12-1-2006

Proteome analysis of pseudomonas putida KT2440 using 2D gel electrophoresis and LC/ESI-Q-TOF mass spectrometry

Archana Pandey

Proteome analysis of *Pseudomonas putida* KT2440 using 2D Gel Electrophoresis and LC/ESI-Q-TOF Mass Spectrometry

A thesis submitted in partial fulfillment of the requirements for the degree of Masters of Science in Chemistry

Approved: _____
Thesis Advisor

_____
Department Head

Archana Pandey
Department of Chemistry
Rochester Institute of Technology
Rochester, NY 14623
December 2006

Proteome analysis of *Pseudomonas putida* KT2440 using 2D Gel Electrophoresis and LC/ESI-Q-TOF Mass Spectrometry

Date: _____

Signature: _____

TABLE OF CONTENT

Acknowledgements

"Success is never ending and failure is never final". Though this statement can be made for any facet of life, it fits most appropriately to the area of research and education where a successful experiment is always a beginning and a failure is just a part of the process in the cycle of learning. This enduring spirit of research and education was the key guiding factor to shape up my research work.

I would like to take this opportunity to thank the people who helped me in reinforcing this spirit and were always there to support, motivate, guide and help me at every stage of my work. Without their support and encouragement, this work would not have been possible. I would like to thank my advisors Dr. Paul Craig and Dr Laura Tubbs for all the guidance, insights, learning and support to complete my work. Thank you for building the right foundation for my work. It has indeed been a tremendous learning experience for me working with you. I would like to thank Dr. Paul Rosenberg and Dr. Terence Morrill for all the motivation, support and encouragement.

I would like to thank Dr Thomas Kim for his suggestions and for allowing me to use his lab, Tom Allston for the help in keeping the mass spectrometer operational and Dave Lake for always helping me with what I needed for my work, Brenda for her constant support and help during my stay at RIT and special thanks to Robert Saccente (Kodak) for helping me calibrate the mass spectrometer.

I would like to thank my colleagues and friends who were always there to help and cheer me up during the course of my work. Thanks for being such wonderful friends. Thank you MaryEd, Jamie Lou, Rachel, Julia, Mark, Peter and Andrew for making me smile everyday. Thanks for the friendship I developed over these 2 years.

Special thanks to Jamie Lou who was there to help me when I needed the most and I couldn't have completed this project without her help.

Lastly; I would like to thank my daughter Snehal, my husband Atul, my parents and my family members for all their love, their constant support and understanding and for making me believe in my potential.

Common Abbreviations

- **2DE : 2-Dimensional Electrophoresis**
- **2D-PAGE: 2-Dimensional Polyacrylamide Gel Electrophoresis**
- **ACN – Acetonitrile**
- **amu – atomic mass unit**
- **API – Atmospheric Pressure Ionization**
- **AUFS - Absorbance Unit Full Scale**
- **BSA – Bovine Serum Albumin**
- **CHAPS – 3-[(3-Cholamidopropyl)-dimethyl-ammonio]-1-proane sulfonate**
- **Da – Daltons**
- **DTT – Dithiothreitol**
- **ES – Electrospray**
- **ESI-Q-TOF - Electrospray Ionization-Quadrupole-Time of Flight**
- **GC – Gas Chromatography**
- **HPLC – High Performance Liquid Chromatography**
- **IAA – Iodoacetamide**
- **IEF – Isoelectric Focusing**
- **IPG – Immobilized pH Gradient**
- **LC – Liquid Chromatography**
- **LC/MS- Liquid Chromatography coupled to Mass Spectrometer**
- **LPO – Lactoperoxidase**
- **MALDI – Matrix Assisted Laser Desorption Ionization**
- **Mb – Myoglobin**

- **MS – Mass Spectrometry**
- **MW – Molecular Weight**
- **m/z – mass-to-charge ratio**
- **ORF - Open Reading Frames**
- **PAGE – Poly Acrylamide Gel Electrophoresis**
- **pI – Isoelectric Point**
- **PMF – Peptide Mass Fingerprinting**
- ***P. putida – Pseudomonas putida***
- **RP-HPLC – Reverse Phase High Performance Liquid Chromatography**
- **SDS – Sodium Dodecyl Sulfate**
- **SOP – Standard Operating Procedure**
- **TFA – Trifluoroacetic acid**
- **THF – Tetrahydrofuran**
- **TIC – Total Ion Chromatogram**
- **TMS – Tandem Mass Spectrometry**
- **UV-Vis – Ultraviolet-Visible**

List of Figures

Page No.

List of Tables

Page No.

Abstract

The proteome can be defined as the complete set of global protein expression by an organism at any given time. It is this gene expression rather than genome itself that is responsible for most of the reactions taking place within a cell. Tools such as two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) and mass spectroscopy (MS) are often utilized for the analysis and identification of the proteins that constitute the proteome. This thesis provides a general overview of proteomics and describes the underlying processes, mechanisms and technologies associated with modern proteomic analysis with specific emphasis on the use of electrospray ionization time of flight mass spectrometry. The use of MS/MS analysis of ESI-Q-TOF data for the identification of proteins is also explained in detail. Other chromatographic approaches such as high performance liquid chromatography (HPLC), where separation is based on the polarity of the mobile phase, and gel filtration chromatography systems, where separation is based on the molecular size, are also described in detail.

The thesis work focuses on identifying the proteomic signature in the soil bacterium *Pseudomonas putida* using these analytical tools, particularly LC/ESI-Q-TOF MS (Electrospray Ionization Time of Flight Mass Spectrometer).

Protein digest analysis was done with lactoperoxidase (LPO), lysozyme, and ribonuclease on HPLC. Whole protein and tryptic peptides were analyzed on the system. Peaks corresponding to the whole proteins as well as the peptides generated from the tryptic digestion proteins were observed.

The goal of the entire project is to isolate proteins from the bacteria *Pseudomonas Putida* strain KT2440 using biochemical techniques to first separate them using 2-D PAGE, subsequently performing in-gel tryptic digest, and finally identifying individual protein spots on the gel using LC / ESI-Q-TOF MS and protein databases.

Introduction

Pseudomonas putida is a metabolically versatile, non pathogenic, gram negative soil bacterium which has the ability to degrade and metabolize various natural and synthetic aromatic compounds and utilize them as the primary source of carbon and energy. *P. putida* also plays a vital role in the elimination of organic waste in polluted soils and is widely used for both genetic and biotechnological applications [1, 2]. The genome of the bacteria is fully sequenced. The 6.18 MB genome of KT2440 strain of *P. putida* is a single circular chromosome with 5427 ORF (open reading frames) and 625 known proteins [20]. An open reading frame (ORF) is a portion of DNA that has the potential of producing proteins.

This thesis work is focused on identifying the proteomic signature of the soil bacterium *Pseudomonas putida* KT2440 cultured on succinate using 2D polyacrylamide gel electrophoresis and mass spectrometry. The proteins were separated by 2D-PAGE followed by in-gel digestion with trypsin. Peptide fragments were analyzed by LC/ESI-Q-TOF MS and identified by peptide mass fingerprinting using online databases. The results were confirmed by tandem mass spectrometry.

This introduction section of the thesis will cover the following subsections:

1. Proteomics Overview
2. Introduction to 2D Gel Electrophoresis
3. High Performance Liquid Chromatography (HPLC)
4. Gel Filtration Chromatography
5. ESI-Q-TOF Mass Spectrometry (LC/MS)
6. Overview and Specific aims

Proteomics

The genome is the complete set of genes (total DNA content) carried by an organism while the proteome is the complete set of all the proteins expressed by a cell at any given time. Since the proteome of an organism is the protein complement of its genome we define proteomics as the study of gene expression at a functional level. In short, it is the protein analog of genomics.

Genes are the unit of heredity that control the development of an individual and provide information for making the proteins that form each of us and our individual characteristics. Although genes have traditionally been the primary area of focus at the molecular level, it is the proteins expressed that do most of the work in carrying out cellular processes and perform most life functions. For this reason, many researchers are turning their attention from genomics to proteomics to study the cellular and molecular basis of health, disease, and other life functions [35].

Information about mechanisms of disease and drug targets, as well as aging can not be obtained by simply studying genes; rather this information can be obtained only from the study of proteins [15]. From the perspective of drug discovery, there is strong belief that proteomics will help elucidate mechanisms of disease that will enable researchers to identify novel drug targets; this approach has already achieved tremendous success with the knowledge and techniques available [35, 50].

The science of proteomics is intrinsically linked to genomics since the information about proteins can be derived from the information contained in the genome of an organism. However, while the genome is static for a particular cell, the proteome which is derived from the genome varies greatly depending upon the cell type as well as its functional state. For example, it is quite amazing that a single insect species can have

one genome but express the two radically different proteomes that are seen in the caterpillar and butterfly phases of its life cycle [23].

The term “proteomics” was coined in 1995 by Wilkins [15, 50] and is a very modern concept defined as the large scale study and global analysis of all the proteins expressed by any cell or tissue under any given conditions, with a focus on their structure and function and the way they work and interact with each other and are modified within the cell or tissue. In short, proteomics is the global and systemic analysis of complete set of proteins expressed by any organism [51].

Proteomics leads us from sequence of a protein to its biological function. It is the science that studies the proteins in general and also specific changes in their expression that result from various environmental conditions and/or disorders. The functional approach of proteomics helps us understand which set of protein is responsible for a given phenotype. While we expect different organisms to express different proteins, the cellular expression of enzymes for the same species can vary widely under different environmental conditions as well. The specific set of proteins expressed by an organism under specific conditions is called the proteomic signature.

There are many different aspects of proteomics. One is the structural approach which focuses on the molecular structure of proteins and relates this information to the database of identified genes. On the other hand, if we are seeking the set of proteins responsible for a certain biological effect, we are addressing the functional aspect of proteomics [52]. Protein-protein interactions also constitute a major aspect of proteomics research.

2D Gel Electrophoresis

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) is the most common technique used in proteomics. For over 20 years, it has been the technique of choice for analyzing the protein composition of cells, tissues and fluids as well as for studying the changes in global patterns of gene expression. The biggest advantage of this technique is its compatibility with mass spectrometry which can then be used to identify the proteins.

The technique provides high resolution separation of the individual proteins of a proteome. 2-D gels can separate hundreds of proteins on a single gel [12].

The steps shown in Figure 1 describe the process of moving from a cellular homogenate to identification of a proteomic signature; clearly 2DE is central to this process this proteomic process.

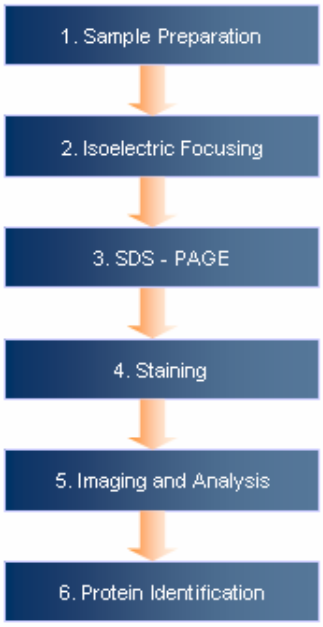


Figure 1. This flow chart describes the proteomic process of moving from a cellular homogenate to identification of a proteomic signature

In 2DE, proteins are separated by two distinct properties: isoelectric point (pI) and molecular weight (MW). The isoelectric point is the pH at which the protein has no net charge. In the first dimension, which is known as isoelectric focusing (IEF), proteins are applied to a mylar strip containing an immobilized pH gradient (called IPG strips) and the proteins migrate to their isoelectric point when an electric field is applied. As proteins migrate to their respective isoelectric points, they pick up or lose protons. As they continue to migrate, the net charge on the proteins and their mobility decreases and

eventually they come to a point where their net charge is zero and they stop moving. This is their isoelectric point. IPG strips are available in different pH ranges (examples: pH 3-10, 4-7, 3-6, 5-8, 7-10) and various lengths (7, 11 and 17cm).

Proteins separated by IEF are further separated orthogonally by SDS-PAGE (Sodium dodecyl sulfate-polyacrylamide gel electrophoresis). The transition from first to second dimension involves 2 steps. The first step involves the equilibration of IPG strips in a buffer containing SDS, a detergent that imparts negative charge to the protein that is proportional to the mass of a protein.

In addition, the proteins that have been separated by IEF are also reduced and alkylated using DTT (dithiothreitol) and iodoacetamide (IAA). DTT reduces the disulfide bonds (R-S-S-R) found in proteins. The resulting –SH groups are then alkylated with iodoacetamide, which attaches an alkyl group to the –SH (R-S-CH_2-$CONH_2$) to prevent reformation of the disulfide bonds as shown in figure below (Figure 2).

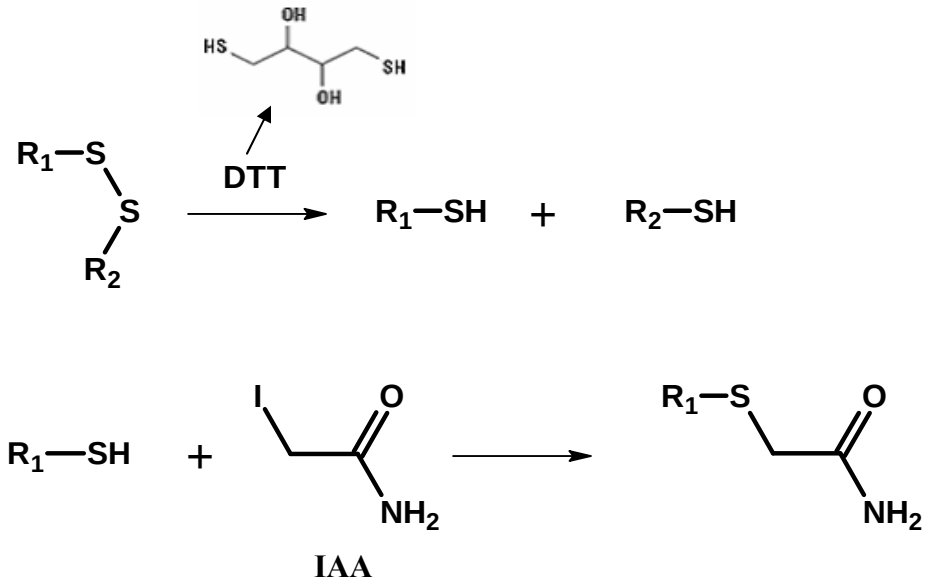


Figure 2. Reduction and alkylation step in the process of 2D gel electrophoresis

The treated IPG strips are then embedded in the top of an SDS-PAGE gel. An electric current is then applied to the system, causing the proteins to migrate from the IPG strip into the SDS-PAGE gel, where they are separated by size. Proteins move through the porous gel with small molecules moving more rapidly than the larger ones.

Polyacrylamide is a cross-linked polymer of acrylamide (Figure 3). Acrylamide is the material used for preparing gels for electrophoresis and separates proteins by their molecular size. Ammonium persulfate is the polymerizing agent which when added to acrylamide gel mixed with bisacrylamide, forms a network of cross-linked polymer. The pore size in the gel is inversely related to the amount of acrylamide used.

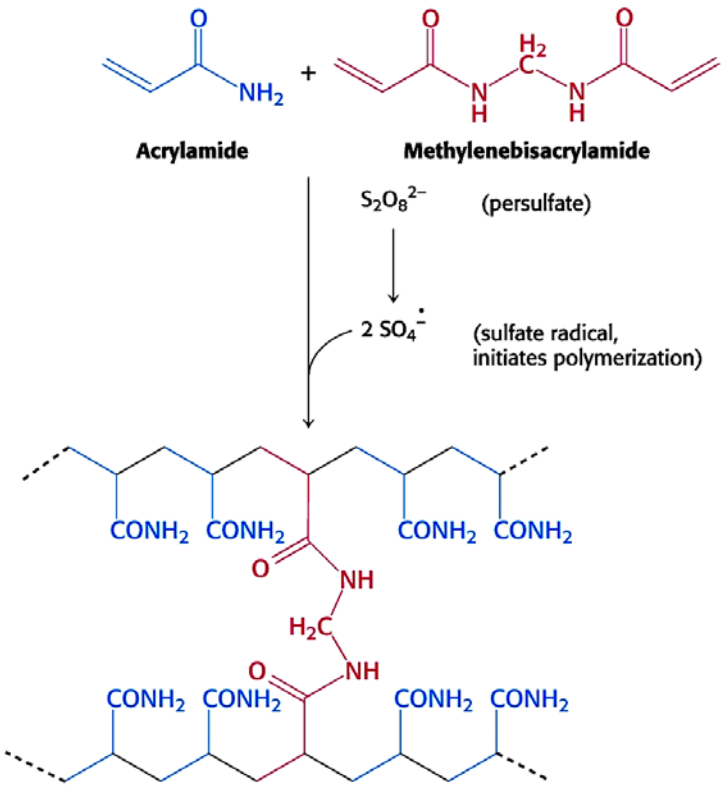


Figure 3. Polyacrylamide, a cross-linked polymer of acrylamide [16]

Gels containing a high percentage of acrylamide are used to resolve small proteins, whereas low percentage acrylamide gels are used to resolve large protein. For example the gel containing 7% acrylamide will have larger pores in the gel and will separate larger proteins than a 12% polyacrylamide gel which would have small pores and thus would separate small size proteins [32, 33]. Typical polyacrylamide concentrations range from 5% to 25%. Gels can also be prepared with a gradient in polyacrylamide concentration (e.g. 8-16%), which broadens the MW range for protein separation on a single gel.

High Performance Liquid Chromatography (HPLC)

Chromatography is a technique used for separating mixtures of compounds. In this technique the separation is based on partitioning of the molecules between the mobile phase and the stationary phase. The mobile phase can be liquid (liquid chromatography) or gas (gas chromatography). Fluid entering the column is called eluent and the fluid emerging from the end of the column is called the eluate. The process of solvent flow through the chromatography column is called elution.

In high performance liquid chromatography (HPLC), the solvent is pumped at high pressure through a column containing stationary phase particles (with diameters of about 3-10 μm). The mixture to be separated and analyzed is forced through a column packed with the stationary phase by the solvent at such a high pressure that it has less time to diffuse in the column resulting in increased resolution in the chromatogram [18].

The column is tubular, packed with silica particles and is available in various dimensions. The stationary phase is most commonly a viscous liquid chemically bonded to the side of the capillary tube or surface of the solid particles packed in the column. Since peptides are small and require a more hydrophobic longer chain length to be

captured, C8 and C18 columns are most often used to separate mixtures of peptides or small molecules [36].

The HPLC instrument consists of a sample / solvent delivery system, pressurized mobile phase, flow controller, sample injection valve, mixing valve, high pressure column, ultraviolet-visible spectrometer as detector, and a recorder / computer to control the system and display results (Figure 4).

In sample delivery systems there are two pumps which deliver the mobile phase to the column. A pump is required for smooth flow of solvents. The mixing valve on the other hand mixes the mobile phases that are delivered from the pump and moves the solvents to the column [18, 27].

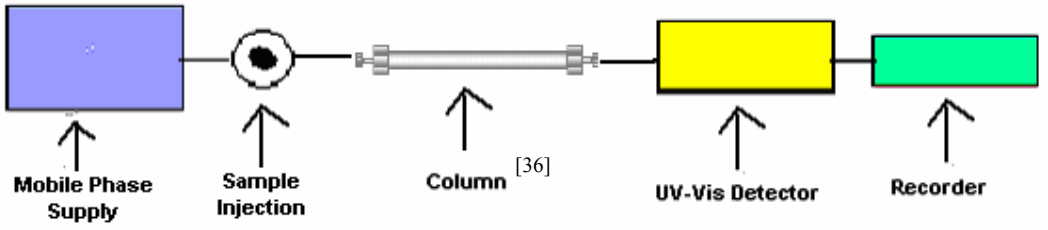


Figure 4. Schematic showing the flow of sample through different components of HPLC [36]

On the way to the column, there is a sample injection valve which allows rapid and precise sample introduction. In most systems there is an auto-injector for this purpose. The sample introduction method is based on sampling loops which are an integral part of liquid chromatography equipment; injectors often have interchangeable small loops, each of which hold a fixed volume [18].

The liquid chromatography column is housed in a chamber where temperature is controlled; the column efficiency increases at elevated temperature. After the components

of a mixture are separated in a chromatography column, solutes eluting from the column are detected so that they could be identified and quantified. Most LC detectors do not identify the compounds, but rather indicate that something is emerging from the column which is visualized in the form of peaks. An ultraviolet detector is the most common HPLC detector as many solutes absorb ultraviolet light. Simple systems commonly employ an intense 254nm emission mercury lamp. Refractive index detectors are the most universal. Other detectors include fluorescence detectors as well as a mass spectrometer. The fraction collector collects time or volume fractions of the eluate that need to be analyzed. The computer which is attached to the pump controls the system and displays results. A chromatogram is a graph/plot showing the response of the detector as a function of elution time.

Chromatography in which the stationary phase is polar and the mobile phase is relatively less polar is referred to as 'normal-phase chromatography'. Normal phase chromatography thus uses a less polar or non- polar solvent and the eluent strength increases as the polarity of the solvent increases. Eluent strength is a measure of the solvent adsorption energy. Strong mobile phase compositions make the solute elute faster. On the other hand chromatography in which the mobile phase is polar than the stationary phase is called 'reversed-phase chromatography' [18]. Reversed-phased chromatography employs a non polar or weakly polar stationary phase (often a hydrocarbon) and a polar mobile phase, which is a mixture of water and an organic solvent such as methanol, tetrahydrofuran or acetonitrile. Solvent strength increases with the percent of the organic component. Its major applications include separation of non-ionic and polar compounds. Most separations of the organic compounds and peptides are done on reverse phased columns.

If the solvent has a fixed composition (e.g., 35% water and 65% methanol), the process is called isocratic elution. In gradient elution, however, the eluent strength is increased during chromatography by increasing the percent of the organic solvent (e.g., 10 – 90% methanol in 30 minutes). In reversed phase chromatography, the most polar components elute first; increasing the polarity of mobile phase increases elution time. For reversed phase chromatography of peptides, acid is sometimes added to these solvents to improve the chromatographic peak shape and to provide a source of protons for LC/MS. The most commonly used acids are acetic acid, formic acid and trifluoroacetic acid [36].

In reversed phase HPLC, peptides can be separated by running a linear gradient of the organic solvent. The majority of small peptides (about 10-30 amino acid long) elute by the time the gradient reaches 30- 35 % organic [36].

Gel Filtration Chromatography

In gel filtration chromatography, molecules are separated by size in a column packed with porous gel beads. Unlike other chromatography there is no attractive interaction between solute and the stationary phase. Instead, the liquid mobile phase passes through the porous gel. The pores are small enough to exclude large molecules but not the small ones. The large proteins pass without entering the pores. The small proteins take longer to pass through the column because they enter the gel and thus flow through large volume before leaving the column (Figure 5). These small proteins have access to the mobile phase inside the beads as well as the mobile phase between beads and elute last in a gel filtration separation. In short, in molecular exclusion chromatography, the fraction of the stationary phase volume available to the solute decreases as the size of the solute increases.

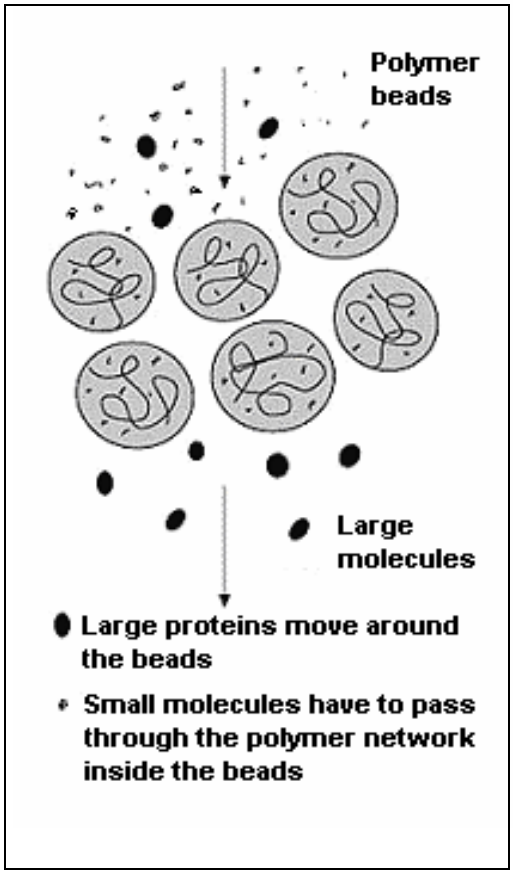


Figure 5. Gel filtration process [34]

The basic components of gel filtration experiments are a chromatography column containing the porous gel beads and the elution buffer. The matrix is the separation media which is the stationary phase of chromatography, and the elution buffer which is the mobile phase of chromatography, flows through the stationary phase. The packed matrix is called the “bed” and the volume the bed occupies is called the “bed volume” [37].

LC/ESI-Q-TOF Mass Spectrometer (LC/MS)

Mass spectrometry (MS) is a powerful technique which is used to determine the mass of a molecule. An MS can be used as the detector in gas and liquid chromatography. For comprehensive analysis of proteins, LC-ESI-Q-TOF MS (Liquid Chromatography Electrospray Ionization Quadrupole Time of Flight Mass Spectrometry) provides the dynamic range needed to analyze a mixture of peptides. Dynamic range refers to the wide range of mass accuracy that ESI-Q-TOF MS provides. The mass accuracy in these systems is stable, even with significant variations in sample concentration. A liquid chromatography system coupled to mass spectrometer is currently the most widely used analytical method for the separation and analysis of proteins and peptides.

