molecular weight information is not available. To increase the confidence level of identification, sequence-specific information obtained by tandem mass spectrometry (MS^n) [5] is often used as another dimension of data input to extract the desired sequence message. The ions of interest are selected as precursor ions, upon dissociation, the spectrum of the product ions, which reflect structural information of the precursor ion (peptide ion) can be obtained. Two tandem mass

spectrometric techniques, MALDI PSD TOF and NanoES ion trap, are employed in our study.

One common problem encountered in gel-separated protein analysis is that peptides extracted from gel inevitably contain surfactants and salts, which will cause ion suppression in subsequent MS analysis. As discussed in Chapter 1, MALDI has the merits of relatively high salt tolerance and low ion suppression. With the reflectron configuration in a modern MALDI TOF spectrometer, a PSD spectrum can therefore be generated to provide fragmentation information. There are, however, several major problems associated with MALDI PSD including poor resolving power and limited control over the fragmentation parameters. Another soft ionization technique, ESI, usually has more severe ion suppression due to its "surface" ionization mechanism. This can be partly compensated for by applying a downscaled interface, namely NanoES, to increase surface/volume ratio and reduce ion suppression [6,7] (see Section 1.2.2). This enables a direct mixture analysis without separation. In addition, the low sample consumption makes it extremely suitable for the analysis of often-precious biomaterials. Ion trap mass spectrometers can provide high resolution and fine control over the CID process. Interfacing NanoES to ion trap can produce high quality MS^n spectra containing rich structural information. It should be pointed out that despite the improved salt tolerance of NanoES over ESI, a cleanup step is still necessary prior to MS analysis.

In this chapter, MALDI TOF and NanoES ion trap MS data from three proteins samples are discussed. For each sample, the gel piece containing proteins was first crushed, an in-gel digestion was then conducted, and the peptide mixtures eluted from the crashed-gel were finally analyzed by MS. The resulting mass data was input for database searching by means of peptide mass mapping. Tandem mass spectrometry data of some peptides by MALDI TOF PSD and NanoES ion trap were obtained to

confirm the identification results. Gel samples of Cold Shock-Like Protein from *Escherichia coli* (*E. coli)* was confidently identified. Human rhinovirus-3C (HRV-3C) protease from HRV-3C protease gel sample was confirmed. A suspected modification form of HRV-3C protease in the gel sample containing HRV-3C protease was also identified. To our surprise, it turned out to be an *E. coli* protein, chloramphenicol acetyltranferance. The result was rationalized with respect to the sample preparation process. Some technical issues, such as comparison between traditional in-gel digestion and our crushing gel method, were also discussed.

3.2 Experimental

3.2.1 Materials

Polyacrylamide/bis (37.5:1), ammonium persulfate, glycine, SDS, silver stain/destain kit, and tris buffer were obtained from Bio-Rad (Hercules, CA). Trifluoroacetic acid (TFA), 2-mercaptoethanol, dithiothreitol (DTT), α-cyano-4-hydroxy-*trans*-cinnamic acid (HCCA) and iodoacetamide were obtained from Sigma (St. Louis, MO). Other reagents included glycerol (Caledon, Georgetown, ON, Canada), bromo-phenol blue (BDH, Toronto, ON, Canada). All solvents used were from Fisher and are reagent grade or HPLC grade.

3.2.2 Sample Preparation

Bacterial Sample Preparation[a]: *Escherichia coli* ATCC (American Type Culture Classification) 9637 cells were prepared as previously described [8]. Cells were

---

[a] Bacterial cells were grown at Edgewood RDE Center, Maryland, USA. Extraction, concentration and purification of proteins from the cells were performed by Ms. Zhengping Wang.

grown overnight (18-24 hours) in nutrient broth with shaking at ambient temperature. Cells were harvested, washed with several volumes of sterile water, lyophilized to dryness and stored at – 4 to 0°C. Samples were shipped by overnight express with dry-ice packaging.

The *E. coli* 9637 extract was prepared by solvent suspension methods. About 100 mg *E. coli* 9637 cells were suspended in 1mL of 0.1% TFA, vortexed for 3-5 minutes and centrifuged. The supernatant was then transferred to a clean vial and this procedure was repeated until no apparent protein signals were detected from the supernatant by MALDI. The supernatants were pooled and filtered using Microcon-3 filter (Amicon) with 3000 Da molecular weight cut-off, and then concentrated to about 0.5 mL in a vacuum centrifuge. This concentrated extract was used for HPLC fractionation.

The protein extract from *E. coli* 9637 was separated by reverse phase HPLC using a 1.6×25 cm C_{18} preparative column. The mobile phases were 0.05% TFA/water (A) and 0.05% TFA/acetonitrile (B), with a gradient of 0-75% B over 300 minutes at a flow rate of 1 mL/min. The column eluent was collected in one-minute increments throughout.

Fraction 166 was concentrated from 500 μL down to about 10 μL. The pH was adjusted by 0.1 M ammonium bicarbonate (NH_4HCO_3) and adding sample buffer and water, the final volume of the sample was about 20μL. 10 μL was loaded into each well of a 15% polyacrylamide gel. Electrophoresis was carried out at 60 mA and 120 V for two gels for about 1-2 hours. The gels were silver-stained.

HRV-3C Protease Sample Preparation[b]: Plasmid pET-3a, which contains the

---

[b] HRV-3C protease gel sample was prepared by Ping Li at Professor David Wishart's laboratory.

HRV-3C protease gene, was transformed into the competent cell line BL21(DE3)PLysS by electroporation. BL21(DE3)PlysS contains a plasmid with the gene of a protein named chloramphenicol acetyltransferace, which assists the pET-3a vector in expressing the proteins of interest. About 10 loops of cells were scraped from the plate and suspended into 1 liter of LB media. The suspension was incubated at 37°C for about 3 hours followed by inducing with 0.5 mM IPTG (isopropylthio-ß-D-galactoside) at 30°C for 6 hours. The cells were harvested bycentrifugation at 3000 rpm for 30 minutes and resuspended in 15 mL of a buffer containing 20 mM imidazole, 1 mM EDTA and 2 mM DTT at pH 6.5. The cells were kept at -80°C overnight and thawed the next day in ambient temperature. To completely break the cells, sonication was applied to the suspension for 20 seconds (four times each) on ice. One milliliter of 6% PEM (polyethylenimine) at pH 7.5 was added to the suspension to precipitate the DNA. After spinning the DNA and cell debris down at 12000 rpm for 30 minutes at 4 °C, the supernatant was dialyzed against imidazole buffer (20 mM, pH 6.5) in 12-14 KD dialysis tubing for 4 hours at 4 °C. The proteins were then purified with a SP sepharose column, and separated by 10% SDS PAGE and stained with coomassie blue.

According to the preparation procedure, HRV-3C protease should have been expressed. In the gel, however, two main bands showed up. One of them had the same molecular weight as HRV-3C protease compared with the HRV-3C protease standard in the next lane. The other band had a higher molecular weight than the HRV-3C protease standard. The band having the same molecular weight with HRV-3C protease standard was assumed to be HRV-3C protease. The band with higher molecular weight, however, is a protein co-expressed with HRV-3C protease and might be a modified form of HRV-3C protease. Identification of the upper band protein and confirmation of the identity of HRV-3C protease (lower band) were carried out by in-gel digestion and mass spectrometry.

3.2.3 Ziptip Cleaning

Samples for NanoES ion trap mass spectrometry were cleaned and concentrated using C_{18} ZipTip pipette tips obtained from Millipore (Cat. No. ZTC18S096). 10 μL of 0.1% TFA was added to the dried samples followed by vortexing for 30 seconds. ZipTips were prewetted with 50% acetonitrile and equilibrated with 0.1% TFA. Samples were then loaded on the ZipTips with 10 –20 cycles of suction and dispensing and washed by 0.1% TFA for 5 cycles. Peptides were eluted using 5 μL of 50% acetonitrile for four cycles. 1 μL of acetic acid was then added to the cleaned sample before NanoES mass spectrometry.

3.2.4 In-Gel Digestion

Protein bands were excised from the gel, put into 0.6-mL siliconized plastic vials and rinsed with 100 mM NH_4HCO_3 to adjust the pH to about 8.5. This NH_4HCO_3 solution was removed by pipette and replaced with 10 μL of the same buffer solution and 2 μL of 90 mM DTT. The gels were crushed using thin 0.25 mL plastic vials and incubated for 30 min at 50°C to allow the proteins to be reduced. Once cooled to ambient temperature, the vials were covered with aluminium foil and 2 μL of 200 mM iodoacetamide was then added. The gel samples were incubated at room temperature for 30 min, and then 1 μL of 0.2 μg/μL trypsin was added. The protein samples were digested at 37°C for 2 to 3 hours. Extraction of the peptides was accomplished with three changes of 20 μL of 75% acetonitrile in 0.25 % TFA / water followed by 10 μL of acetonitrile. For each extraction step, the sample was vortexed for 20 seconds followed by 20 minute of sonication. The pooled extracts were evaporated to dryness in a vacuum centrifuge.

3.2.5 Mass Spectrometry

MALDI MS and MALDI PSD MS: MALDI MS was performed either on a homemade or a Voyager Elite MALDI MS instrument (Voyager Elite, PerSeptive Biosystem, Inc., Framingham, MA). Both instruments are equipped with a delayed extraction (DE) device. A two-layer method was used for MALDI MS analysis. 1 to 2 μL of first layer solution (10 mg of HCCA in 1 mL of 20% methanol/acetone v/v) was deposited onto a probe tip, and evaporated to form a thin matrix layer. 0.5 – 1 μL of gel extract from 50% acetonitrile or 40% methanol saturated by HCCA was then deposited onto the first layer, allowed to air dry, and washed three times with water. The PSD spectra were recorded in the PSD mode of the Voyager Elite instrument.

NanoES Ion Trap MS and MS^n: NanoES MS and MS^n were performed on an Esquire–LC ion trap mass spectrometer (Hewlett-Packard, Palo Alto, CA) with NanoES interface. Spectra were acquired over the m/z range from 200 to 2200.

3.3 Results and Discussion

3.3.1 Identification of *E. Coli* Cold Shock-Like Protein CSPC

Figure 3.1 is the MALDI mass spectrum of the in-gel digestions of a gel band from *E. coli* (HPLC Fraction 166). The resulting peptide masses were entered into the peptide mass mapping program **PeptIdent** at the **Expasy** website (see Chapter 1) and the **Swiss-Prot** database was searched. The protein was identified as *E. Coli* Cold Shock-Like Protein CSPC (SwissProt accession # P36996) with 100% sequence coverage. The results indicate that the bacterial protein was efficiently digested during the course of in-gel digestion in our procedure. The crushing step is believed to facilitate both digestion and extraction by increasing the protein/enzyme and peptides/solvent contact surface respectively.

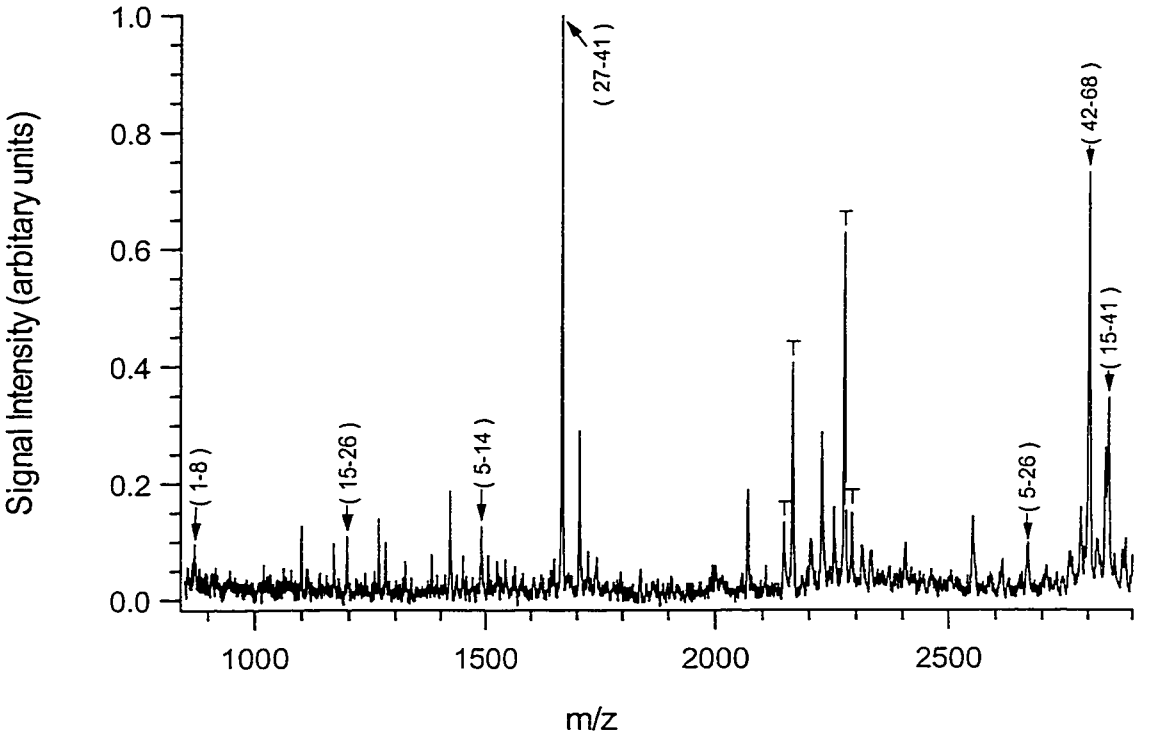


Figure 3.1 MALDI mass spectrum from the in-gel digestion of a gel-separated bacterial protein following HPLC fractionation. Numbers in the brackets refer to the corresponding amino acid sequence of *E. coli* Cold Shock-Like Protein. The peaks marked with letter **T** are those resulted from tryptic autolysis.

