PROTEOMIC STRATEGIES FOR INVESTIGATION OF PROTEIN/LIGAND AND PEPTIDE/LIGAND NON-COVALENT INTERACTIONS BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY

by

ABRAHAM LO

Presented to the Faculty of the Graduate School of

The University of Texas at Arlington in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF SCIENCE IN CHEMISTRY

THE UNIVERSITY OF TEXAS AT ARLINGTON

MAY 2008

ACKNOWLEDGEMENTS

I would like to thank Dr. Kevin Schug and lab members.

APRIL 18, 2008

ABSTRACT

PROTEOMIC STRATEGIES FOR INVESTIGATION OF PROTEIN/LIGAND AND PEPTIDE/LIGAND NON-COVALENT INTERACTIONS BY ELECTROSPRAY IONIZATION MASS SPECTROMETRY

Abraham Lo, M.S.

The University of Texas at Arlington, 2008

Supervising Professor: Kevin A. Schug

Through universal proteomic methods and electrospray ionization-mass spectrometry (ESI-MS), multiple molecular recognition strategies were investigated as a prospective complementary tool for investigating protein-ligand interactions. Three model protein-ligand systems: bovine serum albumin/naproxen, cytochrome c/ATP & ADP, and integrin/RGD peptide were probed based on the categorical difference in the nature of each ligand. It was found that screening for protein binding sites from proteolytic digests was feasible for one of three models. Three different proteolytic digests of cytochrome c featured several examples of binding interaction with the ligands ATP and ADP. The detection of binding between intact cytochrome c and ATP/ADP was also conducted. Both the folded and unfolded state of cytochrome c were observed under ESI-MS conditions with ATP/ADP binding to the native states. Dissociation constants calculated for intact cytochrome c and ATP/ADP and the results were in correlation with existing literature.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Protein-ligand interactions play a crucial role in cellular functions which drive biomedical and biotechnological strategies. Current drug discovery is dependent on understanding protein-ligand interactions which lead to the creation of drugs that either modify or block the site of interaction. The ability to effectively understand the function of a protein's binding site both quantitatively and qualitatively remains a focus of the scientific community.

The method of discovery of non-covalent protein-ligand interactions has gone through several techniques and forms of analysis. The recent emergence of soft ionization mass spectrometry (MS) has added yet another powerful tool to the biochemist's analytical tool-belt. Exploited for its speed and sensitivity, MS has opened a new wave of MS-based non-covalent protein-ligand research. A niche within this field is currently focused on mapping the interaction site where the ligand binds to the protein.

The current MS-based methodology for mapping protein-ligand interfaces incorporates one or a combination of limited proteolysis, genetic engineering, chemical modification, and hydrogen/deuterium exchange strategies. In regards to rapid drug discovery, the limitations of each technique become apparent from the lack of literature regarding the establishment of a high-throughput molecular recognition technique in relation to protein-ligand interaction site mapping. Time-consuming experimental elements and intense data management/analysis of each technique serves as an impetus towards the creation of a high-throughput method.

The goal of this project is to establish a high-throughput molecular recognition technique where electrospray ionization-mass spectrometry (ESI-MS) and universal proteomic methods are used as complementary tools for investigating protein-ligand interactions. The

specific interests related to the goal are the following: screening and identifying possible components in a complex; screening proteins by detecting binding of ligands; detection of binding between protein and ligand; and mapping the interaction interface or identification of protein-ligand contact points. Another goal is to extract quantitative binding information for non-covalent complexes between proteolytically derived peptide fragments and selected ligands.

To identify sequential domains, manipulation of protein denaturation conditions in combination with enzymatic digestion via an assortment of enzymes (i.e., pepsin, trypsin, etc.) will be used to generate varying peptide fragments of different lengths. The main goal is to have the specific binding region of the protein remain intact within the peptide mixture and then screen for specific ligand interactions. The overarching experimental goals stated above might seem ancillary to the currently practiced limited proteolysis techniques, but the novel approach is the high-throughput use of common enzymatic conditions where the imperfect cleavage patterns of the enzyme are incorporated as a tool to obtain binding site containing peptides.

The study covers three model proteins (bovine serum albumin, integrin $\alpha_V\beta_3$, and cytochrome c) that have had their binding site already characterized. These proteins were proteolyzed and had their resulting peptide mixtures submitted to reversed phase-high performance liquid chromatography-ESI-MS. During the chromatographic separation a subsequent screening was accomplished with the respective ligands (naproxen, RGD peptide, and ATP/ADP) through post-column addition. Measurements were conducted on the binding affinity and selectivity between proteins and their ligands.

1.1 Thesis Overview

Chapter 1 provides an introduction and states the goals of this thesis. Chapter 2 lays down the groundwork related to three aspects fundamental to this thesis: proteomics, mass spectrometry, and non-covalent interactions. Next, Chapter 3 provides the experimental materials, instrumentation, and procedures necessary to carry out the ideas implemented in Chapter 1. Moving forward, Chapter 4 presents a more detailed view of protein and peptide

identification through mass spectrometry and the current mapping techniques for protein/ligand interfaces as described above. Research pertaining to the three model proteins: BSA, integrin, and cytochrome c and as well as their respective ligands will be offered within Chapter 4. Subsequently, Chapter 5 provides the data and discussion of the results. Finally, Chapter 6 provides a conclusion and probes the future avenues of research.

CHAPTER 2

BACKGROUND/REVIEW

This chapter presents an overview of proteomics, mass spectrometry, and non-covalent interactions. The underpinnings under which proteomics developed leading up to its relevance to the status quo are explored. Next, the fundamental and practical aspects of mass spectrometry are highlighted. This is followed by the common approaches unifying mass spectrometry and proteomics. Finally, the background of non-covalent interactions and its relevance to mass spectrometry and proteomics is covered.

2.1 Proteomics

2.1.1 What is proteomics?

Proteomics is a relatively new discipline spawned within the field of biochemistry. In contrast to the early origins of biochemistry beginning in the 1800s, the genesis of proteomics can be traced back to the recent 1970s-1980s. The word proteomics was derived by Marc Wilkins, a Macquarie University Ph.D. candidate, in 1994 as “the protein complement of the genome” at a meeting in Siena, Italy [1].

The definition of proteomics is the study of the proteome where the standard definition of the proteome is “the complete set of proteins within a certain organism”. Proteomics not only involves the study and the identification of proteins but also the interactions between proteins that compose various systems, so-called “functional proteomics”. These systems may range from proteins involved in simple biological pathways to more complex systems such as organ systems or the whole organism itself.

Within the short span of several decades, the field of proteomics has become a very hot topic with funding from both the private and public sector ballooning to billions of dollars a year.

A good portion of the NIH's $28.4 billion 2007 yearly budget [2] was dedicated towards proteomic and genomic research. The overwhelming consensus from the scientific community is that we have only begun to scratch the surface within the field of proteomics and that there are historic breakthroughs waiting to be discovered. With proteins being involved in almost all biological processes, the study of proteomics is directly related to questions and problems that plague the 21st century: Can we cure diseases like cancer, Alzheimer's disease, Parkinson's disease, etc.? What effects do environmental toxins have on the human condition? Or even, what type of vaccines do we need to design to defend against the pressing threat of drug-resistant biological pathogens?

In general, proteomics is important for a number of reasons. It allows scientists to identify proteins in normal and disease conditions, to identify pathogenic mechanisms, to improve medicine by matching a person's protein fingerprint to the effective treatment, and to contribute to the understanding of gene function. The current goals of proteomics research are geared towards the determination of the properties of proteins which include sequence, quantity, state of modification, interactions with other proteins, activity, subcellular distribution, and structure [3].

2.1.2 What are proteins?

In proteomics, the protein is the common thread from which everything begins. Without proteins, life would not exist. All functions of living organisms are related to proteins. Proteins are the end result of the "central dogma" of life [4] where DNA is transcribed to RNA and RNA is translated to form proteins.

Proteins are amino acid based macromolecules that are vital in numerous biological processes. These biological processes are often complicated and convoluted, and scientists have only recently begun to unravel the intricate functions of proteins in biological systems. These functions range from enzymatic activity (i.e. pepsin breaking down food products in the stomach [5]) to providing mechanical function (actin and myosin contracting in the muscles [6]).

From an assortment of 20 amino-acids available, proteins are constructed by the cell's machinery in a linear fashion by connecting the C-terminus of one amino acid to the N-terminus of another amino acid. Through a dehydration reaction, a peptide bond forms between the two amino acids. A peptide is created after several amino acids join together to form a linear polymer. The general definition of a peptide is a molecule composed of two or more amino acids. Larger peptides are often referred to as polypeptides or proteins; however, a vital distinguishing factor between a peptide and protein is that a protein has a definitive 3-dimensional structure through the folding of the chain of amino acids.

The 3-dimensional structure of a protein is broken down into four distinct categories: primary, secondary, tertiary, and quaternary structure. The primary structure is the amino acid sequence of the protein. The secondary structures are localized formations stabilized by hydrogen bonds composing ordered arrangements such as alpha helixes and beta sheets. The tertiary structures are the gathering of secondary structures together through hydrogen bonding, disulfide bonds, and salt bridges to form the overall shape of the protein molecule. Finally, the quaternary structures are the aggregation of two or more protein molecules to form a larger multi-protein complex (i.e. hemoglobin).

2.1.3 Development of Proteomics

Four major scientific discoveries pertaining to proteins set the backdrop for the genesis of proteomics. The history of protein chemistry begins first when the word protein was coined by Jons Jakob Berelius in 1838 [7]. Derived from Greek and meaning "of primary importance", the word protein was used for the large organic compounds with equivalent empirical formulas which he studied. In 1926, the next major step in protein research was made by James B. Sumner when he isolated and crystallized the enzyme urease [8]. In 1955, Sir Frederick Sanger determined the entire amino acid sequence of the protein insulin [9]. In 1958, the 3-dimensional structures of the proteins, hemoglobin [10] and myoglobin [11], were elucidated through X-ray diffraction analysis by Max Perutz and Sir John Cowdery Kendrew.

Building upon the momentous discoveries in protein chemistry, the subsequent scientific study of proteins has journeyed through several stages. Patterson and Aebersold breaks down the emergence and maturation of proteomic concepts and technology into three phases [3]: 1. Transition from protein chemistry to proteomics as a platform for scientific advancement; 2. present diversification of proteomic technologies and tools used to ascertain properties of proteins; and 3. comprehensive understanding of the working of biological systems through proteomic data and new science technologies. The first two phases are most relevant to the work described in this thesis.

In the infant stages of proteomics (phase 1), the main goals of protein chemistry conducted in the 1980s and early 1990s were to provide the link between the observed activity of a biochemically isolated protein and the gene that encoded it [3]. These main goals were facilitated by the very important advent of two-dimensional gel electrophoresis (2DE) which allowed for the separation of complex protein matrixes, based primarily on size and isoelectric point. With the separation of proteins, technological developments for analytical protein chemistry were commenced to improve the sensitivity of detection. Numerous achievements were accomplished; however, the reliance on 2DE as the main means for protein research was a bottleneck for advancement. Several ideas were proposed in the 1970s and 1980s to use 2DE to create a database of proteins, but this endeavor failed initially. In the early 1990s, the sequencing of proteins through Edman degradation and PCR provided the link between amino acid sequence and the corresponding protein activity. Much effort was put into the mass DNA sequencing of cDNA derived from mRNA [12] with the well publicized Human Genome Project as its pinnacle. As a result, these endeavors created massive libraries of protein sequence databases.

From the amalgamation of these protein sequence databases, creation of algorithms, and the availability of existing mass spectrometry techniques in concert with a wide variety of separation techniques, the study of proteins was viewed in a new light. While determining the

mass of an unknown protein or peptide with a mass spectrometer is not a unique identifying feature, a protein cleaved with an enzyme of known specificity to create a mixture of peptides could distinctively identify a protein. By extracting peptide masses from the mass spectra after running a digested protein through Reversed Phase-High Performance Liquid Chromatography-Mass Spectrometry (RP-HPLC-MS), one could match the mass list to a predetermined proteolytic sequence database of known proteins [13]. Scientists now had an extremely powerful tool where they could identify and study most if not all proteins (proteome) on a much larger scale than before. This is when proteomics was born.

2.2 Mass Spectrometry

2.2.1 Definition

Mass spectrometry (MS) is the analysis of the mass to charge ratio (m/z) of ions. It is a technique which facilitates the capability of 'weighing' charged molecules. In general, mass spectrometry is a very powerful technique because it allows for the analysis of gas, liquid, and solid samples. Through highly sensitive instrumentation and different analytical techniques, mass spectrometry is capable of elemental and molecular analysis. Mass spectrometry is now considered as an indispensable part of proteomic research. Of particular interest, mass spectrometry is useful for the identification of protein/peptide fragments and the examination of non-covalent interactions.

2.2.2 MS Setup

The general MS instrumental setup consists of 5 major components: the inlet, the source, the mass analyzer, the detector, and the data acquisition system as shown in figure 2.1. The inlet is where the introduction of sample occurs. Once the sample has been in introduced, the formation of ions (ionization) occurs in the source. The two common methods of ionization of proteins/peptides are through electrospray ionization (ESI) and matrix-assisted laser desorption ionization (MALDI). These ionization techniques are "soft in nature, largely preserving ionized species as intact molecular ions. After ionization, the ions are transported to

the analyzer through ion optics via the application of highly controlled electric and/or magnetic fields. The detection of ions and analysis are then accomplished by the detector and the data acquisition systems, respectively.

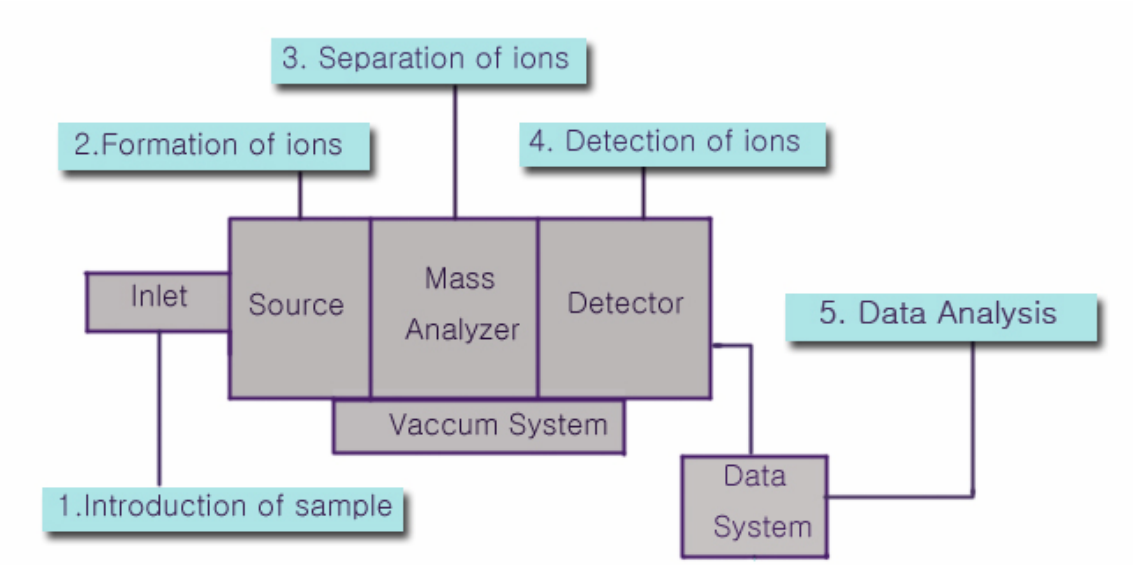


Figure 2.1 General Mass Spectrometry Instrumental Setup

2.2.2.1 Inlet

The inlet is where the introduction of sample occurs. Since the creation of gas phase molecules is a requirement at the source, the inlet must somehow accommodate solution to gas phase transfer at the inlet/source interface. Although there are numerous types of inlets (direct vapor inlet, gas chromatography, liquid chromatography, direct insertion probe, etc.) which interface to the mass spectrometer, the inlet choice depends solely upon the sample. In order to choose which inlet is suitable, one must consider the characteristics of the sample. Samples with high volatility and thermal stability are easily accommodated by a variety of inlets involving thermal vaporization such as GC, direct vapor inlet, and direct insertion probes. Since proteins/peptides have low volatility and are thermally labile, they require direct ionization from the condensed phase [14]. Liquid chromatography provides such an introduction of thermally labile and nonvolatile liquid compounds and mixtures not amenable to gas chromatography.

Liquid chromatography (LC) is a technique where a sample is forced through a column packed with solid particles (stationary phase) by a liquid (mobile phase). The origins of liquid chromatography began in 1900 when Mikhail Tswett coined the word 'chromatography' in his seminal work on chlorophyll [15,16]. Within his research, Tswett was able to separate a mixutre of chlorophylls and xanthophylls by passing a solution in petroleum ether through a glass column packed with calcium carbonate [17]. For several decades, similar methods employed packed columns as such for the separation of mixtures. Moving forward, the 1960s used packed columns with small particles resulting in high performance liquid chromatography or high pressure liquid chromatography (HPLC) where higher pressures were needed to pass liquid through the column. HPLC is highly regarded as an analytical separation technique because of its sensitivity, its high efficiency, its ready adaptability to accurate quantitative determinations, its suitability for separating nonvolatile or thermally fragile species, and its widespread applicability to a variety of substances [18].

With the manipulation of the polarity of the stationary and mobile phases, a sample containing a mixture of analytes can be separated through the partitioning between the two phases. Multiple modes of separation such as reversed phase, normal phase, ion exchange, size exclusion, chiral, etc. are applicable to HPLC. For proteins and peptides, the reversed phase mode of separation is the most widely employed. Reversed phase-HPLC (RP-HPLC) uses a non-polar stationary phase (i.e. C18 bonded silica) and an aqueous/polar organic mobile phase (i.e. acetonitrile or methanol). Through hydrophobic interactions, analytes are retained by the hydrophobic stationary phase in high aqueous mobile phase conditions. Then, they are eluted by substantially increasing the percentage of polar organic solvent in the mobile phase.

Gradient elution is commonly used to separate complex mixtures of peptides in RP-HPLC on a C-18 column. Two reversed phase solvents labeled A and B contain 100% water with a small percentage of acidic modifier (formic acid (FA), trifluoroacetic acid, acetic acid) and 100% Acetonitrile (ACN), respectively. The acidic modifier is used to better the chromatographic

peak shape and also to provide a source of protons for the ionization process. The ratio of solvents is typically varied in a programmed way with a low concentration of organic at the beginning of each run (i.e. 98:2 0.5% FA H_2O:ACN) and increased to a high organic content toward the end of the run(i.e. 98:2 ACN:0.5%FA H_2O). The elution of proteolytically derived peptides (e.g. via enzymatic digestion; see below) from a RP-HPLC C-18 column commonly occurs around the 30-40% organic solvent mark of a typical linear gradient analysis. Peptides that are more hydrophobic are retained longer while those that are more hydrophilic are eluted earlier.

2.2.2.2 ESI Source

The source is where the formation of ions occurs. While interfacing LC to MS, two major hurdles of volatilizing biochemical analytes and removing solvent present in large excess, have to be overcome. In 1989, Fenn's group [19] pioneered a new Atmospheric Pressure Ionization (API) source called Electrospray Ionization (ESI) which addressed the aforementioned issues, allowing for a seamless interface between LC and MS. It is now one of the most commonly used amongst other established interfaces (Atmospheric Pressure Photoionization (APPI), Thermabeam, Matrix-assisted Laser Desorption/ionization (MALDI), Thermospray, Fast Atom Bombardment (FAB), Atmospheric Pressure Chemical Ionization (APCI), Direct Insertion Probe, etc.). Most modern LC-MS instruments can utilize at least two atmospheric pressure ionization modes: electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI).

ESI has the advantages of ionizing low and high molecular weight non-volatile or volatile polar and ionic compounds. It can accommodate conventional HPLC flow rates (≤ 1000 ul/min) and has high sensitivity and high ionization efficiency. As the most commonly employed ion source for LC-MS, ESI is geared towards polar, ionic molecules (i.e. proteins and peptides) and has the advantage of detecting very large biomolecules through multiple charging [20]. For all its virtues, ESI still has several disadvantages: little to no fragmentation with possible adduct formation, no ionization for non-polar compounds, and no universal library.

The exact science behind the mechanisms of ESI is not fully understood at this point, but two prevailing models have been proposed: the ion evaporation [21] and the charged residue model [22,23] Droplet size, surface charge, liquid surface tension, solvent volatility, and analyte ion solvation energies are several factors which must be considered to understand the ESI process. The general mechanism for ESI ion generation, as shown in figure 2.2, can be broken down into four steps. First, the capillary through which the sample solution is passed is held at a high potential (±2-5kV). Second, in response to the potential gradient placed across the atmospheric pressure spray chamber, the liquid solvent disperses into a mist of highly charged droplets. Third, the droplets reduce in size by evaporation of the solvent or by "Coulombic explosion". Coulombic explosion refers to the phenomena where surface tension is overcome by charge repulsion on the shrinking droplet surface. Lastly, the fully desolvated ions resulting from complete evaporation of the solvent are transferred to the mass spectrometer for mass analysis. ESI can be operated in either the positive or the negative ionization mode by varying the potential applied to the capillary. For proteins and peptides, positive mode is ideal.

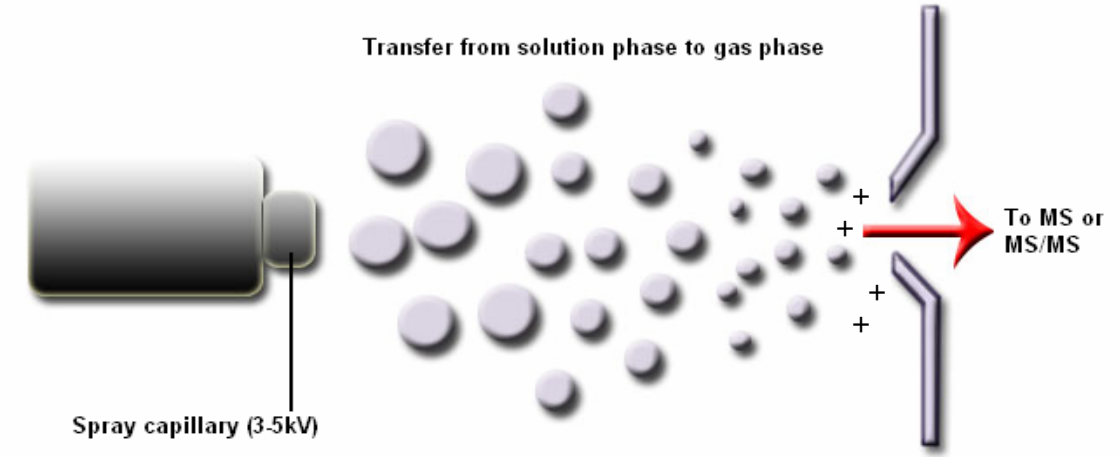


Figure 2.2 General Schematic of the electrospray process

The high sensitivity of ESI, without excessive fragmentation, proves worthy in its application towards protein/peptide research. In contrast to other ionization techniques, where upon ionization significant fragments (offspring ions) occur from molecules, ESI utilizes a 'soft

ionization' technique which preserves the molecular ion (precursor ion) due to the low internal energy imparted onto the molecule. The softness of the ESI technique has been used extensively when considering non-covalent interactions between protein/ligand and peptide/ligand complexes.

As mentioned above, ESI has the potential to produce multiply charged ions allowing higher molecular weight compounds to be observed at lower mass to charge ratio (m/z) values using standard mass analyzers. For example, one might predict that a large protein such as Bovine Serum Albumin (BSA) (MW 66,000 kDa) would be difficult to see on a typical ion trap mass spectrometer with an effective scan range of 100-4000 m/z. However under optimum solvent conditions, BSA appears as an envelope of low m/z multiply charged ions. By applying an averaging algorithm and examining the relative position of isotopic ion abundances, one is able to assign number of charges to the peaks in the ion envelope and estimate the average molecular mass of the protein.

2.2.2.3 Mass Analyzer and Detector

Many types of mass analyzers are available on the market today including magnetic sector, time-of-flight (TOF), quadrupole, ion trap, Fourier transform ion cyclotron resonance (FT-ICR), and orbitrap systems. The choice of mass analyzer is dependent on the analytical goals requiring a certain sensitivity, resolution, mass accuracy, mass range, rate of scanning, ease of use, and cost. Hybrid mass analyzers such as quadrupole time of flight (Q-TOF) or linear ion trap-Fourier transform ion cyclotron resonance (LIT-FTICR) are commonly employed for proteomics but cost can be a limiting factor. Quadrupole ion traps are used in proteomic research due to their advantages of being inexpensive, easily interfaced to many ionization methods, and having MS/MS capabilities. Although resolution, accuracy, mass range, and scanning rate might be inferior in comparison to more expensive mass analyzers, the quadrupole ion trap is adequate for certain realms in proteomics.