The main advantage of mass spectrometers are that they excel in the determination of molecular mass of intact proteins and one of the most common application of LC/ESI-Q-TOF MS is the identification of peptides produced from enzymatically digested proteins. The protein components of any cell or tissue can be identified following enzymatic digestion of the protein. The peptides are fractionated and purified by reversed-phase high-performance liquid chromatography (RP-HPLC) and analyzed by in-line coupling to mass spectrometry (ESI-Q-TOF MS) capable of producing fragment ion spectra. These spectra are then analyzed by comparison to different databases. Instrumental parameters for the subsequent LC/ESI-Q-TOF analysis of these protein mixtures are optimized to generate the maximum number of quality MS spectra from each sample.

In recent years, liquid chromatography systems have been competitive with the 2D-electrophoresis based techniques for protein separation. Since the development and

commercialization of mass spectrometers, electrospray ionization (ESI) and matrix assisted laser desorption ionization (MALDI) are the two dominant methods for introducing samples from a liquid chromatography system into the mass spectrometer. It is commonly accepted that MALDI instruments are ideal for very rapid and accurate determination of peptide mass fingerprinting (PMF) spectra resulting from proteins that have been separated by 2D electrophoresis. A peptide mass fingerprint is obtained when a protein is enzymatically digested and the fragments are analyzed by MS. The resulting fragments sizes (taken from the m/z values produced by the MALDI MS system) can then be compared to expected fragments from known proteins. This approach can be very effective in situations where the sequence of all the proteins in the organism are known, typically as the result of complete sequencing of the genome of the target organism. However, peptide mass fingerprints do not provide any direct evidence for the sequence of the peptides themselves. Therefore, this approach is not effective when dealing with proteins from organisms whose genomes have not yet been sequenced.

MALDI has a high throughput; it is sensitive, specific and forgiving of the presence of the salts and other contaminants that may interfere with the production of ions. On the other hand, ESI coupled instruments can be operated in MS/MS mode and, therefore, give more structurally related information, including the sequence of the peptides being studied. ESI is one of the most sensitive analytical techniques; however the sensitivity decreases in the presence of salts and nonvolatile buffers.

Workings of the mass spectrometer

Mass spectrometers measure mass to charge ratios of ions produced from the molecules which are introduced into the ionization source of the instrument.

To understand the working of MS, it is important to understand that molecules need to be charged and vaporized before they enter the ESI-Q-TOF MS. Once they are in the gas phase, the molecules are ionized by addition of protons (H^+) to create positively charged ions. Cations are accelerated by an electric field to impart the same kinetic energy to each of the ions.

For ions traveling with the same kinetic energy, velocity is inversely proportional to the m/z ratio. Therefore lighter ions reach the detector of the mass spectrometer before heavier ions.

Parts of ESI-Q-TOF MS:

The ESI-Q-TOF MS consists of 3 basic elements:

1. Ionization Source
2. Mass analyzer
3. Detector

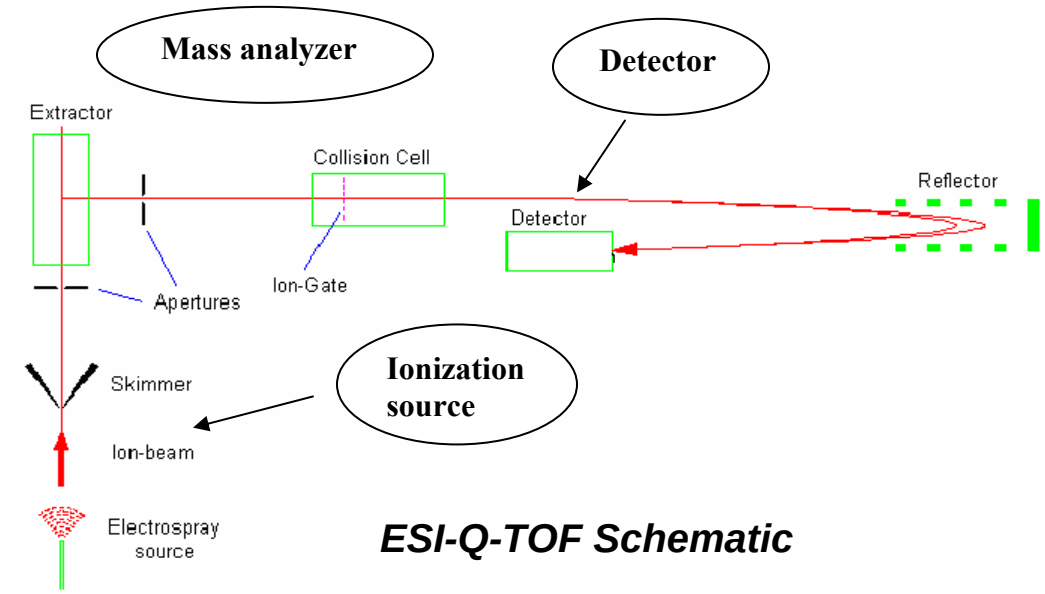


Figure 6. ESI-Q-TOF Schematic [38]

As shown in figure 6, a sample is introduced into the ionization source of the instrument where the gas phase molecules are converted into ions. The mass analyzer then separates these ions on the basis of their mass to charge ratio. The separated ions reach the detector where relative abundance of these charged species is measured [39].

Ionization source

A protein sample can be directly introduced into the ionization source of the instrument or can be introduced via a chromatographic method. The former method involves manual loading of the sample into the ionization source of the instrument which is a very slow and tedious process. To avoid this, HPLC is directly coupled to ESI-Q-TOF MS which automatically delivers the sample into the ionization source of mass spectrometer and also serves to separate the peptides from the sample [39].

Electrospray Ionization Process Overview

The Electrospray ionization process involves spraying a stream of liquid sample under strong electric field into a chamber where ions are produced. The sample is accompanied by flow of nitrogen which results in the formation of an aerosol of charged droplets through the metal spray tip. High voltage is applied to the spray tip which creates an electric field between the spray tip and the inlet to the MS. The sample stream is then dispersed into charge droplets at atmospheric pressure within the spray chamber.

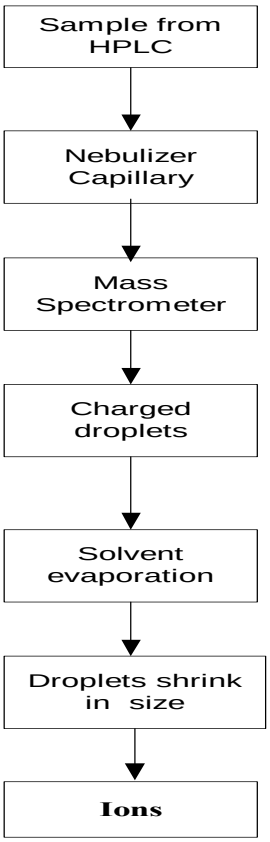


Figure 7. Ionization process

The aerosol is positively charged if the potential of the capillary is positive with respect to the inlet of the ESI-Q-TOF MS and is negatively charged if the potential of the capillary is negative with respect to the inlet of the ESI-Q-TOF MS [18]. In positive ionization mode a small amount of formic or acetic acid is added to help protonation of the sample molecule, whereas in negative ionization mode a trace of ammonia solution or another volatile amine is added to aid in deprotonation of the sample molecule.

Proteins and peptides are usually analyzed under positive ionization conditions [39]. The flowchart shown in figure 7 describes the ionization process of sample moving from HPLC to formation of ions in ESI-Q-TOF MS.

Positive ions entering the capillary move into the skimmer cone where there is a difference in potential of -60 V. The N_2 gas directs the spray emerging from the tip of the capillary into the ESI-Q-TOF MS. The ions here collide with N_2 and break into fragments ions. Charged liquid exiting the capillary forms a cone and then forms a thin filament of liquid which finally breaks into a spray of fine droplets [18].

The droplets shrink to a diameter of 1micrometer (μm) by solvent evaporation. As the charged droplets containing the analytes and the solvent molecules evaporate, ions are emitted from them (Figure 8).The charged droplets evaporate and continue to shrink in size until the repulsive force between like charges overcomes the surface tension of the charged droplets. The solvent evaporates leaving the charged peptide samples as gaseous phase ions [39].

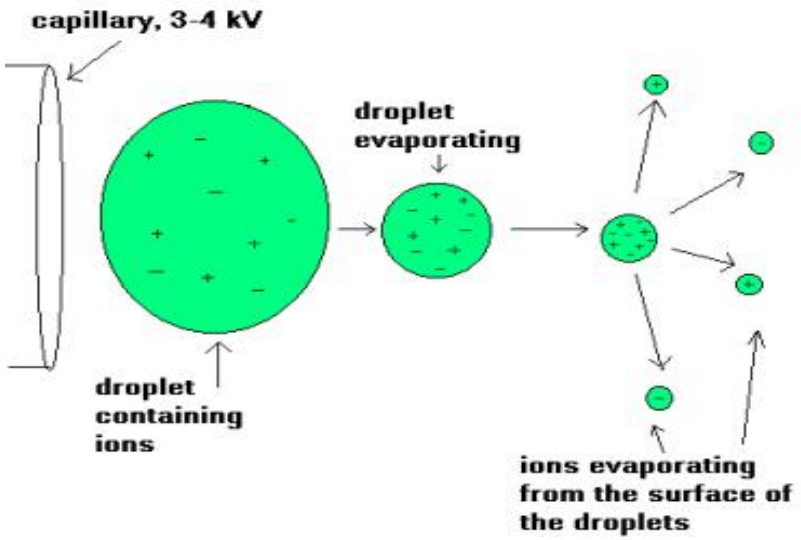


Figure 8. Ion formation [39]

The ions end up in the vapor phase and enter the analyzer region. In short, ESI results in the formation of charged droplets from a stream of solvent which ultimately yield vapor phase ions. The ions accelerated from the same point at the same time at the same potential are separated based on their mass to charge (m/z) ratio.

Mass Analyzer:

The second important component of the mass spectrometer is the mass analyzer which resolves ions on the basis of their mass to charge ratio. There are three different types of mass analyzer: Quadrupole, ion trap and TOF mass analyzer. A TOF mass analyzer measures the m/z ratio of an ion by determining the time required for the ion to traverse the length of the flight tube. It measures how long it takes for each kind of ion to reach a fixed distance to the detector. The mass analyzer separates ions with same kinetic energy (KE) but different m/z which results in lighter ions traveling faster and reaching the detector before the heavier ions. The collision cell is the place where samples collide with inert gas and results in the fragmentation of the ions [27].

A TOF mass analyzer consists of three regions:

a. Source region
b. Drift region
c. Reflectron

A high voltage (around 4-5 KV) is applied to the tip of the capillary located within the ionization source of the mass spectrometer where the molecules in gas phase are converted into ions. Ions are accelerated from the source region and expelled into the drift region of the mass analyzer where there is no magnetic or electric field and hence there is no acceleration. From the drift region, ions enter the reflectron located at the end of the field free drift zone which consists of a series of rings held at increasing positive potential. The reflectron collects ions, realigns them and redirects them to the detector [40]. Here, the ions experience the electric field where ions with the same m/z value are refocused. Ions entering the reflectron are slowed down and turned around and reflected back towards the detector as shown in figure below (Figure 9).

The more KE an ion has when it enters the reflectron, the further it penetrates before it is turned around and reaches the detector. The reflectron almost doubles the time of flight path, increase the resolution and thereby improving mass accuracy [18].

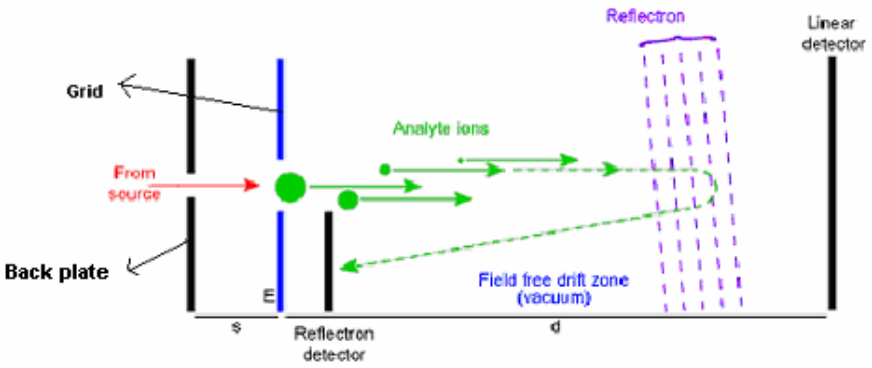


Figure 9. Schematic of a Time-of-Flight mass spectrometer operating in Reflectron Mode [40]

Detector

The third component is the detector. The ESI-Q-TOF MS has a micro-channel plate based ion detector which monitors the ion current, amplifies it and transmits a signal to the data system where it is recorded in the form of a mass spectrum. The m/z values of the ions are plotted against their intensities to show the number of components in the sample, the molecular weight (MW) of each component and the relative abundance of various components in the sample [41].

Research Overview and Specific Aims

The goal of the project is to identify the proteomic signature of the soil bacterium *Pseudomonas putida*. Prior to introduction of the protein sample from this bacterium into the mass spectrometer, sample complexity is reduced by separating the protein by 2D gel electrophoresis followed by excision of individual protein spots from the gel and in-gel digestion with trypsin. Digestion of complex proteins results in formation of small

peptides which are easily ionized in the mass spectrometer and measured based on the m/z value of the ion.

Identification of the proteins from an organism proceeds in a series of steps.

Sample preparation begins when agar plates are streaked with a source culture of *P. putida.* This is followed by performing growth curve analysis of the bacteria in which the lag, log, stationary and death phase of the cell cycle are determined. Cells are then harvested by centrifugation at mid log phase to isolate the protein from a culture that contains the maximum number of healthy cells

The cells are then sonicated to release soluble proteins. Protein concentration is determined using a Bradford protein assay (see Appendix C). This complex protein mixture is resolved by 2-D gel electrophoresis. IPG strips are rehydrated with a protein sample in the presence of a denaturing buffer. The proteins are then separated by isoelectric focusing followed by SDS-PAGE. Protein spots on the resulting gels are visualized by staining with Commassie blue dye. These protein spots are then analyzed using specialized software.

Protein spots of interest are selected and digested using trypsin, a protease. The in-gel proteolytic digestion of the proteins yields peptides. The resulting peptide fragments are fractionated and purified by HPLC which removes any gel contaminants or any other impurities that might interfere with ESI-Q-TOF MS during analysis (i.e. salts, buffers or detergents). Finally, the peptide fragments are analyzed by mass spectroscopy and the mass spectral data obtained from these peptide fragments are compared with the information contained in the databases of the known protein sequence.

Mass spectrometry gives structural information about the proteins. Once the structural information about the proteins is known, databases are utilized to identify the proteins. As was mentioned above, two types of mass spectral data can be obtained for peptide samples. The first, peptide mass fingerprinting (PMF), provides only information about the m/z ratios of peptides in a sample. This method can work in cases where the genome of an organism has been determined. Data obtained by tandem mass spectrometry (TMS) also provides the sequences of the peptides, which can be used to obtain unequivocal proof of the identity of a protein, even if the genome of the source organism is not known

The main focus of this project was to identify the proteomic signature of the soil bacterium *Pseudomonas putida* strain KT2440 cultured on succinate. The proteins were separated by 2D-PAGE followed by in-gel digestion with trypsin. Peptide fragments were analyzed by LC/ESI-Q-TOF MS and identified by peptide mass fingerprinting using online databases. The results were confirmed by tandem mass spectrometry.

Experimental

The following experimental procedure was performed in the laboratory to study the proteomic signature of *Pseudomonas putida* KT2440 grown on succinate.

Isolating protein

Sample preparation includes streaking plates, growing the bacteria, preparing cell culture, centrifugation and sonicating the cells to extract proteins. The Hutner's Growth Media (Appendix A) contains 5mM succinic acid and other minerals required for growth of the bacteria. The streaked plates were incubated at 37°C for 48 hours.

Using sterile techniques, a single colony was picked from the plate and used to inoculate 25-50 mL Hutner's minimal medium as a starter culture. This "starter culture" was then incubated overnight at 30°C in an orbital shaker with an agitation rate of 140 rpm.

After 24 hours of growth the culture was cloudy. A 1% inoculum of the starter culture was transferred to 1 L of Hutner's growth medium in a 2 liter flask which was grown at 30°C until it reached mid-log phase

The 1 liter cell culture was divided equally among four 250 mL centrifuge tubes. Using a Sorvall® RC 5C Plus and the centrifuge SLA-1500 rotor, the cells are spun at 12,000 x gravity (8,200 rpm) at 4°C for 40 minutes. The supernatant was discarded and the pellet was collected in a small pre-weighed 50 mL beaker and the wet weight of the pellet (cell mass) is recorded. If the weight of the pellet was less than 0.5 g, 2 mL of rehydration buffer (Appendix B) was added and if the weight of pellet weighed more than 0.5 g, then 3 mL of the rehydration buffer is added to the beaker containing the pellet. The re-suspended cell solution was placed in a bucket containing crushed ice.

The sonicator (Ultrasonic Heat Systems Mod W-375, Converter Mod C3) was warmed up for about 5 minutes before sonicating and the instrument was set at % Duty cycle 43 – 48 and output control = 3.5. The tip of the sonicator was washed first with deionized water, then acetone and then again with water. The sonicator tip was gently immersed into the beaker containing the protein sample and the pellet was sonicated for 3 minutes, making sure that the sonicator tip did not touch the sides or bottom of the beaker. The sample was then left on ice for 10 minutes. The process was repeated three times which results in a total sonication time of 9 minutes.

The sonicated protein sample was transferred into sterile 2.0 mL disposable microfuge tubes and was then spun down in a microfuge (Biofuge 13 Mod. 7500464/01) at 13,000rpm for 5 hours at 4°C. At the end of 5 hours a tiny black pellet was seen at the bottom of the microfuge tube. The supernatant containing the protein was then transferred to new microfuge tubes without disturbing the unwanted black pellet. The protein was then stored at 4°C.

Determining protein concentration

To find out how much of the sample volume of the extracted protein to be applied to the IPG strips, the concentration of the protein was determined using the Bradford Protein Assay (Appendix C).

First a calibration curve was prepared using bovine serum albumin (BSA) as the standard protein (Table 1 in appendix C), then at least two different dilutions of the sample protein were prepared and the protein concentration was determined by comparison to the calibration curve.

After the concentration of the protein was determined by protein assay, about 2 μg (microgram) of RNAase per mL of protein sample was added to the sample to remove

RNA. The sample was microfuged again for 20 minutes and the supernatant was transferred to another clean microfuge tube.

Rehydration of IPG strips

The protein sample was mixed with rehydration buffer such that the concentration of protein applied to each IPG (Immobilized pH Gradient strip) was around 200 micrograms in rehydration buffer and with a total volume of 185μL.

The entire contents of the tube was then applied to a clean dry well in a rehydration tray along the back corner of the lane leaving about 1cm from each end. The protective plastic backing of the 11cm IPG strip (Bio-Rad, pH 4-7) was removed and the strip was carefully laid over the sample in the well (gel side down), such that the + pH 4-7 was towards the left of the tray and no air bubbles were trapped under the strip (Appendix D).

The strips were rehydrated for about 20-30 minutes at room temperature before overlaying them with ~ 2 – 3 mL mineral oil to prevent evaporation of the sample. The rehydration tray was covered with the tray cover and left overnight for 11-16 hours. As an alternative, samples were also applied by "active rehydration" in the protean IEF cell. The only difference is that the voltage moves the protein across the IPG strip and this rehydration is done at a constant voltage (50V) for 12 hours. IPG strips prepared by active rehydration method demonstrated more spots and less streaking.

Isoelectric Focusing (IEF)

The process involved running a programmed method in a Protean® IEF Cell to focus the proteins on the IPG strips. A clean, dry focusing tray was placed on the lab bench such that the + sign is to the left side. Electrode wicks (Bio-Rad, cat #: 1654071) were placed over wire electrodes in one of the wells and 8μL of nanopure water was

applied to each wick. Each IPG strip was carefully removed from the rehydration tray and held vertically over a filter paper, allowing the mineral oil to drain. The ends of the strips were touched to the filter paper to remove mineral oil; however the strips were not laid on the filter paper.

The IPG strips were placed in the focusing tray with the gel side down. The strips were covered with ~ 2-3 mL of mineral oil. The lid was placed on the tray which was carefully moved into the Protean IEF Cell. The cover of the IEF cell was closed and checked to ensure that the polarities were correct before starting the run. Focusing was done for 16 hours using a linear gradient condition: 6min at 0-300 V, 300 V for 4.5 hours, 300- 4000 V gradient over 5 hours and 4000 V for 6.5 hours

Once the run was complete, the focusing tray was removed from the IEF cell. The strips were removed from the focusing tray and drained of mineral oil. The strips were then transferred into a clean rehydration tray where they were ready for SDS-PAGE. Depending on scheduling, the strips were sometimes stored at -20°C if they could not be used immediately.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE) or 2D Gel Electrophoresis:

10 mL portions of Equilibration buffer stock solution (Appendix E) were placed in two centrifuge tubes labeled “Buffer I” and “Buffer II”. 200mg of DTT (dithiothreitol) was added to the tube labeled buffer I which was shaken to get the DTT into the solution. 250 mg of iodoacetamide was added to the tube labeled buffer II which was shaken until the iodoacetamide dissolved. About 3 – 4 mL of Buffer I was added to each well of the

rehydration tray containing a rehydrated strip, and the rehydration tray was placed on an orbital shaker for 10 minutes at 40 rpm and 30^0C.

While the strip(s) were in the orbital shaker, the time was utilized for the preparation of agarose and running buffer. The bottle of stock agarose (Appendix E) was warmed in the microwave for 30 second increments until it became slightly warm and less viscous. Running buffer was prepared by adding 100 mL of 10X Tris/Glycine/SDS Buffer (Bio-Rad) to 900 mL of nano-pure water in a 1 L graduated cylinder and mixed (as in Appendix E).

After shaking for 10 minutes, the strips were moved into another clean tray. 3-4 mL of buffer II was then added to each tray, and place in the orbital shaker for another 10 minutes at 40 rpm and 30^0C.

Pre-cast SDS-PAGE gels (8-16%) were taken out from the refrigerator and removed from their packaging. The green comb was removed from the top and the white sticker was taken from the bottom of the gel. Wells of pre-cast SDS-PAGE gels were rinsed with nano-pure water.

IPG strips were removed from the shaker and moved into another clean tray. The strips were covered with running buffer and placed in the orbital shaker for ~10minutes at 40 rpm and 30^0C. While the strips were shaking, 1cm wide strips of blotting paper (Bio-Rad, cat. # 1704085) was cut and the well in the pre-cast gel was dried by blotting. The tray was removed from the orbital shaker and the running buffer was drained out.

Using forceps, the IPG strip were carefully placed into the pre-cast SDS-PAGE gel slightly above the well with the gel side facing out and such that the + 4-7 pH sign was on the left side. The strip was pushed gently into the well using forceps. The well was filled to the top with melted agarose using a transfer pipette, and the strip was let to

sit for 5 minutes while the agarose solidified. The process was repeated for the remaining IPG strips, placing each one on a separate SDS-PAGE gel.

The electrophoresis cell was rinsed with distilled water. The pre-cast SDS-PAGE gels were inserted into the electrophoresis cell and running buffer was poured into each cell up to the marked "fill line" making sure that the wires on the green top of the electrophoresis cell are completely submerged in the running buffer, which completes the electric circuit. The top on the electrophoresis cell was placed on the cell. The wires were plugged into an outlet on the power source and the voltage was set to 200 volts, current at 2A and run time at 55 minutes before starting the run .The gel was checked at least every 15 minutes to see the dye-line position on the gel to make sure the sample had not run off the gel. If the dye-line approached the bottom of the gel, the run w stopped immediately.

At the end of the 55 min run, the gel was removed from the plastic casing by keeping the well of the gel on the ridge of the electrophoresis cell lid upside down and gently pressing down until a cracking sound was heard. The pressure on the lid causes the sides to crack open and leave the gel on one side of the plastic casing. The plastic casing can be submerged in distilled water with the gel side facing down towards the water.

Usually the gel fell in the water by itself with gentle prodding. If not, the gel was removed by softly inserting a disposable pipette at the bottom of the gel to ease it out. The gels were rinsed twice in deionized water for 10 minutes with agitation to remove SDS and then stained with sufficient Bio-Safe Commassie blue stain to cover the gels for 90 minutes.

After 90 min, the stain was decanted and the gels were rinsed with deionized water, with agitation, for around 3 minutes. The water was decanted and gel was covered with fresh water for 10 minutes. Rinsing with water was repeated at least three times at

an interval of 10 min. Gels were destained for up to 24 hours by placing them in water after which they are scanned using imaging software. The Coomassie stained protein spots were later excised manually from gels and digested *in situ* with trypsin.

High Performance Liquid chromatography (HPLC)

Tryptic digestion of ribonuclease, lysozyme and lactoperoxidase

Separating and analyzing tryptic digests of protein and peptides by reversed phase HPLC is one of the most commonly used methods for analysis of proteins in research labs and in the pharmaceuticals industry (46).

Lactoperoxidase (MW= 85,000 Da), lysozyme (MW = 14,300 Da) and ribonuclease (MW= 14,700 Daltons) were each digested with trypsin and the resulting peptides were analyzed using reversed phased chromatography on a Hewlett Packard HPLC. Each of the protein samples was digested using trypsin in a mass ratio of 1:50 of trypsin to protein.

Reverse phase chromatography with a very shallow gradient was used to maximize the resolution of the peaks. For LC of the proteins, a Hewlett Packard system (lamp type-G1314A VWD) consisting of degasser, autosampler and UV detector was used. Proteins were separated on a Hewlett Packard C18 (ACE3) reversed phase column (150 mm x 4.6 mm ID), at a flow rate of 1.0 mL/min and injection volumes of 25 µL. 100% water (Solvent A) was used for flushing the system. The mobile phase was 0.1% TFA in water (solvent B) and 0.1% TFA in acetonitrile (solvent C). A gradient ramping from 5% to 30% acetonitrile within 20 min, followed by gradient of 90% acetonitrile within 10 min was used for separation. The spectrometer wavelength was set at 254 nm and 280 nm to monitor absorbance of the effluent.