It is noted that the above case represents only a class of proteins with relatively low molecular weight (~7000 Da). As the molecular weight increases, the sequence coverage is anticipated to decrease. More information, including the masses of intact molecular ions and/or fragment ions, will be required if the peptide mapping fails to give a confident result. Nonetheless, compared with the conventional dehydration/hydration method [9], our in-gel digestion following gel crushing gives rise to better MS sensitivity and peptide coverage and is also more time-effective.

3.3.2 Identification of Chloramphenicol Acetyltransferance

Figure 3.2A shows the MALDI mass spectrum of an in-gel tryptic digest of the upper band from the HRV-3C gel sample. Theoretical peptide masses of HRV-3C protease were obtained by UCSF program **MS-Digest**. The ion masses obtained by MALDI were compared with the theoretical digest results. Surprisingly, only two ions with m/z at 1201 and 1317 match the theoretical digestion results of HRV-3C protease among all 19 obtained masses. Judged by the relative low intensities of these two peaks, HRV-3C protease or its modification cannot be the components of this band.

Since HRV-3C protease was expressed in *E. coli* cell, there is a possibility that this is an *E. coli* protein. Therefore, the peptide masses obtained from the MALDI were entered to UCSF program **MS-Fit** and searched under the specie category *E. Coli*. The searching results are listed below. Only one possible protein, chloramphenicol acetyltransferance, was given.

MS-Fit Search Results

Sample ID (comment): **unknown**

Database searched: **SwissProt.4.18.2000**

Molecular weight search (**21600 - 26400 Da**) selects **7058** entries.

Full pI range: **85824** entries.

Species search (**ESCHERICHIA COLI**) selects **4595** entries.

Combined molecular weight, pI and species searches select **419** entries.

MS-Fit search selects **1** entry.

Considered modifications: | **Oxidation of M** | **Acrylamide Modified Cys** |

Min. # Peptides to Match	Peptide Mass Tolerance (+/-)	Peptide Masses are	Digest Used	Max. # Missed Cleavages	Cysteines Modified by	Input # Peptide Masses
4	**500 ppm**	**monoisotopic**	**Trypsin**	**1**	**carbamidomethylation**	**19**

Result Summary

Rank	MOWSE Score	# (%) Masses Matched	Protein MW (Da)/pI	Species	SwissProt.4.18.2000 Accession #	Protein Name
1	359	6/19 (31%)	25663.3 / 5.91	ECOLI	P00483	chloramphenicol acetyltransferance

Detailed Results

1. 6/19 matches (31%). 25663.3 Da, pI = 5.91. Acc. # P00483. ECOLI.

CHLORAMPHENICOL ACETYLTRANSFERASE.

m/z submitted	MH^+ matched	Delta ppm	start	end	Peptide Sequence
1114.50	1114.54	-41.9665	66	74	(R)LMNAHPEFR(M)
1130.80	1130.54	228.4850	66	74	(R)LMNAHPEFR(M)
1347.30	1347.75	-339.6402	55	65	(K)FYPAFIHILAR(L)
1612.30	1612.91	-379.2072	53	65	(K)HKFYPAFIHILAR(L)

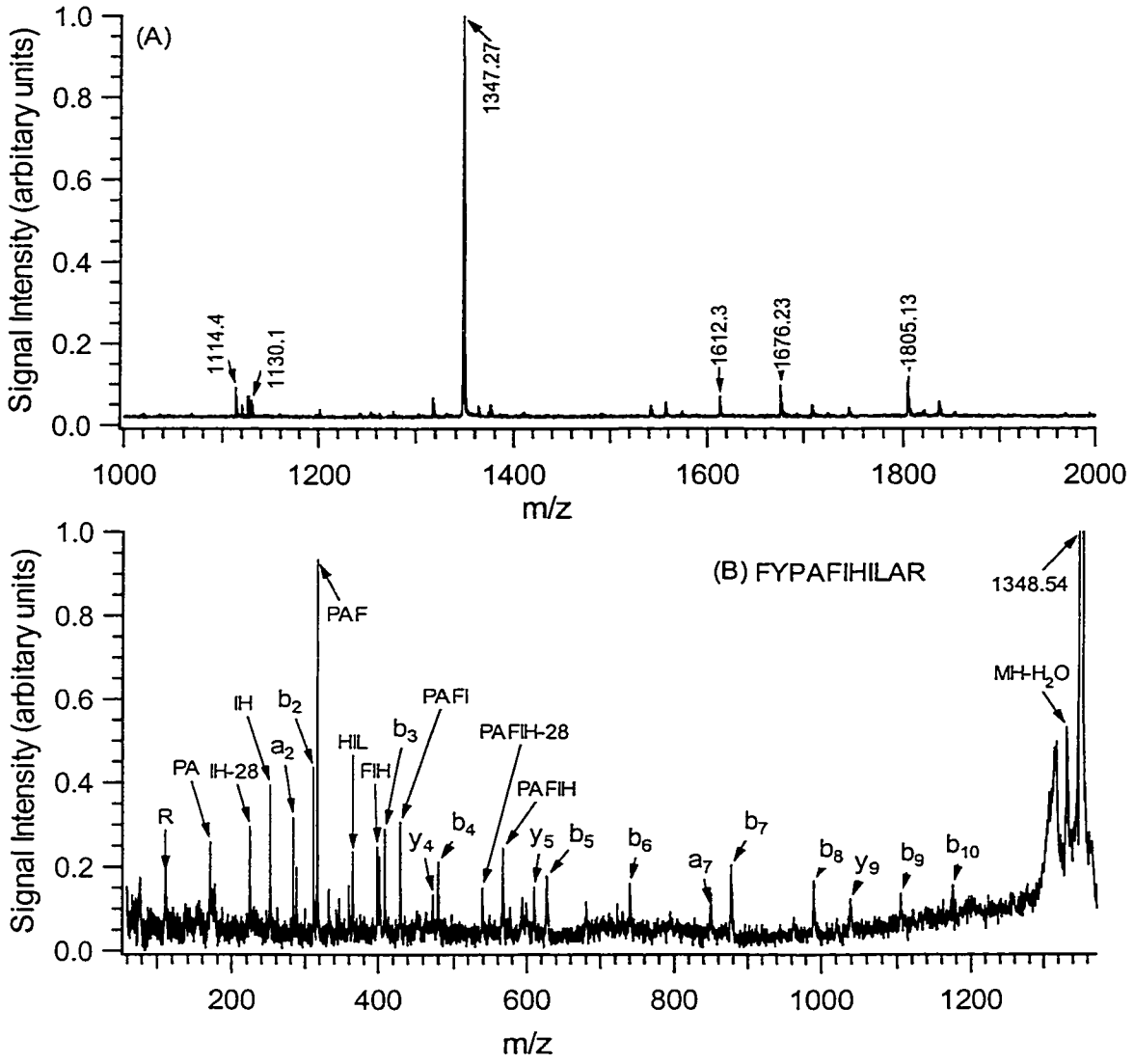


Figure 3.2 Mass spectra from the in-gel digestion of a HRV-3C protease upper band obtained by (A) MALDI TOF. The labeled peaks match with the protein chloramphenicol acetyltransferance. (B) MALDI TOF PSD. The precursor ion appears at m/z 1348.54 Da and the fragment ion peaks are labeled.

1676.60	1676.83	-142.9144	5	18	(K)ITGYTTVDISQWHR(K)
1805.00	1804.93	36.2298	4	18	(K)KITGYTTVDISQWHR(K)
1805.00	1804.93	36.2298	5	19	(K)ITGYTTVDISQWHRK(E)

The matched peptides cover **17%** (38/219 AA's) of the protein.

Although the matched peaks are among those intense ones, the sequence coverage is only 17%. More information is needed to confirm the result. Therefore, a MALDI PSD experiment was further carried out (Figure 3.2B). Ions with m/z 1347 were selected as the parent ion. The masses of the parent ion and the fragment ions were entered into the program **MS-Tag** to search for matching. Following is the searching result:

MS-Tag Search Results

Sample ID (comment): **unknown**
Database searched: **SwissProt.6.17.2000**
Full Molecular Weight range: **87272** entries.
Full pI range: **87272** entries.
Species search (**ESCHERICHIA COLI**) selects **4607** entries.
Number of sequences passing through parent mass filter: **85**
MS-Tag search selects **1** entry.

Parent mass: **1348.5400 (+/- 500 ppm)**
Fragment Ions used: **169.86, 175.78, 223.65, 251.45, 283.24, 288.33, 311.19, 316.25, 346.67, 359.58, 361.47, 398.37, 408.32, 429.40, 472.64, 479.47, 538.35, 566.30, 593.10, 626.45, 679.57, 739.30, 849.33, 876.78, 989.56, 1038.59, 1103.98, 1174.32, 1254.15, 1331.10 (+/- 800 ppm)**

Search Mode	Max. # Unmatched Ions	Peptide Masses are	Digest Used	Max. # Missed Cleavages	Cysteines Modified by	Peptide N terminus	Peptide C terminus
identity	**9**	**average**	**Trypsin**	**1**	**carbamidomethylation**	**Hydrogen (H)**	**Free Acid (OH)**

Result Summary

Rank	# Unmatched Ions	Sequence	MH^+ Calculated (Da)	MH^+ Error (Da)	Protein MW (Da)/pI	Species
1	7/30	(K)FYPAFIHILAR(L)	1348.6396	-0.0996	25663.3 / 5.91	ECOLI

SwissProt. Accession #	Protein Name
P00483	CHLORAMPHENICOL ACETYLTRANSFERASE

Both programs, **MS-Fit** and **MS-Tag,** point to the same protein, chloramphenicol acetyltransferance, independently. Moreover, with the primary structure information, an almost complete **b** ion series with only $\mathbf{b_1}$ missing were recognized in the PSD spectrum. This experiment demonstrates that peptide mass mapping, along with the structural information obtained from PSD experiment, can result in very confident identification.

The upper band can be easily mistaken as a modifed form of HRV-3C protease. But the seemingly unexpected result can be explained in the light of the fact that gene for chloramphenicol acetyltransferance was also contained in the competent cells and that chloramphenicol acetyltransferance was added to the medium (see Section 3.2.2). It therefore was co-expressed with HRV-3C protease.

3.3.3 Confirmation of HRV-3C Protease

To confirm the identity of HRV-3C protease, both MALDI TOF and NanoES ion trap mass spectrometers were used. Tandem mass spectrometry was performed on the NanoES ion trap mass spectrometer.

Figure 3.3A and 3.3B are MALDI and NanoES mass spectra of the in-gel digests of the assumed HRV-3C protease band. Comparing spectrum 3.3A and 3.3B, it

is found that MALDI (Figure 3.3A) is in favor of relatively high mass species (>1000 Da), whereas NanoES mass spectrum (Figure3.3B) is mainly composed of low mass ones (<1000 Da). Using the peptide masses contributed from both spectra for peptide mass mapping, 40% sequence coverage was achieved. Although this is a fairly acceptable coverage, we utilized tandem MS to further increase the confidence in this assignment.