In the 1950s, the quadrupole ion trap was created by the Nobel Prize winning Wolfgang Paul [24]. Finnigan MAT made several breakthroughs during the 1980s in designs based on Paul's research to make the quadrupole ion trap as practical as it is today [25].

The quadrupole ion trap traps incoming ions from its inlet in a chamber composed from several hyperbolic electrodes: an entrance endcap electrode, an exit endcap electrode, and a donut-shaped ring electrode. DC and RF oscillating AC electric fields create a 3D quadrupolar potential energy well within the center of the trap. By varying the voltages on specific electrodes, ions are selectively trapped and ejected out from their oscillating trajectory based on their individual mass to charge (m/z) ratios. A detailed description of quadrupole ion trap technology has been given by Raymond March [26]. The ejected ions are then focused onto the detector to produce the mass spectrum. Most detectors are comprised of a conversion dynode and a channel electron multiplier.

As mentioned above, the advantages of the quadrupole ion trap are that is it inexpensive, easily interfaced to many ionization methods, and has MS/MS capabilities in one analyzer. The disadvantages are that it has low resolution (<4000), low mass accuracy (~±0.5 amu), typically limited mass range (<4000 m/z), moderate scanning rates, and space charging effects can give rise to mass shifts. For our purposes, the quadrupole ion trap is sufficient for the task at hand. The MS^n (tandem MS to the nth degree) capabilities are especially useful. By isolating a particular ion within the trap, ejecting everything else, and then employing Collision Induced Dissociation (CID), one can obtain the relevant fragmentation data of the precursor ion. The process can be repeated by isolating one of the fragment ions and so on, as long as sufficient ion signal remains.

Ion trap mass spectrometers often have the capability to provide full scan followed along with data-dependent MS/MS or MS^n information to obtain data for interacting with proteomics databases. Through specified parameters, data dependent scanning allows the selection of one or more ions of significance for analysis. Thermo's Dynamic Exclusion™

program picks the most intense ion in the full-scan spectrum and extracts MS/MS data from that particular ion. The program then stores the mass into an exclusion list after acquisition and seeks out less intense components that have not yet been examined. This automated process ensures that both high and low abundance proteolytic peptide ions are fragmented to maximize chances of identifying the protein(s) of interest.

2.2.2.4 Data Acquisition System

The data acquisition aspect of mass spectrometry has gone through radical changes in the last 30 years. As strip chart recorders sit around gathering dust as relics of the past, the digital age has brought the necessary advancements for progressive proteomic research. Computer hardware and software play a vital role in the management, categorization, and analysis of data relayed from the mass spectrometer. The internet also plays an important role in the dissemination of information and rapid exchange of ideas, including interfacing collected data with numerous databanks for proteomic-based applications.

2.3 Relation between Proteomics and Mass Spectrometry

As mentioned previously, Patterson and Aebersold have broken down the emergence and maturation of proteomic concepts and technology into three phases. Phase 1 highlights the transition from protein chemistry to proteomics as a coordinated platform for scientific advancement. Phase 2 ascertains the properties of proteins through the diversification of proteomic technologies and tools. Lastly, phase 3 builds towards a comprehensive understanding of the working of biological systems through proteomic data and new science technologies. The following discussion centers largely on methods employed relative to Phase 1. Proteomics is carried out by two predominant experimental setups, the top-down and the bottom-up approach.

2.3.1 Top-down proteomics

The top-down proteomic approach focuses on the protein level analysis where intact protein ions or large protein fragments are subjected to gas phase fragmentation. The

fragmentation pattern from an intact protein is often difficult to interpret since, in ESI, different charge states have different fragmentation patterns [27-30]. As a result, most top-down approaches for determining sequential information rely on high resolution mass analyzers (FT-ICR) to help deconvolute data. Top-down proteomics is often considered not as information rich as bottom-up approach and is still a limited technique. Some of its uses involve the direct determination of molecular mass which can be useful in determining post-translational modifications where sufficient mass resolution and sensitivity are available. A protein's primary structure is also a target of the top-down proteomic approach. *De novo* sequencing is a technique where the sequence of a previously unknown protein is derived through the use of the tandem mass spectrometry data.

2.3.2. Bottom-up proteomics

In bottom-up proteomics, an isolated and proteolytically digested protein of interest yields peptide fragments. Specific cleavage is often accomplished by the protease enzyme trypsin, but other proteases such as chymotrypsin, Glu-C, Asp-N, and others are also utilized. MALDI is commonly utilized for peptide mass fingerprinting of the complex mixture [31]. The masses of peptides are extracted from the mass spectra and matched against a database. Figure 2.3 provides a general schematic of the bottom-up proteomic approach. Just like a human fingerprint, each protein has a distinct pattern of peptide signals for a given protein. Another technique called peptide fragmentation incorporates ESI to take the sequential information based on collisional dissociations of polypeptides from a protein and matches them against a database [32-35]. Rather than providing the exact information on protein sequence, bottom-up approaches are mainly used for protein identification. Various types of collisional dissociation processes can be employed for experiments ranging from protein identification to the localization of post-translational modifications [36,37].

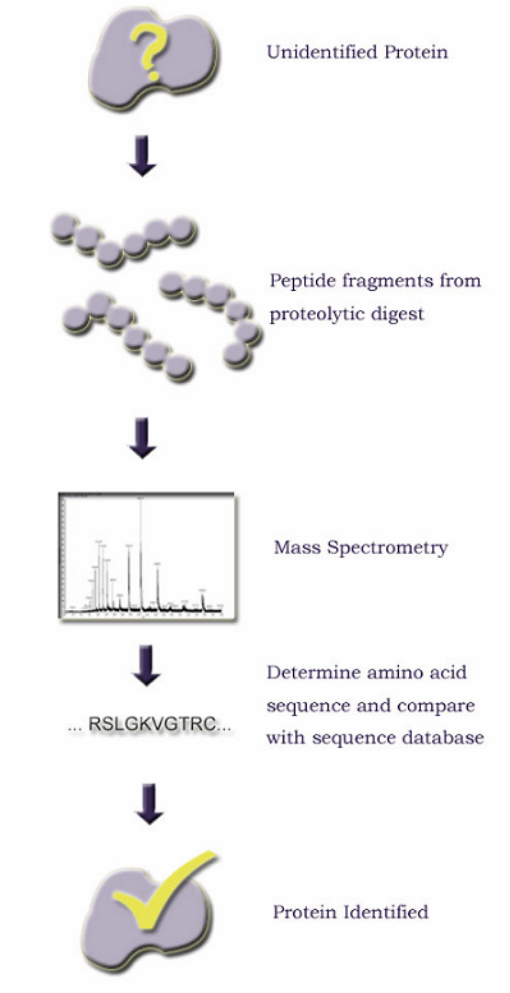


Figure 2.3 Bottom-up Proteomic Approach

2.4 Non-covalent Interactions

2.4.1 Definition

In both biological and synthetic systems, there are many cases where molecules form non-covalent complexes in solution. These non-covalent interactions are often classified as host-guest, selector-selectand, drug-receptor, or receptor-ligand relationships. A non-covalent interaction refers to either one or a combination of electrostatic interactions, ionic bonds, hydrophobic interactions, hydrogen bonds, Van der Waals forces, or dipole-dipole bonds

between two or more molecules. Noncovalent interactions in biological systems are ubiquitous and are essential for controlling many of life's processes.

2.4.2 Methods for detection

The ability to quantitatively and qualitatively determine the nature of non-covalent interactions remains a priority amongst researchers. Several established instrumental methods such as X-ray crystallography, NMR [38,39], ultracentrifugation [25,26], circular dichroism [42-44], light scattering [44,45], and fluorescence [46,47], differential scanning calorimetry [48,49], isothermal titration calorimetry (ITC) [48], surface plasmon resonance (SPR) [48,50], and others have been used to study non-covalent interactions. Although these methods are still being used today because of their virtues, they often have disadvantages [51-53].

Mass spectrometry has emerged recently as another method for detecting non-covalent complexes with several distinct advantages [52,53]. Speed, specificity, and sensitivity are the most obvious advantages [54]. Mass spectrometry allows for the use of small amounts of substrate (low picomole to femtomole levels) instead of the larger quantities of material (milligram) which are often required of NMR and X-ray crystallography. The speed and automation of mass spectrometry permits one to quickly analyze samples and conduct high-throughput screening. Additionally, mass spectrometry circumvents the requirement in other methods where the ligand or target must be labeled or bound to a surface which ultimately may affect the binding characteristics of the system of interest [55].

2.4.3 ESI and non-covalent complexes

2.4.3.1 Discovery

Ganem et al [56] were one of the earliest groups to show that non-covalently associated molecules (FK binding protein - FK506 ligand) in solution could be transferred into the gas phase via ESI and detected as an intact complex. Also that year, Katta and Chait [57] observed the heme-globin complex in native myoglobin by ESI. The importance of these measurements in contrast to previous studies where ESI was used to see adducts and aggregations of various

entities was that these were specific noncovalent interactions as opposed to non-specific interactions resulting from the electrospray process.

2.4.3.2 Questions to be asked

Since most non-covalent interactions of interest exist in physiological, aqueous environments, there is a strong emphasis on establishing correlation between the gas phase ion abundances obtained by of ESI-MS measurements and the corresponding solution phase distribution of species from which the gas phase ions are generated. Most, if not all research pertaining to this subject must be wary of and address this issue before conclusions can be formulated. In order to determine if there is such a correlative relationship several question must be asked. As mentioned above, is the interaction non-specific or specific? Does the solution structure resemble the nature of the gaseous phase structure? And, does a measurement made in the gas phase translate over into the properties of that molecule in the solution phase? There have been many cases in literature which support the correlation between solution phase and gaseous phase non-covalent interactions [58-63].

2.4.3.3 Obtainable Data

From ESI-MS data, the stoichiometry and relative abundance of the complex can be derived by measuring the m/z and abundance of the complex as well as the free associates involved in the equilibrium [58]. In general, stoichiometry measurements and the observed relative abundances from ESI-MS experiments have the most robust correlations to the solution phase [57,64,65]; however, there are always exceptions [66]. The values derived from ESI-MS are often correlated to condensed phase non-covalent interactions and discussed to a great degree within literature. In numerous examples, there has been a positive correlation between the dissociation constants derived from ESI-MS and solution phase equilibrium binding constants. Dissociation constants (K_d) ranging from millimolar (mM) to sub-picomolar (pM) levels have been measured by ESI-MS [67-70]. As mentioned above, specificity can also be

established by surveying a systematic variation of interaction conditions and interaction partners [67].

Daniel et al. [71] effectively breaks down the possibilities of determining non-covalent interaction strength by MS through solution-phase and gas-phase methods. Solution-phase methods operate under the assumptions that there is no partial dissociation of the non-covalent complex in the gas phase and that ion formation does not disturb the solution-phase equilibrium. Solution-phase methods include melting curves [72,73], titration [74-78], determination of transfer coefficients [79-81], and competitive binding analysis [67,84-87]. The determination of host-guest equilibrium binding constants through the titration method is a typical solution-based method. By holding the host concentration constant, varying the guest concentration, and then comparing the ion intensities of the complex to that of the free host by MS, both linearization and non-linear curve fitting can be implemented for the determination of binding constants. Notable host-guest complexes analyzed through the titration method include those between vancomycin antibiotics with tripeptide ligands [74] and albumin protein with oligonucleotides [75]. A number of quantitative evaluations of binding affinity have been performed on protein/enzyme-ligand [89,90], DNA-ligand [91,92], RNA-ligand [93,94], small molecule [95,96], and chiral recognition systems [97].

Gas-phase methods operate under the assumption that the complex undergoes desolvation and is structurally held together through Coulombic, dipolar, hydrogen bonding, and van der Waals forces. Gas-phase methods include cone-voltage-driven dissociation (VC_{50} method) [98,99], collision-induced dissociation (CID) [100-105], guided ion beam tandem mass spectrometry [106,107], blackbody infrared radiative dissociation (BIRD) [108-111], heated capillary dissociation (thermal dissociation in the gas phase) [112,113], among others.

2.4.3.4 Conditions

ESI allows one to produce and mass analyze non-covalent complexes in the gas phase. ESI's 'soft ionization' technique preserves weakly bound complexes due to low internal

energy imparted onto the complex; however, many gas phase complexes are highly sensitive and easily susceptible to dissociative processes. An emphasis is therefore placed on balancing the desolvation of the gas phase complex with the intermolecular forces keeping the complex together. For cases where the non-covalent interactions attributed mainly to electrostatic interactions, the complex will be stabilized in the gas phase when fully desolvated [114,115]. On the other hand, non-covalent interactions primarily held together through hydrophobic interactions are weakened in vacuum [116].

Normally, the ESI process requires specific conditions for optimum operation. This is even more so with the preservation of non-covalent complexes in ESI-MS where instrumental parameters such as low capillary voltage, cone/orifice voltages, source temperature, and probe temperatures are required [117]. Additionally, solution phase conditions are parameters which must be considered. Since most biological systems in the solution phase function under physiological conditions, ESI-MS analysis is often difficult because of the incompatibility between buffer solutions. Minute changes in pH, ionic strength, solute composition, and other factors can adversely affect the native properties of macromolecules. High concentrations of salt lead to suppression of the total ion current and the formation of nonspecific adduct ions which convolute the mass spectrum. Changes in pH away from physiological neutral to acidic conditions can cause shifts in macromolecular structure such as the denaturation of proteins [118-122].

2.4.4 Relation of Proteins to non-covalent complexes

2.4.4.1 Significance

Patterson and Aebersold's phase 2 [3] ascertains the properties of proteins through the diversification of technologies and platforms. One of these technologies is molecular recognition through mass spectrometry. Proteins are molecular machines that interact with other molecules (metal ions, nucleotide cofactors, other proteins, oligonucleotides, polysaccharides) through noncovalent interactions to control physiological processes. These interactions have significant

underpinnings in the comprehension of cellular functions which drive biomedical and biotechnological strategies.

2.4.4.2 Categorization

Studies of non-covalent interactions by ESI-MS have been extensively reviewed [53, 58,61,123,124,125]. Protein-related studies of non-covalent complexes via ESI-MS have included protein interactions with inhibitors [126-130], cofactors, metal ions, nucleotides [131], carbohydrates [132-135], other peptides and protein [136], enzyme-substrate pairings [137], nucleic acid complexes, drugs [130,138-140]. Intermolecular interactions such as these have been studied, but intramolecular non-covalent interactions such as protein folding have also been explored [141,142]. It is impossible to completely cover all protein related non-covalent interactions, but several individuals have made attempts to categorize these interactions into groups. Loo [58] effectively categorizes non-covalent protein complexes into groups highlighted by specific examples: polypeptide-metal ions, protein-small molecule, protein/peptide-peptide, protein subunit complexes, and protein-nucleic acid complexes. Akashi [143] categorizes intact supra-biomolecules into three groups: protein-nucleic acid complex, protein-protein complex, protein-ligand complex, and others. The subsequent sections will highlight the non-covalent complexes: protein-small molecule, protein-peptide, and peptide-peptide as they will be interrelated with the systems addressed later in the thesis.

2.4.4.2.1 Protein-Small Molecule Complexes

In 2004, Doneanu et al coupled nanoscale liquid chromatography with ESI-MS to detect dihydroxybenzoyl-serine binding to a protein called siderocalin [144]. HPLC was used to separate the folded state siderocalin-ligand complex from the unfolded siderocalin. It was demonstrated that certain electrostatic non-covalent protein complexes could survive the denaturing conditions of RP-HPLC and gas phase transfer during the ESI process. A more recent example from the Zenobi lab uses a method based on chip-based nanoelectrospray mass spectrometry in detecting non-covalent ligand binding to the human estrogen receptor α

ligand-binding domain (hERαLBD) [145]. The non-correlation of the relative gas-phase stability of the (hERαLBD)-ligand complex to the solution-phase binding affinity demonstrated that hydrophobic contacts play an important role in solution-phase stability.

2.4.4.2.2 Protein-Peptide Complexes

Veenstra et al. studied non-covalent binding of Calmodulin to two polypeptides, CamK-II and melittin, via ESI-MS [146]. They were able to determine stoichiometry and discovered that detection of complex in ESI-MS is source temperature dependent. Rostom et al. characterized the binding of the protein OppA to 11 peptides with lengths varying from two to five amino acids. They were able to demonstrate highest-affinity complexes formed between OppA and tripeptide ligands and lower-affinity complexes formed between OppA and dipeptides. Tripeptides containing a single D amino acid did not bind to OppA. They were also able to establish binding affinities in good correlation to solution-phase dissociation constant (K_d) measurements [147].

2.4.4.2.3 Peptide-Peptide Complexes

Raji et al. studied the binding affinities of integrin peptide fragments and RGD-based peptides via ESI-MS [148]. The integrin peptide fragments were synthetically created, mimicking the specific binding region of the RGD motif. They were able to show some correlation between gas-phase dissociation constants to solution phase techniques through titration experiments. Poor correlation was shown between ESI-MS-based titration experiments and those performed by affinity capillary electrophoresis. The disconnect was attributed to the inability of determining the relative response of different host-guest complexes, a common problem in mass spectrometry-based titration experiments [97]. Loo et al. used electrospray ionization mass spectrometry to probe antisense peptide interactions [149]. They tested non-covalent interactions between human angiotensin and eight synthetic octapeptides. They were able to achieve relative abundances similar to solution phase studies and calculate dissociation constants in the high micromolar range.

CHAPTER 3

EXPERIMENTAL

3.1 Materials

The proteins bovine serum albumin and cytochrome c (equine heart) were purchased from Sigma (St. Louis, MO), cytochrome c (equine heart) from Calbiochem (Gibbstown, NJ), and integrin α_V (mouse anti-human monoclonal antibody) from Chemicon International (Temecula, CA). For proteolysis of proteins, sequencing grade modified trypsin (porcine) was purchased from Promega (Madison, WI), mass spectrometry grade chymotrypsin from GBioSciences (St. Louis, MO), and proteomics grade pepsin purchased from Amresco (Solon, OH). The ligands employed were purchased as follows: (*S*)-naproxen from Cayman Chemical (Ann Arbor, MI), ATP/ADP from Calbiochem (San Diego, CA) and Alfa Aesar (Ward Hill, MA), respectively, and GRGDSP peptide from Bachem California Inc. (Torrance, CA). Other chemicals and solvents were purchased from the following sources: Dithiothreitol (Sigma), Iodoacetamide (Sigma), Formic Acid from Fluka (Buchs, Switzerland), LC/MS grade H_2O from JT Baker (Phillipsburg, NJ), ACN from EMD (Gibbstown, NJ), NH_4HCO_3 (EMD), and hydrochloric acid (EMD). The proteolytic mixtures were separated by analytical reversed-phased-high pressure liquid chromatography (HPLC) using a optiguard C-18 guard column (1.0mm) from Optimize Technologies (Oregon City, OR) connected to a TSKgel ODS-100V C-18 column (1.0mm x 5.0cm) from TOSOH BioScience (Japan).

3.2 Proteolysis

Two denaturation methods were used. In the first method (Protocol 1), proteins were subjected to thorough denaturation detailed by Medzihradszky[135] prior to proteolytic digestion. Proteins were incubated at 60°C in a solution of 10 mM DTT and 40% acetonitrile (ACN).

Cysteine-containing proteins were then alkylated with 50 mM iodoacetamide (37°C, 1 hr). Prior to proteolytic digestion, the reaction mixtures were diluted 1:4 to provide conditions that were compatible with the proteases that were subsequently employed, trypsin and chymotrypsin. In the second method (Protocol 2), proteins were denatured in accordance to a Promega protocol prior to proteolytic digestion. Proteins were incubated for 60°C for 1 hr in a solution of 8 M urea, 50 mM Tris-HCl (pH 8), 4 mM DTT. After denaturation, the reaction was allowed to cool and 50 mM NH_4HCO_3 (pH 7.8) was added until the urea concentration was below 1 M. After denaturation, the protease enzyme was added. Typically an enzyme/protein ratio of 1:20 (wt/wt) was sufficient for complete digestion when incubated overnight (~16hr) at 37°C.

For pepsin, denaturation and digestion were carried about in acidic conditions requiring a pH below 2.0. For our purposes, a pH of 1.74 was established through the use of hydrochloric acid (HCl). Typically an enzyme/protein ratio of 1:20 (wt/wt) was sufficient for complete digestion when incubated (~24hr) at 30°C.

3.3 Mass Spectrometry

The instrumentation employed in this research is a Thermo LCQ Deca XP (Thermo-Fischer Scientific, West Palm Beach, FL) equipped with a conventional electrospray interface. This instrument features a quadrupole ion trap mass spectrometer capable of high order tandem mass spectrometry through collision-activated dissociation. It is capable of unit resolution full scan and has a mass range capability from 100 – 4000 m/z. The mass spectrometer is equipped with a ThermoSurveyor LC-5 Autosampler and a Surveyor MS Pump as inlet system. The ESI source was operated at a spray voltage of 5.0 kV in the positive ionization mode, a nitrogen nebulizer flow of 20 arbitrary units, transfer capillary temperature of 200°C, transfer capillary voltage of 30 V, and tube lens offset of 15 V. For binding interaction studies, KDS100 syringe pump (KD Scientific, Holliston, MA) with a 2.5 mL syringe (Hamilton, Reno, NV) was used to deliver the solutions of varying concentrations of naproxen, RGD peptide, and ATP/ADP in different buffers at different flow rates through PEEK tubing (internal

volume of 115 µl; Upchurch Scientific, Oak Harbor, WA) interfaced to the normal flow path using a T-junction prior to passing the mixture into the ESI source. Sample mixtures (20µL injection) were introduced into the ESI source through the standard injection valve on the LCQ system at a flow rate of 50µl/min. The experimental setup is shown in figure 3.1.

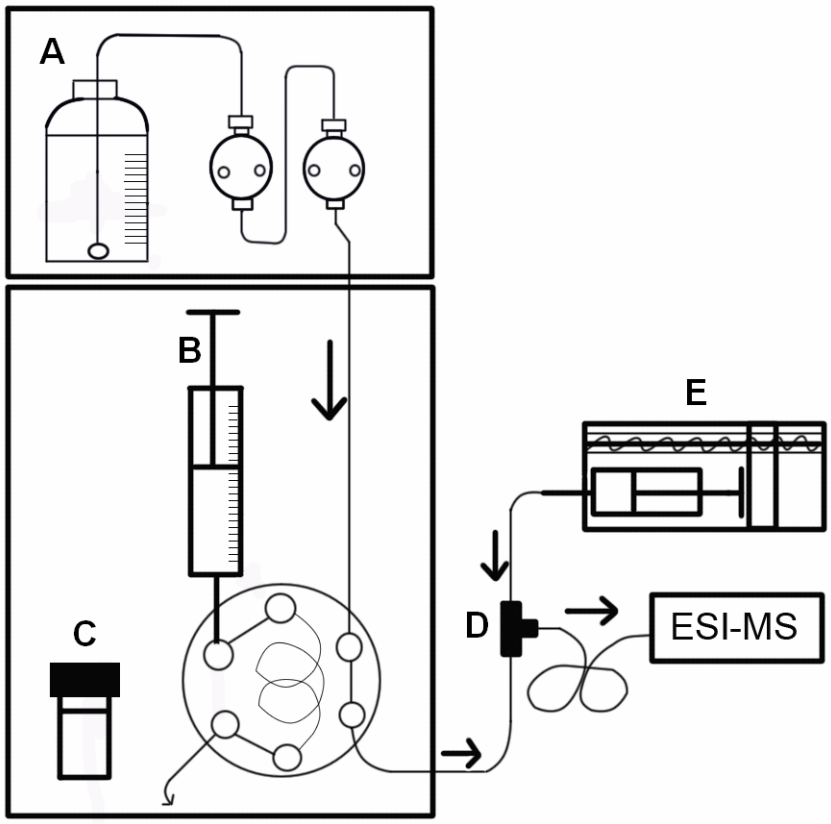


Figure 3.1 Schematic of Experimental Setup
A) Solvent B) Injection of Host C) Sample Carrier D) T-junction
E) Injection of Guest

3.4 Experimental Emphasis

Three protein/ligand models (BSA and naproxen, Integrin $\alpha_V\beta_3$ and RGD peptide, cytochrome c and ATP/ADP) will be explored. The significance and applicability of each model as a determinant for a high-throughput screening technique for protein/ligand interactions will be

investigated. All experimental methods detailed below are geared towards maintaining an efficient high-throughput method. Although ESI-MS analysis is an effective method for studying various non-covalent protein-ligand complexes, there are several limiting factors which must be considered.