Mobile phase:

Pump A: Water

Pump B: 0.1% TFA in HPLC grade water

Pump C: 0.1% TFA in HPLC grade acetonitrile (ACN).

A uniform concentration of TFA (0.1%) was present in the mobile phases (solvent B and solvent C) throughout the gradient.

Conditions:

The parameters of the instrument for analysis were:

Flow rate: 1.000mL/min

Injection volume: 25 µL

Pressure: 88bar

Wavelength: 254 and 280 nm

Gradient: 5%-30% B (ACN) over 20 min and then 30-90% B in 10 min.

Stop time: 40 min

Gradient table:

Table 1. Gradient used

Minutes	% B	% C	mL/min
0 min	5	95	1.000
20 min	30	70	1.000
30 min	90	10	1.000
35 min	5	95	1.000

Materials:

Lactoperoxidase and the other two proteins were dissolved at 1mg/mL in 50 mM tris, pH=8.0, 10mM dithiothreitol (DTT). Trypsin (which is stable but not active in 1mM HCl) was dissolved at 1 mg/mL in 1 mM HCl (46).

Method:

1. 1 mM HCL : 50 mL of 1 mM HCl was prepared from 100 mM HCl stock
2. 50 mM tris HCl: 50 mL of 50 mM tris HCl was prepared (pH = 8.00) from 1.5 M tris HCl stock. pH was adjusted from 9.00 to 8.00 by adding 3-4 drops of 12 M concentrated HCl.
3. 10 mM Dithiothreitol: 10 mM of DTT was made in 50 mM tris- HCl (Add 0.077g DTT to 50 mM HCl.
4. Trypsin solution: 5 mg of trypsin was added to 5 mL of 1 mM HCl.
5. Protein solution: 5 mg of each protein (ribonuclease, lysozyme and lactoperoxidase) was added separately to 5 mL of 50 mM tris HCl.
6. Protein digest: 1:50 mass ratio of trypsin to protein was prepared by adding 40 µL of trypsin solution to 2000 µL of each of the protein solution separately.

Instrument set up:

I. Instrument: Setup pump, gradient table, flow rate and injection volume as mentioned above.

II. Method: Save method as (give method name for each gradient set up) e.g. "LYSO494. M" was given name was lysozyme analysis.

III. Sequence parameter: operator name, subdirectory, prefix and counter

   a. Operator name or sample name (lactoperoxidase / ribonuclease trypt etc)

b. Subdirectory (file name) in which the file is saved.

c. Prefix and counter name and number which together should be 8 digits e.g. prefix (Sep 26) and counter (001)

IV. Sequence table: insert the vial location containing sample, the method name and number of injections/ vial as shown in table below.

Table 2. Sequence table

Location	Sample name	Method name	Inj./location
Vial 0	Lacto/trypt	LACTO494.M	1
Vial 1	Ribo/trypt	LACTO.M	1
Vial 2	Lyso/trypt	LYSO.M	1
Vial 3	Water	SHUTDNPC	1

The whole protein was first analyzed under the gradient condition of 5%-30% B (ACN) over 20 min and then 30-90% B in 10 min to see protein peaks. Trypsin was added to the protein with 1: 50 mass ratio of trypsin to protein (40 μL of trypsin added in 2000uL of LPO / ribo / lyso) for the digestion. The samples were incubated overnight at room temp (18-21 hours) and then peptides were analyzed again by HPLC.

Electrospray Ionization (ESI) Quadrupole Time of Flight Mass Spectrometer

Preparing the calibrating standards

The compounds used for calibration were low MW aspartame (MW 294.3 Da), medium MW bradykinin (MW = 1059.5614 Da), angiotensin I (1295.6775 Da) and neurotensin (1671.9097 Da), and high MW myoglobin (16951.55 Da). After calibration with these compounds, the sample of lysozyme (14,306 Da) was analyzed on the mass

spectrometer and the MW of lysozyme was calculated to ensure the accuracy of calibration.

The instrument was not maintaining the calibration; hence another calibration standard was also used. The ESI-Q-TOF MS was calibrated with singly charged protonated molecular ions of proprietary photographic couplers. This calibration compound was obtained from Kodak which helped solve the calibration issue.

To calibrate the system for low MW compound, aspartame was used. To calibrate the system with medium MW compounds and high MW proteins and peptides, MarinerCalMix (Mariner Standards Kit from Applied Biosystems) was used. This mixture consists of medium MW standards like angiotensin, neurotensin and bradykinin acetate salts and high MW myoglobin. HPLC grade solvents and deionized water were used to prepare the standards.

Bradykinin, Angiotensin I, Neurotensin and Myoglobin were from Perspective Biosystems and the sample were prepared according to the instruction given in V700653 INTERIM Mariner Standards Kit. The samples were analyzed on the ESI-Q-TOF MS.

Aspartame and lysozyme stock solutions were prepared by adding around 5-10 mg of the sample to a 50/50 solution of acetonitrile/water with 1% volume of acetic acid added to it. The sample was filtered and analyzed on ESI-Q-TOF MS.

Prior to injecting the sample through the infusion pump, the syringe was filled with HPLC grade methanol and the mass spectrometer was cleaned with methanol to flush out any samples that might be present from the previous run. After cleaning the system with methanol, the MS was cleaned with 50/50 solution of acetonitrile/ water with 1% volume of acetic acid added to it. The sample was injected for at least 10-15 min.

After running the standards, the process was repeated and system was cleaned with methanol and acetonitrile with 1% acetic acid.

Auto Tuning, taking snapshot and calibrating the mass scale

Calibration is performed to make sure and confirm that the instrument is working correctly and calibrating standards are used to tune and to calibrate the system.

Low MW compounds like aspartame (MW= 294.30), medium MW compounds like angiotensin, neurotensin and bradykinin (MW~ 1000-2000 Da) and high MW myoglobin and lysozyme (MW ~ 10,000-20,000 Da) were directly infused in the ESI-Q-TOF MS without introduction into the LC , prior to calibration to check the MW of these compounds. The molecular weight was off by about 100 Daltons for high MW proteins. The instrument was calibrated using all the above mentioned compounds. The compounds were infused again and MW was calculated to check the accuracy of calibration.

The three different steps used in the calibration process were:

- Autotune
- Taking a snapshot
- Performing mass Calibration

LC analysis of tryptic digest of Lysozyme

Prior to the in gel digestion, a whole known protein (lysozyme) was denatured and digested using trypsin and the peptides resulting from this protein digestion were separated using gel filtration chromatography . Gel filtration chromatography was done to separate the denaturing agent from the protein prior to injecting the sample into the MS.

The MW of the peptides from the m/z spectrum was entered in the database to confirm the identity of the protein.

Sample Preparation:

Buffer: 100 mL of 25 mM ammonium bicarbonate (pH 8.4) was prepared: 0.1975 g NH_4HCO_3 was added to 100 mL of nanopure water in a 100 mL volumetric flask. It was shaken well such that it completely dissolves.

Protein solution in Buffer: 200 pmol/μL of lysozyme was prepared in buffer (0.0286 g lysozyme in 10 mL of the above prepared buffer). The solution was filtered prior to analyzing on LC/MS.

Denaturation of protein: 2 mL of the above prepared protein in buffer is mixed with 2 mL of rehydration buffer containing 8 M urea, CHAPS, DTT and Ampholytes (see appendix B) and kept for 3-4 hrs at room temperature for denaturation. The final concentration of urea in the sample was 4 M.

Control: 20 μL of the above denatured protein was removed to serve as a control and was analyzed with a gradient of 2% B-50%B in 45 min.

Trypsin digestion:

A 1:50 molar ratio of trypsin to protein is required for digestion. 10 mg/mL trypsin solution was prepared in distilled/deionized water and 20 μL of this solution was added to the control (denatured protein) and incubated overnight for (18-20 hrs) at 37^0C. The digested lysozyme sample was filtered prior to injection.

LC analysis of Lysozyme peptides

The sample was first analyzed only on the LC system of LC/ESI-Q-TOF MS (by disconnecting LC from mass spectrometer to avoid flooding the denaturants into

the mass spectrometer). For the analysis of separated peptides, a Shimadzu C18 (ACE3) column 50 cm x 2.1 mm, interfaced to an ultraviolet-visible spectrometer detector. The eluent was monitored by UV at 254 nm and 280 nm. Solvent A was 1% acetic acid in HPLC grade water and solvent B was 1% acetic acid in HPLC grade acetonitrile. A gradient was ramped from 2% B to 40% within 45 min. The flow rate was 0.2 mL/min with an injection volume of 20 μL.

Gel Filtration Chromatography

Gel filtration or size exclusion chromatography was used to separate the peptides from the urea which was used to denature the protein prior to digestion by trypsin. The media used to pack HPLC columns is the TOYOPEARL Size exclusion media HW- 40S [www.sigmaaldirch.com/SUPELCO/ Cat # 807451(16)] with 50-100 um particle size and pore size of 100-7000 Da MW range. TOYOPEARL gels are compatible with a variety of column packing methods and are commonly used in glass columns. Size exclusion chromatography was performed under isocratic conditions, using ammonium bicarbonate buffer solution.

The glass column with media (TOYOPEARL) was mounted on the Waters 650E Advanced Protein Purification System which is a low pressure liquid chromatography system for rapid separation of proteins.

The components used are:

- Manual injector
- Waters AP-1 Column
- Waters 600E Systems controller : Millipore
- Waters 490E Programmable Multi wavelength detector

- Fraction collector : GILSON Model FC-205
- Chart

Procedure:

A. Pouring the media in the column:

1. Open the column from the top and remove the plunger by turning the plunger counterclockwise and without touching the column outlet (bottom) .Wash the column with deionized/distilled water and then with buffer (NH_4HCO_3) twice.
2. Clamp the column upright to a rigid support. Pour water in the column till it is $1/3^{rd}$ distance from the bottom of the column.
3. Stir the gel slurry (TOYOPEARL Size exclusion media Supelco HW- 40S from Sigma Aldrich) thoroughly with a glass rod into a homogenous suspension and carefully transfer the slurry into the reservoir through a small funnel.
4. Allow the slurry to settle until a stable bed is formed. Let the water drain from the outlet drop by drop. If the column is to be left overnight cap the bottom so that the column does not dry out.
5. A layer of water is observed on the top following day. Uncap the bottom, remove the funnel from the top and remove excessive solvent from the top of the column bed without disturbing the bed with a transfer pipette.
6. Add more gel slurry and sufficient buffer to keep the column wet, letting the liquid drain dropwise from the bottom and let column bed settle.

7. Repeat step 5 using the transfer pipette until not more than 1 inches of dead volume is left on the top. At this point, cap the top by carefully and slowly squeezing the plunger clockwise such that there is no air bubble. Liquid will squeeze out from the top.
8. In case any bubbles are seen, remove the top (not entire plunger), add more water and squeeze it in until it sits on the column bed.
9. Tighten the column end, and cap the column inlet and outlet.
10. When you are ready, flush the system with the desired buffer.

B. Preparing the System

- Filter around 3 liters of 25mM NH_4HCO_3 (pH 8.4) through suction filter and transfer it to a clean 2.5 liter storage bottle. Pour it into one of the mobile phase reservoirs (A). Nanopure water was poured into reservoir D.
- Power up the system controller, multi wavelength detector and fraction collector.
- Press the Set up function key and then press “Isocratic” function key in controller front panel.
- Using home key, arrow key and function key set flow rate at 0.5mL/min. Move the cursor to flow field and type 0.5. Press Enter. The flow rate is set to 0.5 mL/min.
- Move the cursor to % A field for the first eluent and type 100. Press Enter. Type 0 and press enter for the remaining % composition for the mobile phase. The sum of the four % composition must be equal to 100.
- Let the system line flush with buffer NH_4HCO_3 (A= 100%) for 30 min.

- While the system is flushing, fill the test tube rack with test tubes and place it in fraction collector. Enter 2min/tube in the fraction collector and press “Yes”.
- Flushing the manual Sample Injector :
  - Make sure the vent vial is in position below the vent tube.
  - Set the handle on the injector to the Load position.
  - Load a syringe with volume of sample buffer approximately 5 times the volume of the loop.
  - Fully insert the syringe needle into needle/filler port
  - Gently discharge the contents of the syringe to completely fill the loop.
  - Discard the contents in the Vent Vial and repeat.

C. Performing a run

- Take the gel filtration standard sample (Bio-Rad Catalogue # 151-1901 www.biorad.com) from the refrigerator. Add 1 mL of ammonium bicarbonate buffer into gel filtration standard: The lyophilized mixture contains thyroglobulin, γ–globulin, ovalbumin, myoglobin and vitamin B_{12} (total protein content is approx. 18mg).
- Fill a syringe with gel filtration standard prepared above and inject 1 mL of the sample by fully inserting the syringe needle into the injector port.

- Turn the injector handle to the "Inject" position to make an injection. Remove the syringe and the run will start.
- For Detector setting :
    - Auto zero the multi wavelength detector by pressing the auto zero function key on the detector,
    - To set the wavelength, press λ key on the detector, type in 280 and press Enter to set the wavelength as 280nm.
    - At the AUFS field, press AUFS key on the detector, type the Absorbance Unit Full Scale as 1.000 and press Enter.
- Press the "start" key on the fraction collector and pull the chart knob down at the same time to start both.
- Enter the following information on the chart :

  Column flow rate = 0.5 mL/min

  Fraction collector = 2 min/tube

  Chart = 0.5 cm/min

  AUFS (Absorbance Unit Full Scale) = 1.000

  Eluent Composition:

  A = Buffer (100%)

  B = base (0%)

  C= Salt (0%)

  D – Nanopure water (0%)
- Since myoglobin and vitamin B_{12} are colored, both are visible when applied to the column and provide a means of ensuring that the column is properly packed and the samples are eluting evenly.

- Samples after eluting from the column to the fraction collector are collected in the test tubes in a Z pattern. Before the samples reach the fraction collector, they move through the detector when their absorbance increases and peaks are observed on the chart.
- After the run is over, flush the column with buffer for 60 minutes. Flush the sample injector with buffer.
- Change the test tubes in the test tube rack with clean tubes.
- Prepare a fresh sample of lysozyme digest (as described above in sample preparation; page 40-41) and inject 1 mL of filtered digested peptide sample and start the run as suggested before. The sample runs for around 2.5 hours until all the 80 test tubes are filled at a rate of 2 min/tube.
- At the end of the run, clean the system with buffer and water again.
- Test tubes containing samples in the fraction collector are observed. Samples corresponding to peaks on the chart because of high absorbance are ready to be analyzed on the MS.
- Similarly inject 1 mL of **rehydration buffer** to check where the urea peak falls on the chart with respect to peptide peaks to make sure there is no urea in the sample to be analyzed on the MS.

LC/ESI-Q-TOF MS identification of Lysozyme peptides

For the separation of peptides, a mass spectrometer consisting of an autosampler and an Electrospray Ionization Quadrupole Time of Flight (ESI-Q-TOF) mass spectrometer was used. Nitrogen was used as a nebulization and desolvation gas.

The instrument was calibrated with singly charged protonated molecular ions of proprietary photographic couplers. The infusion pump was only used for flushing the MS and not for injecting the sample through the LC/ESI-Q-TOF MS, therefore the syringe icon on the MS should not be clicked while the LC/MS is in use. Instead, the system has an autoinjector which serves the purpose of sample uptake.

The sample was prepared in HPLC vials which were kept in the autosampler tray used in the LC system. Autopurge was performed before every run. Before starting the run, the sample was placed in HPLC vials in the LC tray holder making sure the pumps are well connected and there are enough mobile phase solvents. Pump A contained 1% acetic acid in HPLC grade water and pump B had1% acetic acid in HPLC grade acetonitrile.

Databases used to identify proteins

The identification of protein was done by searching lysozyme peptide masses within the NCBI database [55] and comparing these masses with the experimental masses. The peptide masses obtained were used to search protein from the following protein databases: protein Prospector, UCSF, San Francisco, USA [56]; MASCOT search engine [57]; Profound [58]; EMBL [59]; Aldente [60] and Expasy [53].

Optimal search parameters were enzyme trypsin, one missed trypsin cleavage, minimum number of peptide required for match (4-5), peptide charge state as MH^+, unmodified cysteine and methionine and a window of 2 Da mass tolerance using monoisotopic mass.

In-gel digestion

In-gel digestion of unknown proteins isolated from 2D- gel electrophoresis using trypsin is a very common and effective tool used in proteomics. Protein of interest are excised from the gel, digested with trypsin and analyzed by LC/MS with subsequent database searching. Trypsin is a pancreatic serine protease which specifically hydrolyses peptide bond at the carboxyl end of lysine and arginine residues resulting in peptide fragments.

For in gel digestion, the proteins were overloaded on IPG strips. The IPG strips were loaded with 250 µL-350 µL of 13-20 µg /µL (micrograms/microliter) proteins with no rehydration buffer added to it. Once the gels were overloaded with protein, streaking was observed and gels quality declined. To improve the quality of the gels, they were run at 4^0C conditions at 100 volts for a longer time (around 3 hours).

Reagents:

The in-gel tryptic digestion Kit (PIERCE, Product code 89871) was used to perform the digestion which contained the following reagents:

1. Trypsin, Modified, 20 micrograms
2. Trypsin storage solution, 40 µL
3. Acetonitrile, 3 x 24 mL (3 bottles of 24 mL each)
4. Ammonium bicarbonate, 300 mg
5. Iodoacetamide (IAA), 500 mg
6. TCEP (Tris[2-carboxyethyl] phosphine), 500 µL

Additional material required

1. Siliconized Eppendorf tubes(Product code T4691)
2. Centrifugal concentrator (Speed Vac) : Centrivap concentrator (catalog # 78100-00) and Centrivap cold trap (catalog # 78110-00) labconco model ; Vacuum pump (fisher scientific model 4c)
3. Flat nosed tweezers
4. Scalpel (S2771 / S 3021)
5. Ultra pure water (18 megaohm equivalent)
6. Micropipette and 1 % TFA

Material preparation

1. Destaining Solution: The solution is prepared by dissolving 80 mg of ammonium bicarbonate in 20mL acetonitrile and 20mL of ultra pure water. This is used to remove dye from the gel.
2. Digestion Buffer (volume ~ 25 mL): The solution is prepared by dissolving 10 mg of ammonium bicarbonate in 5 mL of ultra pure water.
3. Trypsin Stock: Modified trypsin (20 micrograms) is supplied lyophilized which can be stored at -20^0 C .Trypsin stock solution is prepared by hydrating lyophilized trypsin with 20 µL of trypsin storage solution supplied with the kit. The solution contains components that inactivates the enzyme and protects it from auto digestion. This 20 µL of trypsin stock solution is divided equally into four separate microfuge tubes of 5 µL each which can be stored at -20^0 C in a frost free freezer.

4. Trypsin Working Solution: When required, Trypsin Stock aliquot is thawed on ice and the stock is diluted 10 fold by adding 45 µL of nanopure water. This solution can be stored at -20^0 C without any significant loss in activity.
5. Activated Trypsin (concentration ~10 ng/L): This is prepared just before use in which 1.0 µL of Trypsin Working Solution is diluted with 9 µL of digestion buffer. Since around 10 pieces of gel are digested simultaneously, it is advisable to dilute 10 µL of Trypsin Working Solution with 90 µL of digestion buffer as 10µL is required for each gel piece. The activated trypsin is stored on ice until use.

Procedure for In-gel digestion:

1. The gels were run as described previously. However, the IPG strips were overloaded with 250 µL-350 µL of proteins (2500-3500 µg) with no rehydration buffer added to it, the voltage was set to 100 volts and gels were run under cold conditions (4^{0}C) for 3.5 hours to minimize streaking.
2. Reduction and alkylation of the protein sample is optional since it is done prior to running a 2D gel. It was not performed on these samples to reduce the loss in number of peptides. Gels were stained with Commassie brilliant blue and destained.
3. Excise the band of interest from the 2D gel into 1x1 to 2 x2 mm pieces using a scalpel. Using a cork borer/tweezers, remove the spot from the gel, making sure to include only stained portion of the gels. The gel pieces

were transferred to siliconized eppendorf tubes. Siliconized eppendorf tube reduces the binding of peptides to the tube surface.

4. Cover the gel pieces with 200 µL of destaining solution and incubate at 37^0 C for 30 minutes with shaking.
5. Remove and discard the destaining solution from the tube.
6. Repeat step 4 and 5 once.
7. Shrink gel pieces by adding 50 µL of acetonitrile and incubate the sample for 15 minutes at room temperature. Carefully remove the acetonitrile.
8. Dry the gel piece in centrifugal concentrator (Speed Vac) for 15-20 min.
9. Prepare activated trypsin as described before and swell the gel pieces by adding 10 µL of activated trypsin solution to the Eppendorf tubes containing dry gel pieces. If 10 µL is insufficient to cover fully swelled gel pieces, increase the volume accordingly.
10. Add 25 µL of digestion buffer to the tube.
11. Incubate the sample overnight at 30^0C with shaking. Make sure that the gel piece is at the bottom of the tube and covered with liquid.
12. Remove the digestion mixture from the gel piece using a digital micropipette and transfer it into clear glass conical inserts used for 2mL standard opening vials with size 0.15 mL and bottom spring [21].Since the volume of the sample is too low to be injected by the autosampler, inserts were used. The solution contains the extracted tryptic peptides.
13. 10µL of 1% TFA (trifluoroacetic acid) is added to the gel pieces to further extract peptides and the sample is incubated at room temperature for 5 minutes. Remove this peptide extraction solution from the tube and

combine it with the digestion mixture from step 12. This step serves to inactivate trypsin by stopping additional enzymatic activity.

14. Concentrate the combined sample in a speed vac [Centrivap cold trap and Centrivap concentrator (labconco); vacuum pump (fisher scientific model 4c)] for about an hour and then analyze it by LC/ESI-Q- TOF MS with a gradient of 2% to 40% acetonitrile in 45 min at a flow rate of 0.2 mL/min and injection volume of 10 μL.

Results and Discussion

Three proteins from the soluble fraction of *Pseudomonas putida* KT2440, grown on succinate were identified by proteomics methodologies. The proteins were separated by 2-D gel electrophoresis. Spots of interest were excised and the proteins were digested with trypsin. The resulting peptides were analyzed by LC coupled on line to an ESI-Q-TOF MS. Proteins were identified by comparing the mass spectral data obtained with the information contained in the database of known protein sequence: the NCBI database [55], using EMBL [59], Aldente [60], MASCOT PMF protein identification system [57], and the MS-Fit protein database (protein Prospector, UCSF, San Francisco, USA) [56]. These results were confirmed by tandem MS/MS analysis of the peptides produced by tryptic digestion of each of the three spots.

To accurately identify proteins, it was necessary to generate reproducible gels with dark intense spots, such that the peptides could be identified by the mass spectrometer. It was critical to have consistent growth curves with a reproducible mid log phase to extract a consistent protein sample from healthy cells. Figure 10 is an example of the growth curves of *P. Putida* performed in the lab.

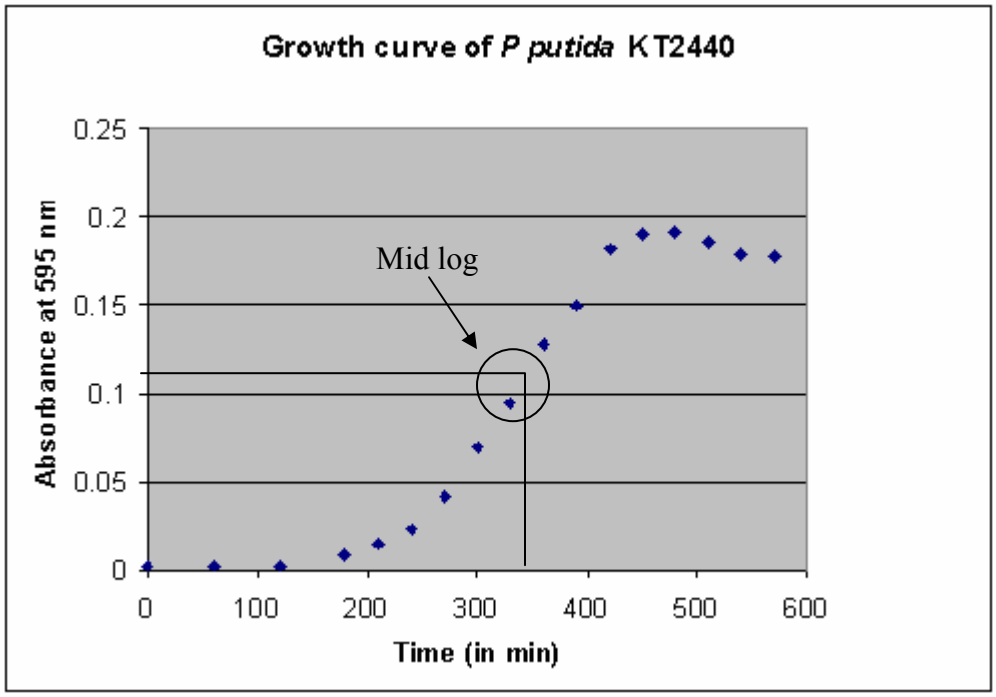


Figure 10. Growth curve of *P. putida* KT2440 grown on 5mM succinate

The mid-log phase of the growth of *P. putida* KT2440 on 5mM succinic acid was 340 min +/- 5min which was virtually identical with the time established by previous investigators in the lab [61].

After generating a reproducible growth curve, 2DE was performed according to the procedure described in experimental section, in duplicates or triplicates to generate reproducible results.

The two gels displayed in figure 11a and 11b are examples of the protein expression observed during the mid log phase of the bacterial growth. Each spot in the gel is a protein on the gels. These proteins increase in pI from left to right and decrease in MW from top to bottom. IEF (Isoelectric focusing) was performed initially on 11 cm pH 3-10 IPG strips; later 11 cm IPG strips with pH 4-7 were used to improve resolution. The

sample (~200µg of protein) was applied to the IPG strip by passive rehydration for 12 hrs (Figure 11a) or by active rehydration (Figure 11b). The 6-fold increase in the protein load in Figure 11b led to more intense spots, suitable for tryptic digestion and MS analysis.

