BIOANALYTICAL APPLICATIONS OF LIQUID CHROMATOGRAPHY MASS SPECTROMETRY – PROTEIN IDENTIFICATION, ENZYME FUNCTION, CARBOHYDRATE ANALYSIS, POST-TRANSLATIONAL MODIFICATIONS, AND TARGETED PROTEOMICS

by

JEREMI DAVID JOHNSON

(Under the Direction of Ron Orlando)

ABSTRACT

The goal of conventional proteomics is typically to identify, in a rapid and efficient manner, as many proteins within a particular extract as possible. Whole protein extracts can contain more than 10,000 proteins. The dynamic range of proteins in a whole protein extract presents but one daunting challenge. Other complications include a wide range of post-translational modifications, the potential for hundreds and even thousands of proteins, as well as the presence of detergents and other chemicals used in the extraction process. As a result, proteome analysis can become exceptionally challenging. Presented here are experiments where only a few proteins within a whole protein extract were of interest. The goal was to identify only those proteins associated with allergic responses.

Post translational modifications (PTMs) are integral parts of many proteins. These modifications often manifest themselves in the form of N-linked glycosylation, O-linked glycosylation, and phosphorylation. A thorough understanding of protein structure and function encompasses the complete structural elucidation and localization of all PTMs. A variety of techniques including mass spectrometry, gel electrophoresis, antibody specific interactions, and targeted labeling, were utilized to tentatively solve all post translational modifications associated with several recombinant Pectate lyases.

ESI-LC-MS and MALDI-MS were utilized to study the effects of point mutations on the specificity of a recombinant pectin methyl esterase. Successful characterization depended on the tandem mass spectrometric sequencing of a methyl esterified hexagalacturonate substrate. Utilizing a series of enzymes mutated to eliminate selected N-linked glycans, the role of N-linked glycosylation was assessed with regards to the enzyme's ability to demethylate the aforementioned substrate. Unique to the substrate were the six potential sites available for enzyme catalyzed de-methylation. This allowed for assessment of the mode-of-action demonstrated by the mutated and wildtype enzymes.

INDEX WORDS: Mass Spectrometry, Tandem Mass Spectrometry, Liquid Chromatography, MALDI-TOF, LC-MS, MS/MS, ESI, Glycosylation, Post Translational Modifications, Protein Identification, Proteomics, Allergens, Fungus, Pectate Lyase, Pectin Methyl Esterase, Gel Electrophoresis, Western Blots, Antibodies, Carbohydrate Sequencing

BIOANALYTICAL APPLICATIONS OF LIQUID CHROMATOGRAPHY MASS SPECTROMETRY – PROTEIN IDENTIFICATION, ENZYME FUNCTION, CARBOHYDRATE ANALYSIS, POST-TRANSLATIONAL MODIFICATIONS, AND TARGETED PROTEOMICS

by

JEREMI DAVID JOHNSON

B.A., Western Carolina University, 1994

M.S. University of Tennessee, Knoxville, 1998

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2003

BIOANALYTICAL APPLICATIONS OF LIQUID CHROMATOGRAPHY MASS SPECTROMETRY – PROTEIN IDENTIFICATION, ENZYME FUNCTION, CARBOHYDRATE ANALYSIS, POST-TRANSLATIONAL MODIFICATIONS, AND TARGETED PROTEOMICS

by

JEREMI DAVID JOHNSON

Major Professor: Ron Orlando

Committee: Jon Amster
John Stickney
Carl Bergmann

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2003

DEDICATION

This dissertation is dedicated to those who supported me throughout my endeavors, those I hold most dear - my family, to whom I am forever indebted.

To My Parents:

M. L. and Ann Johnson..……for your generosity throughout my life. For the love and care you have given me, unconditionally. My admiration and love cannot begin to show – thank you so much.

To My Grandparents:

B.V. and Clovie Putnam…… For the love and care you have provided, the sacrifices you you have made, the wisdom and advice you have so freely given to me – my gratefulness and love cannot begin to be stated.

To Verlin, Tracee, and Tim - Thank you for being the brothers and sister I never had.

ACKNOWLEDGEMENTS

First and foremost my appreciation for God's presence in my life. Through Him all things are possible.

To my family who supported me and encouraged me. I can't explain the gratitude I hold for you in my heart. I will never forget your sacrifices and words of wisdom. God gives us to the earth and places us under the care of people who teach us what life is and how to respond to the many obstacles we will encounter. God placed me under the care of a family whom I hold in the highest regard. I thank God for you everyday.