Most of the pertinent literature regarding protein/lignand complexes focuses on the study of the protein's intact binding site. Each protein's native structure is often characterized by a unique tertiary structure with tightly packed amino acids buried within the protein interior. Oftentimes, multiple binding sites consisting of both high affinity and non-specific binding sites exist throughout the globular surface of each protein. A top-down proteomics approach would seem like the correct approach because of how a protein's primary, secondary, and tertiary structures play an integral role in binding.

However, the current ESI-MS methodology requires protein complexes to be detected within the mass range of the mass spectrometer. Multiple charging allows higher molecular weight proteins and protein complexes to be observed at lower m/z values, but then the resolution and of the instrument becomes a restrictive factor. For many, the resolution and mass range of the mass spectrometer are the limiting parameters in effectively detecting and analyzing large protein-ligand complexes.

Some individuals are therefore constrained to work with digests of proteins where peptide fragments are much more manageable. This bottom-up approach may have its limitations in accurately and quantitatively describing exact binding affinities relating to protein/ligand interactions, but the design of the experiment is proof of principle. The goal is to see the ligand of interest interacting with an identified peptide fragment of a binding site of a protein and consequently develop an effective high throughput analytical method for the elucidation and molecular recognition of protein receptor domains.

To identify sequential domains, manipulation of protein denaturation conditions in combination with enzymatic digestion via an assortment of enzymes (i.e., pepsin, trypsin, etc.)

will be used to generate varying peptide fragments of different lengths. The main goal is to have the specific binding region of the protein remain intact within the peptide mixture. Several ways have already been explored to obtain peptide fragments containing the sequential domain of interest including limited proteolysis techniques, chemical cleavage, genetic engineering, and crosslinking experiments [150-152]. These techniques are often time-consuming and require intense data management/analysis. The experimental emphasis will be therefore placed on basic proteolytic techniques which facilitate a high-throughput process.

Having a peptide fragment that contains the binding region with multiple amino acids flanking that binding region is a main area of focus. In addition to focusing on small peptide fragments (10a.a. to 20a.a.), an effort will be made to have the binding site contained within larger peptide fragments (20a.a.). This would simulate the original nature of the protein, giving better binding data. It has been surmised that peptide fragments greater than 20 amino acids may retain a position of their native structure, and thus participate in specific interactions [153].

CHAPTER 4

MS PROTEOMICS AND MODELS OF STUDY

This chapter provides the specifics relevant to the introduction and the experimental section. First, the pertinent information regarding proteolysis and MS proteomics associated to the experimental section will be presented. Next, the current MS-based methodologies for mapping protein-ligand interfaces mentioned in the introduction will be addressed. Finally, the protein/ligand complexes: BSA/naproxen, Integrin/RGD peptide, and Cytochrome C/ATP/ADP will be covered within this chapter.

4.1 MS Proteomics

4.1.1 Mapping of Primary Structure

Protease mapping is a way to determine the primary structure of a protein. It has gone through an evolutionary change from its early roots in chromatography/gel electrophoresis combined with Edman degradation sequencing to a more streamlined technique where the combination of mass spectrometry, proteolytic techniques, and computer-driven data analysis software (Mascot, PEAKS, OMSSA, SEQUEST, ProteinProspector, etc.) are used to determine the primary structure of a protein. By conducting mass analysis on a peptide mixture resulting from a protein proteolyzed by a sequence specific enzyme, the fragmentation pattern can be statistically compared to databases (NCBInr, Genpept, Swiss Prot, Owl, Ludwignr, Unknome, etc.) populated by already sequenced proteins. For example, NCBI's RefSeq database contains over 5 million proteins. With an appropriately selected protease that generates a significant number of peptide fragments, there is a reasonable probability of identifying the target protein if it is contained in the database.

4.1.1.1 Denaturation

Proteins fold into compact structures stabilized through their secondary and tertiary structures. In order to effectively map the primary structure of a protein, both the hydrophilic surface and hydrophobic core of the protein must be made available to the enzymes for proteolysis. Denaturation of the protein causes the secondary and tertiary structures to destabilize thereby allowing the enzymes to access previously inaccessible sites. Heat, acids, bases, heavy metal salts, detergents, and organic solvents are sometimes employed to achieve denaturing conditions [154]. Urea and guanidinium chloride are commonly used chemicals for the denaturation process. These reagents disrupt intramolecular hydrogen bonds which would otherwise hold the protein in a folded conformation. Tertiary elements such as disulfide bonds connecting cysteines to one another are main points of emphasis when attempting to denature a protein. Dithiothreitol is a commonly used reagent to reduce disulfide bonds. Iodoacetamide is used to cap the cysteines through an alkylation reaction following treatment with dithiothreitol [155]. Please refer to the section above, 3.2 Proteolysis, for two specific methods for denaturing a protein.

4.1.1.2 Proteases

A protease is an enzyme that causes the degradation of a protein. Proteases are exploited in proteomics for their specificity at cleaving certain amino acids within the primary sequence of a protein. Trypsin, Lys C, Lys N, CNBr, Arg C, Asp N, Chymotrypsin, Pepsin, Proteinase K are several site-specific proteases utilized in biochemistry. Although proteases are commonly employed in a standalone fashion, they are occasionally used in tandem for increased proteolysis. The most common and robust of the enzymes is trypsin. Trypsin primarily cleaves at the C-terminal side of the amino acids lysine and arginine, except when either is followed by proline. Chymotrypsin cleaves peptides at the C-terminal side of tyrosine, tryptophan, and phenylalanine. At an acidic pH, pepsin cleaves at the C-terminal side of phenylalanine and leucine.

4.1.2 Tandem Mass Spectrometry

Tandem mass spectrometry (MS/MS) is a commonly employed method for the mapping of primary structure. MS/MS techniques enable fragmentation of a peptide or protein, providing an in-depth analysis of the amino acid sequence and possible post-translational modifications. The fragmentation is accomplished by two methods: in-source fragmentation and post-source fragmentation. The most common post-source fragmentation uses collision-induced dissociation (CID) in a tandem mass analyzer. The idea behind MS/MS is to select a precursor ion, focus the ion in a collision cell, collide the ion with an inert gas, and then mass analyze the resulting product ions. The dissociation of a peptide usually requires an energy of 10-100eV. This low energy process tends to form product ions formed by small neutral losses and cleavage of peptide bonds[153]. Higher energy processes tend to dissociate backbone and side-chain bonds of peptides providing more information but also convoluting the data [156,157].

4.1.2.1 Peptide Fragmentation

The resulting fragmentation ions from MS/MS of a peptide by collisional dissociation are called b-ions and y-ions. a, b, or c-ions are the ions with the charge retained on the N-terminus while x, y, z-ions have the charge retained on the C-terminus. Figure 4.1 depicts peptide fragmentation nomenclature. Roepstorff and Fohlman [158] initially formed the nomenclature for product ion formation and Johnson [159] and Biemann [160,161] subsequently added modifications to the nomenclature. The general mechanism of peptide fragmentation has been studied extensively [162-167] and the process is well understood.

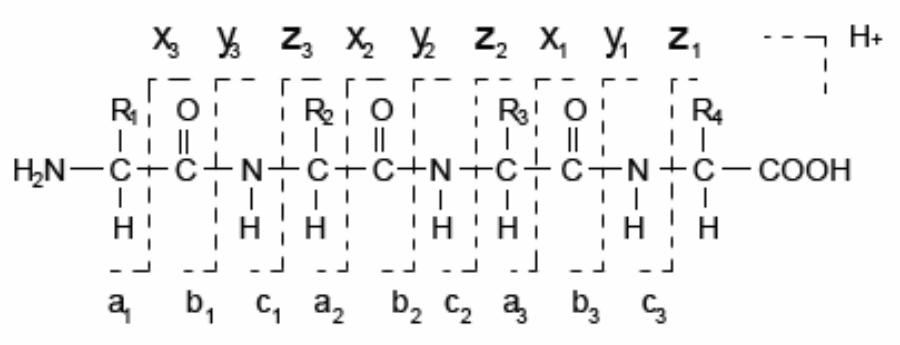


Figure 4.1 Peptide Fragmentation Ions: a, b, or c-ions have charge retained on the N-terminus

4.2 Mapping of Tertiary Structure

The determination of the interface region of the protein in contact with the ligand and the functionally significant amino acid residues is of interest in understanding biochemical function. Previously, x-ray crystallography and NMR have been the common methods for probing protein structure at the molecular level. Recently, mass spectrometry integrated with a combination of hydrogen-deuterium exchange, chemical modification, and limited proteolysis has emerged as an effective tool for identifying interaction regions.

4.2.1 H/D Exchange

Hydrogen-Deuterium (H/D) exchange combined with mass spectrometry can be used to determine interaction regions of a protein [168]. Hydrogen atoms are exchanged with deuterium atoms when the compound is dissolved in deuterium oxide. Those hydrogen atoms which are accessible to the solvent are rapidly exchanged with the deuterium while those located within the hydrophobic core of the protein are left untouched. For complexes, the ligand provides protection from H/D exchange at the interaction region. Protein in the presence and absence of ligand are deuterated, quenched by acid, enzymatically proteolyzed by pepsin, and then mass analyzed. By comparing the masses of peptic fragments, the interaction site can be determined.

4.2.2 Chemical Modification

Two chemical modification techniques include reactions towards specific functional groups on the protein and mutations during protein synthesis. The most common form of

chemical modification is the addition of a crosslinker to covalently link subunits of a protein complex. After crosslinking, factors such as the tertiary structure of the protein, stoichiometry of the noncovalent protein complex [169], and the protein complex interface region [170-172] can be determined.

4.2.3 Limited Proteolysis

In limited proteolysis, the protein is subjected to a protease in non-denaturing conditions. Since certain polypeptide chain regions of the proteins are buried within the core, they are protected from proteolysis. Only the exposed surface of the protein and the flexible sites are accessible to the protease. Like the H/D exchange example above, the unassociated protein is compared to the associated protein by mass analysis and the interaction site is determined. One of the most cited examples of limited proteolysis in relation to non-covalent protein complex is the transcription factor Max and Max-specific DNA complex [173]. Cohen et al were able to effectively probe the solution structure of MAX even though NMR or X-ray crystallography was unable to provide insight at the time.

4.3 Models of Study

In order to successfully study unknown protein/ligand interactions, it is favorable to examine a model protein and its ligand interactions that have already been described and characterized. Three models will be discussed within this section, each with a categorically unique protein/ligand interaction. The first model will pertain to BSA and naproxen, a protein and drug interaction. The second model discussed will be integrin and RGD, a protein and peptide interaction. And the third model will be cytochrome c and ATP/ADP, a protein and small multifunctional nucleotide interaction.

4.3.1 Bovine Serum Albumin and Naproxen

A prime example of a well-researched protein is serum albumin, a ubiquitous protein found in the serum of mammals such as human and bovine. Bovine serum albumin (BSA) is the serum albumin of interest because of its low cost; it is readily available in great quantities,

purified from bovine blood, a byproduct of the beef industry. BSA is a single-chain polypeptide chain consisting of 582 amino acids cross-linked by 17 cysteine residues, 16 of which are involved in Cys-Cys linkages. Belying its mass of 69 kDa, BSA is relatively compact globular protein with a stable helical structure.

Serum albumin plays an important role in the regulation of osmotic pressure between intravascular compartments and body tissues; therefore, it has been extensively studied as a prime target for drug delivery. This is evidenced by the wealth of studies on serum albumin protein interactions with drugs and other biomolecules accomplished by NMR, CD, ROD, Raman, ATR-FTIR, UV-Vis absorbance and fluorescence spectroscopy, equilibrium dialysis, ultrafiltration microdialysis, gel filtration, solid-phase microextraction, batch spectrophotometry and fluorimetry, potentiometry, chemiluminescence, capillary electrophoresis, bio-interaction chromatography, HPLC-MS, and ESI-MS [174-184]. Electrospray ionization mass spectrometry has also been used as a tool for the determination of drug binding sites of human serum albumin by probing non-covalent interactions [185]. Greig et al. determined the dissociation constants for oligonucleotide serum albumin complexes through ESI-MS [75].

Serum albumin is known to have a disposition for binding a wide range of ligands consisting of warfarin, azapropazone, indole, benzodiazepine-like compounds, digitoxin, biliary acids, aspirin, billirubin, fatty acids, and metal ions. These ligands bind to six main binding sites [186-188]. One of these binding sites is called Sudlow site II. Because of its hydrophobic pocket, Sudlow site II has a high affinity for small aromatic ligands such as L-tryptophan, ibuprofen, diazepam, naproxen, and medium-chain fatty acids. Binding at this site is tight and specific with a 1:1 stoichiometry [189]. Naproxen, shown in Figure 4.2, is a non-steroidal anti-inflammatory drug known for its analgesic effects, including reduction in fever, inflammation,

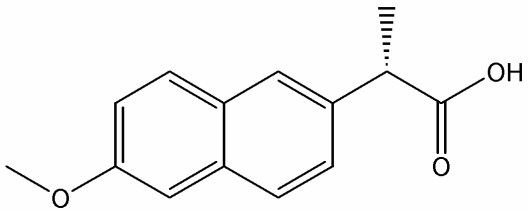


Figure 4.2 Structure of S-Naproxen

and stiffness caused by symptomatic diseases. Naproxen binds with serum albumin mainly through mixed electrostatic and hydrophobic interactions at Sudlow site II. The electrostatic and hydrophobic interactions are mediated by the hydrophobic amino acid residues (Trp214, Leu219, Phe223, Leu234, Leu238, Leu260, Ile264, Ile290, and Ala291) in the Sudlow II binding pocket and the hydrophilic amino acid residues (Arg222, Arg257, His242) on the surface of the pocket [190].

4.3.2 Integrin and RGD peptide

Integrins are integral membrane proteins in the plasma membrane of cells. They play an integral role in the attachment of a cell to the extracellular matrix (ECM) and in signal transduction from the ECM to the cell. Integrin receptors control a multitude of functions such as cell growth, division, signaling, apoptosis, adhesion, and survival [191-193]. Consequently, overexpression of these integrins has been shown to play an important role in cancerous cells [194-198].

The RGD (Arg-Gly-Asp) tripeptide sequence is a common recognition motif for a broad range of integrin receptors [199,200]. Thus, inhibition of integrin through RGD-variants is viewed as a target for the creation of possible anti-cancer agents. The integrin, $\alpha_V\beta_3$, a vitronectin receptor, is a commonly studied integrin due to its ubiquity in several cell types. With $\alpha_V\beta_3$ reported as having a function in tumor angiogenesis and invasiveness [201-203], Dechantsreiter et al. were able to create an angiogensis inhibitor, cyclopeptide N-alkylated cyclo(-RGDf[NMe]V-), that is now in clinical trials [204,205]. Other attempts are currently being made to quantify and characterize the binding of integrins to RGD variants.

Studies on integrin-ligand interactions have included x-ray diffraction, nuclear magnetic resonance spectroscopy, electron microscopy, molecular modeling, immuno-affinity labeling, tagging, and site-directed mutagenesis [12]. More recently, integrin-RGD interactions have been studied by ESI-MS and frontal analysis capillary electrophoresis (FACE) for the determination of absolute binding constants [148]. This study conducted by Raji et al. was able to screen peptide fragments of integrin $\alpha_V\beta_3$ containing the specific binding pocket against synthetic RGD-based peptides.

4.3.3 Cytochrome c and ATP/ADP

Cytochrome c is a protein located on the outer surface of the inner mitochondrial membrane. Its role is to shuttle electrons between complex III and complex IV in the electron transport chain [206]. Found in the mitochondria of all eukaryotic organisms, cytochrome c is one of the best-studied examples of homology. Out of its 100 amino acids (~12.5kDa), 28 amino acids in the polypeptide chain are absolutely conserved across 40 species [207]. These invariant residues play a major role in biological function of the protein and are well studied.

Some key facts are of interest for relevant portions of this thesis. First, cytochrome c contains a heme group, derived from iron protoporphyrin IX, surrounded primarily by hydrophobic amino acid residues [207]. Second, cytochrome c is one of the most studied cytochromes because it is the only one that is water soluble [207]. Third, cytochrome c binds ATP to modulate the electron transfer rate between cytochrome c's redox partners [208]. Lastly, it has been shown that cytochrome c and ATP (shown in figure 4.3) play a role in programmed cell death or apoptosis [209,210].

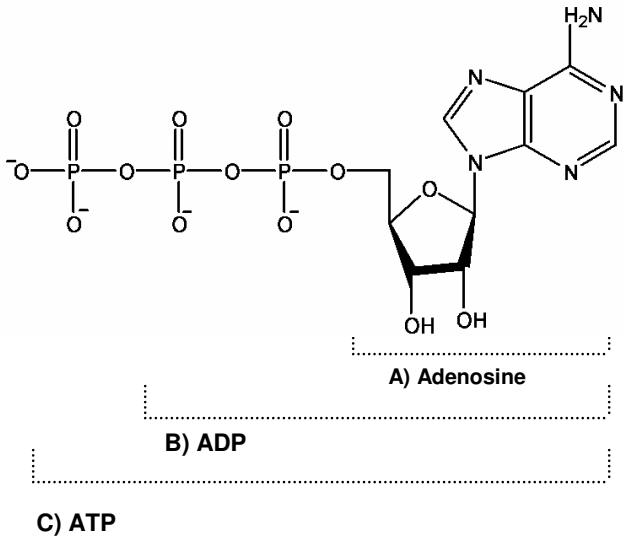


Figure 4.3 Structure of ATP/ADP: A) Adenosine Nucleotide
B) Adenosine diphosphate (ADP) C) Adenosine triphosphate (ATP)

Because cytochrome c and ATP are linked to apoptosis, this has driven the resurgence of numerous studies regarding cytochrome c and ATP interactions. In relation to non-covalent interactions, it was first noted through several analytical techniques that cytochrome c was capable of specifically binding certain anions, particularly phosphate [211-214]. Later it was shown by Corthesy & Wallace that cytochrome c bound ATP at a specific redox-state-dependent binding site, Arg-91, and at two other non-specific binding sites characterized by lower affinity [212]. The binding of ATP to cytochrome c is driven by electrostatic interactions [215] between the basic side chain of Arg-91 and a phosphate group on ATP. Proof of specificity was determined through disrupting the interactions between cytochrome c and ATP by replacing the conserved Arg-91 residue with norleucine in mutagenesis studies [208,214,216].

The linkage of misfolded forms of cytochrome c to apoptosis has also driven the study of protein folding intermediates. These studies have been conducted to probe and understand the mechanisms underlying conformational transitions. Cytochrome c has been shown to

undergo conformation changes at different buffer conditions. In aqueous solutions of low ionic strength, cytochrome c is in the native state at neutral pH. Partially folded states can be induced by acidification [118-121]. At pH 2.0 to 3.0 (25^{o}C), cytochrome c has been shown to undergo a highly cooperative, acid-induced unfolding transition [122]. Unfolded states at low pH conditions ~ pH 2.0 can be stabilized by the addition of salts [122] or glycerol [217]. A destabilized state of the native structure can also be induced by the addition of methanol at pH 3.0 [218,219]. Electrospray ionization mass spectrometry has been used for detecting equilibrium cytochrome c folding intermediates. Grandori was able to detect two partially folded forms that populate the molten-globule state of cytochrome c [220]. Recent literature probes the connection between cytochrome c conformation and ATP in apoptosis. Sinibaldi et al. provided evidence that ATP specifically drives refolding of non-native conformations of cytochrome c [221].

CHAPTER 5

DATA/DISCUSSION

This chapter will first describe three software-based methods for the determination of amino acid sequence coverage. Next, based on sequence coverage the effectiveness of the two digestion protocols is discussed. Then, the detection of the binding site is explored for the three proteins, BSA, integrin, and cytochrome c. Finally, the data and discussion of the results will be presented for each protein/ligand model.

5.1 Determination of sequence coverage

To achieve the greatest probability of obtaining the binding domain within a peptide mixture, sequence coverage and identification of peptide containing the putative ligand binding site (for model systems) were used as the main determinants in choosing the correct proteolytic technique. Several methods of sequence coverage determination and peptide identification were available and explored. The first was a web-based peptide fingerprinting application while the other two were peptide fingerprinting software programs bundled with the Thermo-Finnigan MS equipment.

5.1.1 Protein Prospector

The first method was a peptide mass fingerprinting tool from UCSF's Mass Spectrometry Facility called Protein Prospector MS-FIT [222]. It allows the user to fit mass spectrometry data (mass list) to a protein sequence database as shown in Figure 5.1. Due to a previous familiarity with this web-based peptide fingerprinting software, MS-FIT was first explored as an option for protein identification and sequence coverage. However, this method did not end up being an effective tool for the evaluation of ESI-MS data but still remains a powerful complement for use with MALDI-MS data. This is due to the fact that MALDI-MS is

better suited and efficient in generating accurate mass lists from averaged spectra; whereas, the averaged mass list from an ESI-MS full-scan chromatogram is not as 'clean' and may contain artifacts.

[–] **Results Summary**

Protein Hit Number	MOWSE Score	# pep # mat % mat 150 pks	% Cov	% TIC	Mean Err Da	Data Tol Da	# Hom Prot	MS-Digest Index #	Protein MW (Da)/pI	Accession #	Species	Protein Name
1	2.91e+8	29/37/25	27.8	24.7	0.219	1.05	10	1075960	69294/5.8	1351907 M	BOVIN	Serum albumin precursor (Allergen Bos d 6) (BSA)
2	1.53e+6	23/30/20	16.5	20.0	-0.0194	1.12	10	4217102	65820/8.4	119894366	BOS TAURUS	PREDICTED: hypothetical protein, partial
3	969061	32/40/27	25.7	26.7	0.0213	1.32	3	2070565	65239/5.2	77736151 M	BOS TAURUS	hypothetical protein LOC534150
4	262743	30/37/25	24.5	24.7	0.0367	1.26	No	4237039	67458/7.0	119902570	BOS TAURUS	PREDICTED: similar to semaphorin 6D isoform

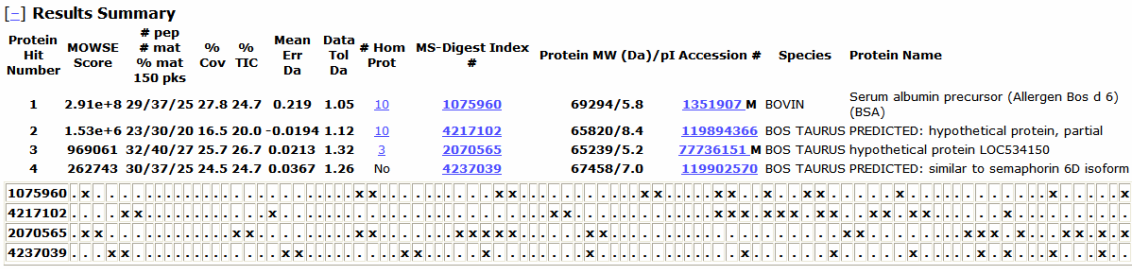


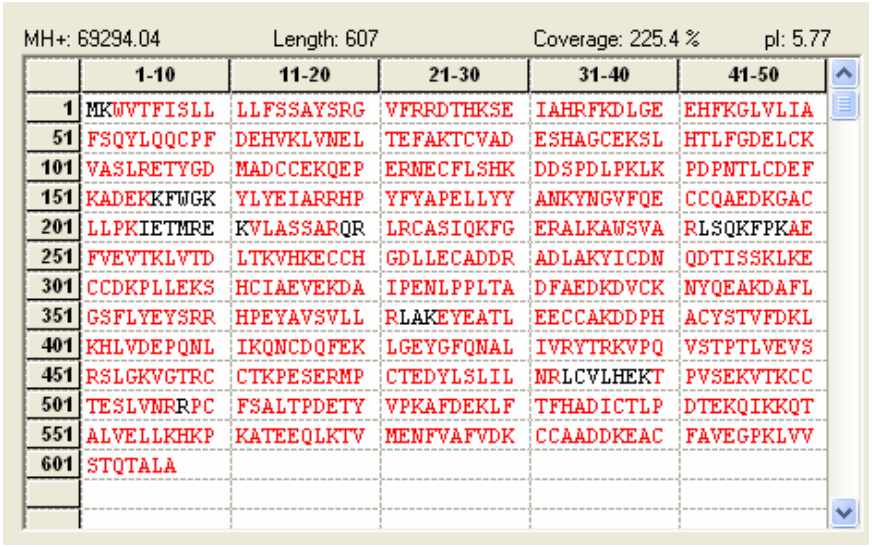
Figure 5.1 Percent Coverage of BSA through Protein Prospector. Several hits result from database search; however, desired BSA is ranked first with a % Coverage of 27.8%.