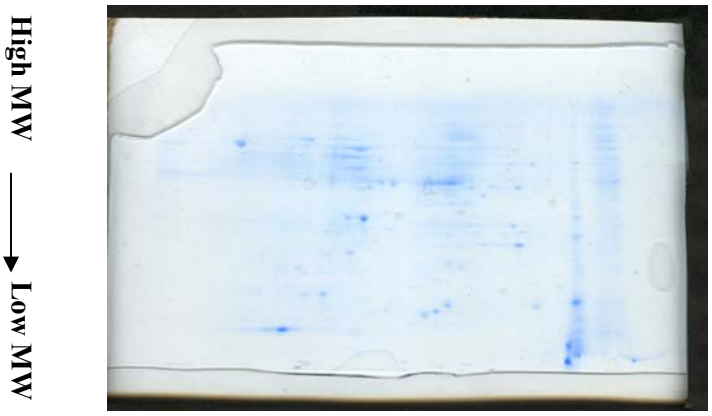


Figure 11a. Commassie stained 2D Gels on pH 3-10 IPG strips with 15µL of protein sample (~195µg) with passive rehydration; image acquired by EPSON scanner (Phoretix 2D Software).

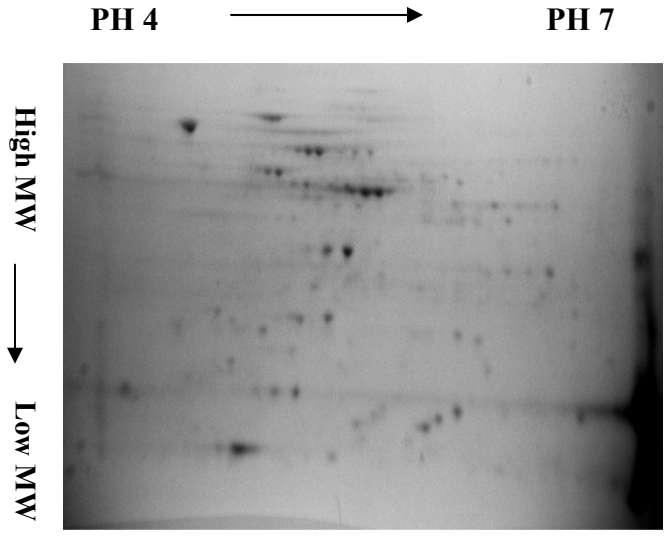


Figure 11b. Commassie stained 2D Gels on pH 4-7 IPG strips with 100µL of protein sample (~1200µg) with active rehydration; image taken by EPSON scanner (Phoretix 2D Software)

Once conditions for reproducible gels were established, 2DE was performed using 11 cm pH 4-7 IPG strips with active rehydration. These gels were used for choosing protein spots for further analysis, since more spots were observed with active rehydration (Figure 12).

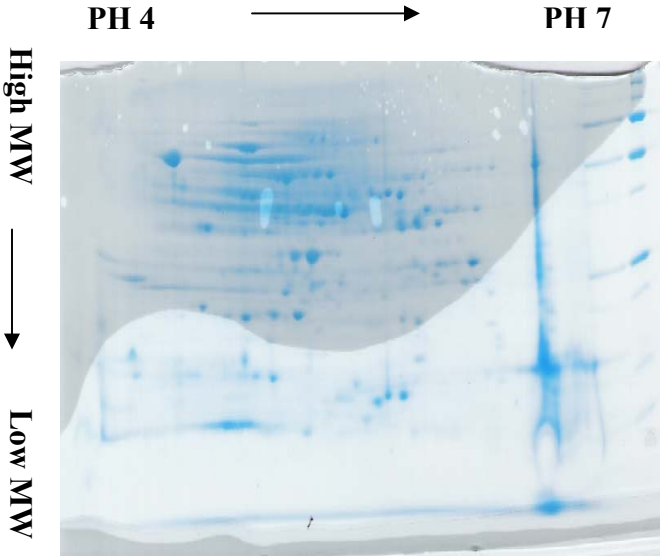


Figure 12. Commassie stained 2D gel image on 11cm pH 4-7 IPG strips with 200-300 µL of protein sample (~ 2000-3000 µg); with active rehydration; image acquired by EPSON scanners (Phoretix 2D Software)

A liquid chromatography system coupled to mass spectrometry is currently the most powerful analytical tool available to research laboratories involved in studying proteins and peptides. The advantage of ESI-Q-TOF MS is that as a result of the multiply-charged ion, the instrument can be calibrated with a low value of m/z using singly-charged compound with known masses. However, there are some limitations to this system such as susceptibility to even low amounts of salts, which can significantly decrease sensitivity, poor compatibility with phosphate and other non volatile buffers and detergents like SDS and /or any other contaminants if present in the samples. The operation of the LC/MS system with such limitations is therefore very complex. Moreover, the system is expensive and requires a lot of maintenance.

Sensitivity of ES is quite good, with detection level of 100 fmol for many peptides. ESI-Q-TOF MS also has a high mass range and can measure the masses of proteins over 50,000 Da and can have upper mass limit of 100,000 Da with mass accuracy of about ± 0.005% [26].

Initially the instrument was calibrated with the low molecular weight compound aspartame (MW ~ 294.3 Da) and the medium molecular weight salts angiotensin, bradykinin and neurotensin (MW ~ 1000-2000 Da). Then the instrument was calibrated with the high MW compound myoglobin (Mb), a protein with MW 16,900 Da. Since lysozyme is another protein in the same MW range as Mb, the accuracy of calibration was checked by analyzing lysozyme on the ESI-Q-TOF MS and calculating its MW. The instrument was also calibrated with singly charged protonated molecular ions of proprietary photographic couplers. The compound used was the sodium adduct of a yellow photographic coupler at m/z 671.3434. A photographic coupler is a compound used in film that when reacted with developer forms a yellow image dye when developed. This compound was obtained from a colleague (Robert Saccente) at Kodak.

For aspartame, the m/z spectrum was consistent with the molecular mass (294.3 Da). The m/z value is singly charged (z =1) and it shows dominant ions at m/z 295.35, which was consistent with the expected protonated ions, (MH^+) i.e. (294.31+ 1.008).

The values for bradykinin were also in agreement with the expected protonated molecular ion (MH^+). Three peaks were observed for Bradykinin (molecular mass-1059.56) with m/z value of 1060.86, 532.07 and 354.62 indicating +1, +2 and +3 charge states respectively.

For lysozyme, the sample was analyzed in a solution of 50:50 acetonitrile: H_20 with 1% volume of acetic acid added to it and the spectrum shows a series of multiply charged ion peaks with m/z value ranging from 1101.5 to 2044.6. The detector of ESI-Q-TOF MS displays two kinds of data images - a total ion chromatogram or TIC (Figure 13a) and a mass spectrum (Figure 13b). The total ion chromatogram or TIC is a plot of relative intensity of all ions against retention time whereas the mass spectrum is a plot of relative abundance against mass to charge ratio (m/z) at a specific time, with % intensity always being on the y axis.

Chromatogram / Spectrum of Hen egg lysozyme (MW= 14,306 Da)

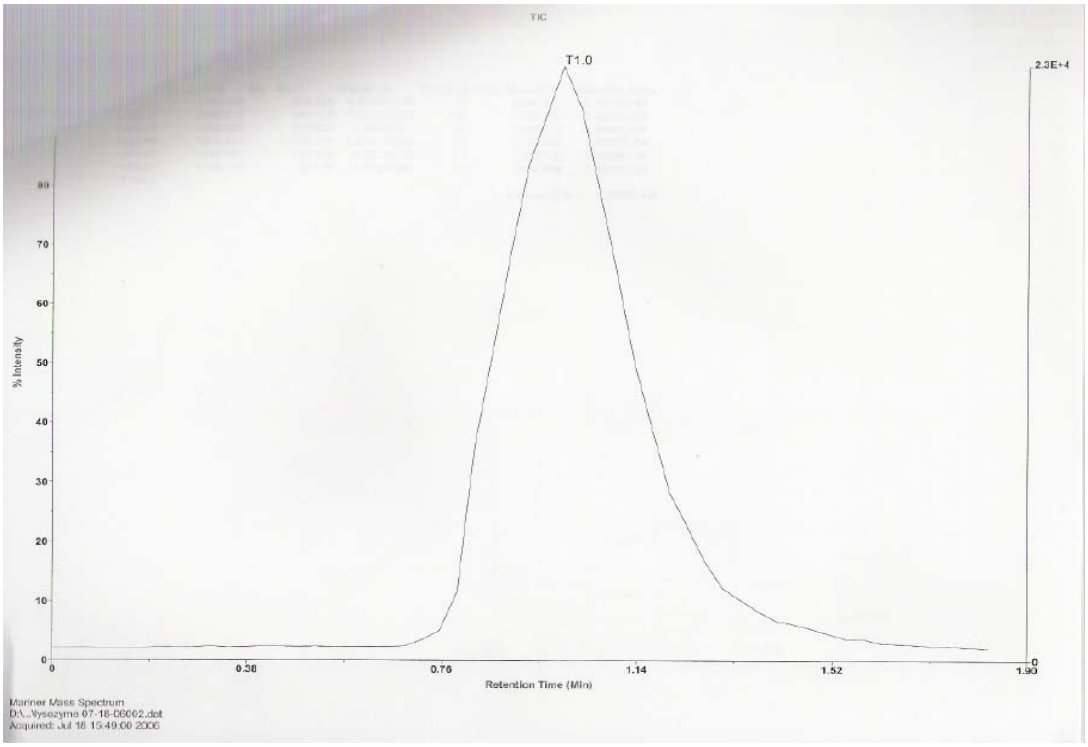


Figure 13a. Total ion chromatogram (TIC) of protein hen egg white lysozyme

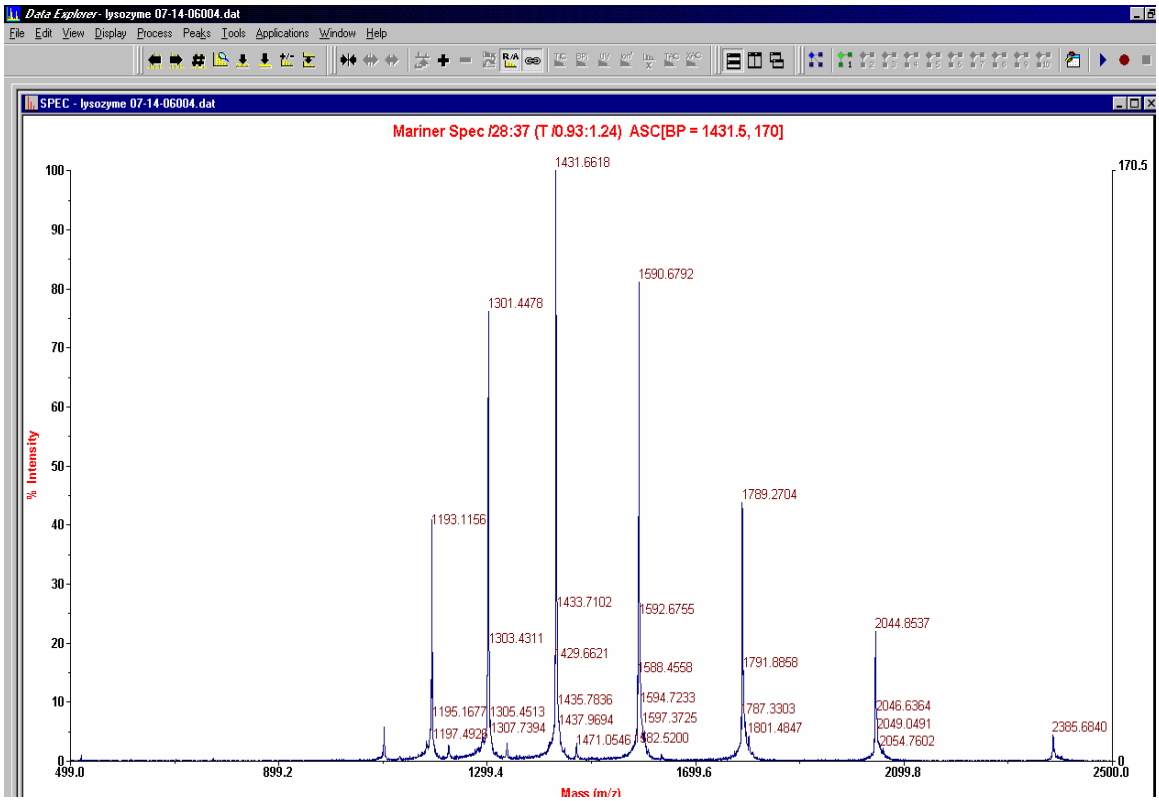


Figure 13b. Electrospray mass spectrum of protein hen egg white lysozyme showing peaks with different number of protons and charges.

In figure 13b, the spectrum has a wide distribution of charge states which shows the mass spectrum of hen egg white lysozyme taken at low resolution with charge states from +7 to +13. In the positive ionization m/z spectrum of the hen egg white lysozyme, each peak represents the intact protein molecule carrying a different number of charges which arises from molecules carrying different numbers of protons. In the multiply charged series, since the molecular weight is the same for all of the peaks, the ion peak with high charge appear at lower m/z values than those with less charge. Any two adjacent peaks differ by a charge of +1. In the lysozyme spectrum shown in figure 13b, if the ion peak with m/z at 1789.2 has n charges, then the ion peak with m/z at 1590.6 will have n+1 charge. The observed molecular mass (MW 14306 Da) is in good agreement

with the average molecular mass of hen egg lysozyme (14306 Da; see details below and in Table 3).

Monoisotopic mass is the sum of the exact atomic mass of the most abundant and stable isotopes of the atoms in a molecule whereas average mass is the average of the isotopic masses of the atoms in a molecule calculated by taking the atomic weights of the element. In short, monoisotopic mass is the mass of the first peak (or the ^{12}C peak) of the isotopic distribution which is the peak with maximum intensity and average mass is the average masses of all of the peaks in the isotopic distribution.

For calculating the molecular mass of low MW compound, the mass is normally calculated by taking the "nominal " mass which is the atomic masses of integers like C=12, H=1, N=14 and O=16. However, such calculations are not useful in mass spectrometry of larger species.

In this case, monoisotopic mass is calculated using the exact atomic masses such as H=1.00782 and C=12.000 etc. Even a small protein of MW 10,000 Da will have around 500-600 hydrogen atoms in it and since the actual mass of the hydrogen is 1.00782, it will show a mass shift 4-5 daltons higher than would be expected by calculating the nominal mass of the compound. So the monoisotopic mass of the compound is calculated based on the accurate masses of the atoms. However, this calculation still does not take into consideration the presence of naturally abundant isotopes and when these are accounted for, an isotopic distribution is produced which is centered at a mass higher than the monoisotopic mass.

For molecules with molecular mass less than 1000 Da, the isotopes can be resolved and the monoisotopic mass is the most abundant. Hence, monoisotopic masses should be used to calculate molecular mass. As the molecular masses increase, the

number of charges also increases and the monoisotopic peak decreases in relative intensity. This in turn leads to an increase in the number of isotopes to be resolved. In this case the average molecular mass of the compound should be considered. However for molecular masses of proteins and peptides, either monoisotopic masses or average masses can be used, depending on the resolution of the instrument and the size of the peptides.

ESI-Q-TOF MS separates ions on the basis of mass to charge ratio (m/z). For any peak with $m/z = m_n$, mass (m) is the sum of the mass of the protein plus n hydrogen atoms attached to it and z is the charge on the peak i.e. m = M + n (1.008) and hence

$$m/z = m_n = M + n\,(1.008)\,/n - (I)$$

where n is the charge on the molecules because of the protons (H^+) attached to it. Any two adjacent peaks in the series of multiply charged ions differ by one charge. Therefore, the next peak at lower mass to charge ratio (m/z) will have n+1 protons and a charge of n+1. For this peak,

$$m/z = m_n+1 = M + (n+1)\,(1.008)\,/\,n+1 - (II)$$

. The charge (n) on each peak can be calculated by equating the above two m/z equations (I and II) and once the charge on the peak is known, then simply by reading the m/z value from the spectrum and solving for the above two equations, the molecular mass of the neutral protein can be calculated as shown in table 3.

Calculation of MW of lysozyme

From the lysozyme spectrum shown in figure 13b, and based on the value of m/z, molecular mass (M) of the protein was calculated as shown in table 3. The molecular mass was found to be in good agreement with the actual molecular mass of hen egg white lysozyme (14306 Da).

Table 3. Calculation of the molecular weight of lysozyme

m/z =Mn	Mn+1-1.008	Mn- Mn+1	Charge (n)= Mn+1-1.008 / Mn- Mn+1	Charge rounded	Mn - 1.008	Molecular mass (MW) =charge x (Mn-1.008)
2044.85	1788.262	256.588	6.969390618	7	2043.842	1.43E+04
1789.27	1589.672	199.598	7.96436838	8	1788.262	1.43E+04
1590.68	1430.652	160.028	8.940010498	9	1589.672	1.43E+04
1431.66	1300.442	131.218	9.910545809	10	1430.652	1.43E+04
1301.45	1192.102	109.348	10.9019095	11	1300.442	1.43E+04
1193.11						
					Average MW	**1.43E+04**

Calibration

Early in the project, the LC/ESI-Q-TOF MS was not retaining its calibration throughout an experiment. It was observed that the m/z value for the spectrum obtained when the protein was analyzed on the MS through the infusion pump differed by about 200 daltons from when the value obtained by injection through the LC interface. Some of the factors considered while solving the calibration problem were the flow rate, sodium/ ammonium ion adduct, inadequate resolution, calibration compound, method of calibration and instrument parameters.

Initially it was thought that difference in the flow rate with which the sample is introduced into the mass spectrometer by direct infusion and the flow rate with which the sample is introduced through the LC interface was responsible for the calibration problem. The flow rate for the infusion pump was 20µL/min whereas the flow rate for the

LC pump was 200μL/min. However it was found that the flow rate effects only the intensity of the signal and not the calibration values.

Another possible source of the calibration problems was the presence of sodium or ammonium ions in the sample. The presence of sodium adduct ions (instead of protons) on a peptide in the MS can lead to a 23 Da increase in the m/z for that peptide. Molecules can similarly be ionized by the addition of other common adduct ions like K^+ (+39) and NH_4^+ (+18). These Na^+, K^+ or NH_4^+ ions get attached to some (and not all) peaks and the averaged spectrum of the peak is taken. However, the buffers used in these studies did not contain appreciable amounts of sodium, potassium or ammonium ions. Therefore, these concerns were dismissed.

Next the calibration compounds were considered. After initially attempting a one point calibration with compounds such as Mb, a two point calibration was chosen to obtain a reliable mass measurement. The two compounds chosen were the photographic compound (MW 671.3434) and lysozyme (MW 14,306). Multiple charging is common for proteins and proteins are considered acceptable calibrants. The use of two calibrants compensated for the inadequate resolution of the Mariner TOF system. An additional concern was isobaric interference, which can result in an interfering ion with the same nominal mass underneath the mass of the calibrant. This can also be overcome by using the two-point calibration approach.

Finally, the instrument parameters were adjusted to improve the stability of the calibration. The major adjustment was a change in mass spectral data collection. Instead of taking a snapshot of one spectrum, snapshots of sequential spectra were accumulated which increased the intensity of signal before the mass calibration step was performed.

The figure below (Figure 14) is the instrument parameter sheet displaying the changed parameters saved after calibrating the mass spectrometer; samples were analyzed using these parameters. Changes made to the API interface settings (einzel lens potential, deflection voltage) and to the TOF analyzer settings (push pulse potential, pull bias, pull pulse potential) impacted the intensity and resolution which in turn improved the stability of the calibration.

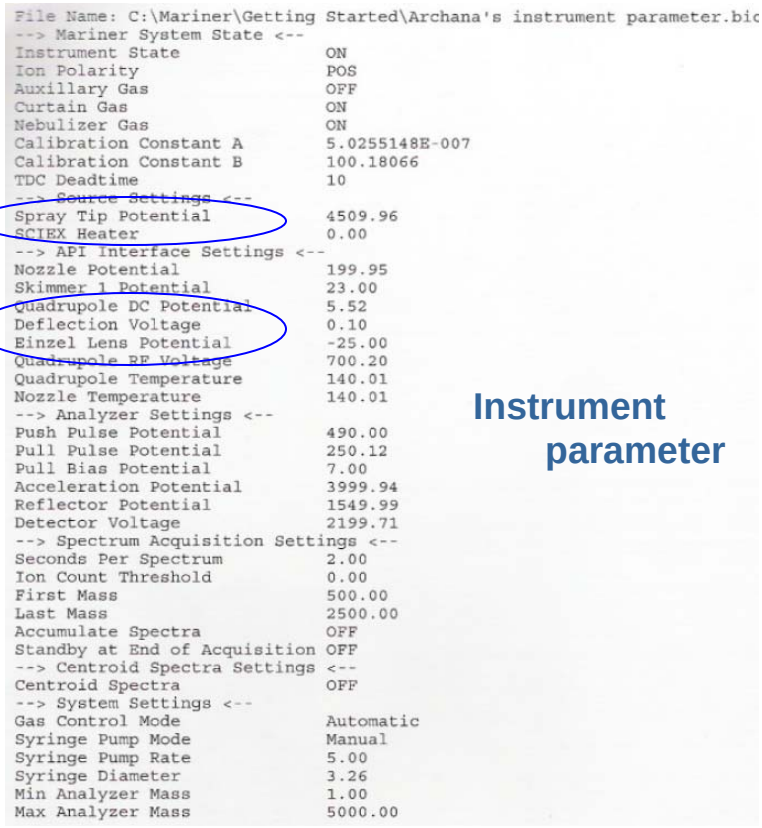


Figure 14. ESI-Q-TOF instrument parameter settings

LC/ESI-Q-TOF MS analysis of lysozyme:

Once the instrument calibration was stabilized, the next step was analysis of a known protein, lysozyme, subsequently using the peptide masses from the analysis to confirm the identity of the protein in online databases. Prior to working on this step, protein digestion was analyzed by HPLC with lactoperoxidase (LPO), lysozyme, and ribonuclease. Both intact protein and tryptic peptides were analyzed on the system.

When the intact proteins were analyzed by HPLC, an individual protein peak was observed on the chromatogram corresponding to each protein. However, once the same protein was digested with trypsin, a series of peaks was observed on the chromatogram corresponding to the tryptic peptides (Figure 15-17). The first few peaks in each chromatogram, observed around 2-3 minutes are "noise" from the solvent. The larger peaks that are displayed after some time represent tryptic peptides in the sample. The peak labeled 24.829 minute in figure 15 is intact lysozyme; following tryptic digestion, seven peptide peaks were observed at 7.759, 8.783, 10.066, 10.361, 13.101, 14.099 and 22.969 minutes (figure 16). The sequence of lysozyme contains 17 lysine and arginine residues; digestion with trypsin should generate 18 peptides. Prior to subsequent digestion, 4 M urea was added as a denaturant. HPLC analysis at 254 nm revealed the presence of approximately the expected number of peptides (Figure 17). Similar results were obtained when lactoperoxidase and ribonuclease were digested with trypsin.