Two peptide ions in Figure 3.3B, 760 and 942, were chosen as the precursor ions. Peak at m/z 942 can be assigned to the peptide with the sequence QYFVEKQ. Figure 3.4 is the MS^2 result of the ion with m/z 942. With the aid of the **MS-Product** program, all the product ion peaks can be easily interpreted (see the labels of Figure 3.4).

Figure 3.5A is the MS^2 result of the ion at m/z 760, which supposedly has the sequence YDYATK. The major fragment ion in this spectrum is the y_4 ion. Most of the energy dumped into the parent ion in the fragmentation step was believed to be consumed by such an ion formation. Accordingly, fragments at even lower mass were barely noticeable. Alternatively, the y_4 ion was chosen as the precursor ion for MS^3 and the spectrum is shown in Figure 3.5B. Both y_3 and y_2 were detected.

3.4 Conclusions.

In this chapter, three proteins separated by gel electrophoresis were identified or confirmed by combining peptide mass mapping and tandem mass spectrometry. It is obvious that there are a number of approaches available for protein identification, among which MALDI, in conjunction with peptide mass mapping is always the first choice [4]. In ideal cases, peptide mass mapping alone can provide confident identification results, as shown in Section 3.3.1.

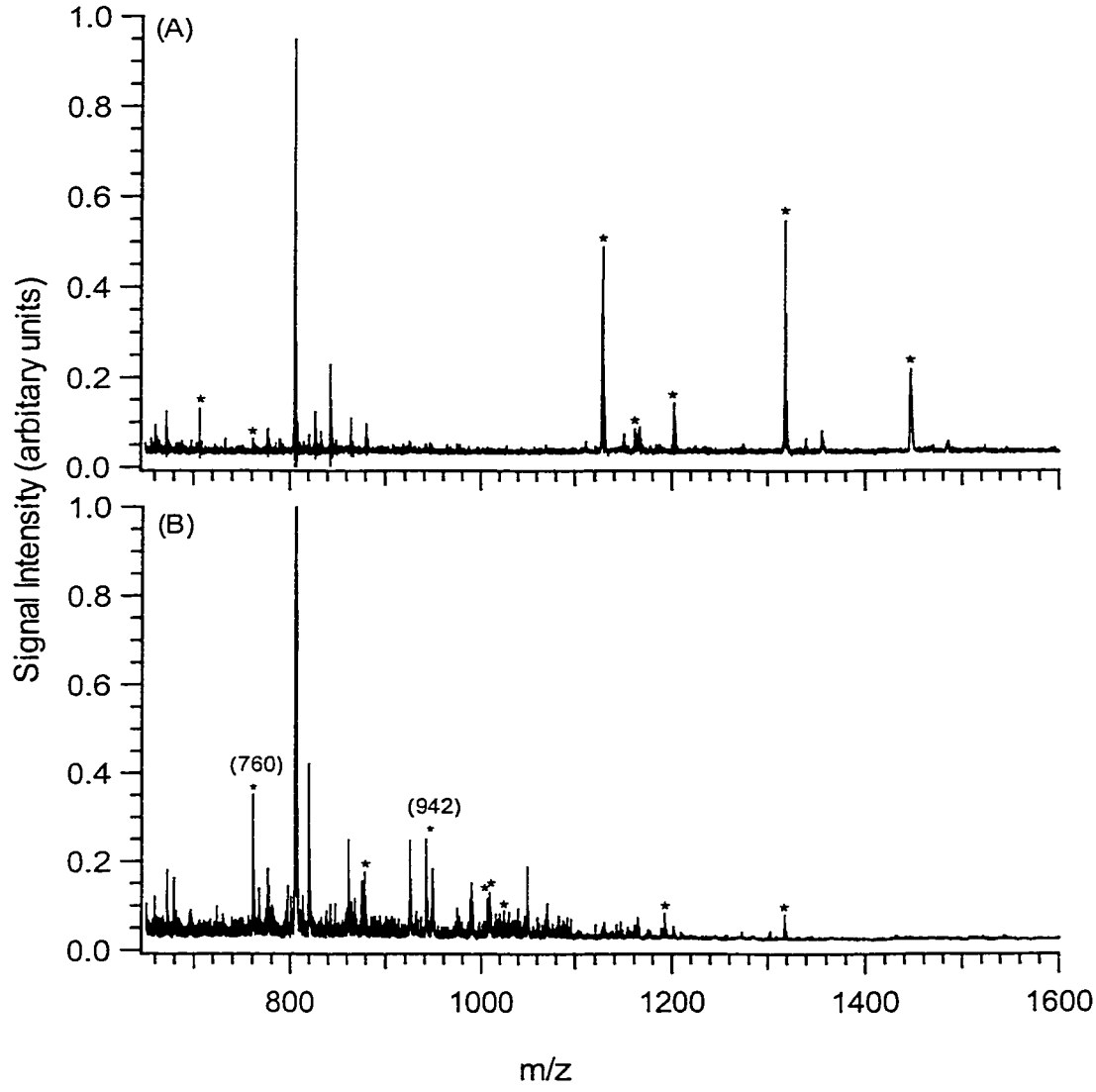


Figure 3.3 Mass spectra from the in-gel digestion of the expected HRV-3C protease gel band from MALDI-TOF (A) and NanoES ion trap (B). The peaks that match with the HRV-3C protease theoretical digest are labeled with asterisks. Ions with m/z 760 and 942 were chosen for tandem mass spectrometry (Figures 3.4 and Figure 3.5).

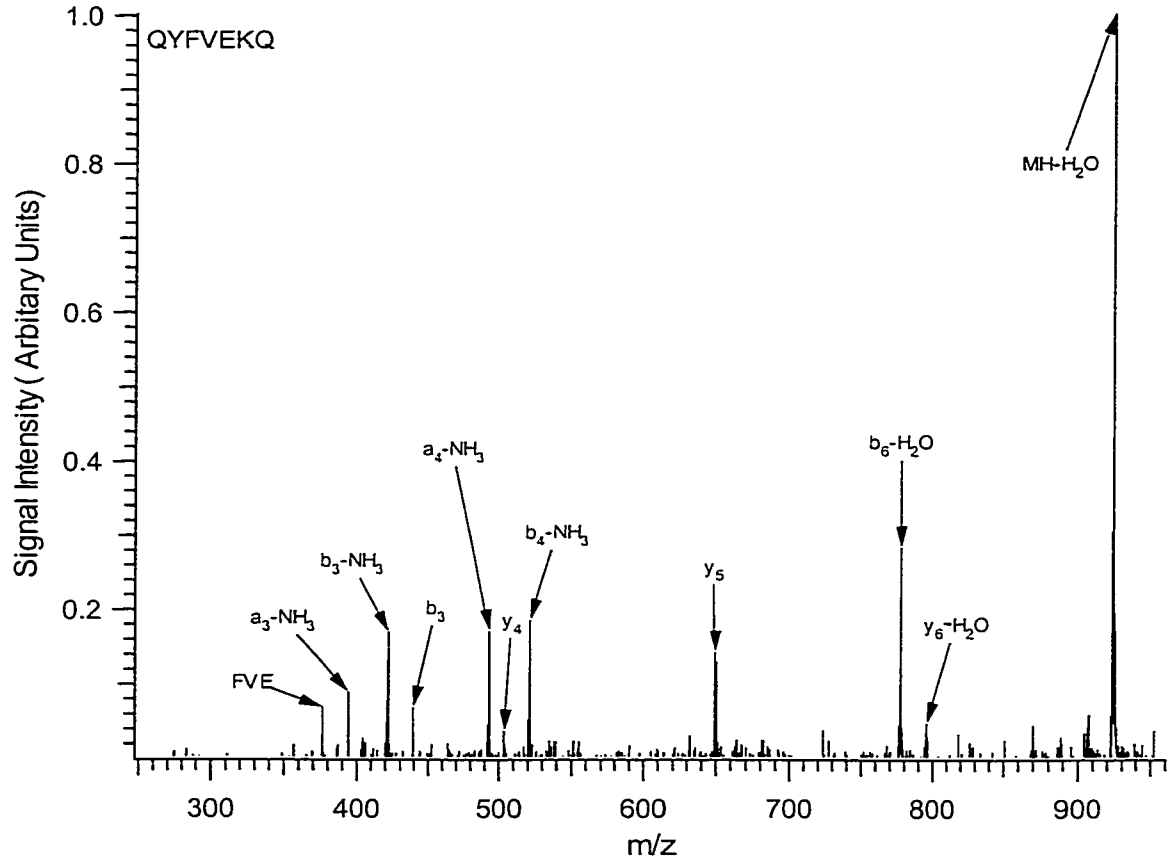


Figure 3.4 MS^2 spectrum of a precursor ion at m/z 942 by NanoES ion trap. Labeled peaks are fragment ions.

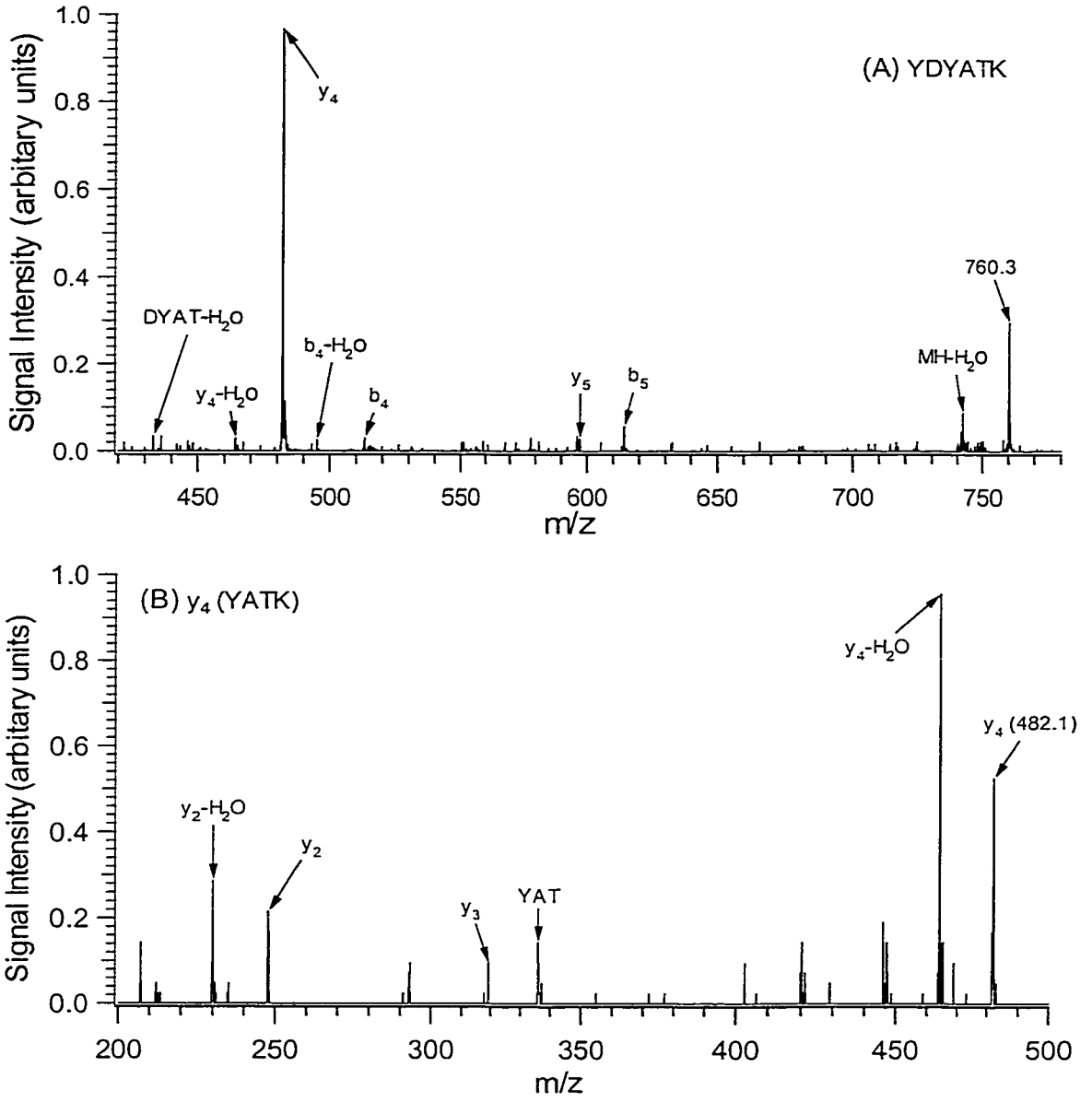


Figure 3. 5 Tandem mass spectra of the precursor ions at m/z 760 (MS^2) and 760/482 (MS^3) by NanoES ion trap. Peaks are labeled with their fragment ions.