5.1.2 Pepmap

Within the suite of applications from Thermo's BioWorks™ was a Pepmap application where a known protein amino acid sequence was inputted with several parameters: type of protease, number of missed cleavages, TIC%, and chemical modifications. The mass list generated from the chromatogram was matched to the parameter-driven theoretical data set and a percent coverage was provided as shown in Figure 5.2. Because of its speed and ease of use, this method was used to determine the effectiveness of several digestion protocols discussed later within this chapter.

5.1.3 TurboSEQUEST

The other BioWorks application, TurboSEQUEST, identifies proteins through a SEQUEST-based search algorithm which takes tandem mass spectrometry (MS/MS) data and compares it to standard protein databases. The probability-based scoring algorithm ranks SEQUEST hits (Green>Yellow>White) and allows the user to determine whether or not there was a false positive as shown in Figure 5.3. Through the strength of tandem mass spectrometry data, this method was used for its accuracy and reliability for confirming potential peptides of interest.



MH+: 69294.04 Length: 607 Coverage: 225.4 % pI: 5.77

	1-10	11-20	21-30	31-40	41-50
1	MKWVTFISLL	LLFSSAYSRG	VFRRDTHKSE	IAHRFKDLGE	EHFKGLVLIA
51	FSQYLQQCPF	DEHVKLVNEL	TEFAKTCVAD	ESHAGCEKSL	HTLFGDELCK
101	VASLRETYGD	MADCCEKQEP	ERNECFLSHK	DDSPDLPKLK	PDPNTLCDEF
151	KADEKKFWGK	YLYEIARRHP	YFYAPELLYY	ANKYNGVFQE	CCQAEDKGAC
201	LLPKIETMRE	KVLASSARQR	LRCASIQKFG	ERALKAWSVA	RLSQKFPKAE
251	FVEVTKLVTD	LTKVHKECCH	GDLLECADDR	ADLAKYICDN	QDTISSKLKE
301	CCDKPLLEKS	HCIAEVEKDA	IPENLPPLTA	DFAEDKDVCK	NYQEAKDAFL
351	GSFLYEYSRR	HPEYAVSVLL	RLAKEYEATL	EECCAKDDPH	ACYSTVFDKL
401	KHLVDEPQNL	IKQNCDQFEK	LGEYGFQNAL	IVRYTRKVPQ	VSTPTLVEVS
451	RSLGKVGTRC	CTKPESERMP	CTEDYLSLIL	NRLCVLHEKT	PVSEKVTKCC
501	TESLVNRRPC	FSALTPDETY	VPKAFDEKLF	TFHADICTLP	DTEKQIKKQT
551	ALVELLKHKP	KATEEQLKTV	MENFVAFVDK	CCAADDKEAC	FAVEGPKLVV
601	STQTALA				

Figure 5.2 Percent Coverage of BSA through PepMap (MS data only)
Red indicates detected regions of amino acid sequence while black indicates undetected regions of amino acid sequence

	Reference				P (pro)	Score	Coverage	MW	Accession	Peptide (Hits)	
	Scan(s)	Peptide	MH+	z	P (pep)	XC	ΔCn	Sp	RSp	Ions	Count
	ABBOS serum albumin precursor [validated] - bovine				7.3e-010	162.3	27.8	69225.6	418694	21 (18 0 1 1 1)	
	100 - 104	K.HLVDEPQNLIK.Q	1305.72	1	6.5e-002	0.797	0.065	8.0	30	8/50	1
	1328	K.SLHTLFGDELCK.V	1362.67	1	5.6e-002	1.396	0.159	6.5	8	8/55	1
	1330	K.SLHTLFGDELCK.V	1362.67	2	9.4e-005	3.357	0.391	606.3	1	22/55	1
	1344 - 1348	K.SLHTLFGDELCK.V	1362.67	2	5.9e-007	3.892	0.437	677.4	1	24/55	1
	1392 - 1400	K.LGEYGFQNALI.I	1111.54	1	1.7e-003	1.481	0.195	98.4	2	18/45	1
	1402 - 1410	K.LVNELTEF.A	964.50	1	4.3e-002	1.265	0.056	110.1	4	12/35	1
	1424 - 1432	K.LGEYGFQNALIVR.Y	1479.80	2	7.2e-007	5.344	0.502	2084.5	1	31/60	1
	1436	K.LGEYGFQNALIVRYTRKVP.Q	2224.22	3	6.4e-003	2.308	0.038	629.1	2	32/180	1
	1438 - 1446	K.DAIPENLPPLTADFAEDKDVCK.N	2401.16	2	7.3e-010	3.928	0.527	411.7	1	27/105	1
	1448	K.LFTFHADICTLPDTEK.Q	1850.90	2	2.3e-008	3.941	0.339	776.2	1	33/75	1
	1448	H.AGCEKSLHTLFGDELCK.V	1850.88	2	1.0	2.331	-	401.3	4	25/80	1
	1452 - 1460	K.DAIPENLPPLTADFAEDKDVCKNYQE.A	2935.37	3	1.6e-004	2.015	0.247	208.9	1	35/250	1
	1480 - 1484	Y.FYAPELLYYANK.Y	1491.75	2	1.8e-007	4.859	0.508	1391.0	1	35/55	1
	1488	Y.FYAPELLYYANK.Y	1491.75	2	5.4e-007	4.721	0.515	1255.3	1	32/55	1
	1490	R.HPYFYAPELLYYANK.Y	1888.93	2	1.5e-008	1.583	0.192	150.1	1	20/70	1
	1538 - 1542	K.TVMENFVAFVDK.C	1399.69	2	1.4e-005	4.202	0.425	735.7	1	24/55	1
	1558 - 1566	K.DAFLGSFLYEYSR.R	1567.74	2	1.3e-006	5.464	0.372	1581.3	1	27/60	1
	1576 - 1584	R.MPCTEDYLSLILNR.L	1667.81	2	2.6e-006	5.224	0.437	864.3	1	28/65	1
	1592 - 1600	L.ECADDRADLAK.Y	1206.54	1	1.0	0.223	0.024	2.8	113	7/50	1
	1614 - 1620	K.GLVLIAF.S	732.47	1	0.2	0.912	0.299	188.8	1	11/30	1
	1622 - 1624	K.DAFLGSFLYEY.S	1324.61	1	1.5e-002	1.783	0.436	202.2	1	17/50	1

Figure 5.3 Percent Coverage of BSA through SEQUEST (MS/MS data)

5.2 Analysis of digestion protocols

Two digestion protocols, protocol 1 and 2, were tested and compared. At similar concentrations, it was determined that protocol 2 provided greater sequence coverage when compared to protocol 1 for both Cytochrome C and BSA as shown in Figure 5.4. Cytochrome C displayed a slight variance in sequence coverage while in contrast BSA exhibited drastic

differences in sequence coverages returning 8% coverage for protocol 1 and 73% coverage for protocol 2. This is attributed to the larger molecular mass of BSA where the denaturation and efficiency of the trypsin digest in protocol 1 is not as robust as that of protocol 2. From these findings, protocol 2 was chosen as the method for obtaining trypsin and chymotrypsin digests of both Cytochrome C and BSA.

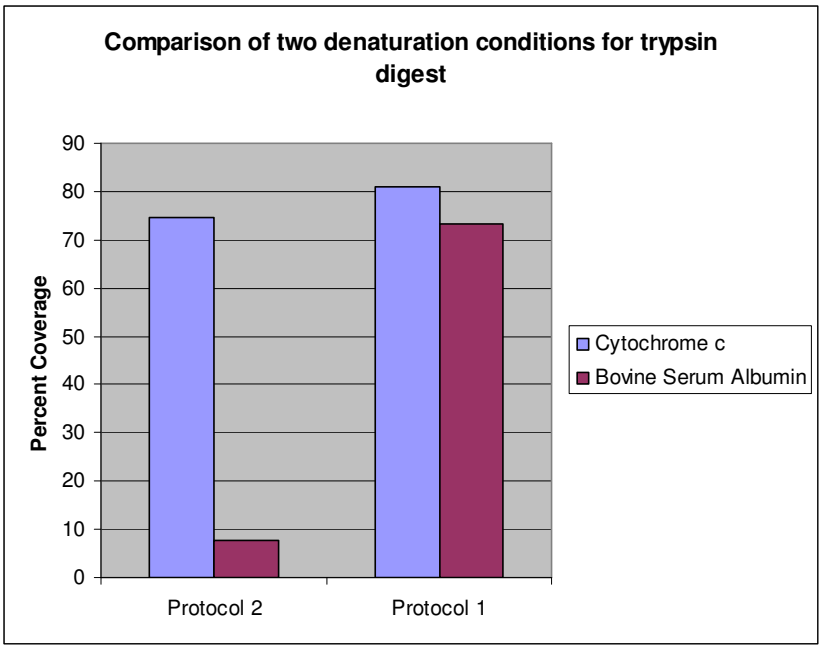


Figure 5.4 Comparison of Denaturation Protocols

Two incubation conditions for denaturation of the protein were performed for protocol 2, 60^{o}C for 1hr and 95^{o}C for 20 minutes. Both temperatures were tested with varying concentrations of protein (10 µM, 25 µM, 50 µM, 100 µM, 150µM). At lower concentrations (10 µM, 25 µM), an additional volume (2.5-5 times greater than the performed volumes) was added to account for imprecision in pipette pickup of small trypsin quantities (<1µL). As shown in Table 5.1, it was found that the optimum starting concentration of protein was 25 µM and the optimum denaturing condition was 60^{o}C for 1 hr. Figure 5.5 provides a view of the percent coverage of trypsin digested BSA at the two denaturing temperatures of 60^{o}C and 95^{o}C. A typical

chromatogram of a protein digest passed through a RP-HPLC C-18 column is demonstrated in Figure 5.6 along with ESI-MS-detected peptide fragments corresponding to each peak.

A) BSA @ 60°C (promega protocol)

Concentration (uM)	Volume (uL)	TIC 1.0%	TIC 3.0%
10	100	59%	9%
10	500	67%	38%
25	100	65%	17%
25	500	96%	66%
50	100	94%	67%
100	100	98%	53%
150	100	89%	55%

C) CytoC @ 60°C (promega protocol)

Concentration (uM)	Volume (uL)	TIC 5.0%
10	200	15%
10	500	36%
25	200	67%
25	500	30%
50	200	66%
100	200	62%
150	200	62%

B) BSA @ 95°C (promega protocol)

Concentration (uM)	Volume (uL)	TIC 1.0%	TIC 3.0%
10	100	51%	13%
10	500	83%	14%
25	100	87%	22%
25	500	99%	49%
50	100	92%	44%
100	100	94%	50%
150	100	89%	60%

D) CytoC @ 95°C (promega protocol)

Concentration (uM)	Volume (uL)	TIC 5.0%
10	200	15%
10	500	32%
25	200	15%
25	500	15%
50	200	31%
100	200	41%
150	200	23%

Figure 5.5 Comparison of two denaturation temperature settings through percent coverage: A) Bovine Serum Albumin at 60°C; B) Bovine Serum Albumin at 95°C; C) Cytochrome c at 60°C; D) Cytochrome c at 95°C. Highlighted portions show optimum percent coverage. TIC% indicates threshold at which the detection of the total ion chromatogram (TIC) is set at. Solutions were all tryptic digests.

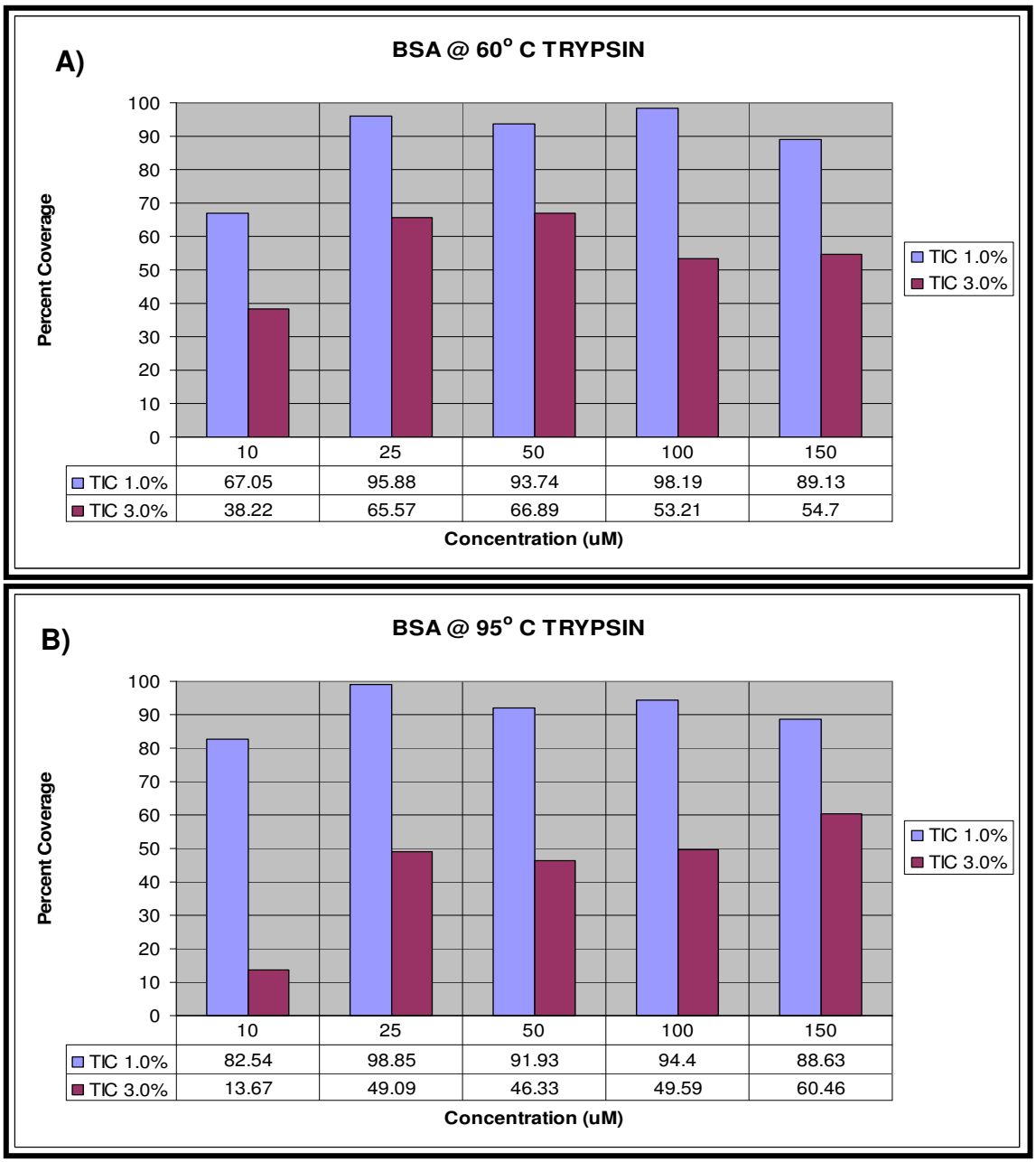


Figure 5.6 Graphical view of percent coverage of BSA digested with trypsin at two denaturing temperatures: A) Bovine serum albumin denatured at 60°C; B) Bovine serum albumin denatured at 95°C. TIC% indicates threshold at which the detection of the total ion chromatogram (TIC) is set at.

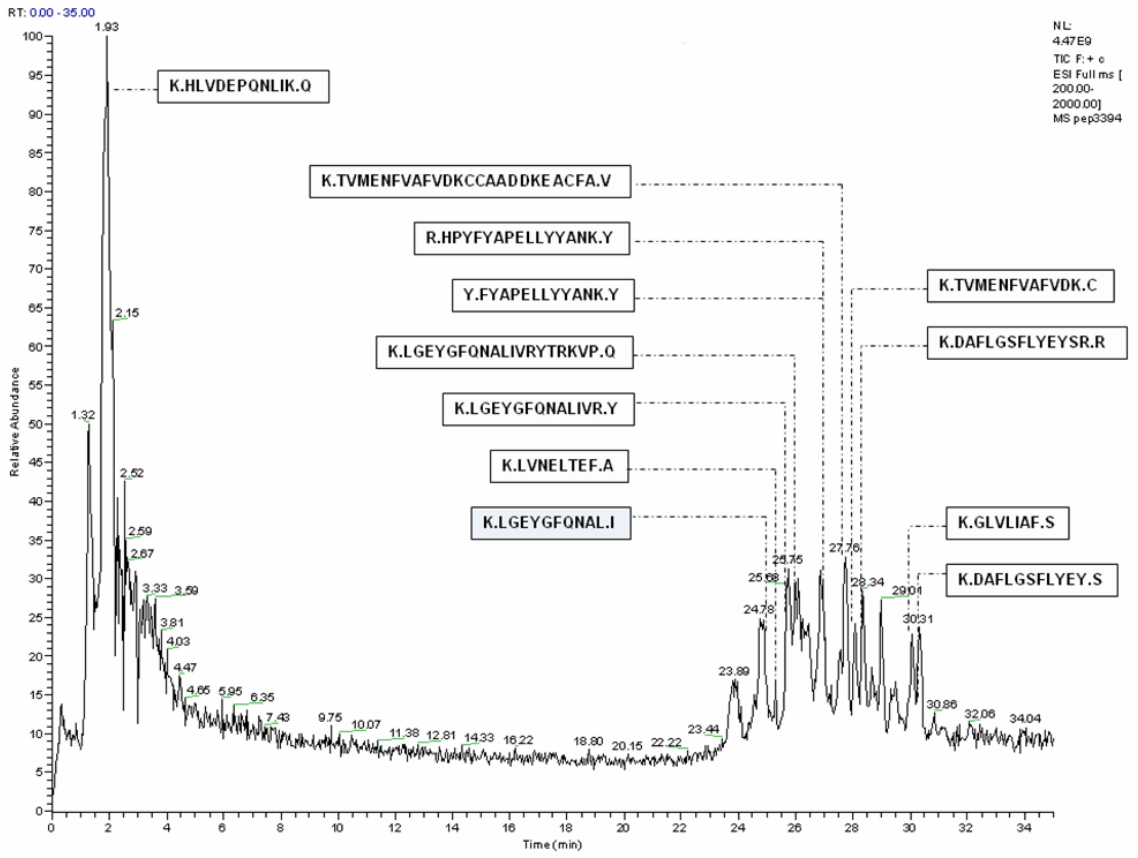


Figure 5.7 Total Ion Chromatogram of a 25 µM BSA trypsin digest with peptide fragments assigned to peaks through full-scan ESI-MS detection

For integrin, protocol 1 was used instead of protocol 2 due to the limited quantity of the protein and the dilute nature of the solution in which the protein was suspended. At a concentration of 1 mg/ml or 9.5 µM, the integrin protein was much harder to work with in comparison to the more abundant solid protein crystals of cytochrome C and BSA in regards to protocol 2. The final digest of integrin provided several unique peptide fragments. Figure 5.7 offers an alternate view of the peptide fragments detected via RP-HPLC-ESI-MS in relation to the amino acid sequence of the protein.

Figure 5.8 Detection of Integrin peptide fragments at 5 %TIC, 3% TIC, and 1% TIC

5.3 Detection of Amino Acid Binding Sites

Reported binding sites of proteins for their respective ligands are shown in Figures 5.8, 5.9, and 5.10. The mixed electrostatic-hydrophobic nature of BSA's IIA binding site is highlighted by the hydrophobic residues (Trp237, Leu242, Phe246, Leu257, Leu261, Leu283, Ile287, Ile313, and Ala314) in Figure 5.10 [190]. Cytochrome c's ATP/ADP binding site at Arg91, an evolutionarily conserved site, is shown in Figure 5.9 [212]. Figure 5.10 provides the sequence

BSA

```
  1 mkwvtfisll llfssaysrg vfrrdthkse iahrfkdlge ehfkglvlia fsqylqqcpf
 61 dehvklvnel tefaktcvad eshagceksl htlfgdelck vaslretygd madccekqep
121 ernecflshk ddspdlpklk pdpntlcdef kadekkfwgk ylyeiarrhp yfyapellyy
181 ankyngvfqe ccqaedkgac llpkietmre kvltssarqr lrcasiqkfg eralkawsva
241 rlsqkfpkae fvevtklvtd ltkvhkecch gdllecaddr adlakyicdn qdtissklke
301 ccdkplleks hciaevekda ipenlpplta dfaedkdvck nyqeakdafl gsflyeysrr
361 hpeyavsvll rlakeyeatl eeccakddph acystvfdkl khlvdepqnl ikqncdqfek
421 lgeygfqnal ivrytrkvpq vstptlvevs rslgkvgtrc ctkpesermp ctedylslil
481 nrlcvlhekt pvsekvtkcc teslvnrrpc fsaltpdety vpkafdeklf tfhadictlp
541 dtekqikkqt alvellkhkp kateeqlktv menfvafvdk ccaaddkeac favegpklvv
601 stqtala
```

Figure 5.9 BSA amino acid sequence. Highlighted and underlined are the amino acids responsible for Naproxen interaction

Cytochrome C

```
  1 mgdvekgkki fvqkcaqcht vekggkhktg pnlhglfgrk tgqapgfsyt danknkgitw
 61 geetlmeyle npkkyipgtk mifagikkkg eredliaylk katne
```

Figure 5.10 Cytochrome c amino acid sequence. Highlighted and underlined are the amino acids responsible for ATP/ADP interaction

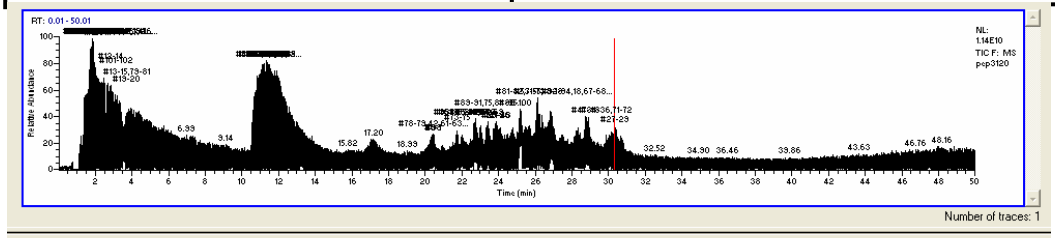
Integrin α_v subunit

```
  1 mafpprrrlr lgprglplll sglllplcra fnldvdspae ysgpegsyfg favdffvpsa
 61 ssrmfllvga pkanttqpgi veggqvlkcd wsstrrcqpi efdatgnrdy akddplefks
121 hqwfgasvrs kqdkilacap lyhwrtemkq erepvgtcfl qdgtktveya pcrsqdidad
181 gqgfcqggfs idftkadrvl lggpgsfywq gqlisdqvae ivskydpnvy sikynnqlat
241 rtaqaifdds ylgysvavgd fngdgiddfv sgvpraartl gmvyiydgkn msslynftge
301 qmaayfgfsv aatdingddy advfigaplf mdrgsdgklq evgqvsvslq rasgdfqttk
361 lngfevfarf gsaiaplgdl dqdgfndiai aapyggedkk givyifngrs tglnavpsqi
421 legqwaarsm ppsfgysmkg atdidkngyp dlivgafgvd railyrarpv itvnaglevy
481 psilnqdnkt cslpgtalkv scfnvrfclk adgkgvlprk lnfqvellld klkqkgairr
541 alflysrsps hsknmtisrg glmqceelia ylrdesefrd kltpitifme yrldyrtaad
601 ttglqpilnq ftpanisrqa hilldcgedn vckpklevsv dsdqkkiyig ddnpltlivk
661 aqnqgegaye aelivsiplq adfigvvrnn ealarlscaf ktenqtrqvv cdlgnpmkag
721 tqllaglrfs vhqqsemdts vkfdlqiqss nlfdkvspvv shkvdlavla aveirgvssp
781 dhiflpipnw ehkenpetee dvgpvvqhiy elrnngpssf skamlhlqwp ykynnntlly
841 ilhydidgpm nctsdmeinp lrikisslqt tekndtvagq gerdhlitkr dlalsegdih
901 tlgcgvaqcl kivcqvgrld rgksailyvk sllwtetfmn kenqnhsysl kssasfnvie
961 fpyknlpied itnstlvttn vtwgiqpapm pvpvwviila vlagllllav lvfvmyrmgf
1021 fkrvrppqee qereqlqphe ngegnset
```