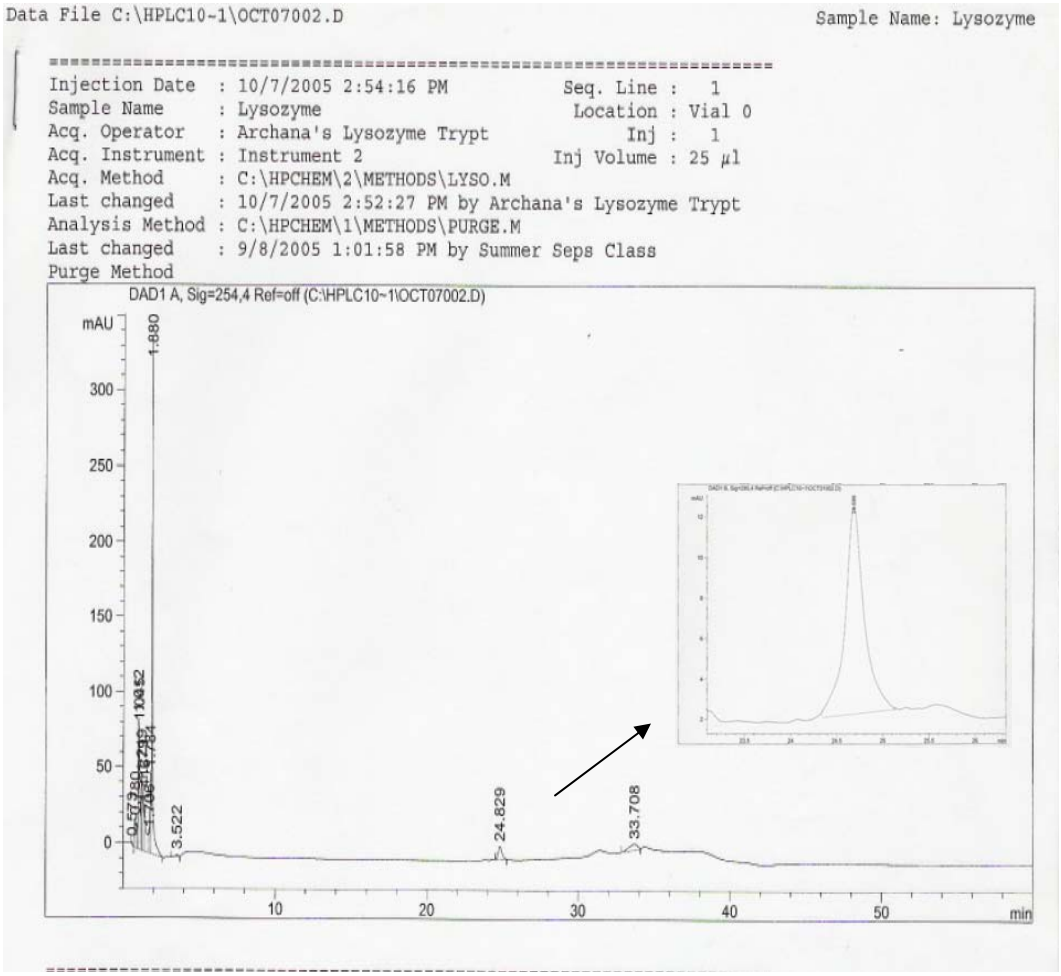


Figure 15. A chromatogram of lysozyme**:** Analysis of peptides obtained from the digestion of proteins was carried out on a Hewlett Packard C18 (ACE3) column 50 cm * 2.1 mm interfaced to an ultraviolet-visible (UV) spectrometer detector set at 254 nm. 25 µL of 1 mg/mL of protein sample prepared in 50 mM tris, pH=8.0, 10mM dithiothreitol was injected. A gradient ramping from 5% to 30% acetonitrile within 20 min, followed by gradient of 90% acetonitrile within 10 min was used for separation with a flow rate of 1.000 mL/min

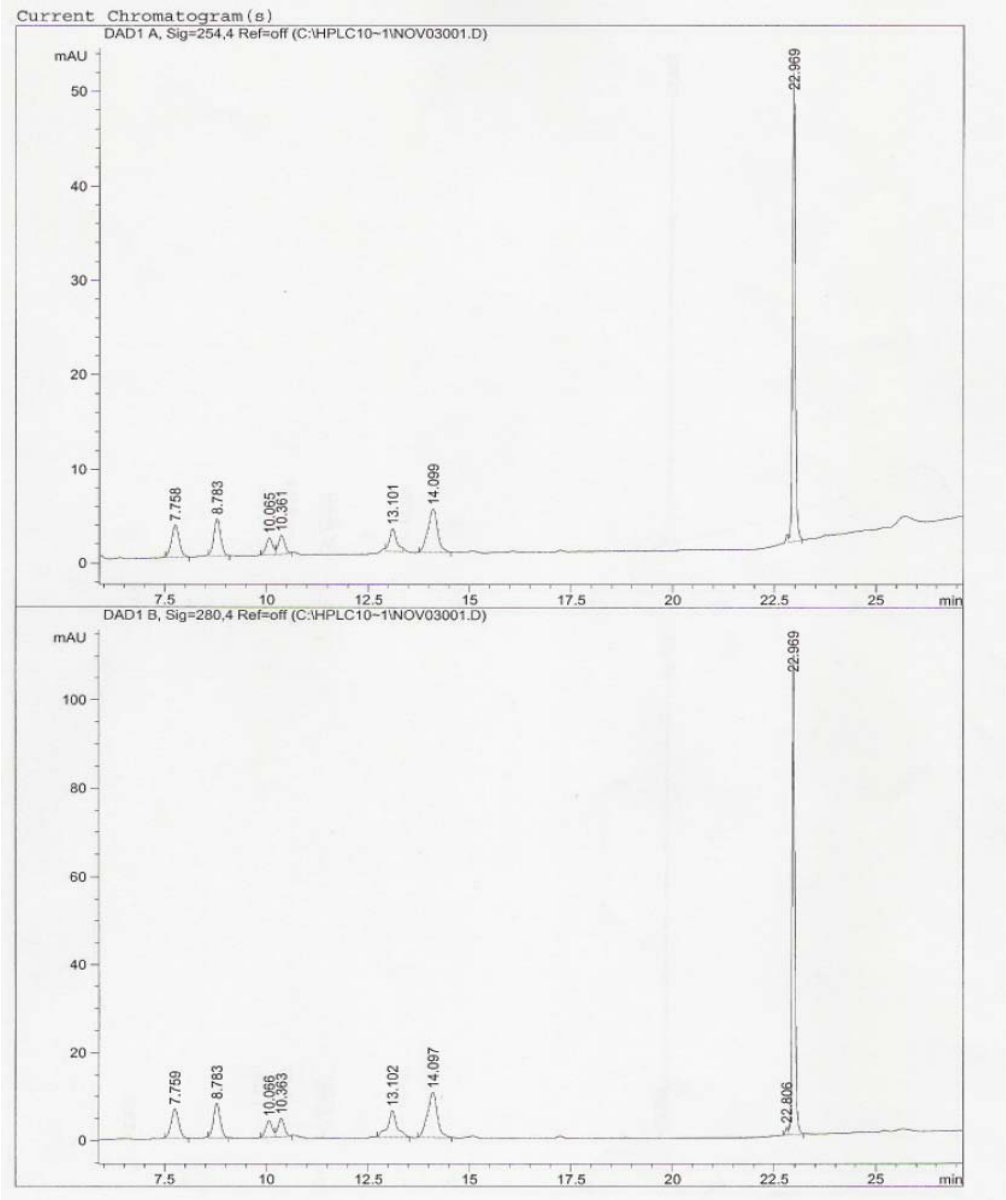


Figure 16. A chromatogram of lysozyme peptides: A 1:50 mass ratio of trypsin to protein prepared in 50 mM tris, pH=8.0, 10 mM dithiothreitol was used for digestion; the sample was incubated overnight and 25µL of this sample was analyzed on HPLC with same separation conditions as described in figure 15.

Various peptide peaks were observed on the LC chromatogram (Figure 17) when the lysozyme was denatured with 4M urea and digested with trypsin. The peptide samples were directly analyzed on the liquid chromatography system.

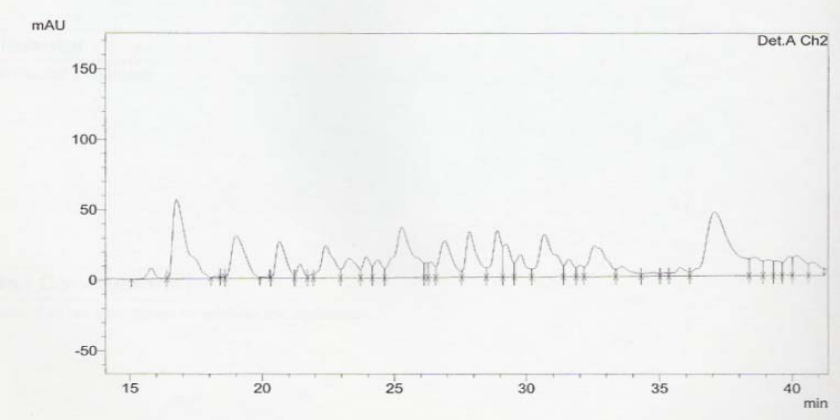


Figure 17. LC chromatogram showing the peaks of lysozyme peptides: 200 pmol/µL of lysozyme was prepared in 25 mM ammonium bicarbonate (pH 8.4) buffer. A 1:1 solution of the protein sample and rehydration buffer containing 8M urea, CHAPS, DTT and ampholytes (Appendix B) was prepared; the protein sample was kept at room temperature for 3-4 hrs for denaturation; A 1:50 molar ratio of trypsin to protein was prepared and incubated overnight for 18-20 hrs at 37^0C for digestion; the sample was analyzed on LC with a gradient of 2% B-50%B in 45 min.

Before analyzing the samples on MS, it was necessary to remove the urea from the protein through a gel filtration column to prevent its entry in the MS. The digested sample was analyzed on gel filtration column. Initially the gel filtration column on the Waters 650E advanced protein purification system contained TOYOPEARL Size exclusion media HW- 55S. This media has a fractionation range of 1000-200,000 Da. Since the lysozyme peptides are in the MW range 150-2000 Da, this media did not give the required separation and was replaced with TOYOPEARL HW- 40S media which has a fractionation range of 100-7000 Da.

Both myoglobin and vitamin B_{12} are visible as they move through the column ensuring that the column is well packed and samples are eluting evenly. Myoglobin displayed a pink band whereas vitamin B_{12} showed a yellow color band as they moved in the column and the respective colors were eventually observed in the test tubes. The plot shown below (Figure 18) was obtained when the standard sample was run through the column. Though there are 5 proteins in the standard, only two peaks were eluted from the gel filtration column. The first four largest proteins had MW > 7000 Daltons and eluted in a single peak with Mb (pink). The second peak contained vitamin B_{12}.

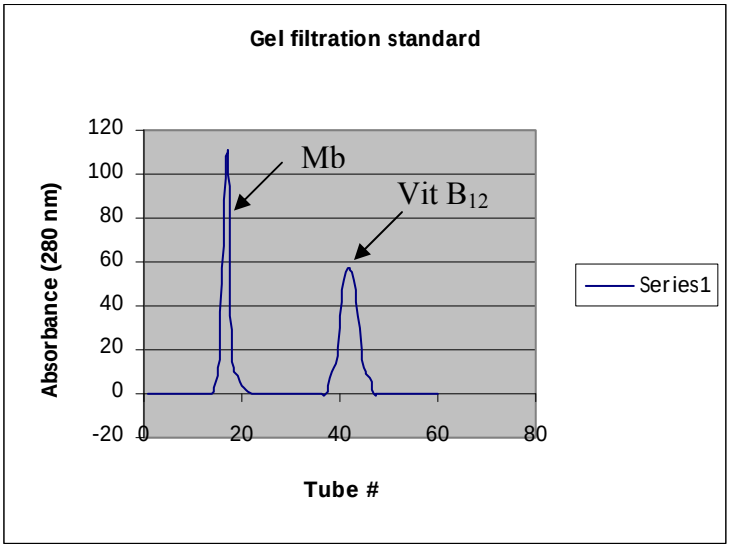


Figure 18. A gel filtration standard sample (Bio-Rad Catalogue # 151-1901: thymoglobulin, γ globulin, ovalbumin, myoglobin and vitamin B_{12}) was injected into the column with a flow rate of 0.5 mL/min in 25 mM NH_4HCO_3, pH = 4.5; absorbance was monitored at 280 nm. The peaks shown above are manually plotted based on the data and the peak observed in the actual chart and the actual chart is available with us for reference.

Once the gel filtration standard was analyzed on the gel filtration column, lysozyme was denatured and digested with trypsin and was injected on the column under the same conditions. Three different peaks eluted at times that were likely to contain peptides, whereas a huge peak observed towards the end of the chart was thought to be the peak for all the denaturants like urea, CHAPS and DTT as shown in the figure 19 below. The denaturant peak was confirmed in an identical run using only the denaturant mixture without the lysozyme peptides as the injected sample.

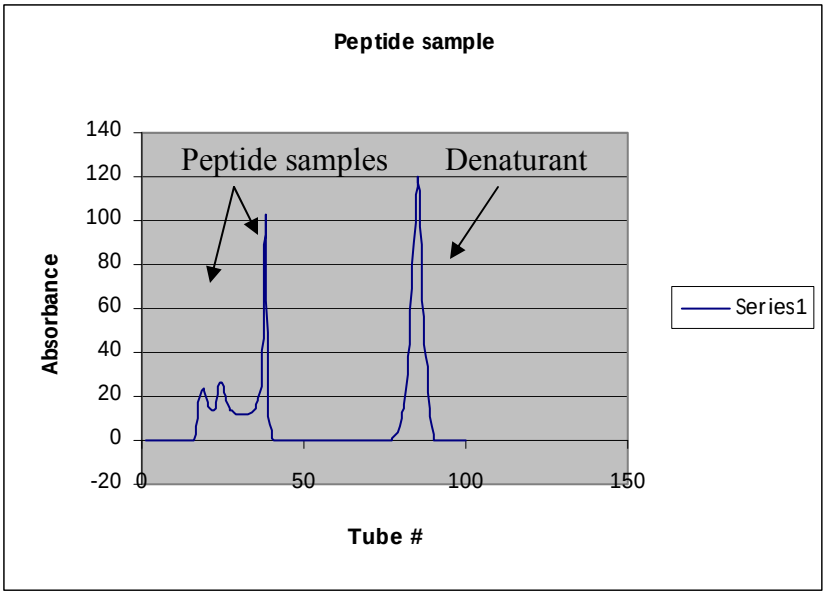


Figure 19. Lysozyme peptide peaks separated by gel filtration chromatography. The lysozyme was denatured and digested, then separated using the conditions described in figure 18.

The peptide samples from tubes # 18, 19, 23, 24, 37 and 38 were analyzed by LC/ESI-Q-TOF MS. The total ion chromatograms and the mass spectra from the tryptic digest of lysozyme from these test tubes are shown below. Different lysozyme peptides were observed on the TIC chromatogram of MS from sample in different tubes (Figure

20 a, b and c). The ESI spectrum of several lysozyme peptides were obtained by taking the averaged spectrum of each peak in the data explorer (Figure 21 a-e).

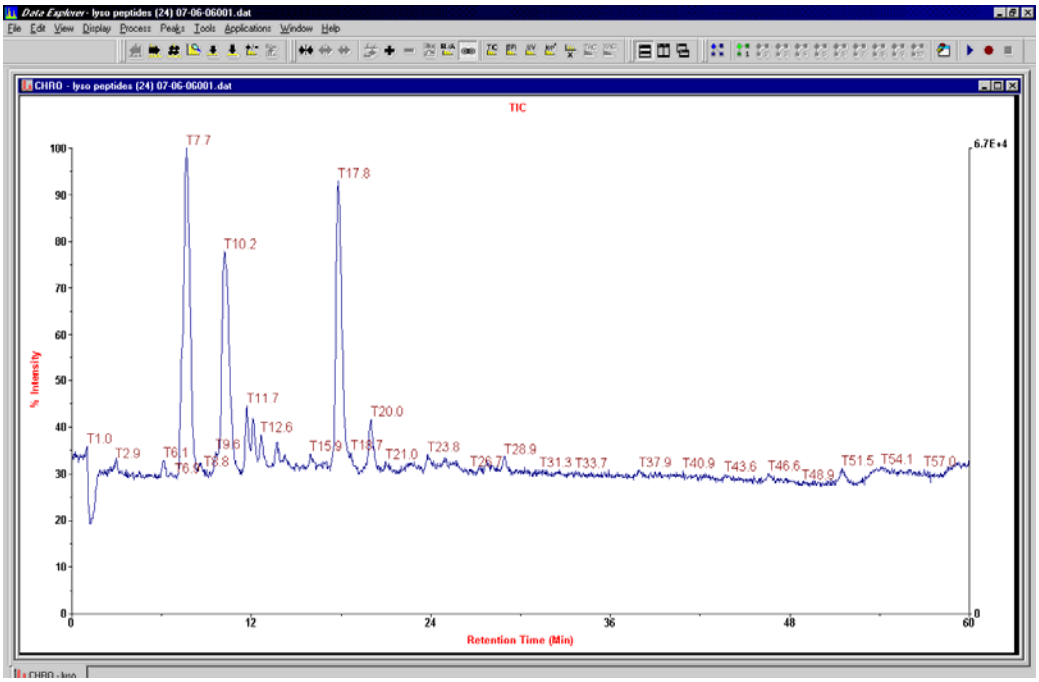


Figure 20 a

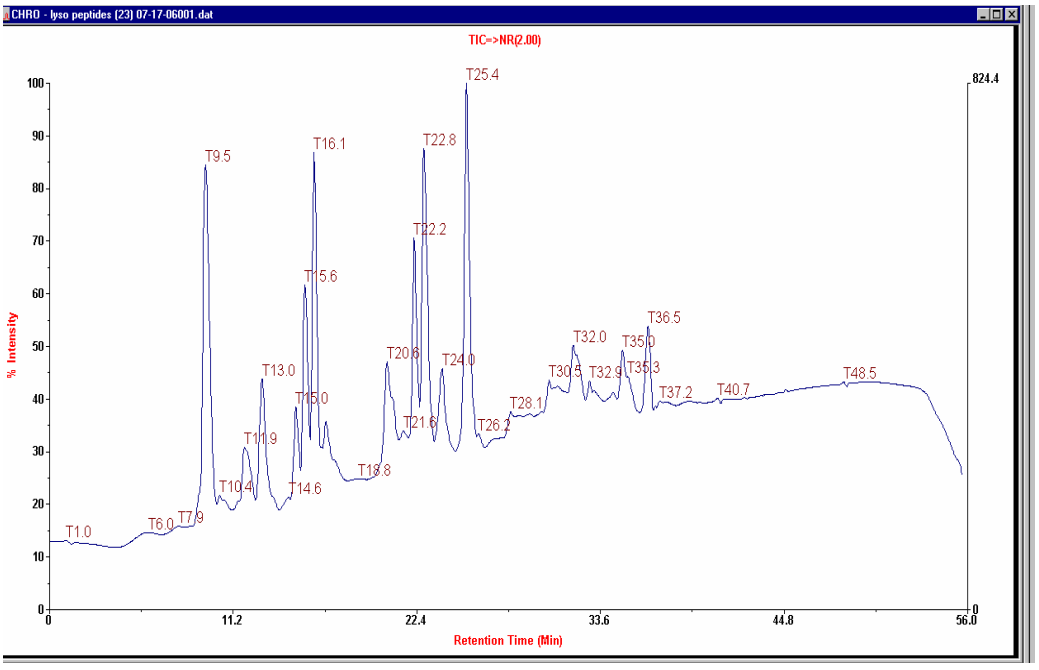


Figure 20 b

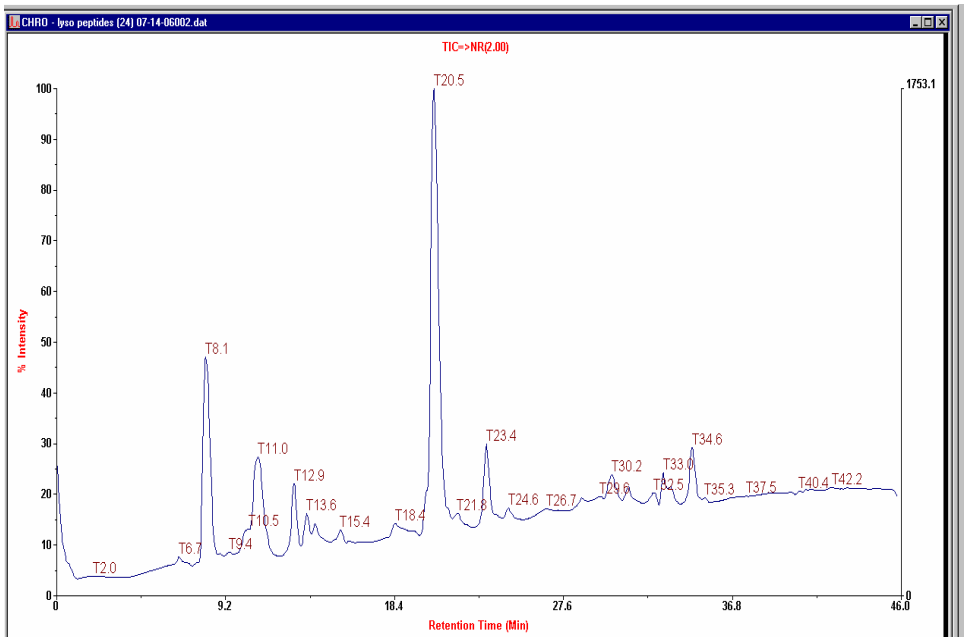


Figure 20 c

Figure 20(a-c). LC-ESI-Q-TOF MS: Total ion chromatograms (TIC) of tryptic digest of lysozyme..TIC of lysozyme peptides from samples in tube 19 (20a), 24 (20b) and 37 (20c). The mobile phase solvents used were water with 1% acetic acid (solvent A) and acetonitrile with 1% acetic acid (solvent B) . The samples were analyzed with a gradient of 2% B-40%B in 45 min, with a flow rate of 0.2 mL/min and injection volume of 20 µL.