However, more often than not, a MALDI peptide mass mapping experiment only generates an ambiguous identification result with low sequence coverage. In this case, sequence-specific fragment ion information generated by PSD and/or CID, along with **MS-Tag** program, can significantly increase the resulting confidence level, as illustrated in Section 3.3.2.

When an acceptable, but not 100%, confidence level is reached by peptide mass mapping, it is always constructive to have some primary structure information as supporting evidence, especially when the information can be gathered from different segments. Although tandem mass spectra alone might not be sufficient to piece together the whole protein, it is still a good compliment to peptide mass mapping, as demonstrated in Section 3.3.3.

It is worthy of note that crushing-gel/in-gel digestion was employed throughout all experiments in this chapter. Our results suggest that this methodology is a simple, efficient and time-effective protein handling technique for identification of real world protein samples.

3.5 Literature Cited

1. Patterson, D. S. *Anal. Biochem.* **1994**, 221, 1-15
2. Jungblut, P., Thiede, B. *Mass Spectrom. Rev.* **1997**, 16, 145-162
3. Gevaert , K., Vandekerckhove, J. *Electrophoresis* **2000**, 21, 1145.
4. Patterson, D. S. *Electrophoresis* **1995**, 16, 1791.
5. Papayannopoulos, I. A., *Mass Spectrom. Rev.* **1995**, 14, 49.
6. Shevchenko, A., Wilm M., Vorm O., Mann M. *Anal. Chem.* **1996**, 68, 850-858
7. Wilm, M., Mann, M. *Anal. Chem.* **1996**, 68, 1.

8. Dai,Y., Li, L., Roser, C. D., Long, S. R. *Rapid Commun. Mass Spectrom.* **1999**, 13, 73.

9. Shevchenko, A., Wilm M., Vorm O., Mann M. *Anal. Chem.* **1996**, 68, 850.

Chapter 4

Proteomic Analysis of Human Skin Fibroblasts Grown on Titanium: a Novel Approach to Study Molecular Biocompatibility[a]

4.1 Introduction

The basis for clinical success of any implant material is its ability to elicit a proper response at the tissue, cellular and molecular levels. Ultimately, the difference between a biocompatible material and a surface that can merely support cell attachment and growth, is determined at the molecular level. Yet there is a lack of consensus on what constitutes molecular biocompatibility.

Among metallic implant materials, commercially pure titanium (cpTi) has been most successfully used in various dental and orthopaedic devices [1-3]. The initial interaction of all known substrates, including titanium [4] with living tissues involves the adsorption of the components of the surrounding fluids, mainly from plasma, to the material surface [5-7]. The deposited proteins form a biofilm, the composition of which is determined by the surface properties of the material. The interaction of cells with this

---

[a] A form of this chapter has been submitted for publication: Kalal Derhami, Jing Zheng, Liang Li, Johan F. Wolfaardt and Paul G. Scott "Proteomic analysis of human skin fibroblasts grown on titanium: a novel approach to study molecular biocompatibility", *Journal of Biomedical Material Research*, submitted. Dr. Derhami did the experiments on protein sample preparation from cell culture to 2-D gel separation, western blotting and protein identification by 2-D database.

conditioning layer determines the cellular response that is manifested in the synthesis of proteins and the organisation of extracellular proteins. The biological-interface between titanium and adherent cells is the junction between this layer of adsorbed serum proteins, extracellular matrix, cell membrane components and the membrane associated cytoskeleton.

To study changes in cells in contact with a surface-adsorbed biofilm requires a method that is capable of resolving subtle changes in protein composition. Proteomics is an emerging area of research that deals with the analysis of gene expression using a combination of techniques to resolve, identify and characterise proteins, and then store this information in comprehensive databases [8]. In the present study, we used some of these techniques to detect changes in the proteins of human skin fibroblasts in response to two different growth-supporting surfaces, namely tissue culture polystyrene (TPS) and commercially pure titanium (cpTi). Whereas TPS is a common substrate for cell culture *in vitro*, cpTi is known to be biocompatible *in vivo*.

To separate and identify individual proteins, we used two-dimensional polyacrylamide gel electrophoresis (2-D PAGE) of extracts from the whole cells and the interface or of material adsorbed from serum. The gels were stained with silver and the spots were analysed by matrix-assisted laser desorption/ionisation mass spectrometry (MALDI MS), database searching, immunoblotting, running a standard or by a combination of these techniques.

In all, forty proteins were identified. We found that many of the proteins collected from the two surfaces were non-selectively deposited from the bovine serum present in the culture medium. The major differences in the proteins collected from the two surfaces were due to this differential adsorption. Among the cellular proteins, fibronectin and a

cytoskeletal protein: non-muscle myosin heavy chain type A (NMMHC-A) were highly expressed in cells grown on the more adhesive TPS surface.

4.2 Experimental

4.2.1 Cell Culture

Fibroblast cultures were established from the dermis of a normal individual as described [9]. Cells were grown in Dulbecco's Minimal Essential Medium containing 10% (v/v) fetal bovine serum and antibiotic-antimycotic (Gibco BRL, N.Y., USA) at 37°C in an atmosphere of 5% CO_2 in air for a period of 1 to 5 days. Materials for cell culture were obtained from Gibco BRL. Cells were used at the passages 2-9. Seven thousand cells/well were seeded in ninety-six-well "strip plates" with or without titanium discs in each well. The well diameter is 6.4 mm and has a growth area of 0.32 cm^2. The strip-wells are made of tissue culture-treated, sterile and non-pyrogenic medical grade polystyrene purchased from Costar Inc. (Pleasanton, CA., USA.)

4.2.2 Adsorption of Proteins from Serum

In the experiments involving analysis of proteins adsorbed from serum, the strip-wells with or without titanium discs were filled with 150 µL of the culture medium with 10% fetal bovine or human serum and incubated for 5 days at 37°C.

4.2.3 Titanium Discs

Commercially pure titanium discs were custom-manufactured (Nobel Biocare AB, Göteborg, Sweden) to be 2.00 mm thick and 6.35 mm in diameter and to fit tightly at the bottom of the wells in the 96-well plates. The discs were cleaned for 30 minutes in 30% nitric acid and then treated according to the standard Brånemark protocol [10], i.e. sonication in butanol for 10 minutes, rinsing twice in 99% ethanol and sonication for 10

minutes in 99% ethanol. The discs were then rinsed and sonicated for 10 minutes in sterile deionised water in sterile glass containers. Subsequent procedures were carried out in a laminar flow-hood and the discs were handled using autoclaved plastic instruments. A plastic suction tip was used to transfer each disc into a well and to tap it down into the bottom.

4.2.4 Extraction of Proteins

Fibroblasts cultured on the different surfaces or proteins adsorbed from serum were extracted using the sample buffer and methods of Hochstrasser [11], with some modifications. After 1 or 5 days in culture, cells were rinsed three times in PBS and extracted using a buffer containing dithiothreitol (0.65 M), CHAPS (3-[(3-Cholamidopropyl) dimethylammonio]-1-propane-sulfonate) (0.65M), urea (8.9M), 2% (v/v) Bio-Lyte™ ampholyte (pH range 3.0-10.0; Bio-Rad, USA), phenylmethane sulfonyl fluoride (1 mM) and butylated hydroxytoluene (0.23 mM) in distilled water. One hundred µL of the extraction buffer was added to the first of 8 wells in each experimental group and after 5 minutes this buffer was transferred to the next well. This process was repeated until all wells in each group had been extracted. The extracts were centrifuged at 100,000 x g for 0.5 h at 10°C in order to remove cellular fragments and then concentrated 5-fold using centrifugal microconcentrators (Amicon, USA) with a molecular weight cut-off of 10,000 Daltons. Concentrated extracts were stored at -80°C until analysed. The same volumes were maintained throughout for extracts made from cpTi and TPS. Proteins in fetal bovine or adult human serum were examined after preparing a 1:500 dilution in the extraction buffer.

4.2.5 2-D Gel Electrophoresis

Separation of the proteins was carried out essentially as described by Hochstrasser

[11], with some modifications. Isoelectric focusing gels, using N,N-methylene-bisacrylamide (Bis) as the crosslinking agent, contained 9.25 M urea, 16.7% acrylamide/BIS (respectively 30% (w/v) and 0.8% (w/v) dissolved in deionised water as stock), 1% (v/v) Bio-Lyte™ (5.0-8.0)(BioRad, USA) and 4.4% (v/v) Bio-Lyte™ (3.0-10.0)(BioRad, USA) and 5.5 % CHAPS/Nonidet (NP40) solution (30% w/v CHAPS and 10% v/v Nonidet NP40 in distilled water). The detergent solution was added after degassing the mixture containing the acrylamide/BIS. Seven μL of TEMED and 15 μL of ammonium persulfate (10%) were added to this mixture to initiate polymerisation.

A Mini-Protean II™ system (Bio-Rad, USA) was used for isoelectric focusing. The upper chamber was filled with degassed NaOH solution (20 mM) (catholyte) and the lower chamber with degassed H_3PO_4 solution (6 mM) (anolyte). With the upper chamber filled with anolyte gentle suction was used to remove the air bubbles from the top of the gel tubes and sample reservoirs. The gels were pre-electrophoresed at 200 V for 10 minutes, 300 V for 15 minutes and 400 V for 15 minutes. Twenty μL of overlay solution (500 μL of extraction buffer: 500 μL of distilled water : 40 μL of 0.05% w/v bromophenol blue) was then added to the top of each gel. The samples (10-30 μL) were introduced beneath the overlay solution and the gels were electrophoresed at 500 V for 10 minutes and 750 V for 3.5 hours. A 1.0 mL syringe filled with equilibration buffer (6% (v/v) of 0.05% bromophenol blue solution, 19% (v/v) of 10% (w/v) SDS solution, 14% (v/v) of 0.5 mol/litre Tris-HCl buffer, pH 6.8 and 49% (v/v) of deionised distilled water) was connected to one end of the gel tube and the gel was slowly extruded into a 1.5 mL Eppendorf tube along with approximately 150 μL of the equilibration buffer and immediately placed at -80°C.

Gels for SDS-PAGE were prepared according to Laemmli [12]. Gels, 7.5% or 10%, were cast in Mini-Protean II dual slab cells (Bio-Rad, USA). The IEF gels were thawed and

placed on a clean glass slide, excess equilibration buffer was removed by suction and the tube gel was transferred to and pushed down between the glass plates. The running buffer was prepared and cooled. The gel apparatus was placed in an ice-bath and the gels were run at 50 V for 15 minutes and then at 160 V for 50 minutes.

4.2.6 Silver Staining

Gels were fixed in 50% methanol and then in 5% acetic acid in water for 20 minutes. They were then washed for 10 minutes with 50% methanol in water and then for 10 minutes in water. The gels were sensitised by a 1-minute incubation in 0.02% (w/v) sodium thiosulphate, and then rinsed with two changes of water for 1 minute each. The gels were then submerged in 0.1% (w/v) silver nitrate and incubated at 4°C for 20 minutes. The silver nitrate was discarded and the gels were rinsed twice with water for 1 minute and then developed in 0.04% formalin in 2% (w/v) sodium carbonate with rapid shaking. As soon as this solution turned yellow it was replaced with fresh developer until the proteins were clearly visible. Development was stopped by soaking the gels in 5% acetic acid. They were then stored in distilled water at 4°C until analysed.

4.2.7 Measurement of Isoelectric Point (pI) and Molecular Weight

Standard proteins for 2D-PAGE (BioRad, USA) were electrophoresed at the same time as each batch of experimental samples. By measuring the position of the standard proteins in the gel, standard curves were generated using a statistical program (GraphPad Prism 2.0, GraphPad Software Inc., USA). Theoretical values for pI and MW were calculated from protein amino acid sequences using the online tool "Compute pI/MW" (http://expasy.cbr.nrc.ca/tools/pi_tool/html) for each SWISS-PROT entry, following removal of any signal peptides and propeptides. Identifications made by comparison to the SWISS-2DPAGE database only considered the apparent pI and MW values.