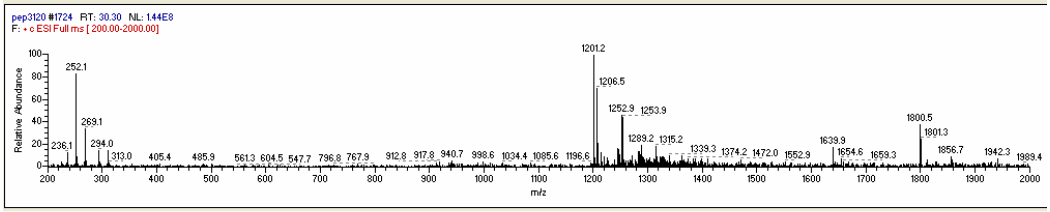
Figure 5.11 Integrin amino acid sequence. Highlighted and underlined are the amino acids responsible for RGD-peptide interaction

A

B



C



D

Modifications:

Fragment	From	To	MH+ (mono)	z	Time	Sequence
22-23	166	195	3255.44			[K]TVEYAPCRSQDIDADGQGFCQGGFSIDFTK[A]
22-24	166	198	3597.60	3	**30.33**	[K]TVEYAPCRSQDIDADGQGFCQGGFSIDFTKADR[V]
23	174	195	2336.01			[R]SQDIDADGQGFCQGGFSIDFTK[A]
23-24	174	198	2678.18			[R]SQDIDADGQGFCQGGFSIDFTKADR[V]
23-25	174	224	5450.63			[R]SQDIDADGQGFCQGGFSIDFTKADRVLLGGPGSFYWQGQLISDQVAEIVSK[Y]
24	196	198	361.18			[K]ADR[V]
24-25	196	224	3133.63			[K]ADRVLLGGPGSFYWQGQLISDQVAEIVSK[Y]
24-26	196	233	4213.16			[K]ADRVLLGGPGSFYWQGQLISDQVAEIVSKYDPNVYSIK[Y]
25	199	224	2791.47			[R]VLLGGPGSFYWQGQLISDQVAEIVSK[Y]
25-26	199	233	3871.00	4	**23.81**	[R]VLLGGPGSFYWQGQLISDQVAEIVSKYDPNVYSIK[Y]
25-27	199	241	4831.47	4	**23.03**	[R]VLLGGPGSFYWQGQLISDQVAEIVSKYDPNVYSIKYNNQLATR[T]
26	225	233	1098.55			[K]YDPNVYSIK[Y]
26-27	225	241	2059.02			[K]YDPNVYSIKYNNQLATR[T]
26-28	225	275	5607.66			[K]YDPNVYSIKYNNQLATRTAQAIFDDSYLGYSVAVGDFNGDGIDDFVSGVPR[A]
27	234	241	979.50	2	**11.57**	[K]YNNQLATR[T]
27-28	234	275	4528.13			[K]YNNQLATRTAQAIFDDSYLGYSVAVGDFNGDGIDDFVSGVPR[A]
27-29	234	278	4826.31	4	**30.30**	[K]YNNQLATRTAQAIFDDSYLGYSVAVGDFNGDGIDDFVSGVPRAAR[T]
28	242	275	3567.66			[R]TAQAIFDDSYLGYSVAVGDFNGDGIDDFVSGVPR[A]
28-29	242	278	3865.83			[R]TAQAIFDDSYLGYSVAVGDFNGDGIDDFVSGVPRAAR[T]
28-30	242	289	5106.45			[R]TAQAIFDDSYLGYSVAVGDFNGDGIDDFVSGVPRAARTLGMVYIYDGK[N]
29	276	278	317.19			[R]AAR[T]
29-30	276	289	1557.81			[R]AARTLGMVYIYDGK[N]
29-31	276	333	6370.96			[R]AARTLGMVYIYDGKNMSSLYNFTGEQMAAYFGFSVAATDINGDDYADVFIGAPLFMDR[G]

Figure 5.12 Detection of integrin α_v peptide fragment containing the putative binding domain: A) % Coverage at different TIC thresholds; B) TIC; C) MS spectra of peptide containing putative binding domain; D) Detection of peptide fragment through Pepmap.

of the extracellular segments of integrin subunit α_v and the amino acid RGD contacts, Asp180, Tyr208, Ala245, and Asp248 [223].

5.3.1 BSA

For trypsin, only one amino acid residue comprising the naproxen binding site of BSA was found via Tandem MS data through SEQUEST. This was a short sequence of 10 amino acids EXADDRADLAK (275-285) with the Leu283 amino acid bolded and underlined for reference. For pepsin, the amino acid sequences of WSVARL (237-242), SQKFPKAEF (243-251), VEVTKL (252-257) were found to contain the one or two of the hydrophobic amino acids of the binding site. With a maximum length of 10 amino acids comprising a single binding site, BSA digests failed to provide peptide fragments of substantial length.

5.3.2 Integrin

The amino acid residues comprising the RGD binding site of integrin α_v were initially not found in the analysis of the trypsin digest solution. An attempt was made to concentrate the solution through speed vacuum and increase the limit of detection to facilitate observation of a possible peptide possessing the binding domain. After concentrating the solution more than 10 fold, a potential peptide of 45 amino acids (234-278) containing the Ala245 and Asp248 contact points, as shown in Figure 5.11, was observed through the Pepmap application but was unable to be confirmed through tandem mass spectrometric data.

5.3.3 Cytochrome C

Tandem MS through SEQUEST confirmed a greater number of peptides hits for cytochrome c. Because of cytochrome c's smaller molecular size, denaturation and proteolysis were more efficient in generating a smaller number of peptides. As a result the chromatographic separation was more efficient, allowing for improved ionization and subsequent detection of various peptides. Trypsin, chymotrypsin, and pepsin were all useful in generating peptides containing Arg91. The largest peptide fragment of 14 amino acids was found through the pepsin digest as shown in Table 5.1. Thus, the success of cytochrome c digests for identifying the proper binding

sites prompted further investigation of this system for studying correlation between protein-ligand and peptide-ligand binding.

Table 5.1 Confirmation of probable binding domains through MS/MS

Protein	Protease	Amino Acid Sequence	MH+	% by Mass	Position	% by AA's
Cytochrome C	Trypsin	KKGEREDLIAYLK	1562.890068	13.51276338	87 - 99	12.5
		KGEREDLIAYLK	1434.795105	12.40525303	88 - 99	11.538462
		KGEREDLIAYLKK	1562.890068	13.51276338	88 - 100	12.5
		GEREDLIAYLK	1306.700142	11.29774268	89 - 99	10.576923
		GEREDLIAYLKK	1434.795105	12.40525303	89 - 100	11.538462
	Chymotrypsin	KGEREDLIAY	1193.616078	10.32001672	88 - 97	9.6153846
	Pepsin	IFAGIKKKGEREDL	1603.916617	13.86747934	81 - 94	13.461538
		FAGIKKKGEREDL	1490.832553	12.88975338	82 - 94	12.5
		AGIKKKGEREDL	1343.764139	11.61819838	83 - 94	11.538462
		AGIKKKGEREDLIA	1527.885317	13.21011195	83 - 96	13.461538
		EREDLIAYL	1121.583716	9.697224179	90 - 98	8.6538462
		REDLIAYL	992.5411224	8.58152061	91 - 98	7.6923077
Bovine Serum Albumin	Trypsin	ECADDRADLAK	1206.541927	1.742891961	275 - 285	1.8121911
	Pepsin	WSVARL	731.4198851	1.056561574	237 - 242	0.9884679
		SQKFPKAEF	1081.567672	1.562362283	243 - 251	1.4827018
		VEVTKL	688.4239674	0.994452469	252 - 257	0.9884679
Integrin	Trypsin	none	none	none	none	none

5.4 Protein/Ligand Studies

5.4.1 BSA/Naproxen

BSA and naproxen were selected as the first protein/ligand model of study. A 20 µL trypsin digest of BSA was injected through RP-HPLC-ESI-MS with the post-column addition of a blank solution containing 50/50:H_2O/ACN. The experiment was then repeated sequentially with 5 µM, 10 µM, 50 µM, and 100 µM naproxen in the same buffer solution. It was found that 12.5

µL/min was suitable for maintaining a constant flow required for the detection of ligand. Scanning was limited to obtaining only full scan data in order to preserve the complex.

For analysis, the trypsin digest of BSA with 0 µM naproxen was used as a blank. Pepmap was used to extract the probable peptide hits. Next, the related data containing retention time, charge state, and spectra mass was exported to Excel. In Excel, the theoretical monoisotopic mass of each peptide in addition to the peptide combined with one naproxen was calculated. All possible protonated and sodiated pseudo-molecular ions up to a 5+ charge state were considered.

The 5 µM naproxen raw data file was then compared to the 0 µM naproxen blank. A mass range filter was applied to both the blank and the 5 µM naproxen spectrum for each peptide and the retention times of the chromatographic peak associated with each peptide was recorded. A peak search was then performed for each possible BSA/naproxen ionic complex through a mass filter function. If a chromatographic peak appeared displaying the appropriate signal for a complex, the retention time was recorded and then compared to the uncomplexed peptide peak retention time for both blank and 5 µM naproxen spectra.

The method for extracting data was extremely time-consuming and laborious. Several key factors should be considered for a future improvement of this process. First, several of the peptide fragments detected by Pepmap are in all likelihood false positives. Additionally, there are likely numerous prospective peptide peaks that were not detected by Pepmap and therefore not considered in the analysis. Second, it is difficult to account for all the permutations regarding salt adduct formation, ligand stoichiometry, peptide charge state, and complex charge state. Finally, one must overcome the 3-dimensional quandary of having to consider chromatographic retention times, spectral mass according to individual scans, and ranges of intensities of ions in order to ascertain complex formation. For this process to become a high-throughput method, an algorithm must be written to sort the data, determine and confirm peptide fragments, and consider permutations with respect to identifying complex formation.

From the collected and analyzed data, there was no evidence of a complex between BSA proteolytic peptide fragments and naproxen. Specifically, there was no indication of complex formation between the Leu283 containing peptide, ECADDRAD**L**AK (275-285), and naproxen. This result was not unexpected because the normal interaction of naproxen and BSA is predominantly mediated through a combination of hydrophobic and electrostatic interactions [224-226]. As mentioned previously, non-covalent interactions through hydrophobic interactions are extremely labile in the gas phase [227]. Having only one or two hydrophobic amino acids preserved in a peptide fragment is not sufficient enough to maintain a stable complex with naproxen in the gas phase since in solution, such hydrophobic interactions are in a large part attributed to the role of the solvent. Additionally, if a peptide containing all 9 hydrophobic amino acids was isolated and bound to naproxen, the relative affinities measured by ESI-MS would probably not correlate to their solution affinities. As demonstrated by Robinson et al [116], ion abundances of complexes mediated through hydrophobic interactions observed in the gas-phase show poor correlation to solution phase binding affinities.

5.4.2 Integrin/RGD

Integrin α_v and the RGD peptide (GRGDSP) were selected as the second protein/ligand model of study. Due to the limited quantity and high cost of integrin, only one set of experiments was completed. A 20 µL speed vacuumed trypsin digest of integrin α_v was injected through RP-HPLC-ESI-MS with the post-column addition of a blank solution (50/50:H_2O/ACN) at 12.5 µL/min. The experiment was then repeated with a post-column addition of 2 µM GRGDSP peptide in the same buffer solution.

The method for extracting data was analogous to the technique used before but this time only the peptide of 45 amino acids (234-278) containing the Ala245 and Asp248 contact points was considered. GRDGSP and complex charge states of 1+ to 5+ were considered along with possible sodium adducts.

From the collected and analyzed data, there was no evidence of a complex formed between the selected peptide and the GRGDSP ligand. The peptide fragment analyzed was a fairly large peptide chain falling under the experimental criteria of being greater than 20 amino acids in length. It was hoped that because of its length, the peptide fragment would structurally mimic some of the native characteristics of the binding site [228]. From the relevant literature, it seemed possible that such an interaction between the featured peptide containing the two binding regions and RGD ligand could be seen. Raji et al [148] were able to observe complex formation between a synthetic peptide Ac-Ala-Gln-**Ala**-Ile-Phe-**Asp**-Asp-Ser-Tyr-Leu-Gly-NH_2 (243-254) and several RGD peptide variants. They were able to determine binding constants by ESI-MS titration for complexes formed between RGD peptides and the synthetic peptide. The positive identification of the complexation between a large integrin peptide fragment and RGD peptide may well have served as a nice compliment to Raji's research.

The lack of complex formation can be attributed to several factors. First, attempts made at confirming the validity of the large peptide fragment through tandem mass spectrometry data were unsuccessful. There is a possibility that the peptide was falsely identified and therefore lacked the binding capabilities for RGD ligand. Second, the studies of Raji et al. consumed pure samples of both synthesized peptide and RGD ligand at very controlled concentrations. In this case, the initial trypsin digest had to be concentrated attesting to the fact that the large peptide fragment was already in very dilute quantities. The lack of sufficient peptide concentration limits the detection of a complex which would be present at even lower concentration. For the detection of complex, future efforts should be made to either obtain more integrin protein or increase the sensitivity of detection through another MS instrument. The role of the solvent may have played a role during the chromatographic separations done in ACN/H_2O mixtures. These solvent conditions may have disturbed the complexes that would normally form under physiological conditions.

5.4.3 Cytochrome c/ATP and Cytochrome c/ADP

Cytochrome C and ATP/ADP were selected as the final protein/ligand model of study. Due to the success of identifying the proper binding site in cytochrome c, further investigation of this system was completed for studying the correlation between protein-ligand and peptide-ligand binding.

5.4.3.1 Peptide/Ligand

First, the peptide-ligand system was studied. A 20 µL trypsin digest of 25 µM cytochrome C was injected through RP-HPLC-ESI-MS with the post-column addition of a blank solution (50/50:H_2O/ACN 0.5% FA). The experiment was then repeated sequentially with 10 µM, 100 µM, and 500 µM of ATP and ADP in the same buffer solution. It was found that 12.5 µL/min was suitable for maintaining a constant flow required for the detection of ligand. Scanning was limited to obtaining only full scan data in order to preserve the complex.

The method for extracting data was similar to the technique utilized above. Only the peptides masses confirmed through tandem MS data were screened for possible binding to ATP and ADP. From the collected and analyzed data, there were numerous cases of complex formation between the peptide fragment and the ATP or ADP ligand. Figure 5.12 presents an example of the peptide [K]KGEREDLIAYLK[K] (88-99) from a trypsin digest exhibiting binding with increasing concentrations of ATP. Table 5.2 provides a view of all the peptides that exhibited complexation with ATP and ADP. Typically, if there was an observation of ATP binding to a peptide fragment, ADP was also observed to bind. As shown in table 5.2, trypsin, chymotrypsin, and pepsin digests each contained at least one case of a peptide containing the binding domain Arg91. However, it was observed that other peptide fragments lacking Arg91 complexed with ATP and ADP. The common thread was that all the peptides contained at least one basic amino acid: lysine, arginine, or histidine which facilitated the binding to ATP or ADP. This is supported by the literature where the phosphate groups of ATP establish strong electrostatic interaction with the basic side chains of Arg91 and Lys88 [221].

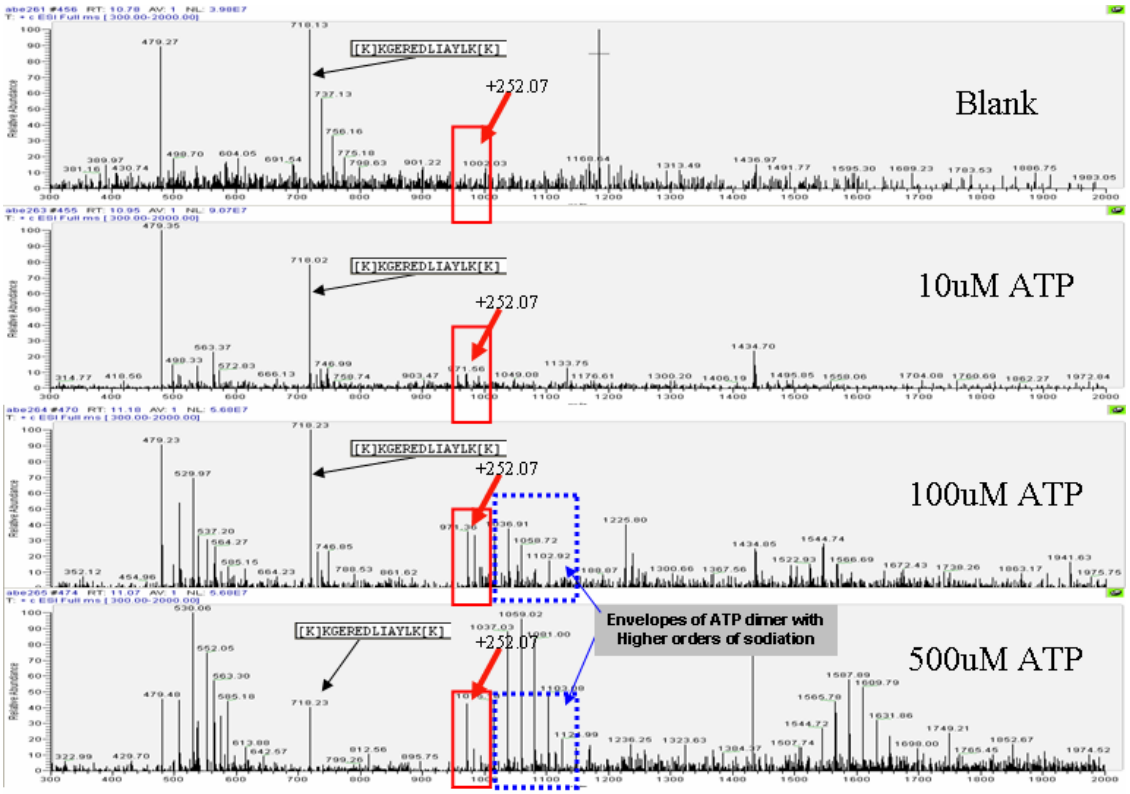


Figure 5.13 Peptide [K]KGEREDLIAYLK[K] (88-99) from trypsin digest exhibiting binding with increasing concentrations of ATP.

As ATP/ADP concentration increases, there is in turn a greater concentration of complex present. Attempts were made to determine if there was a correlative trend of increasing ATP/ADP concentration to complex appearance. Some of the limiting factors may have been the lack of numerous data points at different concentrations. It was observed that the fit to the 1:1 binding model began to deviate at higher concentrations hinting at the possibility of aggregation and non-specific binding.

The inability to effectively determine the concentration of each peptide fragment limits the scope of this approach. However, as proof of principle, this method established a possible screening technique for the binding domain between a peptide mixture and a ligand. A method

where one would be able to see the peptide fragments containing the binding domain show different character from those without the binding domain would have completed this story.

Table 5.2 List of peptides exhibiting complexation with ATP and ADP.
Highlighted sequences are sequences containing the binding site, Arg91.

Protease	Amino Acid Sequence
Trypsin	[K]TGPNLHGLFGR[K]
	[K]GITWGEETLMEYLENPKK[Y]
	[K]MIFAGIKKK[G]
	[R]EDLIAYLK[K]
	[K]KGEREDLIAYLK[K]
Chymotrypsin	[F]AGIKKKGEREDLIAY[L]
	[M]IFAGIKKKGEREDLIAY[L]
Pepsin	[F]AGIKKKGEREDL[I]
	[F]GRKTGQAPGFS[Y]
	[M]IFAGIKKKGEREDL[I]
	[Y]LENPKKYIPGTKMIF[A]
	[M]EYLENPKKYIPGTKMIF[A]

5.4.3.2 Protein/Ligand

For the protein-ligand system, three buffer systems: 50mM NH_4HCO_3, 50mM NH_4HCO_3 with 0.5% FA, and 50/50 H_2O/ACN 0.5% FA were explored for cytochrome c as shown in figure 5.13. These buffer systems were chosen to mimic physiological conditions, physiological conditions in the presence of an acid modifier, and chromatographic conditions, respectively. Folded and unfolded cytochrome c were seen as two charge envelopes where the folded states exhibited lower charge density while the unfolded states exhibited higher charge density. Figure 5.14 provides a closer look at the charge state distribution and assigns the charge to each corresponding peak. The buffer condition of 50mM NH_4HCO_3 with 0.5% FA was selected due to

the higher intensity of the lower charged state species (+6 and +5). These lower charge states were of interest due to the hope that the native conformation of the protein was preserved. An emphasis was subsequently placed on detecting ATP/ADP binding to the native states of cytochrome c.

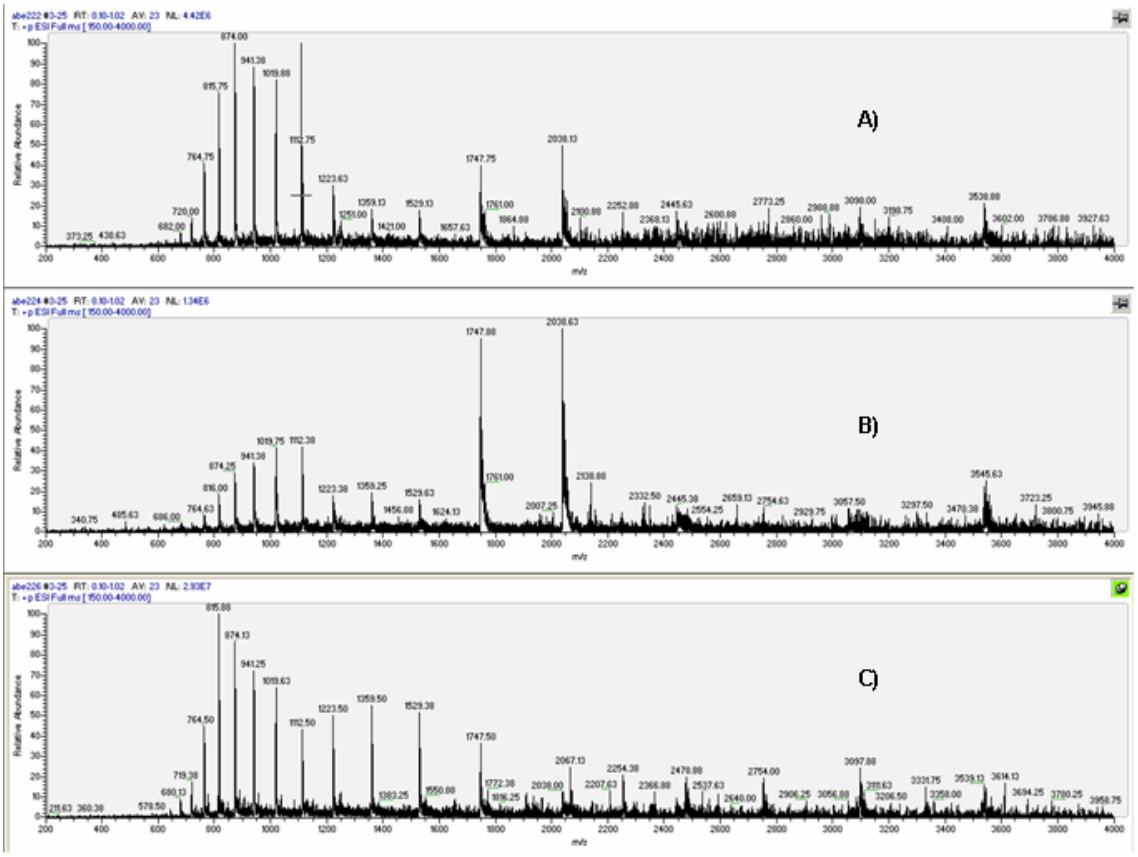


Figure 5.14 Distribution of charge states for Cytochrome c at different buffer conditions:
A) 50mM NH_4HCO_3 B) 50mM NH_4HCO_3 0.5% FA C) 50/50 water/ACN 0.5% FA

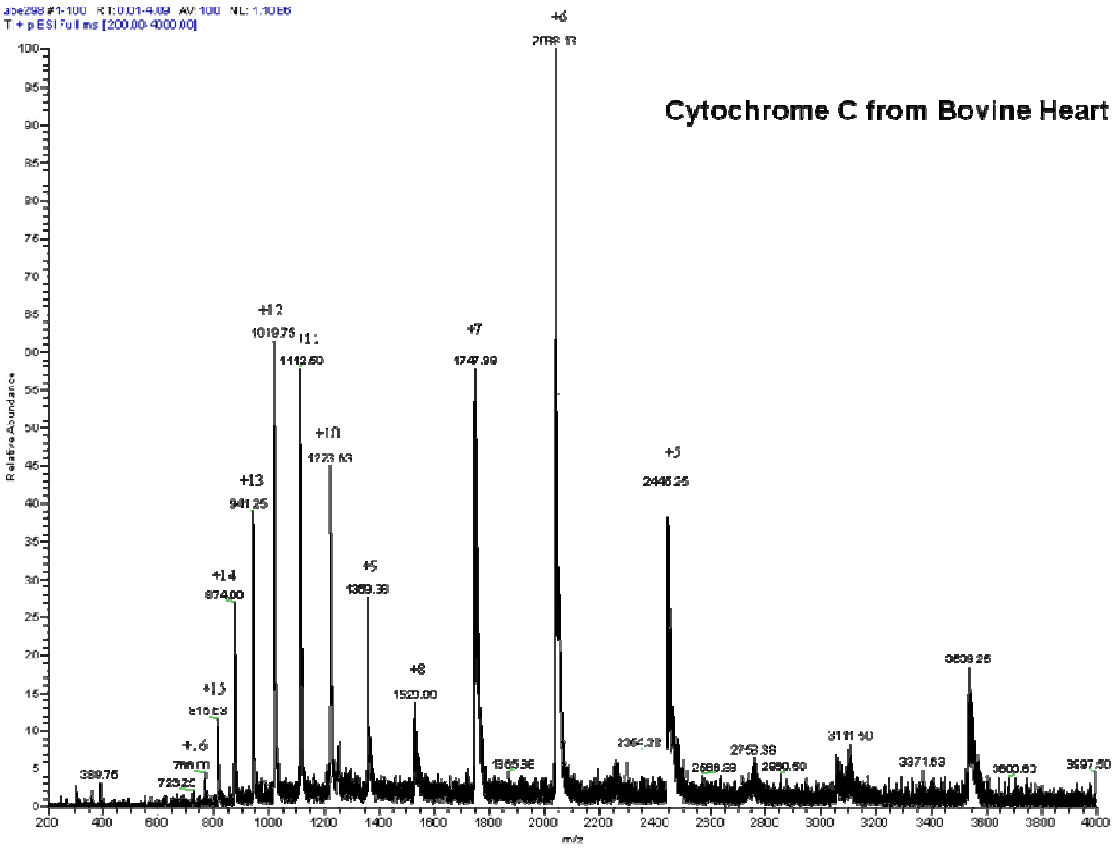


Figure 5.15 Folded and unfolded protein envelopes with charge states assigned to each peak (50mM NH4HCO3 0.5% FA)

25 µM cytochrome c (50mM NH_4HCO_3 with 0.5% FA) was injected into ESI-MS via direct infusion with the post-column addition of a blank solution (50mM NH_4HCO_3 with 0.5% FA). The experiment was then repeated sequentially with the titration of 10 µM, 50 µM, 100 µM, 200 µM, 300 µM, 400 µM, 500 µM, 600 µM, 700 µM, 800 µM, 900 µM, and 1000µM of ATP and ADP in the same buffer solution to the cytochrome c. It was found that 50 µL/min was a suitable flow rate for the detection of ligand and protein. The data were collected in full scan

mode. A 4 minute acquisition window was set aside for each titration experiment. For data analysis, the scans from the 4 minute window were averaged into a single spectrum.