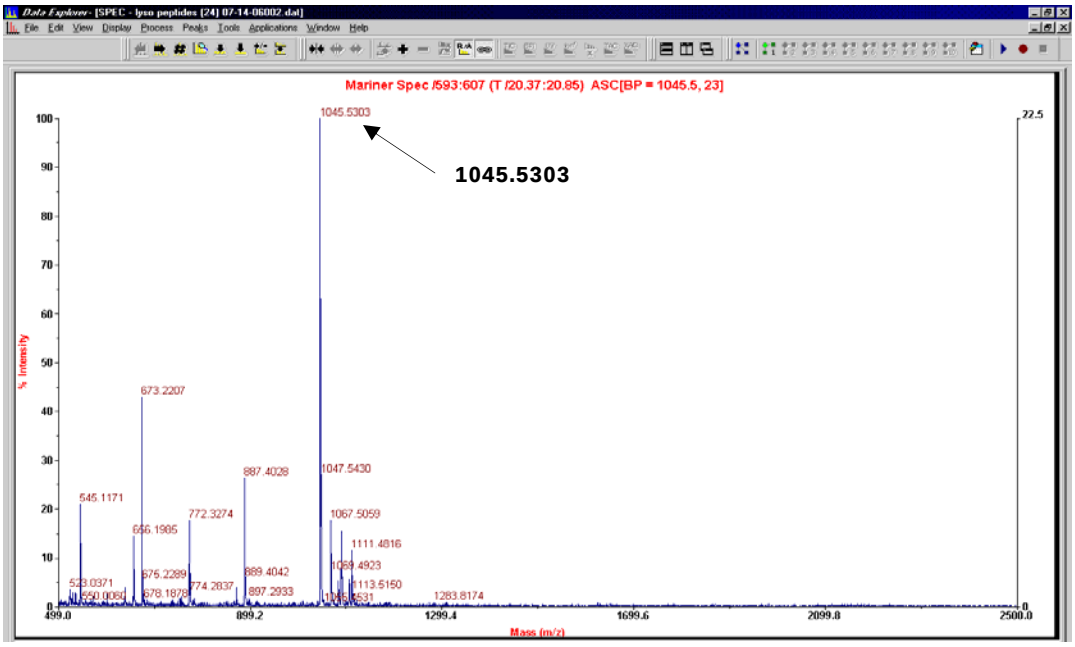


Figure 21 a

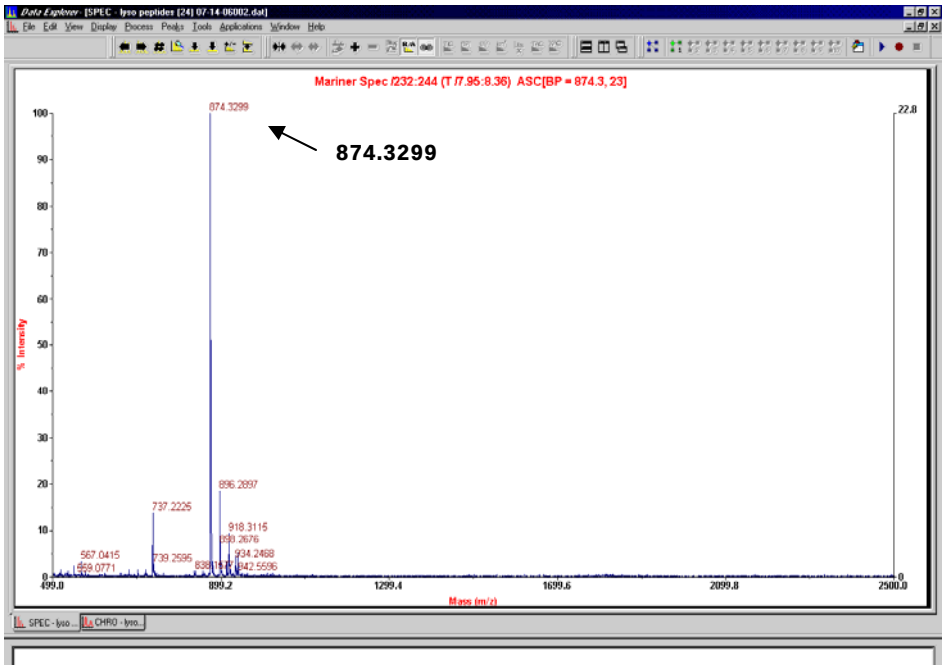


Figure 21 b

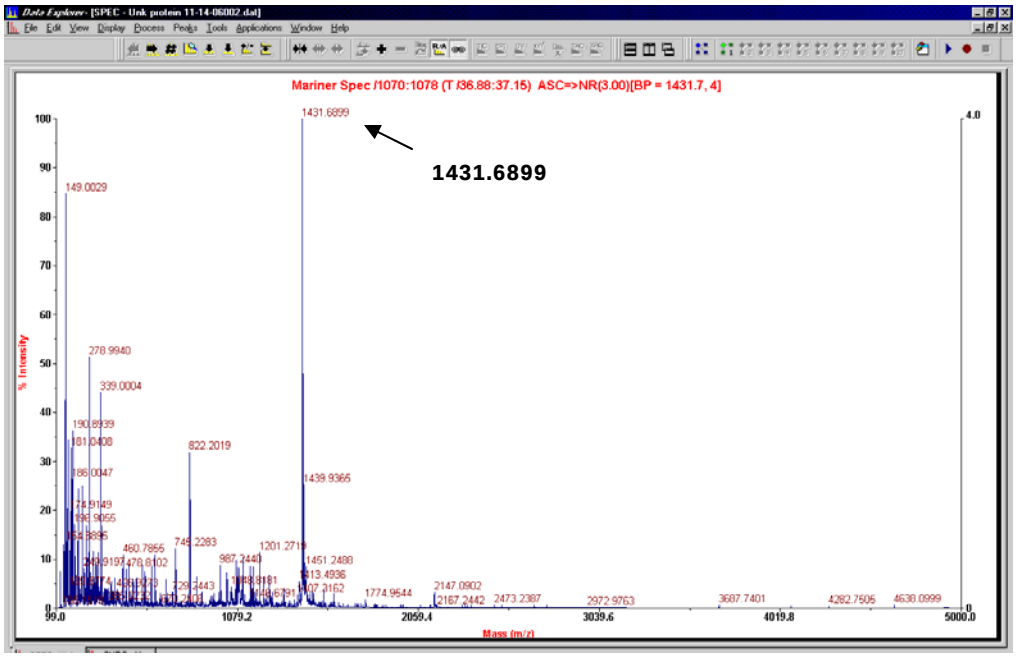


Figure 21 c

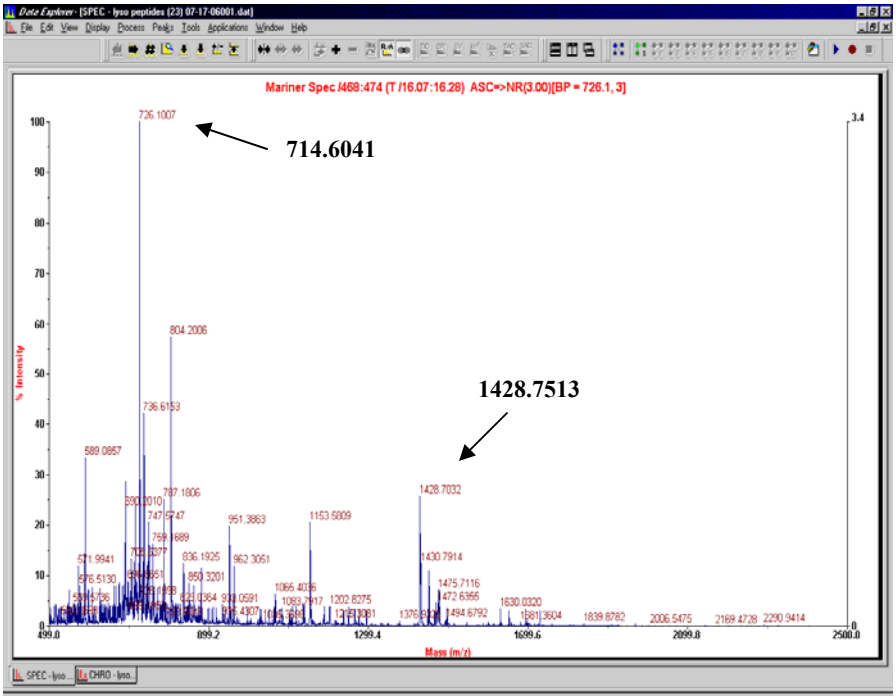


Figure 21 d

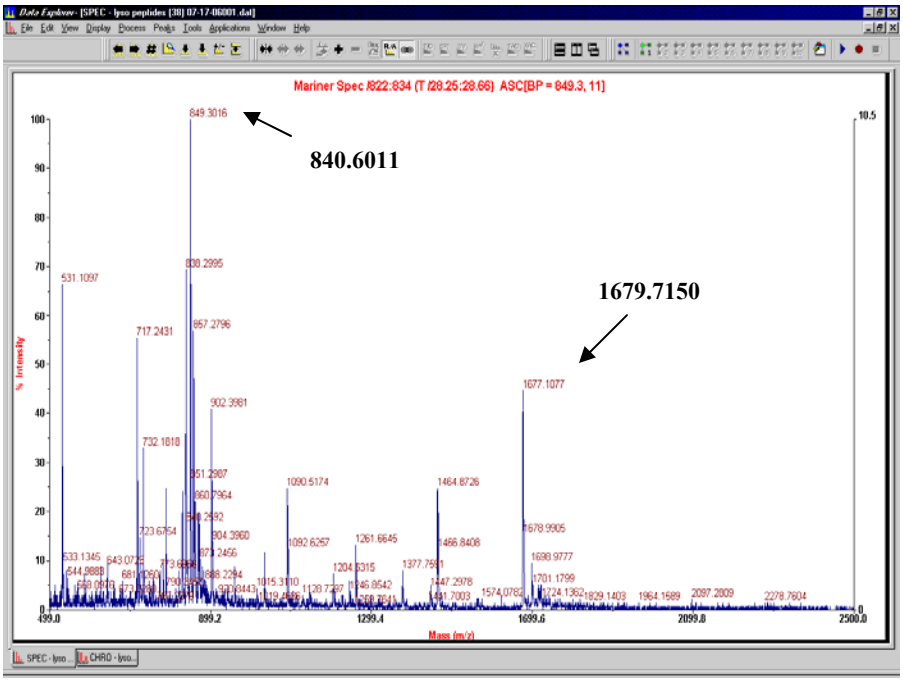


Figure 21 e

Figure 21(a-e). MS Spectra of lysozyme peptides from different samples. The mobile phase solvents used were water with 1% acetic acid (solvent A) and acetonitrile with 1% acetic acid (solvent B) . The samples were analyzed with a gradient of 2% B-40% B in 45 min, with a flow rate of 0.2 mL/min and injection volume of 20 µL.

In the three ESI spectra shown in figure 21a, b, and c, the protonated molecular ion is in the form of MH^+ and in the two ESI spectra shown in figure 21d and e, the protonated molecular ion is in the form of MH^{+2} which indicates m/z spectrum showing a charge state of +1 and +2 were observed respectively; Protein identification was based on m/z spectra with a charge state of +1.

The sequence of hen egg white lysozyme was obtained from NCBI [55] and submitted to the peptide cutter tool on Expasy [53]. The peptide cutter predicted that treatment with trypsin would give 18 peptides; it also provided the expected masses for those peptides, which were saved for comparison with experimental results.

Databases used for protein identification

The results obtained from the spectrum shown in figure 21(a-e) were compared with the database results obtained using Peptide Cutter tool on Expasy. Around 5-6 peptide masses matched the database results. The experimental values were taken from the m/z spectra (MH^+) shown in figure 21 which were then submitted in different database search engines to confirm the identity of the protein (see Table 4).

Table 4. Tryptic digestion of lysozyme: Comparison of experimental and theoretical peptide masses. Data obtained by experimental observation are compared with fragments predicted by the Expasy Peptide Cutter tool [53].

Exp. observation ((MH^+) Da	**Database results (MH^+) Da**
874.3299 Da	874.4166 Da
1045.5317	1045.5425
1428.7513	1427.6662
1679.7150	1675.8009
1754.0928	1753.8351

The five experimental masses generated by LC/ESI-Q-TOF MS were entered in the following databases to identify the protein: MASCOT [57], Profound [58], Aldente [60], (Expasy)/ MS-Fit Protein Prospector, UCSF, San Francisco, USA [56] and EMBL peptide masses [59]. Optimal search parameters were: enzyme trypsin, maximum of one missed trypsin cleavage, a window of 1-2 Da mass tolerance using monoisotopic mass, peptide charge state as (MH^+), and unmodified cysteine and methionine. All the databases identified lysozyme as the primary protein with highest score report and also displayed the matched peptide masses (experimental and database value) along with their sequences.

In Gel digestion

The final objective was to identify different protein spots on the gels using in gel digestion and LC/ESI–Q-TOF MS. To achieve this goal, it was important to generate reproducible gels with dark intense spots such that the peptides could be identified by the mass spectrometer. Prior to gel digestion, the concentration of proteins loaded on IPG strips in the gels was increased by 5-10 fold. However, when this was unsuccessful in producing peptides that could be detected by our instruments, the concentration was further increased. The IPG strips were loaded with 250-350µL of 13-20µg/µL (micrograms/microliter) proteins without any dilution by rehydration buffer. Once the gels were overloaded with protein, streaking was observed and the gels were messy, however intense spots were observed. To further improve separations, the gels were run at 4^0C (instead of room temperature) and the voltage was reduced from 200 V to 100 V resulting in a longer run time (around 3.5 hours). In all these cases, the samples were applied to the IPG strips by active rehydration.

The gels shown in figure 22 were run with the conditions described above and the spots shown by arrow and labeled as 1, 2 and 3 were selected for in gel digestion. The intensity of these spots provides quantitative information about the level of protein expression.

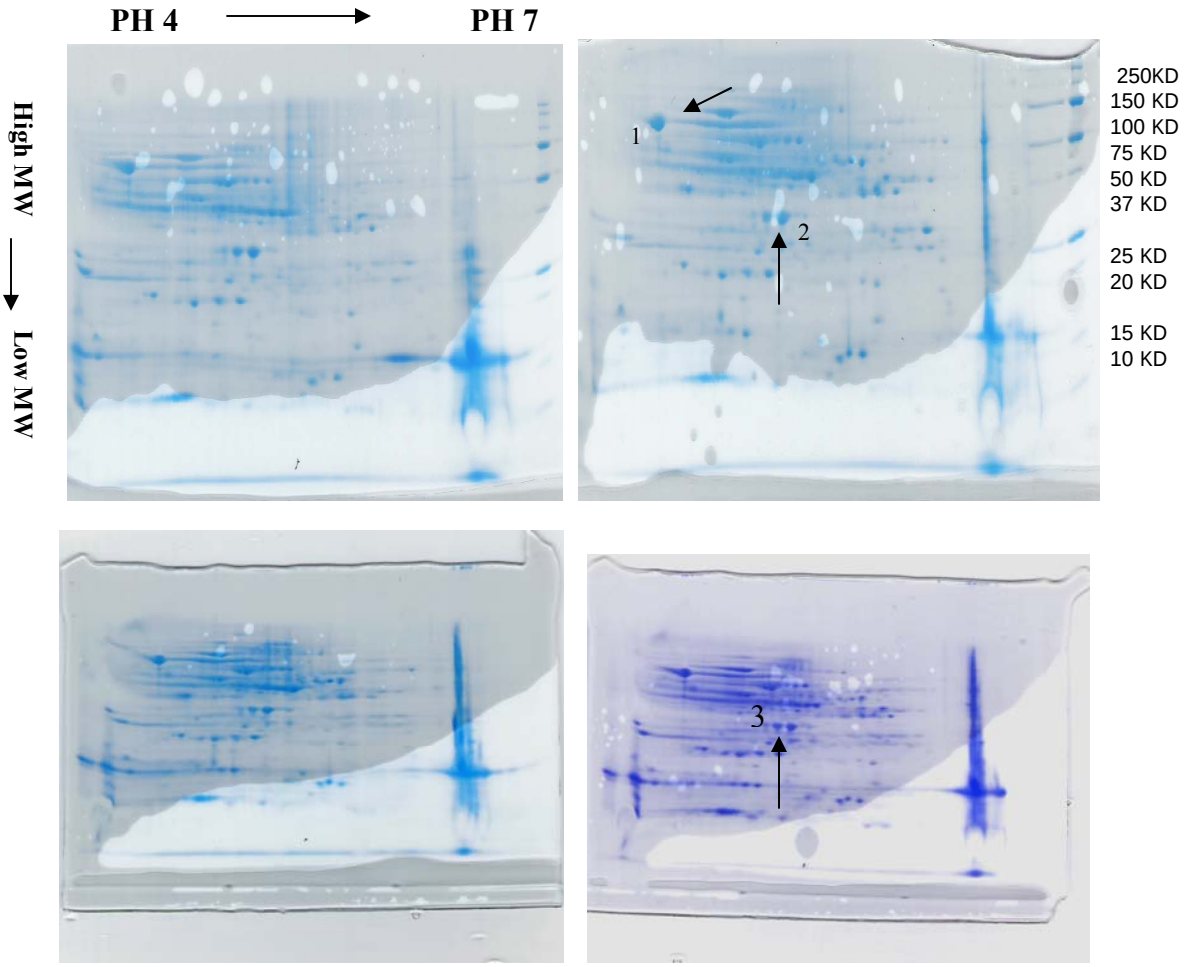


Figure 22. Commassie stained 2D gel images on 11cm pH 4-7 IPG strips with 350-450 µL of protein sample (~ 3000-4000µg); used for in gel digestion; image acquired by PD Quest 2D Analysis Software.

The molecular weights of the unknown proteins in these three spots were estimated by running standard proteins of known molecular weight in a separate lane on each gel. Spot 1 was expected to be in the MW range of 65-75 KD with an estimated pI in the range 4.0-4.5 whereas spot 2 and spot 3 were expected to be in MW range of 25-35 KD with an estimated pI range of 5.0-5.3

The spots labeled 1, 2 and 3 were excised from the gel and in gel digestion was performed as described in the experimental section. The resulting peptides were analyzed

on LC/ESI-Q-TOF MS with a flow rate of 200 µL/min, injection volume of 10µL and gradient of 2%- 50% ACN in 45 min. The samples were analyzed in triplicate and all the chromatograms were very similar. When the samples were analyzed on the LC/ESI-Q-TOF MS, the following chromatograms (Figure 23 a-c) and corresponding mass spectra (Figure 24, 25 and 26) were observed on the ESI-Q-TOF mass spectrometer.

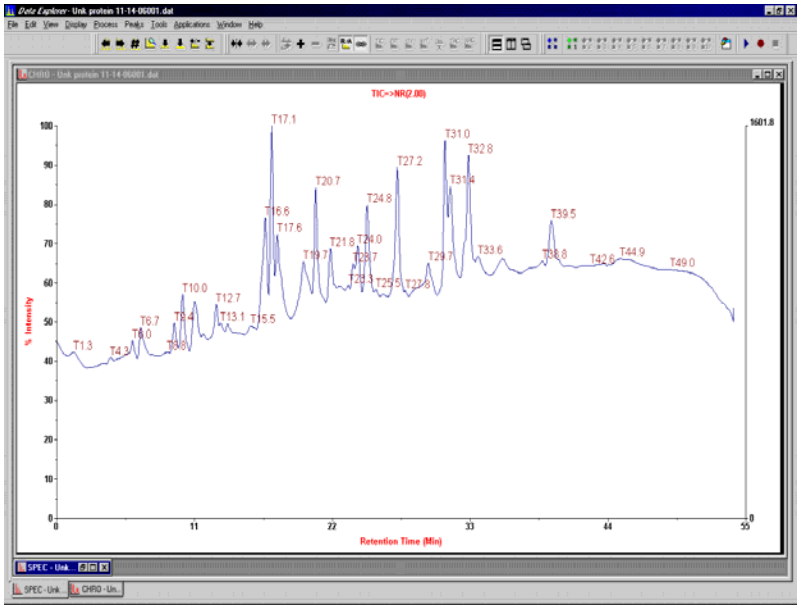


Figure 23 a

Chromatogram from Spot # 1

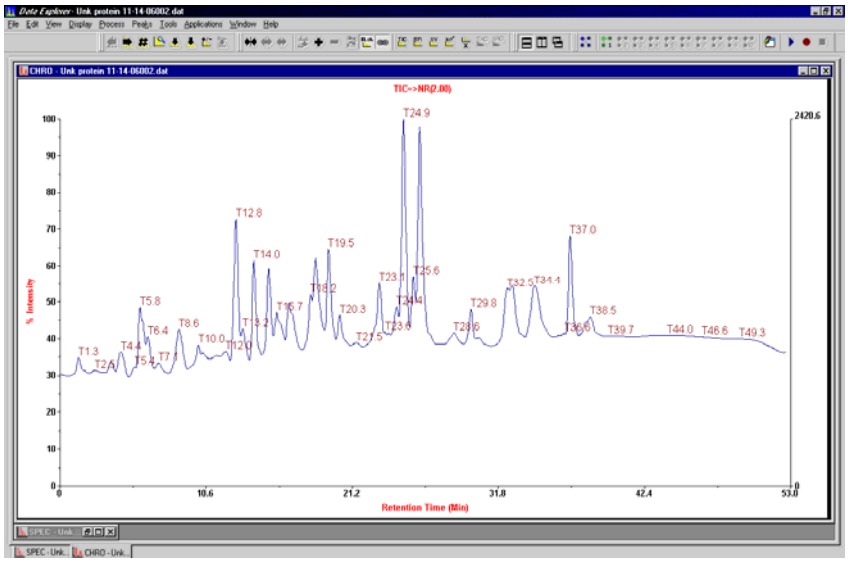


Figure 23 b

Chromatogram from Spot # 2

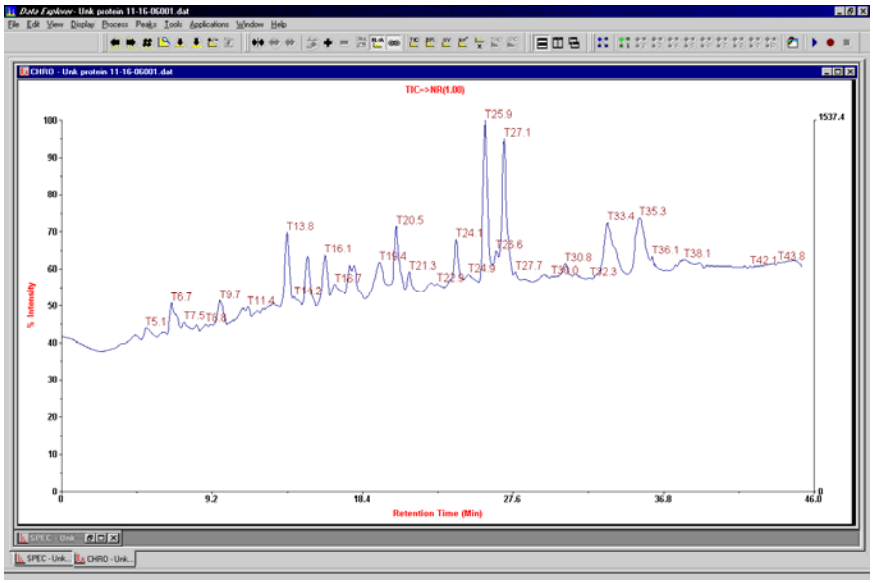


Figure 23 c Chromatogram from Spot # 3

Figure 23(a-c). The chromatograms observed for the tryptic digests of the three selected proteins. Samples were analyzed on the LC/ESI-Q-TOF MS with a flow rate of 200 µL/min, injection volume of 10µL and gradient of 2%- 50% ACN in 45 min. The samples were run in triplicate and all the chromatograms were consistent.

Spectrum from spot 1

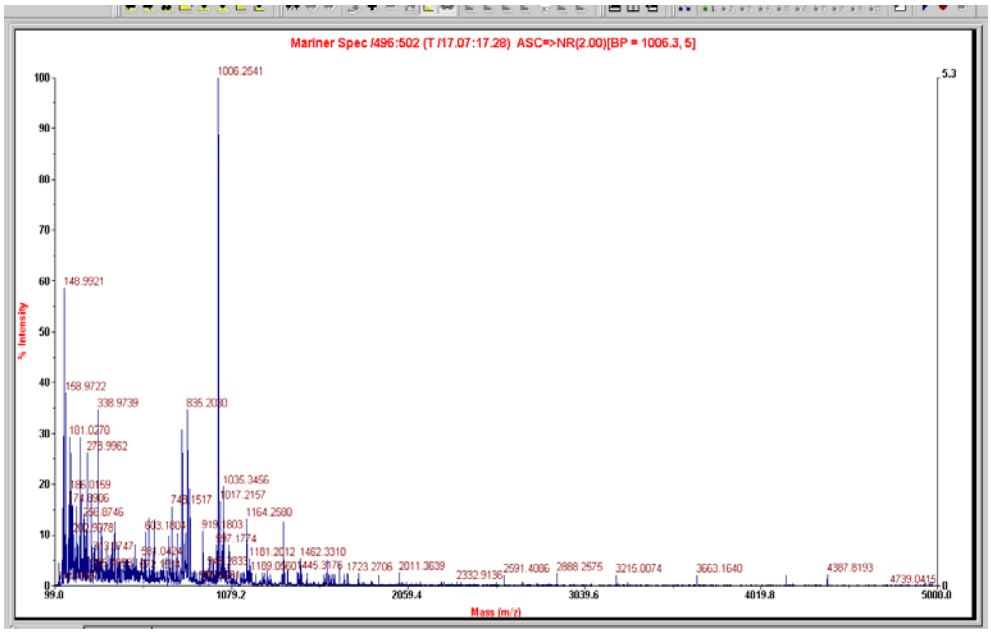


Figure 24 a

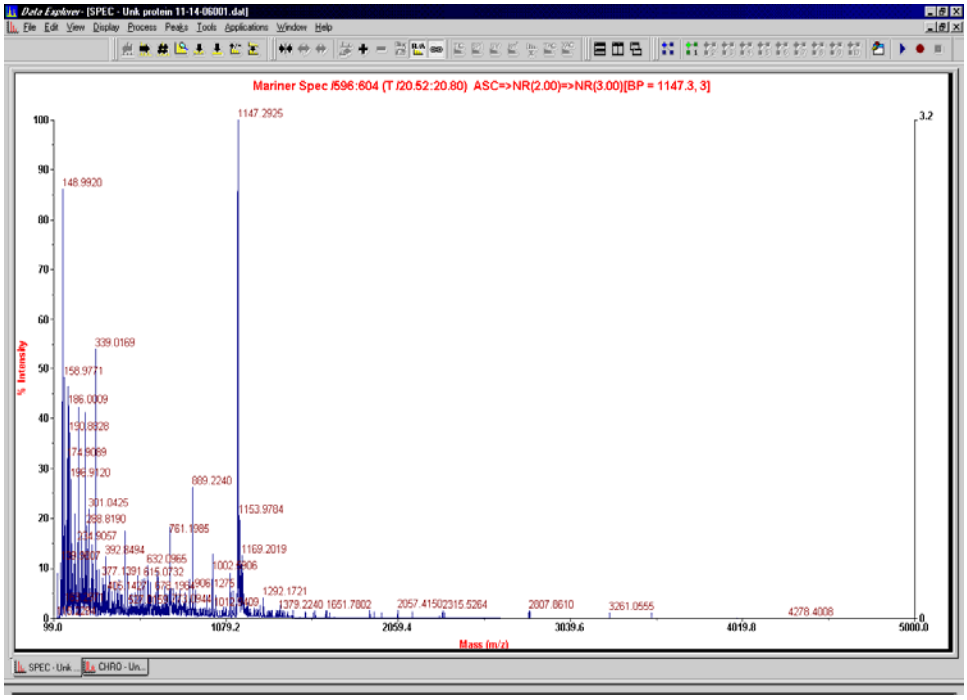


Figure 24 b

Figure 24(a,b). Spectrum from spot 1 — The ESI spectra obtained by taking the averaged spectra of different peaks in total ion chromatograms from spot 1

Spectrum from spot 2

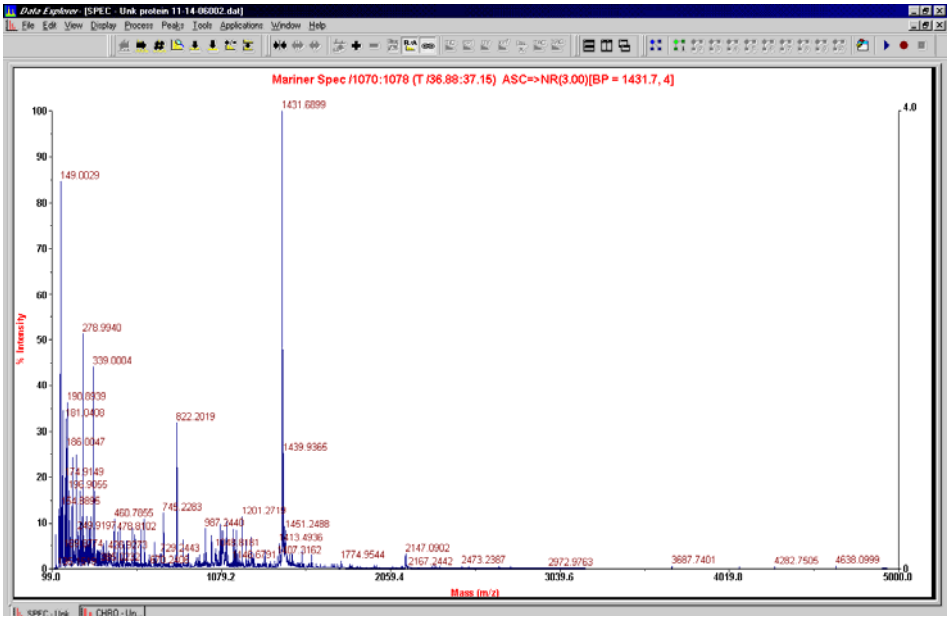


Figure 25. The ESI spectra obtained by taking the averaged spectra of different peaks in total ion chromatograms from spot 2

Spectrum from spot 3

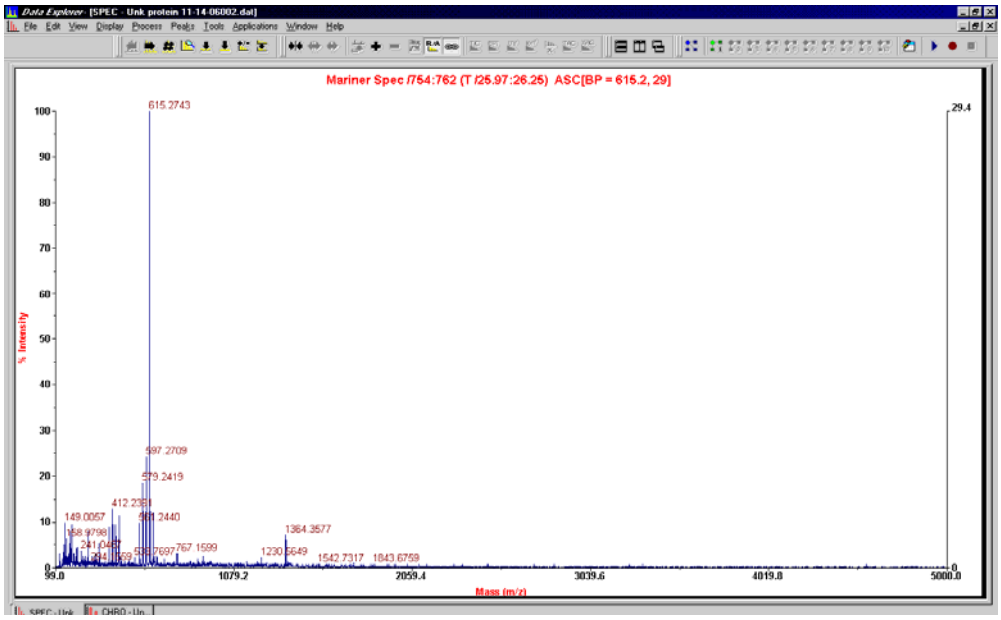


Figure 26. The ESI spectra obtained by taking the averaged spectra of different peaks in total ion chromatograms from spot 3

Database results and analysis for protein identification

The masses obtained by ESI-Q-TOF MS (Table 5) from the three spots were submitted to several databases: EMBL [59], MASCOT [57], MS-Fit [56], Aldente [60] and Profound [58]. The following search parameters were chosen: enzyme trypsin, two missed trypsin cleavage, a window of 2-3 Da mass tolerance using monoisotopic mass, peptide charge state as (MH^+), carbamidomethyl cysteine and oxidized methionine.