4.2.8 Immunochemical Identification of Proteins

Western blotting was used to identify vinculin, actin and fibronectin in the 2D-PAGE gels. Immediately after the second dimension electrophoresis was completed, the gel was placed on a piece of Immobilin-P membrane (Bio-Rad, USA) that had been presoaked in 100% methanol for 15 seconds. Electroblotting was performed for 1 hour at 60 V in a mini transfer-blot apparatus (Bio Rad, USA). The transfer buffer was 25 mM Tris, 192 mM glycine and 10% methanol. The membrane was then air-dried at room temperature. After re-wetting the membranes with 100% methanol, they were incubated with 2% (w/v) non-fat milk dissolved in PBS for 1 hour in order to block non-specific binding. Mouse monoclonal antibodies against vinculin, cellular fibronectin and actin (Sigma, USA) at a 1:500 dilution in 0.1% Tween-PBS, were added to the membranes and incubated for 2 hours at room temperature. After washing 4 times for 5 minutes in PBS, peroxidase-conjugated secondary antibody (1:1000 dilution in 0.1% Tween-PBS) was added, followed by incubation for 1.5 hours at room temperature. The membranes were then washed 4 x 5 minutes in PBS and transferred to a 0.5 mg/mL solution of diaminobenzidine (DAB) in PBS to which 300 μL of 1% cobalt chloride / 10mL of DAB solution and 3 μL of 300% hydrogen peroxidase/mL of DAB was added, until colour was developed. As a control, immunoblotting was performed without the primary antibody. The membranes were washed and dried and then scanned using OFOTO software (USA). The apparent pI and MW of the proteins in the spots were estimated by comparison of positions with those of protein standards. In addition, cellular fibronectin (Sigma, USA) was run as a standard.

4.2.9 In-gel Digestion

The same crushing-gel/in-gel digestion procedure, as described in Chapter 3, was used. Protein spots were excised from the gel. The gel pieces were put into 0.6-mL vials and rinsed with 100 mM NH_4HCO_3 to ensure the pH was about 8.5. This NH_4HCO_3

solution was replaced with 10 μL of the same buffer solution and 2 μL of 90mM dithiothreitol (DTT). The gels were crushed using the tip of a thin 0.25 mL plastic vial. After incubation for 30 min at 50°C, the suspension was cooled to ambient temperature and 2 μL of 200 mM iodoacetamide was added. The gel samples were incubated at room temperature for 30 min. One μL of 0.2 μg/μL trypsin was added for digestion at 37°C for 2 to 5 hours. Peptides were extracted three times into 20 μL of 75% acetonitrile in 0.25% trifluoroacetic acid/water with vortexing for 20 seconds and sonication for 20 minutes, followed by 10 μL of acetonitrile. The pooled extracts were evaporated to dryness in a vacuum centrifuge.

4.2.10 Sample Preparation for MALDI Analysis

A two-layer method was used for coating probes for MALDI mass spectrometry [13]. The first-layer solution was prepared with 10 mg of matrix (α-cyano-4-hydroxy-*trans*-cinnamic acid) in 1 mL methanol/acetone (1:4 v/v). The second-layer solution was a saturated solution of matrix in 40% (v/v) methanol in water. Five μL of the second-layer solution was added to each vial containing dried sample. The vial was vortexed to dissolve the sample. Two μL of the first-layer solution was deposited onto a probe tip to form a thin layer. Onto this layer, 0.5 - 1 μL of the sample/matrix mixture was deposited and allowed to dry. The sample layer was washed three times with water.

4.2.11 MALDI Mass Spectrometry

Mass spectral data on protein digests were collected on a home-built time-lag focusing MALDI TOF mass spectrometer, described in detail elsewhere [14]. The pulse delay was set to 0.25 μs and the pulse voltage was optimized at 2.8 kV for the best resolution in the peptide mass range. In general 60 laser shots were averaged to produce a mass spectrum. Spectra were acquired and processed with Hewlett-Packard supporting

software. The individual spectra were internally calibrated with trypsin autolysis peptide peaks and matrix peaks.

4.2.12 Protein Identification

The peptide mass mapping technique was applied by means of search program **PeptIdent** (http://expasy.cbr.nrc.ca/tools/peptident.html) to identify the proteins from gel. The protein molecular weight and isoelectric point estimated from the position of the protein spot in the gel were entered initially with an assumed accuracy of +/- 10%. If this initial search did not produce a result, the estimated accuracy was relaxed to +/- 20%. The cleavage reagent was known to be trypsin and 1 missed cleavage was permitted in all our searches. Cysteine was designated to be carboxamidomethylated by iodoacetamide and methionine oxidized. The tolerance for peptide masses was set at +/- 500 ppm. Peptides found in a blank spot were subtracted from each sample and the remaining peaks were used in the database search. After identification of the highest scoring proteins, the theoretical monoisotopic and average tryptic peptide masses for each protein were calculated by program **Peptide-Mass** (http://expasy.cbr.nrc.ca/tools/peptide-mass.html) and manually compared with the masses measured by MALDI MS. In order to test the reliability of the peptide mapping procedures, spots containing easily identifiable proteins such as actin and bovine serum albumin were analyzed in the same manner. Some easily distinguishable spots were identified by comparison with the human skin fibroblasts database (http://biosun.biobase.dk/~pdi/2Dgallery/P473mirror_map.html) or by comparison to the SWISS-2DPAGE human plasma database (http://www.expasy.ch/cgi-bin/ch2d-compute-map?PLASMA_HUMAN,P2161). Presently, there is no bovine plasma 2D gel database available on the Internet, therefore serum proteins were compared to the SWISS-2DPAGE human plasma database.

4.3 Results

4.3.1 Adsorption of Serum Proteins onto Titanium and Tissue Culture Polystyrene

In order to distinguish cellular proteins from proteins adsorbed from serum, we incubated the two cpTi and TPS surfaces with culture medium containing 10% fetal bovine serum for 5 days and analyzed the results by 2D-PAGE. Figure 4.1 shows silver-stained gels of these adsorbed biofilms collected from cpTi (Figure 4.1a) and TPS (Figure 4.1b) and Table 4.1 lists the serum proteins that were identified. Several of these proteins (spots 1 – 5) were also seen in the extracts collected from whole cells (Figure 4.2 and 4.3). The adsorption of serum proteins is much greater on cpTi than on TPS. Spots 1-5 were clearly much more abundant in both the adsorbed biofilm and in whole cell-extracts collected from cpTi (Figure 4.1a, 4.2a and 4.3a). It is important to note that the extraction and loading volumes were the same for these samples.

To test whether this adsorption was selective for specific serum components, proteins in total (non-adsorbed) bovine and human serum were investigated by 2D-PAGE (Figure 4.4). Since the same spots appeared in both the biofilms and in the total serum, we conclude that adsorption was non-selective. Protein spots 1-5 were common to both the fetal bovine and human sera, however spots 1, 2 and 3 were present at higher levels in the bovine serum. Overall, more proteins were detected in the human serum (Table 4.1). Human serum proteins 22, 24, 25, 26 and 29 were not detectable in the TPS biofilm, despite their relatively high concentrations in the serum (Figure 4.4a and 4.4c).

Mass spectrometric analysis of the tryptic digests of the separated bovine serum proteins identified spot 1 as fetuin, also known as α-2-HS-glycoprotein (A2HS). This identification was confirmed by repeating the MALDI analysis in two separate experiments and this protein was also identified in the SWISS-2DPAGE human plasma database. The

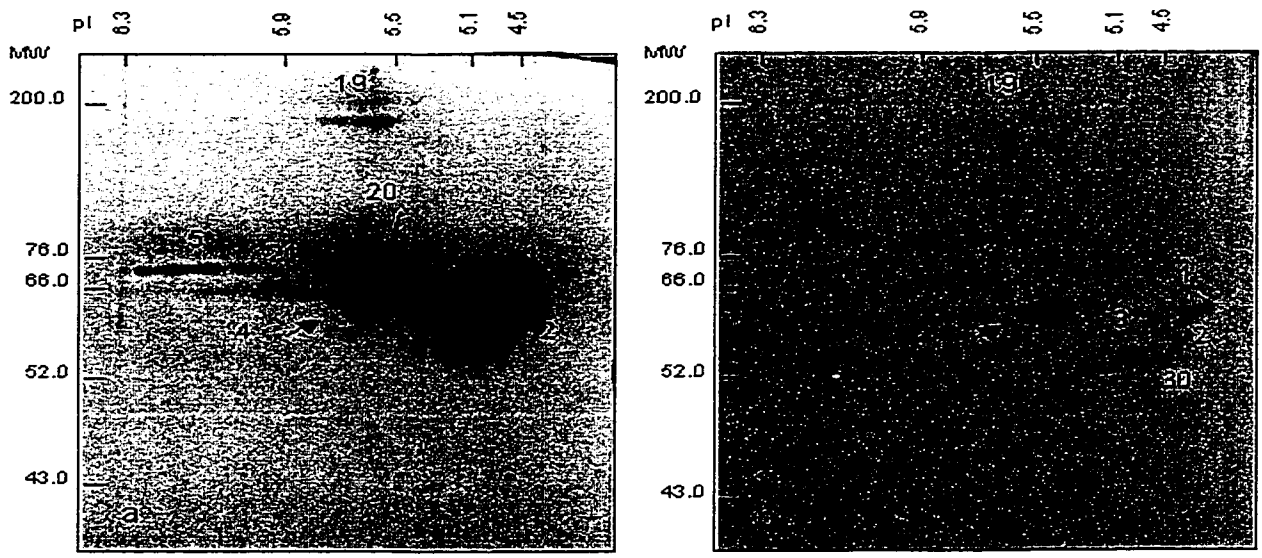


Figure 4.1 Silver stained 2D-PAGE of adsorbed bovine serum proteins extracted from a) cpTi and b) TPS surfaces. The cpTi surface shows significantly higher adsorption of serum proteins (spots 1-5, 19-21). All the spots were identified except spot 19 (see Table 4.1).

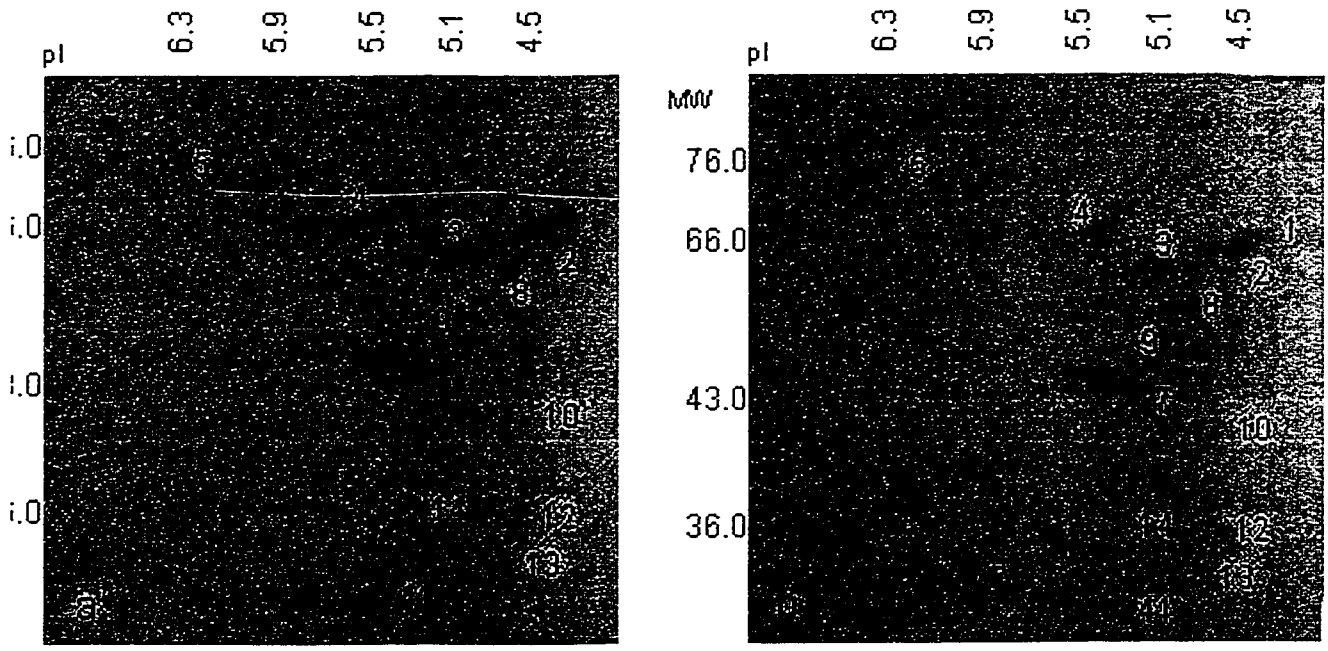


Figure 4.2 Silver stained 2D-PAGE of total cell and substrate extracts from fibroblasts cultured for 1 day on a) cpTi and b) TPS surfaces. Protein spots 1-5 were found at increased concentrations on cpTi. The increased staining of protein spot 8 on TPS is particular to this gel and was not consistently higher in other similar samples.