5.4.3.2.1 Dissociation Constant

A titration approach was adopted to evaluate dissociation constants between host and guest. Figure 5.15 provides an example of the binding of ATP to cytochrome c with increasing

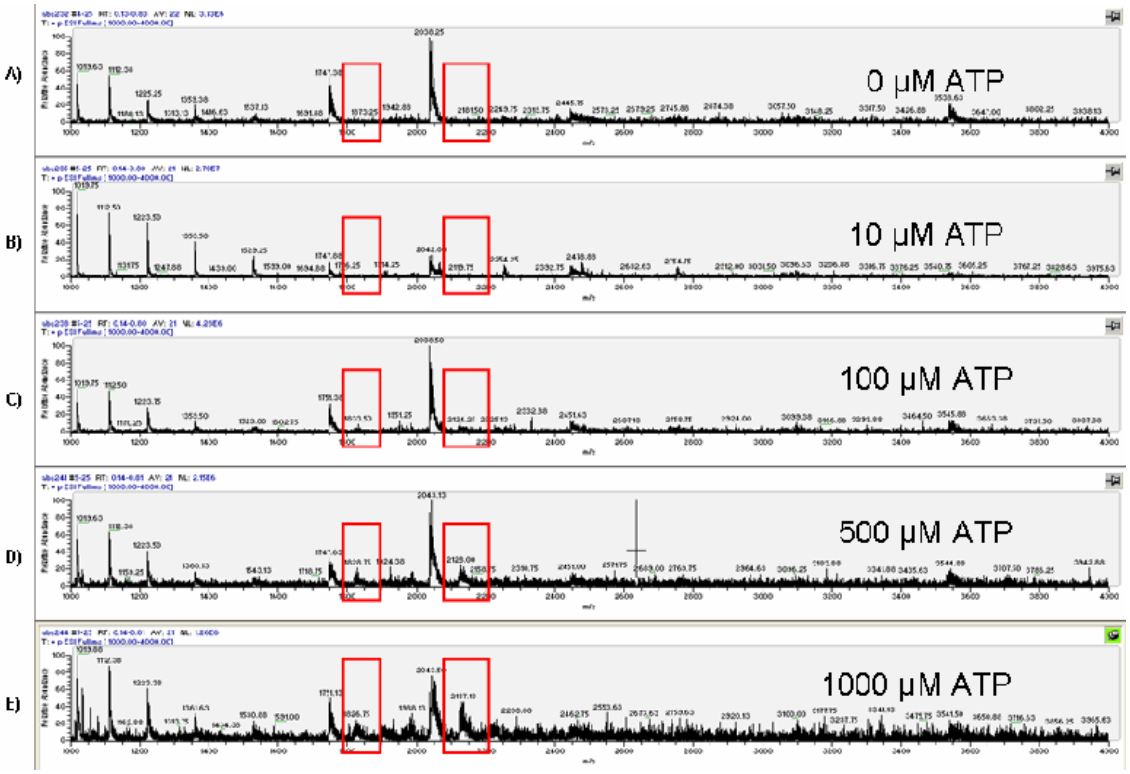


Figure 5.16 Binding of ATP to Cytochrome C with increasing concentrations of ATP with red boxes show areas of complexation
A) 0 µM ATP B) 10 µM ATP C) 100 µM ATP D) 500 µM ATP E) 1000 µM ATP

concentrations of ATP. It was found that ATP and ADP only complexed to cytochrome c at the +5 and +6 charged states, shown in figure 5.16. For more efficient nomenclature, cytochrome c will be denoted as host (H), ATP/ADP as guest (G), and the cytochrome c and ATP/ADP complex as host-guest complex (HG). The favored methodology for determining stability constants was titration curve based on the 1:1 association model. By plotting degree of

association (α) versus the initial guest concentration $C_{O,G}$, a K_d was extracted through the minimization of the sum of the weighted differences between the experimental and fitted data points. The degree of association (α) was modeled after a 1:1 host-guest interaction system

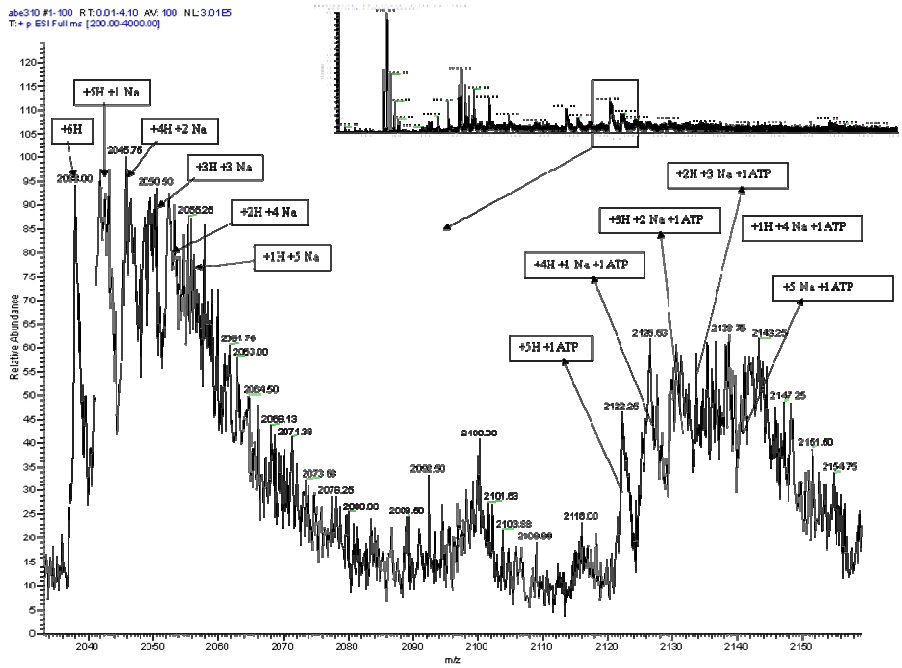
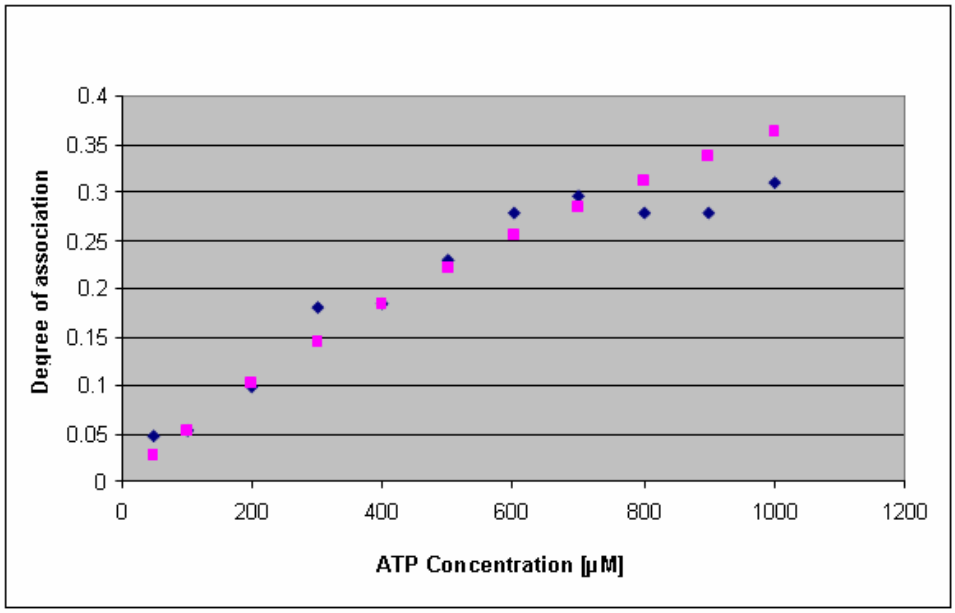


Figure 5.17 Assignment of peaks for the +6 Charge state peak for A) Free cytochrome c and B) ATP bound cytochrome c

where: α = [HG]/[HG]+[H]. Instead of using peak intensities that correspond to concentrations, peak area was utilized. DataExtractor, an in-house application, was employed for extracting peak areas from the mass spectrum. K2, another in-house application, was used for plotting α versus $C_{O,G}$ for the determination of K_d.

The dissociation constants established from the ESI-MS based titrations for cytochrome c and ATP and cytochrome c and ADP were 1.75 mM and 4.98 mM respectively. An example of how the dissociation constant was established is shown in figure 5.17. Several dissociation constants reported in the literature are in agreement with the K_d established here. Sinibaldi et al [221] obtained a binding constant between cytochrome c and ATP of 0.17 mM through docking simulation experiments. Craig and Wallace [229] used equilibrium dialysis at physiological ionic

strength and found a K_d of 3.3 mM for ferricytochrome c and ATP and a K_d of 4.0 mM for ferrocytochrome c and ATP. The higher K_d observed for cytochrome c and ADP over cytochrome c and ATP are in agreement with the literature that



K1 fit	**1.75E+03**
delta / n	1.54E-01

Figure 5.18 Dissociation constant established from the ESI-MS based titrations for cytochrome c and ATP

supports cytochrome c's preference for ATP over ADP. This statement is also supported by the simple observation that in most spectra, the relative ratios of HG to H are greater for ATP as well. The strong agreement of dissociation constants between these initial gas-phase measurements and solution-phase measurements seen in literature for the cytochrome c and ATP provides a good indication there is acceptable correlation between solution phase equilibrium concentrations and gaseous phase ion abundances for this system.

The K_d values obtained for cytochrome c and ATP/ADP still require some supporting data to confirm their validity. Some factors to consider are the lack of safeguards to monitor nonspecific interactions, cooperative and non-cooperative binding, as well as binding beyond

the 1:1 system. The 1:1 binding prerequisite is especially troublesome since there was evidence of 1:2 binding and possibly 1:3 binding within the spectra. Although observed at relatively low intensities, these bindings may be signs of possible nonspecific interactions and aggregation not observed at physiological conditions. Supporting evidence in the literature shows that an average of 2.47 and 1.64 molecules of ATP bind to oxidized and reduced horse cytochrome c respectively at an ATP concentration of 1mM in 5mM-Tris/cacoylate buffer. It was found that at saturated conditions the ATP was bound to the binding site and at two other non-specific sites of lower affinity [230].

CHAPTER 6

CONCLUSION/FUTURE WORK

6.1 Conclusion

The goal of this thesis was to demonstrate the potential for a high-throughput molecular recognition technique where electrospray ionization-mass spectrometry (ESI-MS) and universal proteomic methods can be used as complementary tools for investigating protein-ligand interactions. Several goals delineated at the outset of this thesis were accomplished: screening and identifying possible components in a complex, mapping the interaction interface or identification of protein-ligand contact points, and detection of binding between protein and ligand.

Through the study of three model proteins, it was found that screening for protein binding sites from proteolytic digests was feasible for one of three models. It was established that three different proteolytic digests of cytochrome c featured several examples of binding interaction with the ligands ATP and ADP. Within several cases of binding interaction, the binding site was discovered to be present. Although the specificity or quantitation of the interactions was not fully realized, proof of principle was established. It was found that the two models, integrin/RGD and BSA/naproxen, failed to produce strong evidence of interaction. These models are challenging due to the size of the proteins and a significant amount of further research would be necessary to optimize conditions for analysis in the context of the goals of this work.

The detection of binding between protein (cytochrome c) and ligand (ATP/ADP) was also conducted. Both the folded and unfolded state of cytochrome c were observed under ESI-MS conditions with ATP/ADP binding to the native states. Dissociation constants calculated for

intact cytochrome c and ATP/ADP were in correlation with the existing literature. The observation of deviation from 1:1 stoichiometry and greater binding affinity of ATP over ADP was also supported by the literature.

6.2 Future Work

6.2.1 Limited Proteolysis

One goal of generating longer peptide fragments containing the binding domain was met with a limited degree of success. Methods to increase peptide length through limited proteolysis could also be an avenue of future study. Digestion of protein in the presence of ligand, avoidance of denaturing conditions, and time-dependent proteolysis could all provide longer peptide fragments. With the wealth of information concerning the various techniques associated with limited proteolysis, there can undoubtedly be a marriage these techniques with a high-throughput method to screen for protein/ligand interactions.

6.2.2 Metallo-proteases

Another goal, extracting quantitative binding information for non-covalent complexes between proteolytically derived peptide fragments and selected ligands, was left unfulfilled. The ability to effectively determine the concentration of each peptide fragment is a prerequisite for the accurate determination of binding affinity and but does not within the scope of using conventional proteomic techniques. However, quantitative proteomic techniques are available.

The use of artificial metallo-proteases for discriminative cleavage was explored as an option to obtain quantifiable amounts of proteolytic peptides from a known concentration of protein. Kostic and co-workers pioneered the use of palladium based complexes that cleave very specific amino acid sequences at high efficiencies. It was first shown that the Pd(II) complexes, $[Pd(OH)(H_2O)_3]^+$ and $[PdL_2(H_2O)_2]^{2+}$ (L_2 = en, dtcol, dtco), selectively cleaved the protein cytochrome c at the amide bond His18-Thr19 forming two peptide fragments [231]. Another study showed that Pd(II) complexes consistently cleaved bovine serum albumin at the sites Glu57-Ser58, Gly85-Asp86, Leu103-Ser104, and Lys285-Ser286 [232]. Copper(II) and

zirconium(IV) complexes have also been used to hydrolyze peptide bonds with high specificity [233-234].

Several attempts were made to quantitatively cleave cytochrome c with a Pd reagent, $[Pd(H2O)_4](ClO_4)_2]$, generously donated by Prof. Nenad Kostic (Texas A&M-Commerce). The procedure involves equine heart cytochrome c where an initial concentration of ~4 mM is oxidized with a small molar excess of potassium ferricyanide ($K_3Fe(CN)_6$). The oxidizing agent is then removed through ultrafiltration and a relatively concentrated solution (<2.0 mM) of stock ferric cytochrome c is prepared. Concentration of a sufficiently diluted aliquot of stock is precisely measured through ultraviolet-visible spectroscopy by taking into account the molar absorptivities of 27,500 $M^{-1}cm^{-1}$ for Fe(III) and 9100 $M^{-1}cm^{-1}$ for Fe(II) at 550 nm. The original concentration of the stock solution is then back-calculated from the dilute solution. Using 0.10 M $HClO_4$ as a solvent, a solution of 0.20 to 0.50 mM cytochrome is prepared. Next, an equimolar to five fold excess of the Pd reagent is added. The pH of the final mixture should fall within the range of 2.0 to 2.5. The reaction mixture is left at 40°C for 2 days or 50°C for 1 day in the dark. The reaction is stopped by adding a ≤ 20 fold molar excess of sodium diethyldithiocarbamate. The Pd reagent is then removed by precipitation through centrifugation leaving a clear solution behind containing the protein fragments.

Initial attempts with this procedure were successful in producing the larger of the two peptide fragments. Figure 6.1 provides a view of the fragment which appears to have a bimodal distribution similar to intact cytochrome c. Whether these two areas of charge distribution are similar to the unfolded and folded states of cytochrome c remains to be seen. With the Pd reagent's cleavage efficiency of over 70% as reported in the literature, the peptide fragment could then possibly be purified and quantified leading to the determination of dissociation constants between this peptide fragment and ATP/ADP. This would provide a supplemental piece to see if there is a relationship to the intact protein.

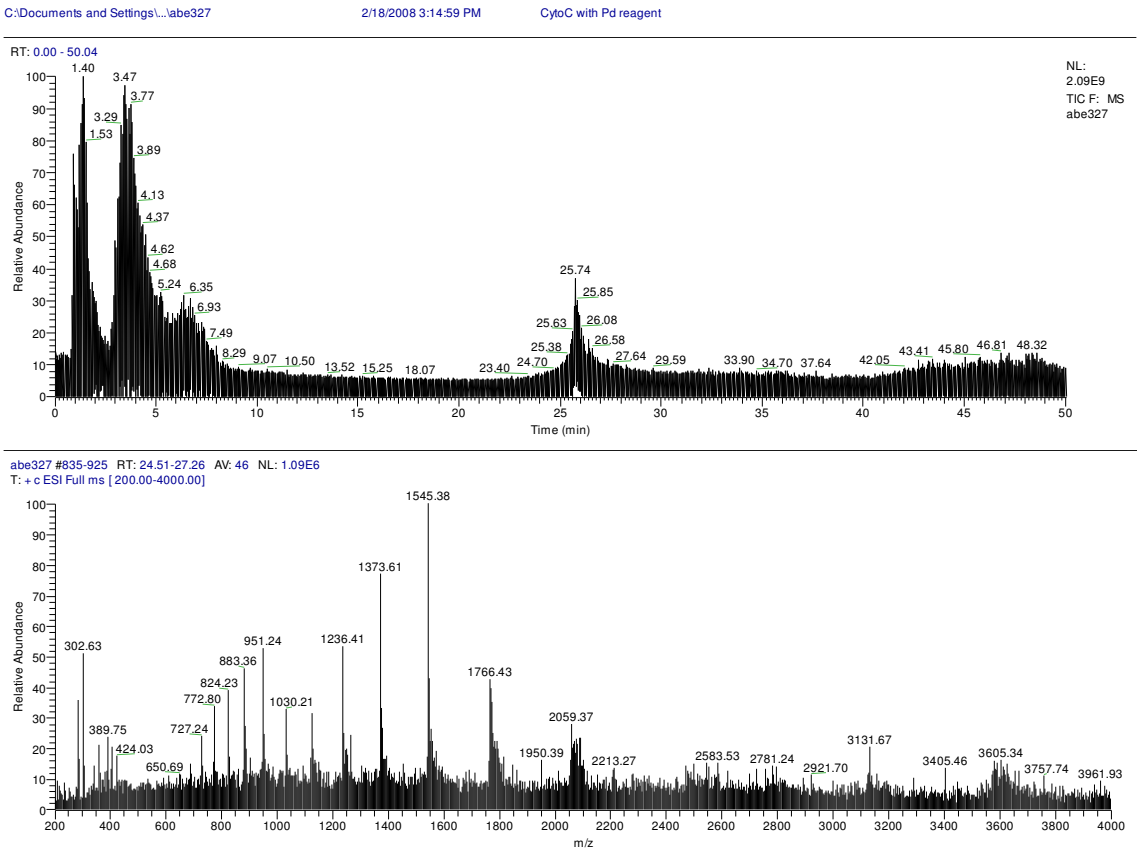


Figure 6.1 Cytochrome c large peptide fragment obtained from metallo-protease $[Pd(H_2O)_4](ClO_4)_2]$

6.3 Closing remarks

It can be said that the goals established within this thesis were satisfied reasonably well. Several new pieces of information yet unpublished within the literature were presented and will undeniably provide strong supporting evidence to the existing body of research. With new ideas and techniques left unexplored, there is such a hope that one day such methods described above will become routine in the study of protein/ligand interactions.

REFERENCES

1. Wilkins, M.R.; Sanchez J.C.; Gooley A.A. Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnol Genet Eng Rev* **1996**, 13, 19-50.

2. Zerhouni, E.A. http://www.nih.gov/about/director/budgetrequest/fy2007directorsbudgetrequest.htm (accessed April 3, 2008)

3. Patterson, S.D.; Aebersold, R.H. Proteomics: the first decade and beyond. *Nat Genet* **2003**, 33, 311-323.

4. Crick, F. Central Dogma of Molecular Biology. *Nature* **1970**, 227, 561-563.

5. Florkin, M. Discovery of pepsin by Theodor Schwann. *Rev Med Liege* **1957**, 12(5), 139-44.

6. Adams, R. J.; Pollard, T. D. Propulsion of organelles isolated from Acanthamoeba along actin filaments by myosin-1. *Nature* **1986**, 322, 754-756.

7. Wisniak, J. Jöns Jacob Berzelius A Guide to the Perplexed Chemist. *The Chemical Educator* **2000**, 5(6), 343-350.

8. Dounce, A.L. Prof. James B. Sumner. *Redox Report* **1955,** 176(4488), 859.

9. Sanger, F. Chemistry of Insulin. *Br Med Bull.* **1960**; 16, 183-188.

10. Muirhead, H.; Perutz, M. Structure of hemoglobin. A three-dimensional fourier synthesis of reduced human hemoglobin at 5.5 A resolution. *Nature* **1963**, 199(4894), 633-8.

11. Kendrew, J.; Bodo, G.; Dintzis, H.; Parrish, R.; Wyckoff, H.; Phillips, D. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **1958**, 181(4610), 662-6.

12. Adams, M.; Kelley, J.; Gocayne, J. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **1991**, 252, 1651-1656.

13. Chait, B.T. Chemistry. Mass spectrometry: bottom-up or top-down. *Science* **2006**, 314(5796), 65–6.

14. Van Bramer, S.E. *An Introduction to Mass Spectrometry*. Widener University,1997; 1-37.

15. Leicester, H.M. *Source Book in chemistry, 1900-1950*; Cambridge: MA, Harvard, 1968; 316-223.

16. Teich, M. *A Documentary History of Biochemistry, 1770-1940;* Fairleigh Dickinson University Press: Rutherford, NJ, 1992.

17. Zechmeister, L. Mikhail Tswett -- The Inventor of Chromatography *Isis*, **1946**, 36, (2), 108-109.

18. Skoog, D.A; Holler, F.J.; Nieman, T.A. *Principles of Instrumental Analysis*. 5th; Brooks/Cole Thomson Learning: USA, 1998; p 726.

19. Fenn, J.B.; Mann, M.; Meng, C.K.; Wong, S.F. Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989**, 246(4926), 64-71.

20. Yamashita, M.; Fenn, J.B. Electrospray ion source. Another variation on the free-jet theme. *J Phys Chem* **1984**, 88, 4451-4459.