Table 5. Peptide masses in daltons obtained from 3 spots:

Spot 1 (Da)	Spot 2 (Da)	Spot 3 (Da)
173.9974	**615.2480**	**615.2743**
1006.2541	**656.2663**	**656.2725**
1060.8237	**788.2068**	**730.2047**
1095.2413	**1149.5596**	**1193.7680**
1147.2925	**1249.3674**	**1364.8288**
1261.3066	**1364.3509**	
1385.3818	**1431. 6899**	
1515.4917		

All of these databases except EMBL and MS-Fit identified proteins from species other than *Pseudomonas putida*. EMBL identified four proteins from *Pseudomonas putida* KT2440 with MW in the expected range (based on protein migration on the gels). MS-Fit did identify one of the 4 proteins from *P. putida* KT2440 that was identified by EMBL, but only if the taxonomy id of the species (Id: 160488) was entered; otherwise no information about proteins from *P. putida* was obtained from this search engine.

The MW of the protein identified by MS-Fit was in the expected range; however the pI of the protein was 5.6, which was not in the estimated range (4-5). When the accession number of these proteins obtained from the EMBL database was used, it was found that the pI values for all four proteins were outside the expected ranges. All the proteins displayed a pI range of 5.0-5.6 and none of the proteins was considered a match to spot #1 from the gel.

Based on this information from these peptide mass fingerprinting databases, the identities of these protein spots were considered unreliable.

Samples (spots from the gels) were then sent to the mass spectrometry facility at The Ohio State University. There, the samples were digested with trypsin and the peptides were analyzed by liquid chromatography coupled to tandem mass spectrometry (TMS), which was used for the identification of proteins. Using the MS/MS ion search tool in MASCOT, a list of proteins from *Pseudomonas putida* KT2440 were identified (Figure 27), along with detailed information about the proteins – MW, peptide masses, and the amino acid sequence of all the peptide as shown in figure 28. When all these proteins identified from *Pseudomonas putida* KT2440 were closely analyzed, based on the probability score and MW of these proteins and the pI and MW values obtained from the 2D gels, the first protein on the list (flagellin FliC) was considered the best candidate for spot 1. The values for the other proteins from *Pseudomonas putida* were not in estimated MW and/or pI range.

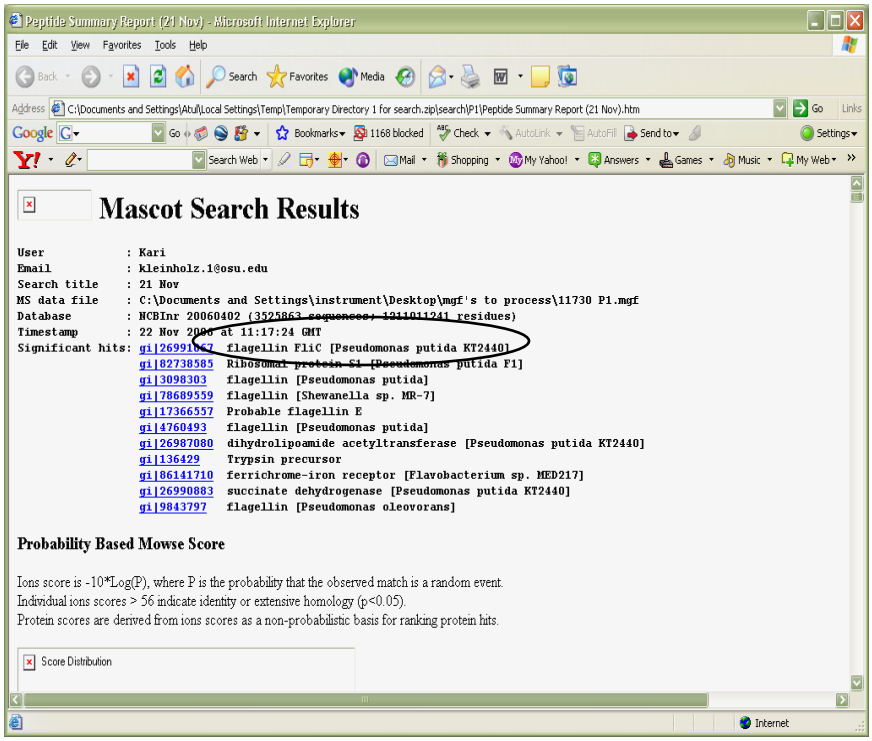


Figure 27. Mascot search engine results based on tandem mass spectrometry analysis of spot #1, with list of candidate proteins.

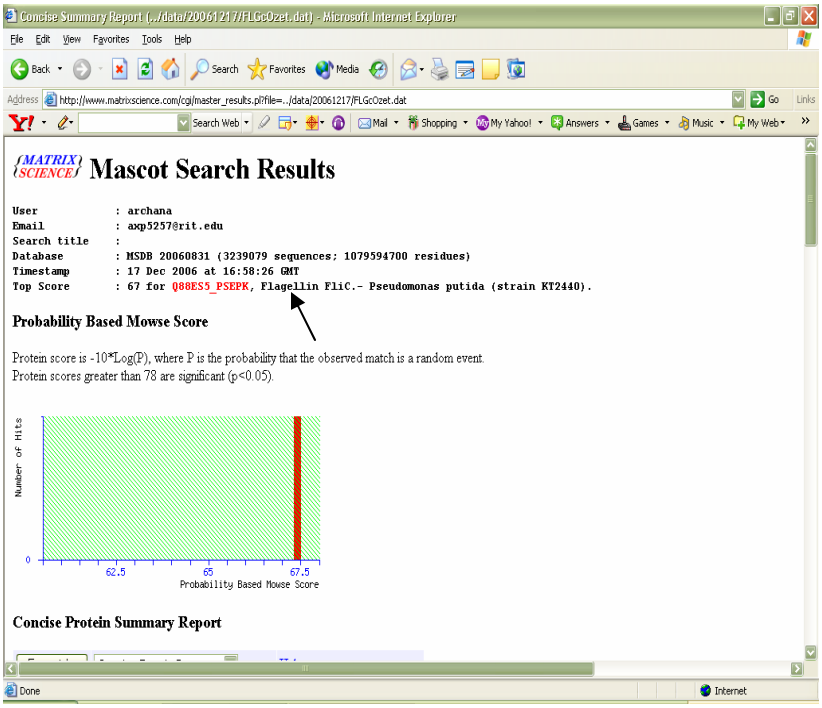
In figure 28, the observed masses are from the doubly or triply charged ions, the experimental mass is the mass calculated as neutral, and the calculated mass is the mass calculated based on the theoretical mass of the amino acid sequence.

Query	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Score	Expect	Rank	Peptide
131	673.5300	672.5227	672.4170	0.1057	0	22	2.4e+002	5	K.GLSVAVK.N
166	420.0390	838.0634	837.3902	0.6733	0	41	1.6	1	K.DVSATSMK.G
204	486.5690	971.1234	971.5148	-0.3914	0	(49)	0.44	1	R.SQLGAVQNR.F
205	972.5300	971.5227	971.5148	0.0079	0	(33)	16	1	R.SQLGAVQNR.F
206	972.5800	971.5727	971.5148	0.0579	0	(27)	61	1	R.SQLGAVQNR.F
207	487.1140	972.2134	971.5148	0.6986	0	51	0.24	1	R.SQLGAVQNR.F
228	1037.5600	1036.5527	1036.5301	0.0226	0	(21)	1.8e+002	1	R.IANTTTFGGR.N
229	1037.6300	1036.6227	1036.5301	0.0926	0	(32)	18	1	R.IANTTTFGGR.N
230	520.1040	1038.1934	1036.5301	1.6633	0	(72)	0.0016	1	R.IANTTTFGGR.N
231	520.1540	1038.2934	1036.5301	1.7633	0	(70)	0.0026	1	R.IANTTTFGGR.N
232	520.1590	1038.3034	1036.5301	1.7733	0	80	0.00028	1	R.IANTTTFGGR.N
233	520.2040	1038.3934	1036.5301	1.8633	0	(43)	1.2	1	R.IANTTTFGGR.N
262	548.1340	1094.2534	1094.5026	-0.2491	0	(80)	0.00024	1	K.SSDALGTTMGR.L
263	548.1540	1094.2934	1094.5026	-0.2091	0	85	8.5e-005	1	K.SSDALGTTMGR.L
264	548.6540	1095.2934	1094.5026	0.7909	0	(80)	0.00024	1	K.SSDALGTTMGR.L
265	548.7240	1095.4334	1094.5026	0.9309	0	(71)	0.0018	1	K.SSDALGTTMGR.L

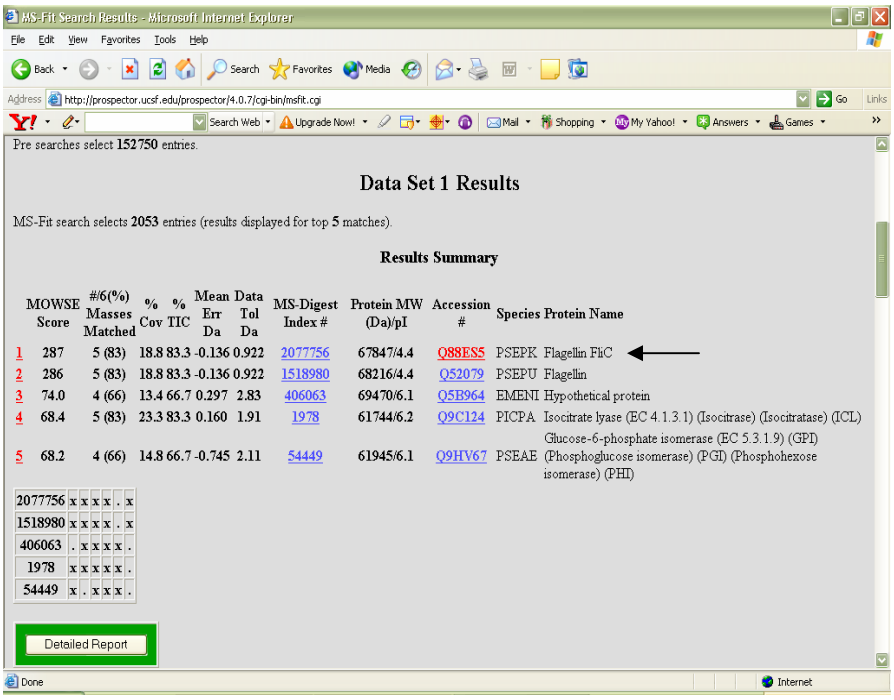
Query	Observed	Mr (expt)	Mr (calc)	Delta	Miss	Score	Expect	Rank	Peptide
268	555.9640	1109.9134	1110.4975	-0.5841	0	(68)	0.0038	1	K.SSDALGTTMGR.L + Oxidation (M)
271	556.4890	1110.9634	1110.4975	0.4659	0	(71)	0.0023	1	K.SSDALGTTMGR.L + Oxidation (M)
291	580.6140	1159.2134	1158.5629	0.6506	0	(89)	3.7e-005	1	K.DDAAGLQISNR.L
292	580.6240	1159.2334	1158.5629	0.6706	0	90	3.3e-005	1	K.DDAAGLQISNR.L
293	580.9290	1159.8434	1158.5629	1.2806	0	(89)	3.4e-005	1	K.DDAAGLQISNR.L
347	1258.5800	1257.5727	1257.6928	-0.1201	0	(32)	17	1	R.LVLTSANGQDIK.L
348	1258.6200	1257.6127	1257.6928	-0.0801	0	(35)	9.8	1	R.LVLTSANGQDIK.L
349	629.8440	1257.6734	1257.6928	-0.0194	0	49	0.41	1	R.LVLTSANGQDIK.L
350	629.8440	1257.6734	1257.6928	-0.0194	0	(44)	1.4	1	R.LVLTSANGQDIK.L
352	630.5740	1259.1334	1257.6928	1.4406	0	(43)	1.7	1	R.LVLTSANGQDIK.L
515	803.1990	1604.3834	1605.8396	-1.4561	0	(115)	9.1e-008	1	M.ALTVNTNITSMSVQK.N
516	803.3440	1604.6734	1605.8396	-1.1661	0	125	8.7e-009	1	M.ALTVNTNITSMSVQK.N
517	804.0040	1605.9934	1607.7678	-1.7744	0	(98)	4.4e-006	1	R.IQDADFAAETAELSK.Q
518	804.3090	1606.6034	1605.8396	0.7639	0	(80)	0.00027	1	M.ALTVNTNITSMSVQK.N
519	804.3440	1606.6734	1605.8396	0.8339	0	(66)	0.0069	1	M.ALTVNTNITSMSVQK.N

Figure 28. Mascot search engine results based on tandem mass spectrometry analysis of spot #1, with peptide masses and amino acid sequence data.

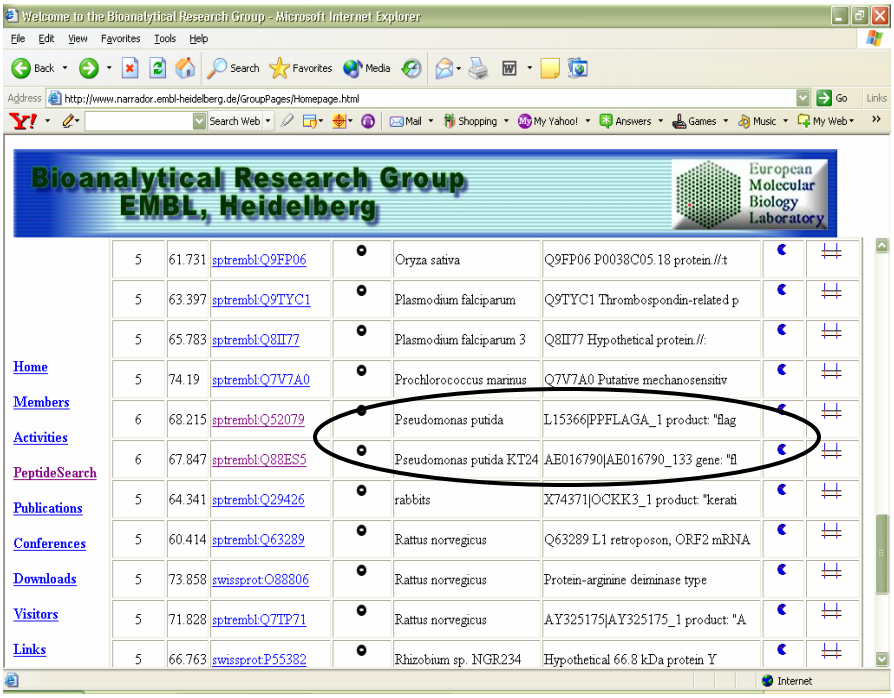
When the masses obtained from LC/ESI-Q-TOF MS were compared with the masses obtained by tandem MS/MS analysis, some of the observed masses from the LC/ESI-Q-TOF MS spectrum matched the masses obtained from Mascot search results. It was then realized that the masses from the LC/ESI-Q-TOF MS analysis were entered in the database assuming a charge of +1 on each fragment. Based on the TMS data, many of these fragments actually had charges of +1, +2 or even +3 in some cases. There was no way the charge on the fragment from the LC/ESI-Q-TOF MS could be known, unless all the three different peaks showing different charge states were observed on a single spectrum, which did not occur with our samples. After finding the charge state of the peptide masses (obtained from ESI-Q-TOF MS) based on the Mascot results (from TMS analysis), masses of the peptides were calculated as neutral. When these masses were submitted to Mascot, Prospector and EMBL, all these databases identified the protein "flagellin Flic" with the MW and pI were close to the estimated value, as shown below in figure 29(a-c).



29a : MASCOT search result



29b: MS-Fit search result



29c : EMBL search result

Figure 29(a-c). Results from MASCOT, MS-Fit and EMBL search engines for the peptide fragments from the LC-ESI-Q-TOF MS analysis, after correction for fragment charge based on the TMS analysis.

To further confirm the MW and the pI of the protein, the sequences of each flagellin Flic peptide (obtained from the Tandem MS analysis at The Ohio State University) were used to search the proteome of *P. putida* in an electrophoresis simulation program [63] as shown in figure 30. All the proteins from *Pseudomonas putida* KT2440 can be loaded and separated (based on their theoretical pI and MW values) in this simulation program. The simulation was set to perform isoelectric focusing with pH 4-7 IPG strips and SDS-PAGE on a gel containing 8-16% acrylamide.

IEF and 2D PAGE were performed as shown below in figure 30a and all the proteins in *Pseudomonas putida* was displayed in the pI range of 4-7. To identify protein based on sequence similarity, the "search protein" icon in figure 30a was pressed and the window shown in figure 30b was observed. The sequence of each of the matched peptide masses from flagellin from Mascot search result were entered and "search" icon was pressed again. In each case, a single protein spot was identified by the simulation program based on sequence identity (Figure 30c) and this spot was "flagellin FliC" protein as shown in figure 30d. The experimentally determined MW and pI obtained were very close to the expected value (MW 67856.94 and pI 4.36).

To further confirm the identify of spot #1 as flagellin Flic, the sequences of other candidate proteins from *Pseudomonas putida* KT2440 that were identified by Mascot (based on TMS analysis) and PMF (MASCOT, MS-Fit and EMBL) of our LC/ESI-Q-TOF MS data were collected and submitted to the electrophoresis simulation. In some cases, no proteins were identified by the electrophoresis simulation program [63]. Any proteins that were identified by the simulation were not in the right MW and pI range.

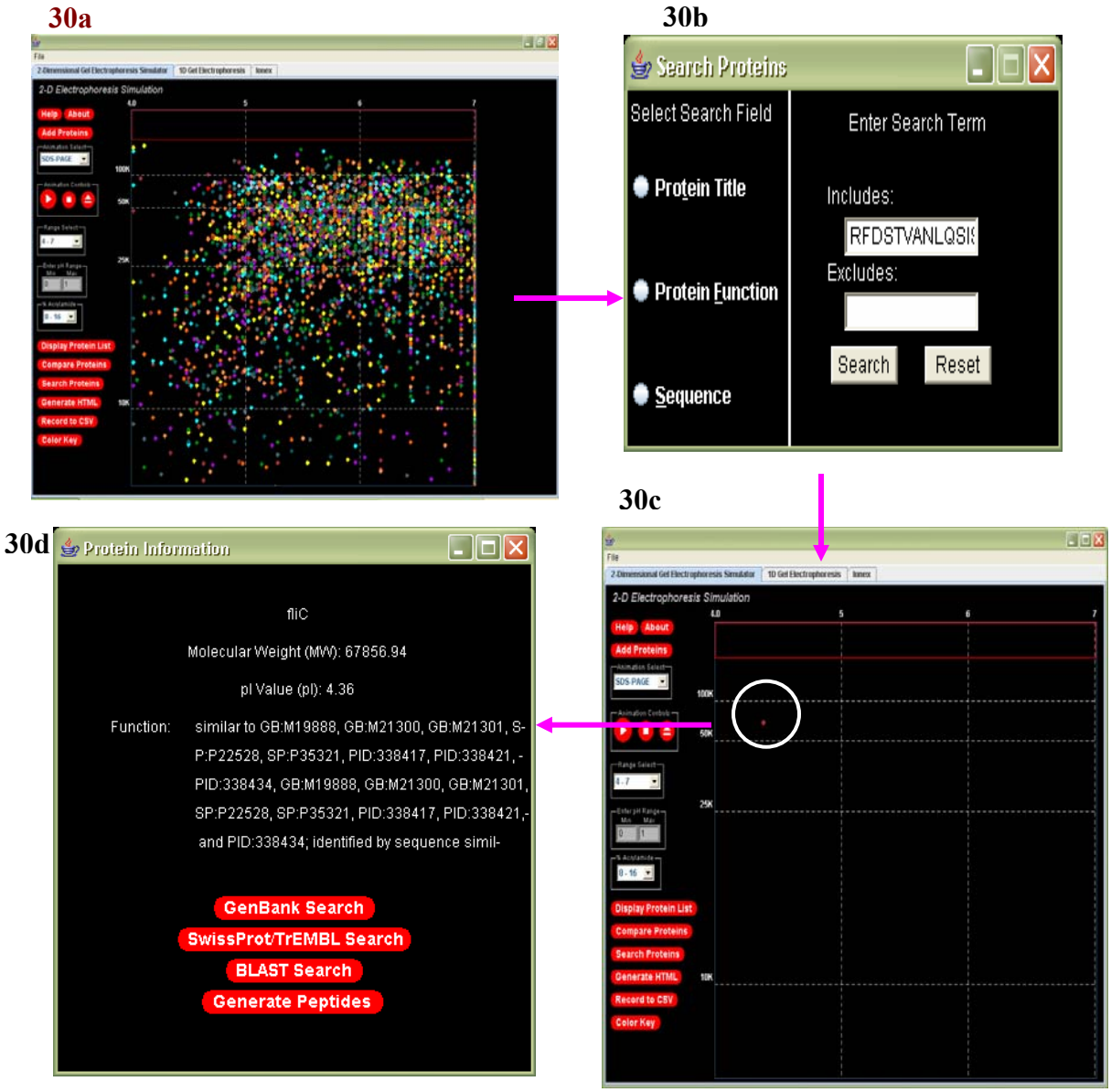


Figure 30(a-d). Results obtained by searching the electrophoresis simulation program [63] using the peptides generated by tandem MS-MS of spot 1. Figure 30a displays all the proteins in *Pseudomonas putida* in the pI range of 4-7 when simulation was set to perform isoelectric focusing with pH 4-7 IPG strips and SDS-PAGE on a gel containing 8-16% acrylamide. When the "search protein" icon in figure 30a was pressed to identify proteins based on sequence similarity, the window in figure 30b was observed. The sequence of each of the matched peptide masses from flagellin from Mascot search result were entered and "search" icon was pressed as shown in figure 30b. A single protein

spot was identified by the simulation program based on sequence identity (Figure 30c) and this spot was "flagellin FliC" protein as shown in figure 30d.

Based on these results from the PMF (from the LC/ESI-Q-TOF MS data) and Mascot search engine result (from the tandem MS-MS data) , and based on MW, pI and probability score, the most probable choice for spot 1 protein from *Pseudomonas putida* KT2440 was "**flagellin FliC"** with **MW 67856.94 and pI 4.36.**

The same procedure was applied to the analysis of spots 2 and 3. Initially reliable results could not be obtained from the PMF search engine that were consistent with our experimentally determined MW and pI values of the proteins, due to the confusion over the charge states on peptide fragments from the LC-ESI-Q-TOF MS results.

For spot 2, when the masses (Table 5) were entered in the database, EMBL and MS-Fit identified only one protein from *P. putida* KT2440 (hydroxyl methyl pyrophosphate reductase). The pI value for the protein was in the expected range; however the MW of the protein was observed to be somewhat higher than the experimentally determined MW from the gel. No other database gave any results from *Pseudomonas putida* for spot #2.

Analysis of spot #2 by tandem MS/MS followed by a Mascot search pointed to elongation factor Ts as the best candidate. The charge states of the peptide masses obtained from ESI-Q-TOF MS were then corrected based on the TMS search results and the mass of each peptide was calculated as neutral. When these masses were submitted to the databases, spot 2 was identified as elongation factor Ts with MW 30426.11 Da` and pI 5.18, which agreed with the experimentally determined values.

The sequences from the tandem MS-MS were then submitted to the electrophoresis simulation program [63] to confirm the MW and pI of protein. Based on highest score (according to matched peptide masses), MW and pI of protein, the best matching protein for spot 2 is **elongation factor Ts** (**MW 30426.11 Da` and pI 5.18)** from *Pseudomonas putida* KT2440.

A similar approach was used with peptides extracted from spot 3. Based on TMS search engine results and electrophoresis simulation program [63] and also according to the peptide score match, MW and pI value of protein, and spot 3 was identified as **electron transfer flavoprotein with MW 31241 Da and pI 5.12.**

Based on the experimentally determined MW and pI values and the results from PMF search engines tools (MASCOT, MS-Fit and EMBL) and TMS search engines (Mascot search engine result from the tandem MS-MS data), and highest score of matched peptide masses, the most probable choice for these proteins from the three different spots in *Pseudomonas putida* KT2440 are: Spot 1: **Flagellin FliC** with MW 67856.94 and pI 4.36; Spot 2: **Elongation factor Tsf** (MW 30426.11 Da` and pI 5.18); Spot 3: **Electron transfer flavoprotein** (MW 31241 Da and pI 5.12).

Conclusions

Proteomic analysis using 2DE and HPLC coupled to ESI-Q-TOF MS (LC/MS) is a very popular method for protein identification. In the past few years, the introduction of electrospray ionization has made mass spectrometry a vital tool in the research for studying proteins and peptides for proteomics.

The objective of this research work was to analyze proteins from the soluble fraction of the bacteria *Pseudomonas putida* KT2440 strain using 2-D gel electrophoresis coupled to electrospray ionization quadrupole time of flight mass spectrometry (ESI-Q-TOF MS) mass spectrometry. Protein spots were excised from the 2DE gels and digested using trypsin. The resulting peptides were analyzed by LC/ESI-Q-TOF MS with subsequent online database analysis using the peptide mass fingerprinting (PMF) search engine to identify the protein.