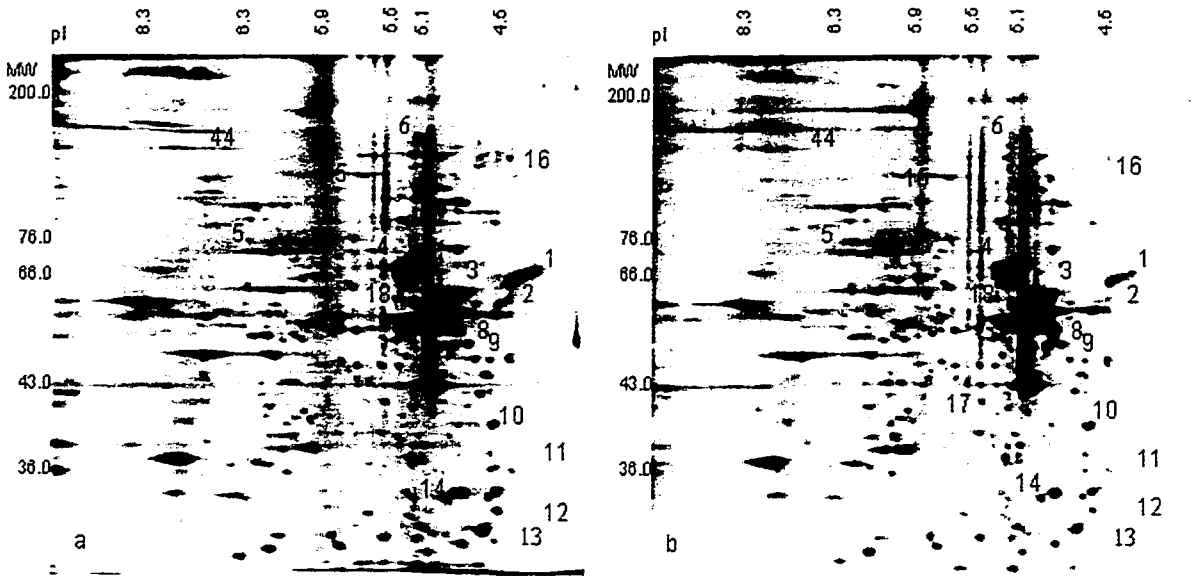


Figure 4.3 Silver stained 2D-PAGE of total cell and substrate extracts from fibroblasts cultured for 5 days on a) cpTi and b) TPS. Protein spots 6 and 44 show increased expression on TPS.

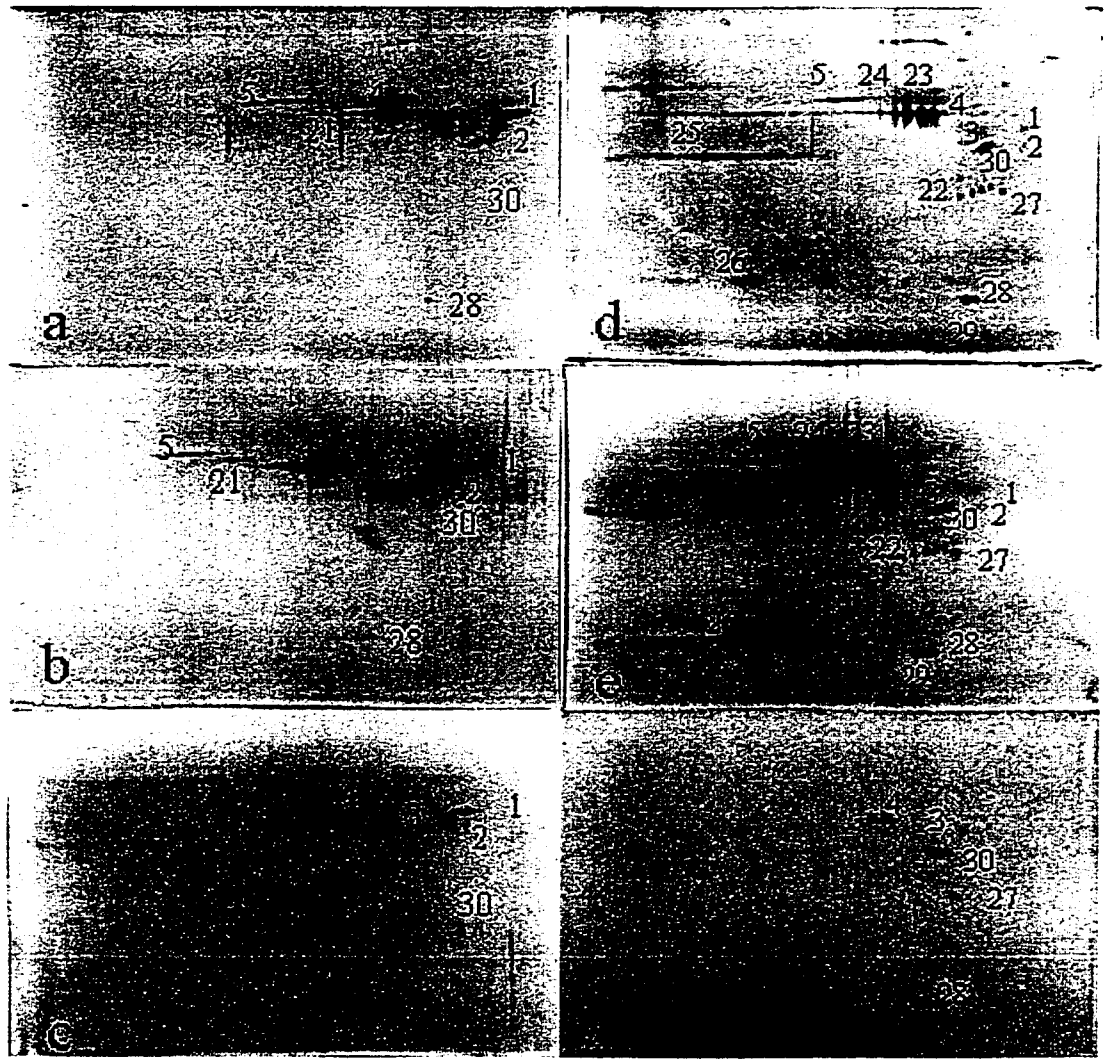


Figure 4.4 Comparison by silver stained 2D-PAGE of total proteins in bovine and human serum and proteins adsorbed from serum onto cpTi and TPS. a) total proteins in fetal bovine serum, b) proteins adsorbed from whole fetal bovine serum onto cpTi , c) proteins adsorbed from whole fetal bovine serum onto TPS. d) total proteins in whole adult human serum, e) proteins adsorbed from whole adult human serum onto cpTi, f) proteins adsorbed from whole adult human serum onto TPS. The adsorption onto either surface appears to be non-selective.

Table 4.1 Proteins adsorbed from bovine and human serum

Spot No.[1]	Protein name (species)[2]	Accession Code[3]	Identification Method[4]	Apparent[5]		Theoretical[6]	
				pI	MW	pI	MW
1	Fetuin (b), α2 HS glycoprotein (h)	P12763	Swiss 2D, MALDI (9/27)	4.00	60.0	5.10	36.4
2	Thrompospondin 1 (b,h)	Q28178	Swiss 2D, MALDI (5/29)	4.50	57.0	4.22	52.1
3	α-fetoprotein (b, h)	P02772[7]	Swiss 2D, MALDI (5/28)	5.10	58.0	5.54	65.3[7]
4	Serum albumin (b, h)	P02769	Swiss 2D, MALDI (15/42)	5.50	65.0	5.60	66.4
5	Plasminogen Chain A (b, h)	P06868	Swiss 2D, MALDI (11/21)	6.10	72	6.40	63.1
19	Unknown (b)			5.40	>100.0		
20	Serum albumin (b)	P02769	Swiss 2D, MALDI (12/21)	5.50	72	5.60	66.4
21	Serotransferrin (b)	Q29443	Swiss 2D, MALDI (5/13)	5.80	66.0	6.50	75.9
22	Haptoglobin 1 β-chain (h)	P00737	Swiss 2D	5.50	43.0	6.30	27.3
23	Serum albumin (h)	P02768	Swiss 2D, MALDI (9/25)	5.65	65.0	5.67	66.5
24	Serum albumin (h)	P02768	Swiss 2D, MALDI (12/39)	5.75	65.0	5.67	66.5
25	IgG heavy chain γ (h)	P99006	Swiss 2D	6.40	51.0		
26	IgG light chain (h)	P99007	Swiss 2D	6.40	25.0		
27	Complement 3 (C3DG) (h)	P01024	Swiss 2D, MALDI (4/16)	5.00	39.0	5.00	38.9
28	Apo A-I (h)	P02647	Swiss 2D	5.20	23.0	5.27	28.1
29	Haptoglobin 1 α chain (h)	P00737	Swiss 2D	5.40	17.0	5.23	9.2
30	Vitamin D-binding protein (h)	P02774	Swiss 2D	5.00	58.0	5.20	51.2
48	Serum amyloid P (h)	P02743	Swiss 2D	5.30	26.0	6.12	23.3

[1]Spot numbers used here are the same as those in the figures. [2] b = bovine, h = human. [3] SWISS-PROT accession codes. [4] Swiss 2D and MALDI refer to identification from the SWISS-2DPAGE human plasma database and by peptide mass mapping, respectively. Numbers in parentheses are the numbers of peptide masses found relative to the total numbers submitted to the search. [5] pI and MW from calibration with standard proteins. Where the protein was present as a series of spots the centre of the group was used for the estimation of pI and MW. [6] pI and MW were calculated from the sequence provided under the given accession code. Where 2 species are shown in column 2 the theoretical pI and MW are for the bovine homologue. [7] The amino acid sequence of bovine -fetoprotein is not available for peptide mass mapping, therefore this spot was identified from the sequence for the mouse homologue. Polyclonal immunoglobulin G has no unique sequence, precluding calculation of theoretical values for pI and MW.

protein appears in Figures 4.1, 4.2 and 4.4 as a series of 4 or more spots, the largest of which is also the most acidic. This pattern is commonly seen for proteins with several glycosylation sites where addition of oligosaccharides generates larger and more acidic forms. Both fetuin and FETA are major components of fetal serum. Bovine fetuin contains three N–linked and two O-linked oligosaccharides with carbohydrates accounting for nearly 30% of the total weight [15,16]. The sialic acid in these oligosaccharides contributes to the acidic nature of this protein and its removal raises the pI to 5.25[17].

Spot 2 was tentatively identified as thrombospondin 1 (TSP1), however only 5/29 peptide masses matched and we could not confirm this spot in the human plasma 2D-PAGE database. The search in the bovine database for the peptide masses obtained for spot 3 did not show any matches, however a search in the mammalian database showed 4 peptide mass matches for the mouse vitamin D binding protein (VTDB). Since VTDB is a homolog of α-fetoprotein (FETA), which is the second most abundant protein in the fetal bovine serum after serum albumin, we suspected that spot 3 might contain FETA. Indeed, manual peptide mass matching yielded 5 matches with mouse FETA and, in addition, the theoretical pI (5.54) and MW (65.3 kDa) of FETA are close to our measured values. The SWISS-PROT database does not include a sequence for bovine FETA, which explains why our first search did not identify this protein. Spot 4 was identified as serum albumin and this was confirmed by comparison to the human plasma 2D-PAGE database. The series of spots with apparent pIs of 5.9 to 6.3, (labeled "5") were found to contain bovine plasminogen chain A. This identification was confirmed in two independent experiments.

The less extensively adsorbed bovine serum proteins (spots 19, 20, 21 and 30: Fig. 1) were analyzed by MALDI MS. We identified spot 20 as bovine albumin and spot 21 was found to contain serotransferrin. We were not able to extract sufficient material from

spot 30 for peptide mass mapping but comparison to the SWISS 2D-PAGE human plasma database suggests that it contains the vitamin D binding protein (VTDB) and corresponds to spot 30 from human serum (Figure 4.4).

Spots 22 and 29 in Figure 4.4 were identified as haptoglobin 1 β and α chains, respectively, by comparison to the SWISS 2D-PAGE human plasma database. As expected, MALDI/MS analysis of spots 4, 23 and 24 confirmed that they contained isoforms of human serum albumin. The two components of immunoglobulin G: the heavy chain (spot 25) and light chain (spot 26), were identified by comparison to the SWISS 2D-PAGE human plasma database. Analysis by MALDI MS showed spot 27 to contain complement 3. This was confirmed by comparison to the database. Apolipoprotein A-1 (Apo A-1) (spot 28) and serum amyloid P (spot 48) were also identified in the database. Spot 19 could not be identified by either method.