21. Tal'roze, G.V. Molecular Beam of Macroions *J. Phys. Chem.* **1968**, 42, 2240-2249.

22. Iribarne, J.V.; Thomson, B.A. On Evaporation of Small Ions from Charged Droplets. *J. Chem. Phys.* **1976**, 64, 2287-2294.

23. Thomson, B. A.; Iribarne, J. V. Field Induced Ion Evaporation from Liquid Surfaces at Atmospheric Pressure. *J. Chem. Phys.* **1979**, 71, 4451-4463.

24. Paul, W., Steinwedel, H. Ein neues Massenspektrometer ohne Magnetfeld. *RZeitschrift für Naturforschung A* **1953**, 8 (7), 448-450.

25. Stafford, G.C.; Kelley, P.E.; Syka, J.E.P.; Reynolds, W.E.; Todd, J.F.J. Recent improvements in and analytical applications of advanced ion trap technology. *International Journal of Mass Spectrometry and Ion Processes* **1984**, 60(1), 85-98.

26. March, R.E.; Todd, J.F.J. *Ion Trap Mass Spectrometry: Volume II: Ion Trap Instrumentation (Modern Mass Spectrometry)*: 1st; CRC-Press:Boca Raton, FL, 1995.

27. Hogan, J.M.; McLuckey, S.A. Charge state dependent collision-induced dissociation of native and reduced porcine elastase. *J. Mass Spectrom*. **2003**; 38, 245.

28. Reid, G.E.; Wu, J.; Chrisman, P.A.; Wells, J.M.; McLuckey, S.A. Charge-state-dependent sequence analysis of protonated ubiquitin ions via ion trap tandem mass spectrometry. *Anal. Chem*. **2001**, 73, 3274.

29. Newton, K.A.; Chrisman, P.A.; Reid, G.E.; Wells, J.M.; McLuckey, S.A. Gaseous apomyoglobin ion dissociation in a quadrupole ion trap: [M + 2H]2+ - [M + 21H]+21. *Int. J. Mass Spectrom. Ion Process.* **2001**, 212: 359.

30. Wells, J.M.; Stephenson, J.L.; McLuckey, S.A. Charge dependence of protonated insulin decompositions. *Int. J. Mass Spectrom. Ion Process*. **2000**, 203: A1.

31. Bienvenut, W.V.; Deon C.; Pasquarello C. Matrix-assisted laser desorption/ionization-tandem mass spectrometry with high resolution and sensitivity for identification and characterization of proteins. *Proteomics* **2002**, 2, 868-876.

32. Wood, T.D.; Chen, L.H.; White, C.B.; Babbitt, P.C.; Kenyon, G.L.; McLafferty, F. Sequence verification of human creatine kinase (43 kDa) isoenzymes by high-resolution tandem mass spectrometry. *FW. Proc. Natl. Acad. Sci. USA* **1995**, 92, 11451-5.

33. Amunugama, R.; Hogan J.M.; Newton K.A.; McLuckey S.A. Whole protein dissociation in a quadrupole ion trap: identification of an a priori unknown modified protein. *Anal Chem* **2004**, 76, 720-727.

34. Kelleher, N.L.; Lin H.Y.; Valaskovic G.A.; Aaserud D.J.; Fridriksson E.K.; McLafferty F.W. Top down versus bottom up protein characterization by tandem high-resolution mass spectrometry. *J Am Chem Soc* **1999**, 121, 806–812.

35. Reid, G.E.; McLuckey S.A. 'Top down' protein characterization via tandem mass spectrometry. *Journal of Mass Spectrometry* **2002**, 37, 663-675.

36. Zubarev, R. Protein primary structure using orthogonal fragmentation techniques in Fourier transform mass spectrometry. *Expert Rev Proteomics* **2006**, 3, 251-261.

37. Kjeldsen, F.; Haselmann K.F.; Budnik B.A.; Sorensen E.S.; Zubarev R.A. Complete characterization of posttranslational modification sites in the bovine milk protein PP3 by tandem mass spectrometry with electron capture dissociation as the last stage. *Anal Chem* **2003**, 75, 2355-2361.

38. Fielding, L. Determination of Association Constants (Ka) from Solution NMR Data. *Tetrahedron* **2000**, 56, 6151-6170.

39. Hirose, K. A Practical Guide for the Determination of Binding Constants. *J. Incl. Phenom. Macrocyc Chem.* **2001**, 39, 193-209.

40. Schuster, T.M.; Laue, T.M. *Modern Analytical Ultracentrifugation: Acquisition and Interpretation of Data for Biological and Synthetic Polymer Systems*; Birkhäuser: Boston, MA, 1994; p 351.

41. Harding, S.E.; Rowe, A.J.; Horton, J.C. *Analytical Ultracentrifugation in Biochemistry and Polymer Science*; The Royal Society of Chemistry; Cambridge: UK, 1992.

42. Formisano, S.; Johnson, M.L.; Edelhoch, H. Thermodynamics of the self-association of glucagon. *Proc. Natl. Acad. Sci. USA* **1977**, 74, 3340–3344.

43. Formisano, S.; Johnson, M.L.; Edelhoch, H. Effects of Hofmeister salts on the self-association of glucagon. *Biochemistry* **1978**, 17, 1468–1473.

44. Lu, H.S.; Ratzkin, B. Studies on the structure and function of glycosylated and nonglycosylated neu differentiation factors. Similarities and differences of the alpha and beta isoforms. *J. Biol. Chem.* **1995**, 270, 4784–4791.

45. Wyatt, P.J. Light scattering and the absolute characterization of macromolecules. *Analytica Chimica Acta* **1993**, 272, 1–40.

46. Otto-Bruc, A.; Antonny, B.; Vuong, T.M.; Chardin, P.; Chambre, M. Interaction between the retinal cyclic GMP phosphodiesterase inhibitor and transducin. Kinetics and affinity Studies. *Biochemistry* **1993**, 32, 8636–8645.

47. Wensel, T.G.; Stryer, L. Activation mechanism of retinal rod cyclic GMP phosphodiesterase probed by fluorescein-labeled inhibitory subunit. *Biochemistry* **1990**, 29, 2155–2161.

48. Hensley, P. Defining the structure and stability of macromolecular assemblies in solution: the re-emergence of analytical ultracentrifugation as a practical tool. *Structure* **1996**, 4, 367–373.

49. Doyle, M.L.; Hensley, P. Experimental dissection of protein-protein interactions in solution *Adv. Mol. Cell. Biol.* **1996**. 22.

50. Szabo, A.L.; Stolz, L; Granzow, R. Surface plasmon resonance and its use in biomolecular interaction analysis (BIA) *Curr. Opin. Struct Biol* **1995**, 5, 699-705.

51. Smith, D.L.; Zhang, Z. Probing noncovalent structural feature of protein by mass spectrometry. *J. Am. Chem. Soc.* **1994**, 116, 7937-7938.

52. Smith, R.D.; Zhang, Z. Probing noncovalent structural features of proteins by mass spectrometry. *Mass Spectrometry Review*, **1994**, 13, 411-429.

53. Smith, D.L.; Deng, Y.; Zhang, Z. Probing the Non-Covalent Structure of Proteins by Amide Hydrogen Exchange and Mass Spectrometry. *J. Mass Spectrom.* **1997**, 32**,**135–146.

54. McLafferty, F.W. Tandem Mass Spectometry. *Science* **1981**, 214(4518), 280-287.

55. Hofstadler, S.A.; Sannes-Lowery, K.A. Applications of ESI-MS in drug discovery: interrogation of noncovalent complexes. *Nature Reviews Drug Discovery* **2006**, 5.

56. Ganem, B.; Li, Y.T.; Henion, J.D. Detection of Noncovalent receptor-ligand complexes by mass spectrometry. *J. Am. Chem. Soc.* **1991**,113, 6294-6296.

57. Katta, V.; Chait, B.T. Observation of the heme-globin complex in native myoglobin by electrospray-ionization mass spectrometry. *J. Am. Chem. Soc.* **1991**, 113, 8534-8535.

58. Loo, J.A. Studying noncovalent protein compelxes by electrospray mass spectrometry. *Mass spectrometry Reviews* **1997**, 16, 1-23.

59. Przybylski, M.;Glocker, M. Electrospary mass spectrometry of biomacromolecular complexes with noncovalent interactions: new analytical perspectives for supramolecular chemistry and molecular recognition processes. *Angewandte Chemie.International edition in English* **1996**, 35, 807-826.

60. Smith, R.D.; Bruce J.E.; Wu Q.; Lei Q.P. New mass spectrometric methods for the study of noncovalent associations of biopolymers. *Chem Soc Rev* **1997**, 26, 191-202.

61. Veenstra, T.D. Electrospray ionization mass spectrometry in the study of biomolecular non-covalent interactions. *Biophys Chem* **1999**, 79, 63-79.

62. Krishnaswamy, S.R.; Williams E.R.; Kirsch J.F. Free energies of protein-protein association determined by electrospray ionization mass spectrometry correlate accurately with values obtained by solution methods. *Protein Sci* **2006**, 15, 1465-1475.

63. Mathur, S.; Badertscher, M.; Scott, M.; Zenobi, R. Critical evaluation of mass spectrometric measurement of dissociation constnts: accuracy and cross-validation against SPR and CD for the calmodulin-melittin system. *Phys. Chem. Chem. Phys.* **2007**, 9, 6187-6198.

64. Loo R.R.; Goodlett, D.R.; Smith, R.D.; Loo, J.A. *J. Am. Chem. Soc.* **1993**, 115, 1201.

65. Fabris, D.; Fenselau, C. Characterization of Allosteric Insulin Hexamers by Electrospray Ionization Mass Spectrometry. *Anal Chem* **1999**, 71, 384-387.

66. Wigger, M.; Eyler, J.R.; Benner, S.A.; Li, W.; Marshall, A.G. Fourier transform-ion cyclotron resonance mass spectrometric resolution, identification, and screening of non-covalent complexes of hck src homology 2 domain receptor and ligands from a 324-member peptide combinatorial library. *J. Am. Soc. Mass Spectrom.* **2002**, 13, 1162–1169.

67. Jorgensen, T.J.D.; Roepstorff, P.; Heck, A.J.R. *Anal Chem.* **1998**, 70, 4427-4432.

68. Sinz, A.; Jin, A.J.; Zschornig, O. Evaluation of the Metal Binding Properties of a Histidine-Rich Fusogenic Peptide by Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry *J. Mass Spectrom.* **2003**, 38, 1150-1159.

69. Zenobi, R.; Bovet, C; Jecklin, M; Mathur, S; Touboul, D.; Wortmann, A. Proc. Of the 55th ASMS Conference on Mass Spectrometry and Allied Topics, Indiannapolis, IN June 7, 2007.

70. Griffey, R.H.; Sannes-Lowery, K.A.; Dader, J.J.; Mohan, V.V.l Swayze, E.E.; Hofstadler, S.A., Characterization of Low-Affinity complexes Between RNA and Small Molecules Using Electrospray Ionization Mass Spectrometry. *J. Amer. Chem. Soc.* **2000**, 122, 9933-9938.

71. Daniel, J.M.; Friess, S.D.; Rajagopalan, S.; Wendt,S.; Zenobi, R.J.M. Quantitative determination of noncovalent binding interactions using soft ionization mass spectrometry. *International Journal of Mass Spectrometry* **2002**, 216, 1-27.

72. Cheng, X.; Gao, Q.; Smith, R.D.; Jung K.E.; Switzer, C. Comparison of 3'3'- and 2',5'-linked DNA duplex stabilities by electrospray ionization mass spectrometry. *Chem. Commun.* **1996**, 747-748.

73. Fändrich, M.; Tito, M.A.; Leroux, M.R.; Rostom, A.A.; Hartl, F.U.; Dobson, C.M.; Robinson, C.V. Observation of the noncovalent assembly and disassembly pathways of the chaperone complex MtGimC by mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, 97(26), 14151-14155.

74. Lim, H. K.; Hsieh, Y.L.; Ganem, B.; Henion, J. Recognition of cell-wall peptide ligands by vancomycin group antibiotics: Studies using ion spray mass spectrometry *J. Mass Spectrom.* **1995**, 30(5), 708-714.

75. Greig, M.J.; Gaus, H. J.; Cummins, L.L.; Sasmor, H.; Griffey, R.H. Measurement of Macromolecular Binding Using Electrospray Mass Spectrometry. Determination of Dissociation Constants for Oligonucleotide: Serum Albumin Complexes. *J. Am. Chem. Soc.* **1995**, 117, 10765-10766.

76. Ayed, A.; Krutchinsky, A.N.; Ens, W.; Standing, K.G.; Duckworth, H.W. Quantitative evaluation of protein-protein and ligand-protein equilibria of a large allosteric enzyme by electrospray ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **1998**, 12, 339-344.

77. Griffey, R.H.; Hofstadler, S.A.; Sannes-Lowery, K.A.; Ecker, D.J.; Crooke, S.T. Determinants of aminoglycoside-binding specificity for rRNA by using mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, 96(18), 10129-33.

78. Carte, N.; Legendre, F.; Leize, E.; Potier, N.; Reeder, F.; Chottard, J.C.; Van Dorsselaer, A. Determination by electrospray mass spectrometry of the outersphere association constants of DNA/Platinum complexes using 20-mer oligonucleotides and ($[Pt(NH_3)_4]^{2+}$, $2Cl^-$) or ($[Pt(py)_4]^{2+}$, $2Cl^-$). *A. Anal. Biochem.* **2000**, 284, 77-86.

79. Johnstone, R.A.W.; Lewis, I.A.S. Crown ether complexes of metallic cations investigated by fast atom bombardment. *Int. J. Mass Spectrom. Ion Phys.* **1983**, 46, 451-454.

80. Johnstone, R.A.W.; Lewis, I.A.S.; Rose, M.E. Fast atom bombardment of crown ether/cation complexes in solution Inferences on mechanisms of energy transfer. *Tetrahedron* **1983**, 39,1597-1603.

81. Leize, E.; Jaffrezic, A.; Van Dorsselaer, A. Correlation Between Solvation Energies and Electrospray Mass Spectrometric Response Factors. Study by Electrospray Mass Spectrometry of Supramolecular Complexes in Thermodynamic Equilibrium in Solution. *J. Mass Spectrom.* **1996**, 31, 537-544.

82. Young, D.S.; Hung, H.Y.; Liu, L.K. Estimation of Selectivities and Relative Cationization Efficiencies of Different [Crown+M]+ by Electrospray Mass Spectrometry. *J. Mass Spectrom.* **1997**, 32, 432-437.

83. Ayed, A.; Krutchinsky A.N.; Ens W.; Standing K.G.; Duckworth H.W. Quantitative evaluation of protein-protein and ligand-protein equilibria of a large allosteric enzyme by electrospray ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **1998**, 12, 339-344.

84. Jørgensen, T.J.D.; Staroske, T.; Roepstorff, P.; William, D.H.; Heck, A.J.R. Subtle differences in molecular recognition between modified glycopeptide antibiotics and bacterial receptor peptides identified by electrospray ionization mass spectrometry. *J. Chem. Soc., Perkin Trans.* **1999**, 2, 1859-1863.

85. Kempen, E.C.; Brodbelt, J.S. A Method for the Determination of Binding Constants by Electrospray Ionization Mass Spectrometry. *Anal. Chem*. **2000**, 72, 5411-5416.

86. Gokel, G.W.; Goli, D.M.; Mingati, C.; Echegoyen, L. Clarification of the hole-size cation-diameter relationship in crown ethers and a new method for determining calcium cation homogeneous equilibrium binding constants. *J. Am. Chem. Soc.* **1983**, 105, 6786-88.

87. Goff, C.M.; Matchetter, M.A.; Shabestary, N.; Khazaeli, S. Complexation of caesium and rubidium cations with crown ethers in *N,N*-dimethylformamide. *Polyhedron* **1996**, 15, 3897-3903.

88. Lim, H.K.; Hsieh, Y.L.; Ganem, B.; Henion, J. Recognition of cell-wall peptide ligands by vancomycin group antibiotics: Studies using ion spray mass spectrometry. *J. Mass Spec.* **1995**, 30, 708-714.

89. Veros, C.T.; Oldham, N.J. Quantitative determination of lysozyme-ligand binding in the solution and gas phases by electrospray ionisation mass spectrometry. *Rapid Commun Mass Spectrom*. **2007**, 21, 3505-10.

90. Annis, D.A.; Shipps, G.W.; Jr Deng, Y.; Popovici-Müller, J.; Siddiqui, M.A.; Curran, P.J.; Gowen, M.; Windsor, W.T. Method for quantitative protein-ligand affinity measurements in compound mixtures. *Anal Chem.* **2007**,15, 4538-42.

91. Gupta, R.; Beck J.L.; Ralph S.F.; Sheil M.M.; Aldrich-Wright J.R. Comparison of the binding stoichiometries of positively charged DNA-binding drugs using positive and negative ion electrospray ionization mass spectrometry. *J Am Soc Mass Spectrom* **2004**, 15, 1382-1391.

92. Monchaud, D.; Allain C.; Bertrand H. Ligands playing musical chairs with G-quadruplex DNA: A rapid and simple displacement assay for identifying selective G-quadruplex binders. *Biochimie,* , In Press, Uncorrected Proof.

93. Swayze, E.E.; Jefferson E.A.; Sannes-Lowery K.A. SAR by MS: a ligand based technique for drug lead discovery against structured RNA targets. *J Med Chem* **2002**, 45, 3816-3819.

94. Hofstadler, S.A.; Sannes-Lowery K.A.; Crooke S.T. Multiplexed screening of neutral mass-tagged RNA targets against ligand libraries with electrospray ionization FTICR MS: a paradigm for high-throughput affinity screening. *Anal Chem* **1999**, 71, 3436-3440.

95. Sharma, J.; Besanger, T.R.; Brennan, J.D. Assaying Small-Molecule-Receptor Interactions by Continuous Flow Competitive Displacement Chromatography/Mass Spectrometry. *Anal Chem.* **2008**.

96. Blom, K.F.; Larsen B.S.; McEwen C.N. Determining affinity-selected ligands and estimating binding affinities by online size exclusion chromatography/liquid chromatography-mass spectrometry. *J Comb Chem* **1999**, 1, 82-90.

97. Schug, K.A. Solution phase enantioselective recognition and discrimination by electrospray ionization--mass spectrometry: state-of-the-art, methods, and an eye towards increased throughput measurements. *Comb Chem High Throughput Screen* **2007**, 10, 301-316.

98. Fändrich, M.; Tito, M.A.; Leroux, M.R.; Rostom, A.A.; Hartl, F.U.; Dobson, C.M.; Robinson, C.V. Observation of the noncovalent assembly and disassembly pathways of the chaperone complex MtGimC by mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, 97, 14151-14155.

99. Rogniaux, H.; Van Dorsselaer, A.; Barth, P.; Biellmann, J.F.; Barbanton, J.; Van Zandt, M.; Chevrier, B.; Howard, E.; Mitschler, A.; Potier, N.; Urzhumtseva, L.; Moras, D.; Podjarny, A. Binding of aldose reductase inhibitors: correlation of crystallographic and mass spectrometric studies. *J. Am. Soc. Mass Spectrom*. **1999**, 10, 635-647.

100. Li, Y.T.; Hsieh, Y.L.; Henion, J.D.; Ocain, T.D.; Schiehser, G.A.; Ganem, B. Analysis of the Energetics of Gas-Phase Immunophilin-Ligand Complexes by Ion Spray Mass Spectrometry. *J. Am. Chem. Soc.* **1994**, 116, 7487-7493.

101. Penn, S.G.; He, F.; Green, M.K.; Lebrilla, C.B. Analysis of the Energetics of Gas-Phase Immunophilin-Ligand Complexes by Ion Spray Mass Spectrometry. *J. Am. Soc. Mass Spectrom*. **1997**, 8, 244-252.

102. Chen, Y.L.; Campbell, J.M.; Collings, B.A.; Konermann, L.; Douglas, D.J. Stability of a highly charged noncovalent complex in the gas phase: Holomyoglobin. *Rapid Commun. Mass Spectrom.* **1998**, 12, 1003-1010.

103. Apostol, I. Assessing the Relative Stabilities of Engineered Hemoglobins Using Electrospray Mass Spectrometry. *Analytical Biochemistry,* **1999**, 272, 8-18.

104. Jørgensen, T.J.D.; Delforge D.; Remacle J.; Bojesen G.; Roepstorff P. Collision-induced dissociation of noncovalent complexes between vancomycin antibiotics and peptide ligand stereoisomers: evidence for molecular recognition in the gas phase. *International Journal of Mass Spectrometry,* **1999**, 188, 63-85.

105. Van der Kerk-van Hoof, A.; Heck A.J. Covalent and non-covalent dissociations of gas-phase complexes of avoparcin and bacterial receptor mimicking precursor peptides studied by collisionally activated decomposition mass spectrometry. *J Mass Spectrom* **1999**, 34, 813-819.

106. Rodgers, M.T.; Armentrout, P.B. Noncovalent metal-ligand bond energies as studied by threshold collision-induced dissociation. *Mass Spectrom Rev* **2000**, 19, 215-247.

107. Rodgers, M.T.; Armentrout P.B. Noncovalent Interactions of Nucleic Acid Bases (Uracil, Thymine, and Adenine) with Alkali Metal Ions. Threshold Collision-Induced Dissociation and Theoretical Studies. *J Am Chem Soc* **2000**, 122, 8548-8558.

108. Gross, D.S.; Zhao Y.; Williams E.R. Dissociation of heme-globin complexes by blackbody infrared radiative dissociation: Molecular specificity in the gas phase? *Journal of the American Society for Mass Spectrometry,* **1997**, 8, 519-524.

109. Schnier, P.D.; Klassen J.S.; Strittmatter E.F.; Williams E.R. Activation Energies for Dissociation of Double Strand Oligonucleotide Anions: Evidence for Watson-Crick Base Pairing in Vacuo. *J Am Chem Soc* **1998**, 120, 9605-9613.

110. Klassen, J.S.; Schnier P.D.; Williams E.R. Blackbody infrared radiative dissociation of oligonucleotide anions. *Journal of the American Society for Mass Spectrometry,* **1998**, 9, 1117-1124.

111. Strittmatter, E.F.; Schnier P.D.; Klassen J.S.; Williams E.R. Dissociation energies of deoxyribose nucleotide dimer anions measured using blackbody infrared radiative dissociation. *Journal of the American Society for Mass Spectrometry,* **1999**, 10, 1095-1104.

112. He, F.; Ramirez J.; Garcia B.A.; Lebrilla C.B. Differentially heated capillary for thermal dissociation of noncovalently bound complexes produced by electrospray ionization. *International Journal of Mass Spectrometry,* **1999**, 182-183, 261-273.

113. Garcia, B.; Ramirez J.; Wong S.; Lebrilla C.B. Thermal dissociation of protonated cyclodextrin-amino acid complexes in the gas phase. *International Journal of Mass Spectrometry,* **2001**, 210-211, 215-222.

114. Loo, J. A.; He, J. X.; Cody, W. L. Higher Order Structure in the Gas Phase Reflects Solution Structure. *J. Am. Chem. Soc.* **1998**, 120, 4542-4543.

115. Sannes-Lowery, K. A.; Hu, P.; Mack, D. P.; Mei, H. Y.; Loo, J. A. HIV-1 Tat Peptide Binding to TAR RNA by Electrospray Ionization Mass Spectrometry. *Anal. Chem.* **1997**, 69, 5130-5135.

116. Robinson, C. V.; Chung, E. W.; Kragelund, B. B.; Knudsen, J.; Aplin, R. T.; Poulsen, F. M.; Dobson, C. M. Probing the Nature of Noncovalent Interactions by Mass Spectrometry. A Study of Protein-CoA Ligand Binding and Assembly. *J. Am. Chem. Soc.* **1996**, 118, 8646-8653.

117. Beck, J.L.; Colgrave, M.L.; Ralph S.F.; Sheil M.M. Electrospray ionization mass spectrometry of oligonucleotide complexes with drugs, metals, and proteins. *Mass Spectrom Rev* **2001**, 20, 61-87.

118. Theorell, H.; Åkesson, A. *J Am Chem Soc*. **1941**, 63, 1812–1820.

119. Knapp, J.A.; Pace, C.N. Guanidine hydrochloride and acid denaturation of horse, cow, and Candida krusei cytochromes c. *Biochemistry* **1974**, 13, 1289-1294.

120. Babul, J.; Stellwagen, E. Participation of the protein ligands in the folding of cytochrome c. *Biochemistry (N Y)* **1972**, 11, 1195-1200.

121. Moore, G. R.; Pettigrew, G. W. *Cytochromes C: Evolutionary, structural, and physicochemical aspects;* Springer-Verlag: Berlin,1990.