The results obtained from the databases were not in the expected MW and pI range and could not be used to identify the proteins from *P putida* KT2440. Therefore, samples were sent to The Ohio State University for analysis by tandem mass spectroscopy (TMS), which yielded peptide masses and sequence information. These results were used to search the Mascot database, yielding a series of protein candidates. It was interesting to note that some of the observed masses from the TMS data were in agreement with the experimental peptide masses obtained from ESI-Q-TOF MS. Later it was realized that the masses entered in the database which were obtained from ESI-Q-TOF MS (which were originally assumed to have a charge of +1) actually had a mixture of charge states (+1, +2, and even +3 in some cases). The ESI-Q-TOF mass spectrometer did not have sufficient resolution to separate different peaks with different charge states on the spectra. The sensitivity and resolution of LC/ESI-Q-TOF MS was also not

adequate to distinguish peaks and provide confirmed results. The resolving power of LC/ESI-Q-TOF MS is 5000 (50% centroid) whereas newer and more capable ESI-Q-TOF instruments easily gets a resolution of 10,000 (100% centroid) [17] and instruments with this resolution are able to assign different charge states to the peptides [26].

After calculating the charge state of the peptide masses (obtained from ESI-Q-TOF MS) as neutral, the masses were submitted to the databases, which were then able to identify the proteins from *P. putida* KT2440 with MW and pI values close to the experimentally determined values from the 2DE gels. These results were further confirmed by an electrophoresis simulation program, which included a search tool for sequence identity. These results were also confirmed based on the pI value when compared to the Expasy database [53]. Based on the results from the PMF and the TMS search engines and based on the estimated MW, pI and highest score of matched peptide masses, the most probable choice for the proteins from the three different spots in *Pseudomonas putida* KT2440 are: Spot 1: **Flagellin FliC** with MW 67856.94 and pI 4.36; Spot 2: **Elongation factor Tsf** (MW 30426.11 Da` and pI 5.18); Spot 3: **Electron transfer flavoprotein** (MW 31241 Da and pI 5.12).

The key learning in this project was to understand the working of LC/ESI-Q-TOF MS, analyze the spectrum and interpret the data coming from the mass spectrometer. The operation of the LC/ESI-Q-TOF MS instrument and analysis of its output were first explored using the known protein lysozyme prior to working with unknown proteins in the 2D gel.

It has been found that poor confidence in peptide mass fingerprinting identities can occur when a mixture of proteins or other contaminants is present in the original sample. On the other hand, when the sensitivity and resolution of the instrument is

sufficient to obtain reliable mass accuracy with samples that have been properly prepared and are relatively pure, then peptide mass fingerprinting can yield high confidence protein identifications. In cases where the results from peptide mass fingerprinting data do not provide a high level of confidence, it is necessary to use tandem mass spectrometry (MS/MS) to validate the results [64].

Future developments in ESI-Q-TOF MS (LC/MS) techniques should focus on improving resolution, sensitivity and dynamic range. Development of such methods is expected to help in direct and rapid analysis of complex protein mixtures of the entire proteome of *P. putida* KT2440. Success in such method development will lead to improvements in the accuracy and speed with which the data can be analyzed. Although complete coverage of the proteins for any organism has yet not been accomplished, it is believed that continued advancement of such techniques one day will enable researchers to identify all proteins expressed in a cell.

In general, it is good strategy to use peptide mass fingerprinting as the first step when attempting to establish protein identity by using mass spectrometry. Poor confidence in the PMF search results can occur either because of contaminants in the sample or because of poor resolution of the instrument. In that case, tandem MS/MS must be able used to validate the result.

References

1. Y. H Kim, K. Cho, S.H Yun, J.Y. Kim, K. H. Kwon, J.S.Yoo and S. II. Kim , "Analysis of aromatic catabolic pathways in *Pseudomonas putida* KT2440 using a combined proteomic approach: 2DE/MS and cleavable isotope coded affinity tag analysis", *Proteomics*, 2006, **6**, 1301-1318.
2. L. Kurbatov, D.Albrecht, H. Herrmann and L. Petruschka in , "Analysis of the proteome of *Pseudomonas putida* KT2440 grown on different sources of carbon and energy", *Environmental Microbiology,* 2006, **8**, 466-478.
3. E. A. Kim, J.Y. Kim, S. J. kim, K. R. Park, H. J. Chung, S. H. Leem, S. II. Kim, "Proteome analysis of *Acinetobacter lwoffii* K24 by 2D gel electrophoresis and electrospray ionization quadrupole- time of flight mass spectrometry," *Journal of Microbiological Methods*, 2004, **57 ,** 337-349.
4. C. Delahunty and J. R. Yates III, "Protein identification using 2D-LC-MS/MS," *Methods,* 2005*,* **35**, 248-255.
5. Y. Ishihama, "Proteomic LC–MS systems using nanoscale liquid chromatography with tandem mass spectrometry", *Journal of Chromatography A ,* 2005*,* **1067**, 73-83.
6. J. P. Chervet, G. Mitulovic, E. Varesio and R. Locher, "Comprehensive 2-D capillary/nano LC/MS for complex proteomics", *Pathology - Research and Practice"*, 2004, **200**, 281-282.
7. R. L Beardsley and J. P Reilly, "Quantitation Using Enhanced Signal Tags : A technique for Comparative Proteomics", *Journal of Proteome Research,* 2003, **2**, 15-21.
8. S. P. Gygi and R. Aebersold. "Mass spectrometry and proteomics", *Current Opinion in Chemical Biology,* 2000, **4**, 489-494.

9. S. L. Cockrill, K. L. Foster, J. Wildsmith, A. R. Goodrich, J.G. Dapron, T. C. Hassell, W.K. Kappel and G. B. I. Scott, " Efficient microrecovery and guanidination of peptides directly from MALDI target spots," *Biotechniques* , 2005, **38**, 301-304.

10. H. J. Issaq, "The role of separation science in proteomics research", *Electrophoresis,* 2001, **22,** 3629-3638.

11. D. E. Garfin, "Two-dimensional gel electrophoresis: an overview", *Trends in Analytical Chemistry*, 2003, **22**, 263-272.

12. S. B. Giorgianni, "Proteome analysis by two-dimensional gel electrophoresis and mass spectrometry: strengths and limitations", *Trends in Analytical Chemistry*, 2003, **22**, 273-281.

13. S. J. Fey and P. M. Larsen, "2D or not 2D", *Current Opinion in Chemical Biology*, 2001, **5**, 26 – 33.

14. 2D Electrophoresis for Proteomics: A methods and product manual. BioRad Laboratories, publication 2641, Bulletin 2651, US/EG, revision B, 01-769, Feb 2002.

15. G. Paul and H. Timothy, "Molecular biologist's guide to proteomics", *Microbiology and Molecular Biology Review*, 2002, **66**, 39-63.

16. J. M. Berg, J. L. Tymoczko and L. Stryer, *Biochemistry*, 5th edition, W.H Freeman and Company; New York, NY, 2001.

17. Mariner ™ biospectrometry ™ workstation users guide, copyright 1999, 2001, Applied Biosystems; http://docs.appliedbiosystems.com/pebiodocs/00104518.pdf (accessed 05-17-2006).

18. D.C. Harris, *Quantitative Chemical Analysis* (sixth edition), W.H Freeman and Company; New York, NY, 1999.

19. M.T. Davis, J. Beierle, E.T. Bures, M.D. McGinley, J. Mort, J. H. Robinson, C.S. Spahr, W. Yu, R. Luethy and S.D. Patterson, "Automated LC–LC–MS–MS platform using binary ion-exchange and gradient reversed-phase chromatography for improved proteomic analyses", *Journal of Chromatography B: Biomedical Sciences and Applications,* 2001, **752**, 281-291.
20. K.E. Nelson, C. Welnel, I.T. Paulsen, R.J Dodson, H. Hillbert, V.A.P. Martins dos Santos, D.E. Fouts, S.R. Gill, M. Pop, M. Holmes, L. Brinkac, et al, "Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440", *Environmental Microbiology* , 2002, **4** , 799-808.
21. Inserts for 2 mL standard opening vial obtained from Sigma-Aldrich; http://www.sigmaaldrich.com/catalog/search/ProductDetail/SUPELCO/24707 (accessed 05-17-2006).
22. Toyopearl® Size Exclusion Media HW-40C for Gel filtration Chromatography applications obtained from Sigma-Aldrich; http://www.sigmaaldrich.com/catalog/search/ProductDetail/SUPELCO/807449 (accessed 06-23-2006).
23. Concept of Genomics and Proteomics ; http://www.bio-pro.de/en/life/thema/01950/index.html (accessed 05-14-2005).
24. Matrices used in Matrix assisted Laser Desorption Ionization (MALDI) ; http://www.jic.bbsrc.ac.uk/services/proteomics/maldi.htm (accessed 06-26-2006).
25. Application of Mass Spectrometry in proteomics; http://www.bmsf.unsw.edu.au/training/BIOT3061/3061MS02.pdf (accessed 07-05-2006).

26. Mass Spectrometry Tutorial – Basic concepts of mass spectrometry; http://www.mc.vanderbilt.edu/msrc/tutorials/ms/4.htm (accessed 07-08-2006).

27. D.A. Skoog , F.J. Holler , T.A. Niemann, *Principles of Instrumental Analysis* (fifth edition) ; Harcourt Brace college ; Philadelphia, PA, 1998.

28. Structure of amino acids used to synthesize proteins; http://chemed.chem.purdue.edu/genchem/topicreview/bp/1biochem/amino2.html (accessed 07-18-2005).

29. ESI-Q-TOF Schematic; http://www.sciner.com/MCP/c_mcp.htm (accessed 09-15-2006).

30. EMBL - Database for protein identification by peptide mass data - prepared by EMBL bioanalytical research group; http://www.narrador.embl-heidelberg.de/GroupPages/PageLink/peptidesearchpage.html (accessed 09-05-2006).

31. Latest trend in the area of Proteomics; http://www.sciencemag.org/products/proteomicsnew.dtl (accessed 07-28-2006).

32. An introduction to Electrophoresis; http://homepages.gac.edu/~cellab/chpts/chpt4/intro4.html (accessed 08-01-2006).

33. In gel digestion kit from PIERCE ; http://www.piercenet.com/Proteomics/browse.cfm?fldID=21518847-2D72-475F-A5B9-B236EC5B641E#gel%20electrophoresis (accessed 08-01-2006).

34. Flow of biological information form DNA to proteins ; http://www.schoolscience.co.uk/content/5/chemistry/proteins/images/p40fig2.gif (accessed 08-09-2006).

35. From the Genome to the Proteome : Basic Science ; http://www.ornl.gov/sci/techresources/Human_Genome/project/info.shtml (accessed 09-05-2006).
36. Reverse Phase HPLC basic for LC/MS - an ion source tutorial; http://www.ionsource.com/tutorial/chromatography/rphplc.htm (accessed 10-10-2006).
37. Principles of Gel Filtration Chromatography; http://www.edvotek.com/pdfs/108.pdf (accessed 09-15-2006).
38. Curved microchannel plates in reflectron Time of Flight mass spectrometer provides better mass resolution http://www.sciner.com/MCP/c_mcp.htm (accessed 10-16-2006).
39. An introduction to mass spectrometry; http://www.astbury.leeds.ac.uk/facil/MStut/mstutorial.htm (accessed 08-25-2006).
40. Time-of-Flight (TOF) mass analysis ; http://www.chm.bris.ac.uk/ms/theory/tof-massspec.html (accessed 10-13-2006).
41. Mass spectrometry from Wikipedia; http://en.wikipedia.org/wiki/Mass_spectrometry (accessed 11-21-2006).
42. Protein Digests: a Comparative Analysis; Separating Protein (Lactoperoxidase) by reversed- phase HPLC; http://www.vydac.com/vydacpubs/AN9602/AN9602.html (accessed 10-17-2005).
43. General overview and theory of mass spectrometry; http://docs.appliedbiosystems.com/pebiodocs/00104518.pdf (accessed 07-08-2006).

44. Waters 650E Advanced Protein Purification System (Quick Start / user's guide) Millipore Corporation, 34 Maple street Milford, MA 01757. Manual Number 175-02TP/04TP, Revision 0 September 1993.

45. X.Chen, S. W. Cushman, L. K. Pannell and S. Hess, " Quantitative proteomic analysis of the secretory proteins from the rat adipose cells using a 2D liquid chromatography-MS/MS Approach", *Journal of Proteome Research*, 2005, **4,** 570-577.

46. B.K. Choi, D.M. Hercules, T. Zhang and A.I. Gusev, "Comparison of Quadrupole Time-of-Flight, and Fourier Transform Mass analyzers for LC-MS Applications", *Current Trends in Mass Spectrometry*, 2003, **18**, 24-39.

47. E. P. Romijin, J. Krijgsveld and A.J.R. Heck, " Recent liquid chromatographic-(tandem) mass spectrometric applications in proteomics", *Journal of Chromatography A,* 2003, **1000**, 589-608.

48. R.S. Johnson, M.T. Davis, J.A. Taylor and S.D. Patterson, " Informatics for protein identification by mass spectrometry", *Methods,* 2005, **35**, 223-236.

49. K. Fujii, T. Nakano, H. Hike, F. Usui, Y. Bando, H. Tojo and T. Nishimura, "Fully automated online multi-dimensional protein profiling system for complex mixtures", *Journal of Chromatography A,* 2004, **1057**, 107-113.

50. V. Dhingra, M.Gupta, T.Andacht and Z. F Fu , "New frontiers in proteomics research : A perspective", *International journal of pharmaceutics,* 2005, **299**, 1-18.

51. I. Lefkovits, "Functional and structural proteomics: a critical appraisal", *Journal of Chromatography B,* 2003, **787**, 1-10.

52. T.C. Hunter, N.L. Andon, A. Koller, J.R. Yates and P.A. Haynes, "The functional proteomics toolbox: method and applications" *Journal of Chromatography B,* 2002, **782**,165-181.

53. The Expasy (Expert Protein Analysis System) proteomics server of the Swiss Institute of Bioinformatics (SIB); http://www.expasy.org/tools/peptidecutter (accessed 12-03-2006).

54. Protean IEF cells, Ready Strip IPG strip , rehydration/equilibration and focusing tray and electrode wicks from Bio-Rad; http://www.biorad.com/LifeScience/pdf/Bulletin_2426.pdf (accessed 12-03-2006).

55. The National Center for Biotechnology Information (NCBI), U.S National Library of Medicine; http://www.ncbi.nlm.nih.gov/ (accessed 12-03-2006).

56. Protein Prospector- UCSF Mass Spectrometry Facility , UCSF tools for peptide masses data (MS-Fit, MS-Pattern, MS-Digest, etc.); http://prospector.ucsf.edu/ (accessed 12-05-2006).

57. MASCOT- Peptide mass fingerprint from Matrix Science Ltd., London- a search engine that uses mass spectrometry data to identify proteins from primary sequence databases ; http://www.matrixscience.com/search_form_select.html (accessed 12-05-2006).

58. Profound – A protein database that search known protein sequences with peptide mass information from Rockefeller and NY Universities; http://prowl.rockefeller.edu/prowl-cgi/profound.exe (accessed 12-10-2006).

59. EMBL (European Molecular Biology Laboratory) Bioanalytical Research Group: Database used for protein identification by peptide mass data, http://www.embl-heidelberg.de/ (accessed 12-10-2006).

60. ALDENTE – Database that identify protein with Peptide mass fingerprinting (PMF) data; http://www.expasy.org/tools/aldente/ (accessed 12-12-2006).

61. K. Fowlkes: "A Proteomic Study of *Pseudomonas putida* by Two-Dimensional Gel Electrophoresis: Establishing Quantitative Standards for Intra-Laboratory Results", M.S. Thesis, Dept of Chemistry, Rochester Institute of Technology, September 2005.

62. Mass Spectrometry and Proteomic Facility; The Ohio State University Columbus, OH 43210; www.ccic.ohio-state.edu/MS (accessed 12-05-2006).

63. Jill Zapoticnyj, Adam Bazinet, Janine Garnham, Matthew Conte, and Paul Craig, 2DE – a simulation of two dimensional electrophoresis, submission pending.

64. Mass accuracy of ESI-Q-TOF Mass Spectrometer; http://www.research.ucsf.edu/Core/cores.asp (accessed 12-17-2006).

Glossary

- **Active Rehydration: A low electrical current flows to the IPG strip which helps pull larger proteins into the strip**
- **Centrifugation – process that involves centrifugal force for separating mixtures. The process is used for extracting proteins from the cell.**
- **Electrophoresis: migration of charged particles or ions in solution under the influence of an electric field**
- **Eluate - Fluid emerging from the end of the column**
- **Eluent - Fluid entering the HPLC column**
- **ESI-Q-TOF MS - Electrospray Ionization Quadrupole Time of Flight Mass Spectrometer that gives structural information about the protein by identifying the masses of the peptides**
- **Gel Filtration Chromatography: A type of column chromatography method in which stationary phase is packed in a glass column and where separation is based on the size of the molecule.**
- **Gradient – Changing solvent concentration over time Proteomics: Global and systematic analysis of the complete set of proteins expressed by an organism.**
- **HPLC – High Performance Liquid Chromatography, a technique for separation, fractionation and purification of protein**
- **Isocratic – Uniform concentration of solvent**
- **Isoelectric Point –Point at which net charge on the protein molecule is zero**
- **LC/MS – Liquid chromatography coupled to mass spectrometer and used in this study for analysis of proteins and peptides.**

- **ORF – Open reading frame is a short stretch or a portion of DNA that has the potential of producing proteins**
- **Passive Rehydration: Proteins are absorbed into the dehydrated gel on the IPG strip and no electrical current is applied to the strips.**
- **Polyacrylamide Gel Electrophoresis (PAGE): technique used to separate molecules on the basis of size and charge; used to separate proteins in this study.**
- **Proteomic Signature: Specific set of proteins expressed by an organism under specific conditions.**
- ***Pseudomonas putida* KT2440: the organism of interest due to its ability to degrade and metabolize various organic compounds.**
- **RP-HPLC – Reverse Phase High Performance Liquid Chromatography in which mobile phase is more polar than the stationary phase**
- **Sodium Dodecyl Sulfate (SDS): a detergent that becomes an anion in solution and bind to protein in proportion to their masses.**
- **Sonication- A process of isolating proteins by disrupting cell membrane using high frequency sound waves which in turn releases cellular components into the solution**
- **Succinic Acid: Benign carbon source for growth of *P. putida***
- **TOYOPEARL HW 40S – Size exclusion media used for gel filtration chromatography used to fractionate and separate mixtures of protein and other compounds over a wide range of size**

Appendix A: Hutner's Medium

Hutner's Medium is made by dissolving Solution A, B, C and 5mM succinic acid in water. The proportions are given in Table 1. The recipes for solution A, B and C are given below.

Solution A:	per liter deionized H_2O:	
	Na_2HPO_4	141.2 g
	KH_2PO_4	136 g
	(pH = 7.25)	

Solution C:	per liter deionized H_2O:	
	$(NH_4)_2SO_4$	200 g

Solution B:	per liter deionized H_2O:	
	Nitrilotriacetic acid	10 g
	MgSO4:	4.45 g (or 29.56 g $MgSO4 \cdot 7H_20$)
	$CaCl_2 \cdot 2H_20$	3.33 g
	$(NH_4)_6Mo_7O_{24} \cdot 4H_20$	9.25 mg
	$Fe\ SO_4 \cdot 7H_20$	99 mg
	Metals 44	50 mL

Metal 44 Solution is prepared by mixing the following:
In 800 mL deionized water, dissolve the following:

EDTA	2.50 g
$ZnSO_4 \cdot 7H_20$	10.95 g
$Mn\ SO_4 \cdot 7H_20$	1.54 g
$Fe\ SO_4 \cdot 7H_20$	5.00 g
$Cu\ SO_4 \cdot 5H_20$	392 mg
$Co\ (NO_3)_2 \cdot 6H_20$	248 mg
$Na_2B_4O_7 \cdot 10H_20$	177 mg

A few drops of concentrated H2SO4 are added to retard the precipitation and the solution is diluted with deionized water to make a total volume of 1 liter. The solution appears lime green in color and clears in appearance, and can be stored in a brown glass bottle at room temperature for an indefinite period.

Table 1: Volume of solutions to be mixed to form various volumes of Hutner's medium

Solution A	10mL	20mL	40mL
Solution B	5mL	10mL	20mL
Solution C	1.25mL	2.5mL	5mL
Succinic acid (5M)	2.5mL	5mL	10mL
Distilled water	231.25mL	462.5 mL	925 mL
Total Volume	250 mL	500 mL	1000 mL

Note: Medium is prepared in a flask that is at least 4 times the volume of liquid to provide adequate aeration of sample.

Appendix B: Rehydration Buffer

Rehydration buffer is prepared by mixing the following:

48.1 g of Urea (8M)
4.0 g of CHAPS (4%)
0.2 mL of Ampholytes
0.77 g of DTT (50mM)
The solution is dissolved in 100 mL water.
Store at -20^0 C

Appendix C: Procedure for Bradford protein assay

Label 10 test tubes 1 – 10.

1. Take the Bio-Rad Protein Assay Dye Reagent Concentrate (catalog #: 500-0006) out of the refrigerator, shake it well, and make a 1:4 dilution. (**Note:** Normal dilution volumes contain around 20mL of dye in 80mL of distilled water, total volume = 100mL or 10mL of dye into 40mL of distilled water, total volume = 50mL)
2. Filter the dye through Whatman #1 filter paper in a funnel. While filtering, begin step 3.
3. Create calibration samples containing the following:

Table 1: Standard Calibration curve sample concentrations

Tube #	**Bio-Rad Protein Assay Standard Protein (Catalog # 500 – 0005)**	**Distilled water**
1	0 µL	100 µL
2	5 µL	95 µL
3	10 µL	90 µL
4	15 µL	85 µL
5	20 µL	80 µL
6	30 µL	70 µL
7	40 µL	60 µL
8	50 µL	50 µL
9	60 µL	40 µL
10	70 µL	30 µL

4. Mix the contents of tubes by pipetting in and out several times or vortex mixing them.
5. Add 3mL of the filtered and diluted dye to tubes 1– 10.
6. Add 80µL of 0.1N HCl stock solution to each tube. (Tube #s: 1 – 10). Mix using a vortex mixture.
7. Transfer all of the solution in test tubes 1 – 10 to clean disposable cuvettes.
8. Measure UV absorbance (@ 595nm, slit width = 1nm.

9. For each of the unknown sample proteins make solutions that fall within the concentration/absorbance ranges of the calibration curve.

(Note: This step will be different depending on the concentration of your protein and your sample protein may have to be diluted. To start off it is a good idea to make 2 dilutions – one containing 15µL and one of 30µL of your sample protein.

10. These sample protein solutions should be treated exactly like the standard protein samples used to make the calibration curve.
11. Transfer all of the sample solutions to clean disposable cuvettes and measure the absorbance as in step 8.

Calculation of Protein Concentration

Take equation of line from the standard calibration curve.
Sample equation: **y = 0.0014x + 0.474**
Y = absorbance, X = protein concentration (µg) Solve for X.

X = (Sample absorbance – 0.474)/0.0014

X*dilution factor/µL used = Protein concentration (µg /µL)

Sample calculation:
A sample containing 20µL of protein gave an absorbance of 0.610.

(0.610 – 0.474)/0.0014 = X = 97.14µg of Protein
97.14µg *1/20µL = 4.857 µg/µL

Appendix D: Rehydration Procedure

1. Take out the number of strips to be run (pH 4 – 7, 11cm) from the freezer.
2. Let the protein sample thaw, vortex to resusupend urea.
3. Take the rehydration/sample buffer stock solution out of the -20 freezer. (Note: Make sure the solution is thoroughly thawed and everything is back in solution before you add it to your samples.)
4. Aliquot enough rehydration/sample buffer (from stock solution) into 1.5mL centrifuge tubes so that total volume (buffer and protein) equals 185 µL.
5. Add protein sample and mix. (**Note:** The concentration of isolated protein should be previously determined from the protein assay step. The concentration of protein applied to each strip should be approximately 200µg in rehydration buffer with a total volume = 185µL.)

6. Take up entire contents of the tube and apply to one well in a clean dry rehydration tray along the back corner of the well. Leave about 1cm at each end.
7. Gently peel off plastic backing of the strip using forceps.
8. Lay the strip, gel side down, in the well such that the "+ pH 4-7 "is legible and towards the left of the tray (when the sloped edge is on the right). Take care not to trap any air bubbles under the strip.
9. Let the strips rehydrate approximately 20 minutes before overlaying them with ~ 2 – 3 mL mineral oil, cover tray and let sit at room temperature overnight (~ 11 – 16 hours).

Active rehydration was also performed when the focusing tray was used for rehydrating the strips which was transferred in IEF protean cell and active rehydration program was selected. This voltage moves the protein across the IPG strip and this rehydration is done at a constant voltage (50V) for 12 hours.

Appendix E Reagent Preparation

Equilibration buffer stock solution: prepared by mixing the following
36 g of 6M urea

10 mL of 20% SDS

3.3 mL of 1.5M Tris/HCl pH = 8.8

40 mL of 50% glycerol mixed in a total volume of 100 mL ddH20

Preparation of running buffer solution

Mix 100mL of Tris/Glycine/SDS Buffer (from Sigma Aldrich) with 900mL of nano-pure water in a 1L graduated cylinder. Mix thoroughly

Preparation of agarose

0.83 mL of 1.5M Tris buffer (from Bio-Rad),

0.721 g of glycine (from Sigma Aldrich),

0.25 g of low melting point agarose (from Sigma Aldrich),

0.25 mL of 80% SDS (concentration should be 0.1% SDS),

Trace amount of bromomphenol blue,

All the above was diluted to a total volume of 50 mL with D.I. water)