For several spots (or series of spots) there were significant discrepancies between apparent and theoretical pI and MW values, usually with the former being larger and more acidic. In four cases (haptoglobin α and β chains, vitamin D binding protein and serum amyloid protein) these discrepancies could be explained by the existence on these proteins of potential sites for N-glycosylation.

4.3.2 Fibronectin and Non-muscle Myosin Heavy Chain A are Upregulated on Tissue Culture Polystyrene

Both spots 6 and 44 were more abundant in extracts from cells cultured on TPS for 5 days compared to those from cells on cpTi (Figure 4.3). Analysis of spot 6 by MALDI MS showed it to contain non-muscle myosin type A heavy chain. This identification was confirmed in two separate experiments and the protein was also identified in the human

MRC5 fibroblast database (http://biosun.biobase.dk/~pdi/2Dgallery/P473mirror-map.html).

The identification of spot 44 as cellular fibronectin was confirmed by immunoblotting (Figure 4.5), which also identified spot 15 as vinculin and spot 7 as actin. The position of cellular fibronectin in the standard 2D-PAGE also matched the spot detected in the immunoblot (data not shown). Non-muscle myosin heavy chain A and fibronectin were the only cellular proteins that we have so far identified that are differentially expressed by fibroblasts cultured on the two surfaces. These observations were confirmed in two completely independent experiments.

4.3.3 Identification of Other Fibroblast Proteins

Table 4.2 lists the proteins from human skin fibroblasts that were identified by various procedures. The spot-numbers in Figures 4.2 and 4.3 correspond to those in the master map (Figure 4.6). In all, twenty-eight cellular proteins were identified, most of which could be confirmed by comparison to various 2D-PAGE databases. Five proteins were identified by MALDI MS. Results of MALDI/MS showed spot 7 to contain the two actin isoforms (spot 7 and spot 7a) of non-muscle type protein. The occurrence and positions of these spots in 2D gels of many cell lines are well known. Vimentin and its degradation products (spots 8, 9a, 9b, 9c), tropomyosins (spots 10, 11,12, 13), annexins (spots 14, 42), plasminogen activator inhibitor-1 (spot 17) and α and β-tubulin (spots 18, 38) were identified by comparison to the human skin fibroblast database on the basis of their isoelectric points, molecular weights and their proximity to other protein spots. Thrompospondin 4 (spot 16) could only be identified by MALDI MS analysis.

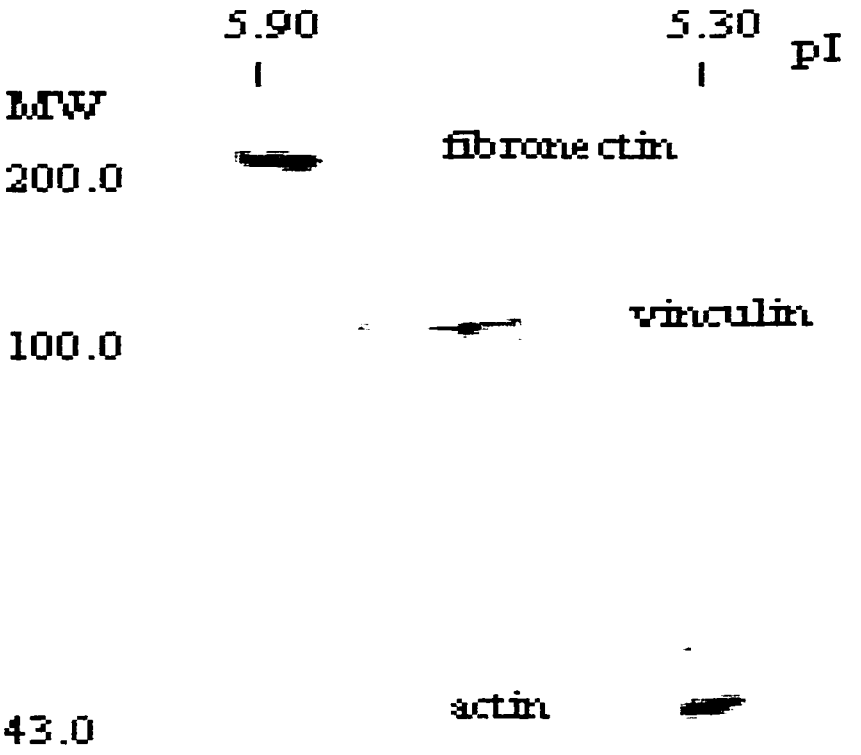


Figure 4.5 Immunoblot of 2D-PAGE gel of a cell extract stained with antibodies directed against cellular fibronectin, vinculin and actin.

Table 4.2 Proteins from human skin fibroblasts

Spot No.[1]	Protein name	Accession Code[2]	Identification Method[3]	Apparent[4]		Theoretical[5]	
				pI	MW	pI	MW
6	Non muscle myosin heavy chain type A	P35579	Human-2D, MALDI (15/27)	5.50	200	5.50	226
7	Actin-γ	P02571	Human-2D, MALDI, IB (18/57)	5.30	43.0	5.31	41.7
7a	Actin-β	P02570	Human-2D, MALDI, IB (17/57)	5.30	43.0	5.29	41.6
8	Vimentin	8520	Human-2D	5.00	52.0	5.06	53.5
9	Vimentin variant	8422	Human-2D	5.00	50.0		
9a	Vimentin degradation product	1301	Human-2D	5.00	48.0		
9b	Vimentin degradation product	304	Human-2D	4.85	46.0		
9c	Vimentin degradation product	9307	Human-2D	4.80	44.0		
10	Tropomyosin	9227	Human-2D	4.80	38.5	4.69	32.7
11	Tropomyosin 2	9119	Human-2D	4.80	37.5	4.63	32.9
12	Tropomyosin 4	9119	Human-2D	4.80	33.3	4.68	32.8
13	Tropomyosin 5	9118	Human-2D	4.80	31.8	4.48	28.3
14	Annexin V	8216	Human-2D	5.00	32.4	4.83	35.7
15	Vinculin	P18206	Human-2D, MALDI (10/48), IB	5.80	110	5.83	116
16	Thrompospondin 4	P49746	MALDI	4.40	110.0	4.43	104
17	PAI-2	6307	Human-2D	5.50	40.0	5.46	46.6
18	α -tubulin	7513	Human-2D	5.10	61.2	5.02	50.2
33	Hsp 90	1608	Human-2D	5.00	87.0	4.94	85.0
34	BiP (grp 78)	8619	Human-2D	5.00	75.0	5.03	72.1
35	Hsp 70	6602	Human-2D	5.30	65.0	5.37	71.0
36	t-plastin (fimbrin)	6513	Human-2D	5.40	63.0	5.52	70.4
37	Hsp 60	7517	Human-2D	5.30	55.0	5.70	61.0

38	β-tubulin	8511	Human-2D	4.80	55.0	4.75	49.8
39	Rho GDI	8123	Human-2D	4.80	25.0	5.03	23.2
40	Translationally controlled tumor protein	8113	Human-2D	4.80	23.0	4.84	19.6
41	Neuron cytoplasmic protein (PGP 9.5)	6104	Human-2D	5.40	27.0	5.33	24.8
42	Annexin I	5209	Human-2D	6.10	35.0	6.64	38.6
43	α-enolase	2418	Human-2D	7.20	47.0	6.99	47.0
44	Fibronectin	P02751	STD, IB	5.80	250	5.45	263
45	Myosin light chain	8002	Human-2D	4.50	17.0	4.56	16.8
46	Inititiation factor 4D	7006	Human-2D	5.00	17.0	5.08	16.7

[1] Spot numbers used here is the same as those in the figures. [2] SWISS-PROT accession codes begin with a letter. Those from the SWISS-2DPAGE database begin with a number. [3] Human 2D, MALDI, IB and STD refer to identification from the database of human fibroblast proteins, by peptide mass mapping, by immunoblotting and by comparison to positions of standard proteins, respectively. Numbers in parentheses are the peptide masses found, compared to the total numbers submitted to the search. [4] pI and MW from calibration with standard proteins. [5] pI and MW calculated from the sequence provided under the given accession code, where available.

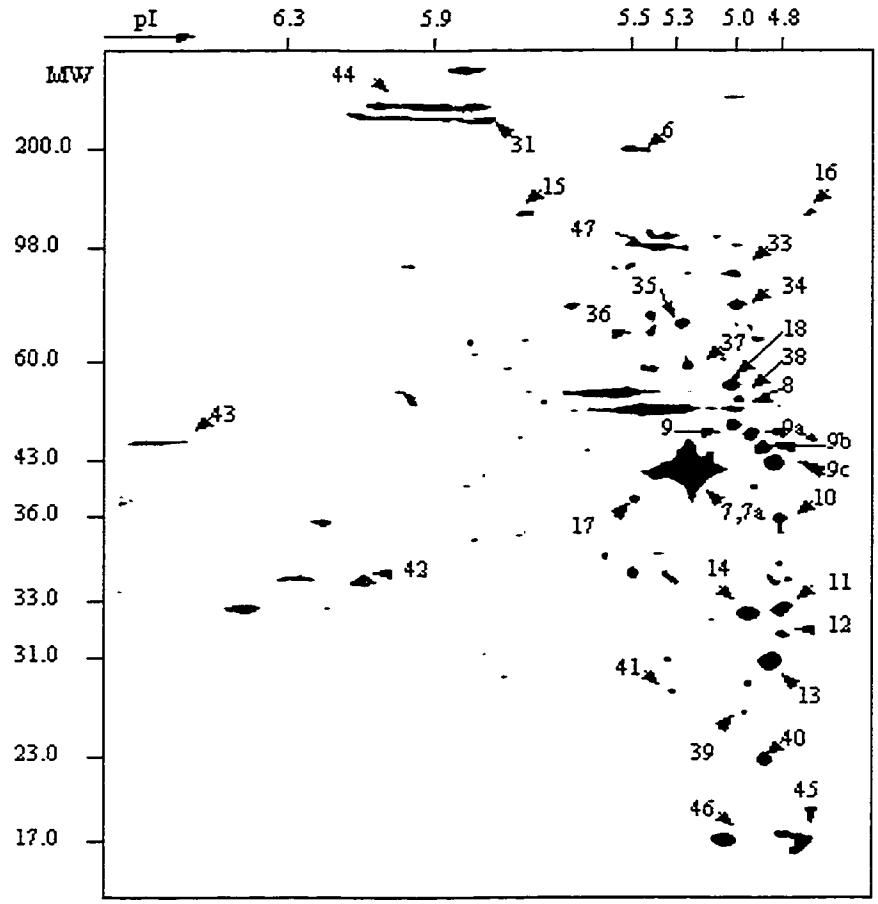


Figure 4.6 A schematic map of proteins found in 2D-PAGE gels of whole cell lysates of skin fibroblast

4.4 Discussion

Initial interaction of most implant materials with living tissues begins with the adsorption of plasma proteins onto the substrates [18-21]. The physical and chemical composition of this adsorbed biofilm presumably determines the biocompatibility of a material. We found that plasma proteins adsorbed to titanium surfaces are consistently present as a significant proportion of the extracellular matrix in the extracts of whole cell layers in both early (1 day) and late (5 day) cultures. This is important since it suggests that fibroblasts remain in contact with the adsorbed serum proteins for at least 5 days - the duration of our previously published experiments on fibroblast adhesion [22]. Other investigators have reported the adsorption of serum albumin [2,21], thrombospondin [23] and other serum components from blood to titanium surfaces. The present study confirms these findings and adds several additional serum proteins to this list (Table 4.1). These remain firmly attached to cpTi and may therefore play a major role in determining cell and tissue response to titanium.

We found adsorption of serum proteins onto either surface to be non-selective, however, significantly more adsorbs to cpTi. This higher concentration of each protein, and/or the thickness of the adsorbed biofilm, may be important in the initial interaction of cells with cpTi. Some of the adsorbed proteins may act as targets for secondary binding by other adhesive proteins and thus modulate cell adhesion. For example, plasminogen and serotransferrin (a known mitogen for fibroblasts [24]) could directly reduce (the strength of) cell attachment to cpTi while the increased serum albumin on cpTi could reduce fibronectin adsorption and hence indirectly reduce cell attachment [25].