122. Goto, Y.; Hagihara Y.; Hamada D.; Hoshino M.; Nishii I. Acid-induced unfolding and refolding transitions of cytochrome c: a three-state mechanism in H2O and D2O. *Biochemistry* **1993**, 32, 11878-11885.

123. Miranker A.D. Protein complexes and analysis of their assembly by mass spectrometry. *Curr. Opin. Struct. Biol.* **2000**, 10, 601–606.

124. Heck, A.J.R.; van den Heuvel, R.H.H. Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev*. **2004**, 23, 368–389

125. Rodgers, M.T.; Armentrout, P.B. Noncovalent metal-ligand bond energies AS studied by threshold collision-induced dissociation. *Mass Spectrom.* **2000**, 19(4), 215-247.

126. Wendt, S.; Mccombie, G.; Daniel, J.; Kienjofer, A.; Hilvert, D.; Zenobi, R. Quantitative evaluation of noncovalent chorismate mutase-inhibitor binding by ESI-MS. *J. Am. Soc. Mass Spectrom* **2003**, 14(12), 1470-1476.

127. Silke, W.; McCombie, G.; Daniel, J.; Kienho, A.; Donald, H.; and Zenobi, R. Quantitative Evaluation of Noncovalent Chorismate Mutase-Inhibitor Binding by ESI-MS. *J Am Soc Mass Spectrom*. **2003**, 14, 1470–1476.

128. Loo, J.A.; Holler, T.P.; Foltin, S.K.; McConnell, P.; Banotai, C.A.; Horne, N.M.; Mueller, W.T.; Stevenson, T.I.; Mack, D.P. Application of electrospray ionization mass spectrometry for studying human immunodeficiency virus protein complexes. *Proteins*, **1998,** 2, 28-37.

129. Kraunsoe, J.A.; Aplin, R.T.; Green, B.; Lowe, G. An investigation of the binding of protein proteinase inhibitors to trypsin by electrospray ionization mass spectrometry. *FEBS Lett.* **1996**, 396(1),108-12.

130. Daniel, J.M.; McCombie, G.; Wendt, S.; Zenobi, R. Mass spectrometric determination of association constants of adenylate kinase with two noncovalent inhibitors. *J. Am. Soc. Mass Spectrom.* **2003**,14, 442–448.

131.White, H.D.; Ashcroft, A.E. Real-Time Measurement of Myosin-Nucleotide Noncovalent Complexes by Electrospray Ionization Mass Spectrometry. *Biophysical Journal*, **2007**.

132. Conway, M.C.; Whittal R.M.; Baldwin M.A.; Burlingame A.L.; Balhorn R. Electrospray mass spectrometry of NeuAc oligomers associated with the C fragment of the tetanus toxin. *J Am Soc Mass Spectrom* **2006**, 17, 967-976.

133. Fermas S.; Gonnet F.; Varenne A.; Gareil P.; Daniel R. Frontal Analysis Capillary Electrophoresis Hyphenated to Electrospray Ionization Mass Spectrometry for the Characterization of the Antithrombin/Heparin Pentasaccharide Complex. *Anal. Chem.*, **2007**, 79(13), 4987 -4993.

134. McClements, D.J. Non-covalent interactions between proteins and polysaccharides. *Biotechnology Advances*, **July 7, 2006**.

135. Wang, W.; Kitova, E.N.; Klassen J.S. Influence of Solution and Gas Phase Processes on Protein-Carbohydrate Binding Affinities Determined by Nanelectrospray FTICR-MS. *Anal. Chem.* **2003**, 75, 4945-4955.

136. Krishnaswamy, S.R.; Williams, E.R.; Kirsch, J.F. Free energies of protein–protein association determined by electrospray ionization mass spectrometry correlate accurately with values obtained by solution methods. *Protein Sci.* **2006**, 15(6), 1465–1475.

137. Dennhart, N.; Letzel, T. Mass spectrometric real-time monitoring of enzymatic glycosidic hydrolysis, enzymatic inhibition and enzyme complexes. *Analytical and Bioanalytical Chemistry* **2006**, 386(3), 689-698.

138. Jørgensen, T.J.D.; Roepstorff, P.; Heck, A. Direct determination of solution binding constants for noncovalent complexes between bacterial cell wall peptide analogues and vancomycin group antibiotics by electrospray ionization mass spectrometry. *Anal. Chem.* **1998**, 70, 4427–4432.

139. Kempen, E.C.; Brodbelt, J.S.A method for the determination of binding constants by electrospray ionization mass spectrometry. *Anal. Chem.* **2002**, 72, 5411–5416.

140. Powell, K.D.; Ghaemmaghami, S.; Wang, M.Z.; Ma, L.; Oas, T.G.; Fitzgerald, M.C. A general mass spectrometry-based assay for the quantitation of protein-ligand binding interactions in solution. *J. Am. Chem. Soc.* **2002**,124, 10256–10257.

141. Chowdhury, S.K.; Katta, V.; Chait, B.T. Probing conformational changes in proteins by mass spectrometry. *J. Am. Chem. Soc.* **1990**, 112, 9012–9013.

142. Iavarone, A.T.; Williams, E.R. Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.* **2003**,125, 2319–2327.

143. Akashi, S. Investigation of molecular interaction within biological macromolecular complexes by mass spectrometry. *Med Res Rev* **2006**, 26, 339-368.

144. Doneanu, C.E.; Strong R.K.; Howald W.N. Characterization of a noncovalent lipocalin complex by liquid chromatography/electrospray ionization mass spectrometry. *J Biomol Tech* **2004**, 15, 208-212.

145. Bovet, C.; Wortmann A.; Eiler S. Estrogen receptor-ligand complexes measured by chip-based nanoelectrospray mass spectrometry: an approach for the screening of endocrine disruptors. *Protein Sci* **2007**, 16, 938-946.

146. Veenstra, T.D.; Tomlinson A.J.; Benson L.; Kumar R.; Naylor S. Low temperature aqueous electrospray ionization mass spectrometry of noncovalent complexes. *J Am Soc Mass Spectrom* **1998**, 9, 580-584.

147. Rostom, A.A.; Tame J.R.; Ladbury J.E.; Robinson C.V. Specificity and interactions of the protein OppA: partitioning solvent binding effects using mass spectrometry. *J Mol Biol* **2000**, 296, 269-279.

148. Raji, M.; Frycak P.; Beall M. Development of an ESI-MS screening method for evaluating binding affinity between integrin fragments and RGD-based peptides. *International Journal of Mass Spectrometry* **2007**, 262, 232-240.

149. Loo, J.A.; Holsworth D.D.; Root-Bernstein R.S. Use of electrospray ionization mass spectrometry to probe antisense peptide interactions. *Biol Mass Spectrom* **1994**, 23, 6-12.

150. Sorensen, P.; Kussmann M.; Rosen A. Identification of protein-protein interfaces implicated in CD80-CD28 costimulatory signaling. *J Immunol* **2004**, 172, 6803-6809.

151. Back, J.W.; Sanz M.A.; De Jong L. A structure for the yeast prohibitin complex: Structure prediction and evidence from chemical crosslinking and mass spectrometry. *Protein Sci* **2002**, 11, 2471-2478.

152. Taverner, T.; Hall N.E.; O'Hair R.A.; Simpson R.J. Characterization of an Antagonist Interleukin-6 Dimer by Stable Isotope Labeling, Cross-linking, and Mass Spectrometry. *J Biol Chem* **2002**, 277, 46487-46492.

153. Quinkert, G.; Egert, E.; Griesinger, C. Peptides: Synthesis, Structure, and Applications, Gutte, B., Ed.; Academic Press: San Diego, **1995**.

154. Ohio State. http://class.fst.ohio-state.edu/FST822/lectures/Denat.htm (accessed April 4, 2008)

155. Yen, T.Y.; Yan H.; Macher B.A. Characterizing closely spaced, complex disulfide bond patterns in peptides and proteins by liquid chromatography/electrospray ionization tandem mass spectrometry. *J Mass Spectrom* **2002**, 37, 15-30.

156. Johnson, R. S.; Martin, S. A.; Biemann, K. Collisioninduced fragmentation of (M+H)+ ions of peptides: side chain specific sequence ions. *Int. J. Mass Spectrom. Ion Proces*. **1988**, 86, 137–154.

157. Johnson, R. S.; Martin, S. A.; Biemann, K.; Stultz, J. T.;Watson, J. T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal. Chem.* **1987**, 59, 2621–2625.

158. Roepstorff, P.; Fohlman, J. Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed. Mass Spectrom*. **1984**, 11, 601.

159. Johnson, R.S.; Martin S.A.; Biemann K.; Stults J.T.; Watson J.T. Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Anal Chem* **1987**, 59, 2621-2625.

160. Biemann, K. Contributions of mass spectrometry to peptide and protein structure. *Biomed Environ Mass Spectrom* **1988**, 16, 99-111.

161. Biemann, K. Nomenclature for peptide fragment ions (positive ions). *Methods Enzymol.* **1990**, 193, 886–887.

162. Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. The structure and fragmentation of Bn (n>3) ions in peptide spectra. *J. Am. Soc. Mass Spectrom*. **1996**, 7, 233–242.

163. Yalcin, T.; Khouw C.; Csizmadia I.; Peterson M.; Harrison A. Why Are B Ions Stable Species in Peptide Spectra? *J Am Soc Mass Spectrom* **1995**, 6, 1165-1174.

164. Nold, M.J.; Wesdemiotis C.; Yalcin T.; Harrison A.G. Amide bond dissociation in protonated peptides. Structures of the N-terminal ionic and neutral fragments. *International Journal of Mass Spectrometry and Ion Processes* **1997**, 164, 137-153.

165. Nair, H.; Somogyi A.; Wysocki V.H. Effect of alkyl substitution at the amide nitrogen on amide bond cleavage: electrospray ionization/surface-induced dissociation fragmentation of substance P and two alkylated analogs. *J Mass Spectrom* **1996**, 31, 1141-1148.

166. Summerfield, S.G.; Whiting, A.;Gaskell, S. J. Intraionic interactions in electrosprayed peptide ions. *Int. J. Mass Spectrom. Ion Proces.* **1997**, 162, 149–161.

167. Tsaprailis, G.; Hair, H.; Somogyi, Á.; Wysocki, V.H.; Zhong, W.; Futrell, J.H.; Influence of secondary structure on the fragmentation of protonated peptides. *J. Am. Chem. Soc*. **1999**, 121, 5142–5154.

168. Zhang, Z.; Smith D.L. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci* **1993**, 2, 522-531.

169. Farmer, T.B.; Caprioli R.M. Determination of protein–protein interactions by matrix-assisted laser desorption/ionization mass spectrometry. *Journal of Mass Spectrometry* **1998**, 33, 697-704.

170. Sinz, A. Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes. *J Mass Spectrom* **2003**, 38, 1225-1237.

171. Back, J.W.; de Jong L.; Muijsers A.O.; de Koster C.G. Chemical cross-linking and mass spectrometry for protein structural modeling. *J Mol Biol* **2003**, 331, 303-313.

172. Nadeau, O.; Carlson G. Chemical cross-linking in studying protein-protein interactions. *Protein-Protein Interations: A Molecular Cloning Manual;* Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York 2002; 75–92.

173. Cohen, S. L.; Ferre-D'amare, A. R.; Burley, S.K.; Chait, B. T. Probing the solution structure of the DNA-binding protein Max by a combination of proteolysis and mass spectrometry. *Protein Science* **1995,** 4,1088-1099.

174. Kim, H.S.; Austin, J.; Hage, D.S. *Methods Mol. Biol.* **2004**, 276, 169.

175. Tanenbaum, R.;Sulitzeanu D. Enrichment of memory cells carrying receptors for a protein antigen (HSA). II. Improved enrichment technique, using BSA density gradient separation. *Immunology* **1975**, 29, 409-422.

176. Zlotos, G.; Oehlmann M.; Nickel P.; Holzgrabe U. Determination of protein binding of gyrase inhibitors by means of continuous ultrafiltration. *J Pharm Biomed Anal* **1998**, 18, 847-858.

177. Kuchimanchi, K.R.; Ahmed M.S.; Johnston T.P.; Mitra A.K. Binding of cosalane--a novel highly lipophilic anti-HIV agent--to albumin and glycoprotein. *J Pharm Sci* **2001**, 90, 659-666.

178. Vestemar, V.; Medic-Saric M.; Rendic S. Binding of (+)- and (-)-isomers and racemate of etomoxir to human serum albumin and effect of stearic acid and stanozolol. *J Pharm Sci* **1994**, 83, 485-490.

179. Ohkawa, T.; Norikura R.; Yoshikawa T. Rapid LC-TOFMS method for identification of binding sites of covalent acylglucuronide-albumin complexes. *J Pharm Biomed Anal* **2003**, 31, 1167-1176.

180. Curry, S.; Mandelkow H.; Brick P.; Franks N. Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat Struct Biol* **1998**, 5, 827-835.

181. Il'ichev, Y.V.; Perry J.L.; Ruker F.; Dockal M.; Simon J.D. Interaction of ochratoxin A with human serum albumin. Binding sites localized by competitive interactions with the native protein and its recombinant fragments. *Chem Biol Interact* **2002**, 141, 275-293.

182. Dalvit, C.; Flocco M.; Stockman B.J.; Veronesi M. Competition binding experiments for rapidly ranking lead molecules for their binding affinity to human serum albumin. *Comb Chem High Throughput Screen* **2002**, 5, 645-650.

183. Frostell-Karlsson, A.; Remaeus A.; Roos H. Biosensor analysis of the interaction between immobilized human serum albumin and drug compounds for prediction of human serum albumin binding levels. *J Med Chem* **2000**, 43, 1986-1992.

184. Gu, C.; Nikolic D.; Lai J.; Xu X.; van Breemen R.B. Assays of ligand-human serum albumin binding using pulsed ultrafiltration and liquid chromatography-mass spectrometry. *Comb Chem High Throughput Screen* **1999**, 2, 353-359.

185. Benkestock, K.; Edlund P.O.; Roeraade J. Electrospray ionization mass spectrometry as a tool for determination of drug binding sites to human serum albumin by noncovalent interaction. *Rapid Commun Mass Spectrom* **2005**, 19, 1637-1643.

186. Carter, D.C.; Ho J.X. Structure of serum albumin. *Adv Protein Chem* **1994**, 45, 153-203.

187. Herve, F.; Urien S.; Albengres E.; Duche J.C.; Tillement J.P. Drug binding in plasma. A summary of recent trends in the study of drug and hormone binding. *Clin Pharmacokinet* **1994**, 26, 44-58.

188. Sudlow, G.; Birkett D.J.; Wade D.N. Further characterization of specific drug binding sites on human serum albumin. *Mol Pharmacol* **1976**, 12, 1052-1061.

189. Fielding, L.; Rutherford S.; Fletcher D. Determination of protein-ligand binding affinity by NMR: observations from serum albumin model systems. *Magn Reson Chem* **2005**, 43, 463-470.

190. He, X.M.; Carter D.C. Atomic structure and chemistry of human serum albumin. *Nature* **1992**, 358, 209-215.

191. Ruoslahti, E.; Pierschbacher M.D. New perspectives in cell adhesion: RGD and integrins. *Science* **1987**, 238, 491-497.

192. Albelda, S.; Buck C. Integrins and other adhesion molecules. *FASEB J* **1990**, 4, 2868-2880.

193. Travis, J. Biotech gets a grip on cell adhesion. *Science* **1993**, 260, 906-908.

194. Eliceiri, B.P.; Cheresh D.A. The role of αv integrins during angiogenesis: insights into potential mechanisms of action and clinical development. *J Clin Invest* **1999**, 103, 1227-1230.

195. Hynes, R.O. Integrins: A family of cell surface receptors. *Cell* **1987**, 48, 549-554.

196. Hynes, R.O. Integrins Bidirectional, Allosteric Signaling Machines. *Cell* **2002**, 110, 673-687.

197. Hynes, R.O. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* **1992**, 69, 11-25.

198. Gottschalk, K.E.;Kessler H. Strukturen von Integrinen und ihren Ligandkomplexen– Implikationen fuer das Medikamenten-Design und die Signaltransduktion. *Angewandte Chemie* **2002**, 114, 3919-3927.

199. Thumshirn, G.; Hersel U.; Goodman S.L.; Kessler H. Multimeric cyclic RGD peptides as potential tools for tumor targeting: solid-phase peptide synthesis and chemoselective oxime ligation. *Chemistry* **2003**, 9, 2717-2725.

200. Ruoslahti, E. RGD and other recognition sequences for integrins. *Annu Rev Cell Dev Biol* **1996**, 12, 697-715.

201. Strömblad, S.; Cheresh D.A. Integrins, angiogenesis and vascular cell survival. *Chem Biol* **1996**, 3, 881-885.

202. Brooks, P.C.; Montgomery A.; Rosenfeld M. Integrin alpha v beta 3 antagonists promote tumor regression by inducing apoptosis of angiogenic blood vessels. *Cell* **1994**, 79, 1157-1164.

203. Felding-Habermann, B.; Mueller B.M.; Romerdahl C.A.; Cheresh D.A. Involvement of integrin alpha V gene expression in human melanoma tumorigenicity. *J Clin Invest* **1992**, 89, 2018-2022.

204. Dechantsreiter, M.A.; Mathä, B.; Jonczyk, A.; Goodman, S.L.; Kessler; H. in Peptides 1996 (Proc. 24th European Peptide Symposium, September 8-13, 1996, Edinburgh, Scotland) (Eds.: R. Ramage, R. Epton), Mayflower Scientific, England, 1998, pp. 329-330

205. Dechantsreiter, M.A.; Planker E.; Matha B. N-Methylated cyclic RGD peptides as highly active and selective alpha(V)beta(3) integrin antagonists. *J Med Chem* **1999**, 42, 3033-3040.

206. Pettigrew, G.W.; Moore, G.R. *Cytochromes c: Biological aspects*; Springer, Berlin. 1987.

207. Garrett, R.H.; Grisham, C.M. *Biochemistry*, Third Edition; Brooks Cole. 2006.

208. Craig, D.B.; Wallace C.J. Studies of 8-azido-ATP adducts reveal two mechanisms by which ATP binding to cytochrome c could inhibit respiration. *Biochemistry* **1995**, 34, 2686-2693.

209. Liu, X.; Kim C.N.; Yang J.; Jemmerson R.; Wang X. Induction of Apoptotic Program in Cell-Free Extracts: Requirement for dATP and Cytochrome c. *Developmental Cell* **2001**, 1, 515-525.

210. Kluck, R.M.; Ellerby L.M.; Ellerby H.M. Determinants of cytochrome c pro-apoptotic activity. The role of lysine 72 trimethylation. *J Biol Chem* **2000**, 275, 16127-16133.

211. Margalit, R.; Schejter A. Ion-binding linked to oxidation state of cytochrome c. *Eur J Biochem* **1973**, 32, 500-505.

212. Corthesy, B.E.; Wallace C.J. The oxidation-state-dependent ATP-binding site of cytochrome c. A possible physiological significance. *Biochem J* **1986**, 236, 359-364.

213. Margoliash, E.; Barlow G.H.; Byers V. Differential binding properties of cytochrome c: possible relevance for mitochondrial ion transport. *Nature* **1970**, 228, 723-726.

214. Brautigan, D.L.; Ferguson-Miller S.; Tarr G.E.; Margoliash E. Definition of cytochrome c binding domains by chemical modification. II. Identification and properties of singly substituted carboxydinitrophenyl cytochromes c at lysines 8, 13, 22, 27, 39, 60, 72, 87, and 99. *J Biol Chem* **1978**, 253, 140-148.

215. Craig, D.B.; Wallace, C.J.A. ATP binding to cytochrome c dimished electron flow in the mitochondrial respiratory pathway. *Protein Science* **1993**, 2, 966-976.

216. McIntosh, D.B.; Parrish J.C.; Wallace C.J. Definition of a nucleotide binding site on cytochrome c by photoaffinity labeling. *J Biol Chem* **1996**, 271, 18379-18386.

217. Kamiyama, T.; Sadahide Y.; Nogusa Y.; Gekko K. Polyol-induced molten globule of cytochrome c: an evidence for stabilization by hydrophobic interaction. *Biochim Biophys Acta* **1999**, 1434, 44-57.

218. Bychkova, V.E.; Dujsekina A.E.; Klenin S.I.; Tiktopulo E.I.; Uversky V.N.; Ptitsyn O.B. Molten globule-like state of cytochrome c under conditions simulating those near the membrane surface. *Biochemistry* **1996**, 35, 6058-6063.

219. Kamatari, Y.O.; Konno T.; Kataoka M.; Akasaka K. The methanol-induced globular and expanded denatured states of cytochrome c: a study by CD fluorescence, NMR and small-angle X-ray scattering. *J Mol Biol* **1996**, 259, 512-523.

220. Grandori, R. Detecting equilibrium cytochrome c folding intermediates by electrospray ionisation mass spectrometry: two partially folded forms populate the molten-globule state. *Protein Sci* **2002**, 11, 453-458.

221. Sinibaldi, F.; Mei G.; Polticelli F. ATP specifically drives refolding of non-native conformations of cytochrome c. *Protein Sci* **2005**, 14, 1049-1058.

222. Jacob, R. J.; Baker P. R.; Huang, L.; Baldwin M. A.; Burlingame A.L. Maximinizing Proteomic Information from MS Data: Enhancements to Protein Prospector, a Suite of Programs for Mining Genomic Databases, 48th ASMS Conference of Mass Spectrometry and Allied Topics, Long Beach, California May 27th - May 31st 2000.

223. Arnaout, M.A.; Goodman S.L.; Xiong J.P. Coming to grips with integrin binding to ligands. *Curr Opin Cell Biol* **2002**, 14, 641-651.

224. Colmenarejo, G.; Alvarez-Pedraglio A.; Lavandera J.-. Cheminformatic Models To Predict Binding Affinities to Human Serum Albumin. *J Med Chem* **2001**, 44, 4370-4378.

225. Ermondi, G.; Lorenti M.; Caron G. Contribution of Ionization and Lipophilicity to Drug Binding to Albumin: A Preliminary Step toward Biodistribution Prediction. *J Med Chem* **2004**, 47, 3949-3961.

226. Banerjee, T.; Singh S.K.; Kishore N. Binding of naproxen and amitriptyline to bovine serum albumin: biophysical aspects. *J Phys Chem B* **2006**, 110, 24147-24156.

227. Wu, Q.; Gao J.; Joseph-McCarthy D. Carbonic anhydrase-inhibitor binding: From solution to the gas phase. *J Am Chem Soc* **1997**, 119, 1157–1158.

228. Starovasnik, M.A.; Braisted A.C.; Wells J.A. Structural mimicry of a native protein by a minimized binding domain. *Proc Natl Acad Sci US A* **1997**, 94, 10080-10085.

229. Craig, D.B.; Wallace C.J. The specificity and Kd at physiological ionic strength of an ATP-binding site on cytochrome c suit it to a regulatory role. *Biochem J* **1991**, 279 (Pt 3), 781-786.

230. Wallace, C.J.; Corthesy B.E. Protein engineering of cytochrome c by semisynthesis: substitutions at glutamic acid 66. *Protein Eng* **1986**, 1, 23-27.

231. Zhu, L.; Qin L.; Parac T.N.; Kostic N.M. Site-Specific Hydrolytic Cleavage of Cytochrome c and of Its Heme Undecapeptide, Promoted by Coordination Complexes of Palladium (II). *J Am Chem Soc* **1994**, 116, 5218-5224.

232. Zhu, L.; Kostić N.M. Sequence-dependent cleavage of albumins with palladium (II) complexes: role of serine residue in controlling the high regioselectivity of protein cleavage. *Inorg Chim Acta* **2002**, 339, 104-110.

233. Hegg, E.L.; Burstyn, J.N.; Hydrolysis of unactivated peptide bonds by a macrocyclic copper(II) complex: Cu([9]aneN3)Cl2 hydrolyzes both dipeptides and proteins. *J. Am. Chem. Soc.* **1995**, 117, 7015-7016.

234. Cepeda, S.S.; Grant, K.B. Hydrolysis of insulin chain B using zirconium(IV) at neutral pH. *New J. Chem*. **2008**, 32, 388-391.

BIOGRAPHICAL INFORMATION

Abraham Lo was a former 1997 Welch Summer Scholar at the University of Texas at Arlington. After graduating in 2003 from University of California Berkeley with a B.S. in Biochemistry, he went on to earn his Masters in Chemistry at the University of Texas at Arlington. His research interests include proteins and medical related research. He plans on pursuing a degree in medicine and hopes to one day become a physician.