Next to serum albumin, fetuin and FETA appear to be the most abundant serum proteins adsorbed to the substrates. The amino acid sequence of fetuin includes an EF hand

motif that is suggested to bind Ca^{2+}. In fact, high amounts of fetuin consistently occur in the mineralized phase of bone and teeth [26] and the 56 kDa bone sialoprotein (BSP) has been shown to be dephosphorylated fetuin [27]. It is therefore suggested that fetuin is important in the remodeling of hard tissues [15]. The mean concentration of fetuin in fetal bovine serum (10 - 22 mg/mL) is considerably higher than the concentration of its human homolog (α2-HS glycoprotein) in human serum (0.4 - 0.6 mg/mL) [15,17]. Therefore, since the adsorption is non-selective, the concentration of bovine fetuin on cpTi will be much greater than that of the human α2-HS glycoprotein. This raises a concern about the clinical relevance of *in vitro* investigations of titanium implants using bovine serum as a blood substitute: it is probably more appropriate to use human serum in this experimental system, to better reflect clinical conditions.

Serum albumin and FETA are two of the most prominent members of the albuminoid gene family that also includes vitamin D-binding protein. Albuminoid proteins are synthesized in the liver and secreted into the blood and are thought to have a multitude of biological activities [28]. It is believed that proteolytic cleavage of serum albumin and FETA generates regulatory peptides capable of eliciting a variety of biological responses [27]. Indeed, FETA has been shown to be a regulator of growth in many cell types [29]. It has also been proposed as an adhesion molecule for epithelial cells [28]. The high degree of identity (30-47%) between the amino acid sequence of FETA and several extracellular cellular matrix (ECM) adhesion molecules such as syndecan, fibronectin, collagens and integrins [29], also implicate it in cell-cell and cell-ECM adhesion. Since fibronectin is a major part of the extracellular matrix collected from the substrates, FETA could be suggested to act as a binding-site for secondary interaction with fibronectin.

We have previously described a dense network of fibrillar fibronectin at the interface of fibroblasts and TPS [22], whereas we found that fibroblasts growing on cpTi

organize a less dense matrix of cellular fibronectin. A dense fibrillar matrix of fibronectin is also seen associated with fibroblasts under mechanical strain *in vitro* [30]. It is conceivable that the thick biofilm of serum proteins adsorbed to cpTi results in a decrease in the affinity of cellular fibronectin for the substrate, thus causing a decrease in intracellular tension. A reduction in cellular contractility may in turn cause a down-regulation of fibronectin and the formation of a less dense fibrillar matrix on cpTi. It is known that mechanical tension leads to an increase in synthesis of fibronectin and a coordinated change in the cytoskeleton [30]. An increase in the rigidity of the extracellular matrix results in a strengthening of the cytoskeletal linkages [31]. In the present investigation we found that modulation in the expression of cellular fibronectin is accompanied by a corresponding change in the expression of myosin heavy chain (NMMHC-A). This cytoskeletal protein is known to play a role in stress fiber contractility. Its down-regulation on cpTi could conceivably reduce cellular contractility and promote fibroblasts to organize a less dense fibrillar matrix of fibronectin. This in turn could result in a decrease in the strength of cell adhesion to titanium. This possibility is consistent with the results of our previous study in which we demonstrated that fibroblasts adhere less strongly to cpTi than to TPS after several days in culture [22]. We proposed that assembly of a less dense fibrillar matrix of fibronectin by fibroblasts cultured on cpTi might be a consequence of a reduction in the cytoskeletal stiffness. It remains to be seen if a difference in the adsorbed biofilm alone can explain the observed down regulation of myosin heavy chain in cells grown on cpTi.

Together, the findings presented in this paper suggest that differences in the quantity of protein adsorbed from serum may have resulted in a change in the expression and/or adhesion of fibronectin and a decrease in fibroblast contractility. This might then explain the lower strength of adhesion of fibroblasts to cpTi compared to TPS, reported

previously [22]. In comparison to TPS, the surface of cpTi appears to promote the formation of a more concentrated carpet of plasma proteins. The adsorbed serum proteins could potentially modulate cell adhesion and thus elicit changes in the cytoskeleton and consequently affect the ability of cells to organize their extracellular matrix. We have also demonstrated here that the proteomic approach to investigate biocompatibility may be of value in the search for the specific protein(s) involved in the biological response to a material.

4.5 Literature Cited

1. Adell, R., Eriksson, B., Lekholm, U., Branemark, P-I., Jemt, T. *Int. J. Oral Maxillofac. Implants*, **1990**, 5, 347.
2. Parel, S.M., Branemark, P-I., Tjellstrom, A., Gion, G., *J. Prosthet. Dent.*, **1986**, 55, 600.
3. Hagert, C.G., Branemark, P-I., Albrektsson, T., Strid, K.G., Irstam, L., *Scand. J. Plast. Reconstr. Surg.*, **1986**, *20*, 207.
4. Scarborough, D.E., Mason, R.G., Dalldorf, F.G., Brinkhous, K.M., *Lab. Invest.*, **1969**, 20, 146.
5. Ellingsen, J.E., *Biomaterials*, **1991**, 12, 593.
6. Sevastianov, V.I., *High Performance Biomaterials*, Szcher, M. (ed), Technomic Publishing, Lancaster, PA, **1991.**
7. Baier, R.E., Dutton, R.C., *J. Biomed. Mater. Res.* **1969**, 3, 191.
8. Celis, J.E., Ostergaard, M., Jensen, N.A., Gromova, I., Rasmussen, H.H., Gromov, P., *FEBS Letts*. **1998**, 430, 64.
9. Nakano, T. and Scott, P.G., *Biochem. Cell. Biol.* **1986**, 64, 387.
10. Hobo, S., Ichida, E., Garcia, L.T., *Osseointegration and Occlusal Rehabilitation*,

Quintessence, Chicago, **1989**, 3.

11. Hochstrasser, D.F., Harrington, M.G., Hochstrasser, A.C., Miller, M.J., Merril, C.R., *Analyt. Biochem*, **1988**, 173, 424.
12. Laemmli, U.K., *Nature*, **1970**, 227, 680.
13. Dai, Y., Whittal, R.M., Li, L., *Analyt. Chem.*, **1996**, 68, 2494.
14. Whittal, R.M., Li, L., *Analyt. Chem.*, **1995**, 67, 1950.
15. Brown, W.M., Saunders, N.R., Møllgård, K., Dziegiewska, K.M., *BioEssays*, **1992**, 11, 749.
16. Demetriou, M., Binkert, C., Sukhu, B., Tenebaum, H.C., *J. Biol. Chem.*, **1996**, 271, 12755.
17. Nie, Z., *Am. J. Physiol.*, **1992**, 263 (3 Pt 1) C551.
18. Grinnell, F., *Int. Rev. Cytol.*, **1978**, 53, 65.
19. Imamoto, G.K., Winterton, L.C., Stoker, R.S., Van Wagner, R.A., Andrade, J.D., Moser, D.F., *J. Colloid Interface Sci.*, **1985**, 106, 459.
20. Vroman, L., Adams, A.L., **1986**, 111, 391.
21. Liu, F., Zhou, M., Zhang, F., *Applied Radiation & Isotopes*, **1998**, 49, 67.
22. Derhami, K., Wolfaardt, J., Wennerberg, A. and Scott, P.G., *J. Biomed. Materials Res.* In Press (2000).
23. Kanagaraja, S., Lundstrom, I., Nygren, H., Tengvall, P., *Biomaterials*, **1996**, 17, 2225.
24. Dawes, K.E., Gray, A.J., Laurent, J.G., *Eur. J. Cell Biol.*, **1990**, 61, 126.
25. Hannan, G.N., Reilly, W., *Exp. Cell Res.*, **1988**, 178, 343.
26. Schinke, T., Amendt, C., Trindl, A., Pöscke, O., Müller-Esterl, W., Jahen-Dechent, W., *J. Biol. Chem.*, **1996**, 271, 20789.

27. Ohnishi, T., Nakamura, O., Ozawa, M., Arakaki, N., Muramatsu, T., Daikuhara, Y., *J. Bone Min. Res.*, **1993**, 8, 367.

28. Gersten, D.M., Hearing, V.J., *Med. Hypoth.*, **1989**, 30, 135.

29. Mizejewski, G.J., *Proc. Soc. Exp. Biol. Med.*, **1997**, 25, 353.

30. Meazzini, M.C., Toma, C.D., Schaffer, J.L., Gray, M.L., Gerstenfeld, L.C., *J. Ortho. Res.* **1998**, 16, 170.

31. Choquet, D., Felsenfeld, D.P., Sheetz, M.P., *Cell,* **1997**, 88, 39.

Chapter 5

Conclusions and Future Work

The combination of electrophoretic separation with mass spectrometric analysis provides a very powerful tool for protein analysis. Information on molecular weight, peptide masses from an enzymatic digestion, as well as fragment ion masses obtained from mass spectrometry and tandem mass spectrometry can be used to successfully determine/confirm the identity or modification of a protein.

In Chapter 2, we report a passive elution method for the analysis of proteins from copper-stained/destained gels using a minimum number of steps. This method minimizes SDS interference in MALDI analysis, thus providing the opportunity of discerning other factors causing peak broadening from the SDS effect. With the use of this gel-extraction method and the time-lag focusing MALDI TOF mass spectrometer, accurate molecular weight determination for SDS-PAGE separated proteins is possible for some proteins, particularly those containing no cysteine residues. However, the accuracy may be greatly reduced as a result of gel-induced protein modifications. In these cases, the molecular weight of the unmodified species cannot be determined with a high degree of accuracy without *apriori* knowledge of the number of modifications. However, in such cases, a reasonable estimate of the molecular weight can be made, and such a "window" may still be a useful complement to tryptic digest data for protein identification. This work clearly illustrates that covalent protein modifications, not the SDS detergent, limit the accuracy of protein molecular weight determination by MALDI TOF MS. It is further shown that the use of a short gel-separation time and pre-electrophoresis with an acrylamide scavenger did not reduce the extent of protein modifications. Finally, the detection sensitivity of the passive extraction/MALDI

method is demonstrated to be in the low pmol for several proteins with molecular weights up to 67,000.

In Chapter 3, three proteins separated by gel electrophoresis were identified or confirmed by combining peptide mass mapping and tandem mass spectrometry. There are a number of approaches available for protein identification, among which peptide mass mapping is always the first choice because of the advantages such as ease of manipulation and high through put. In ideal cases, peptide mass mapping alone can provide confident identification results, as shown in Section 3.3.1. However, more often than not, a peptide mass mapping experiment generates an ambiguous identification result with low sequence coverage. In this case, sequence-specific fragment ion information generated by PSD and/or CID, along with the **MS-Tag** program, can significantly increase the resulting confidence level, as illustrated in Section 3.3.2. When an acceptable, but not 100%, confidence level is reached by peptide mass mapping, it is always constructive to have some primary structure information as supporting evidence, especially when the information can be gathered from different segments. Although tandem mass spectra alone may not be sufficient to piece together the whole protein, they are still a good complement to peptide mass mapping, as demonstrated in Section 3.3.3.

It is worthy to note that the crushing-gel/in-gel digestion methodology, introduced in this study, is a simple, efficient and time-effective protein handling technique for identification of real world protein samples.

In Chapter 4, human skin fibroblast proteins were separated by 2-D gel electrophoresis. After separation, the identification of the proteins was attempted employing techniques which included peptide mass mapping, 2-D database searching, immunoblotting and the comparison with a standard in the gel. The identified proteins indicated whether or not there were any changes in cells, which were in contact with

different implant materials. Therefore examination of changes in the entire proteome of cells in response to different growth substrates was possible. This investigation of biocompatibility is valuable in the search for specific proteins involved in biological response to an implant material.

Although the two mass spectrometry techniques, MALDI and NanoESI, have the ability to analyze mixtures of analytes, ion suppression is still an issue. Coverage of peptide masses in a mass spectrum is limited by the difference of the peptides on their abundances caused by enzymatic digestion, the ionization efficiency of one peptide versus another, and other factors including matrix and solvent effects in the MALDI process and limited surface to charge ratio for ion generation in the NanoESI process. Therefore, a separation step prior to the mass analysis might be a solution to these problems. On-line capillary LC interfaced ESI has the advantages of low sample consumption and no need for clean-up, making it particularly suited to the gel-separated biological samples. With the ability of auto MS/MS, a better coverage of peptide masses as well as tandem mass spectrometry information are expected to be obtained at the same time.