My gratitude and respect is extended to the faculty and staff of the Department of Chemistry and my committee members for your efforts and involvement in the successful completion of my degree.

My thanks go to the scientists, students, and staff of the Complex Carbohydrate Research Center for providing many of the resources used in this dissertation, for advice, and friendship. I am especially grateful to Dr. Carl Bergmann and Dr. Kumar Koli for their advice, expertise, patience, and friendship.

Less often than not we tend to forget those teachers who helped to shape and expose us to the 'elements' which we will eventually use in many parts of our life. The foundations and understanding provided through grade school and secondary education are among the most important and help prepare us, in many ways, for the paths we shall choose and the obstacles we shall face. Among the many teachers I have encountered throughout life I would like to give my gratitude and respect to Larry Poston, my biology teacher, friend, and mentor. I know undoubtedly that the education and friendship he has provided to so many has lead only to improvements within our lives.

To Dr. Ron Orlando – my deepest respect and admiration. Thank you for the countless hours of helpful discussion, for the knowledge and abilities learned through your guidance and support while completing a very memorable research experience.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

Characterization of bio-molecules has been propelled to new levels in recent years thanks to advancements in mass spectrometry and various hyphenated mass spectrometric techniques. Included in the range of bio-molecules being studied by mass spectrometry are proteins, peptides, and carbohydrates (simple and complex). Most notable of the successful developments in mass spectrometry are the relatively soft ionization methods 'Matrix Assisted Laser Desorption Ionization' (MALDI)[1] and Electrospray Ionization (ESI)[2]. In addition to mass spectrometry, several established techniques such as liquid chromatography and gel electrophoresis have played important roles in the study of proteins, carbohydrates, and other biomolecules.[3-7] However, the molecular weight of a molecule can harbor a wealth of information and provide insight into the identity and structure of the molecule(s) being studied.

Post translational modifications (PTMs) are integral parts of proteins which require thorough characterization above that of the amino acid sequence alone. Some of the more common manifestations of PTMs are N-linked glycosylation, O-linked glycosylation, and phosphorylation, for example.[8] Classical methods of PTM characterization, such as permethylation and β-elimination, still find routine use in laboratories. However, today, characterization of the resulting modified protein or peptide is largely handled by a variety of mass spectrometric techniques.

The fundamental principle of mass spectrometry is the measurement of mass/charge (m/z). It has been described as the most efficient, effective, and accurate method for the assessment of biomolecule mass. Recent reports have demonstrated the use of mass spectrometry for the successful measurement of biomolecule masses in excess of 500,000 Da.[9] Perhaps more useful to the field of protein characterization are the low sample requirements (femtomoles), speed, and accuracy associated with mass spectrometry. Even the simplest of mass analyzers have mass accuracies two orders of magnitude better than gel electrophoresis (the accepted alternative method for determining the mass of large biomolecules).

The completion of several genomes has fueled the development of a relatively new field of study designed to identify and functionally characterize the direct products of genes. This new field, known as proteomics, seeks to unravel the precise roles of proteins in a variety of biological phenomenon such as physiological development and process, disease pathology, cellular response, and structural integrity. Currently MS serves three major areas within the field of proteomics. Mass spectrometry serves biotechnology as the preferred technique for the structural characterization and quality control of recombinant proteins. MS also finds extensive use in protein identification within both classical biochemical projects and high-throughput proteomics. Finally, MS has become the method of choice for the structural elucidation and localization of all post-translational modifications.[10]

In conventional proteomics it is often the goal of the researcher(s) to identify, in a rapid and efficient manner, as many proteins within an extract as possible. However, the dynamic range of proteins in a whole protein extract is but one daunting aspect of proteome characterization. When coupled with the potential for very complex samples and the presence of detergents and others chemicals used in the extraction process proteome analysis can become very challenging. Thus a variety of separation techniques have been coupled with mass spectrometry to help circumvent many of the aforementioned difficulties.

The research presented in this dissertation focuses on the application of mass spectrometry to the study of proteins and carbohydrates. Chapter III focuses on a somewhat conventional proteomics project in which whole proteome extracts are characterized utilizing a series of biochemical techniques, most notably mass spectrometry. This work differs from conventional proteomics by virtue of its focus on the characterization of key proteins present within the extract as opposed to all proteins. Chapters IV and V focus more closely on characterizing the functional aspects of proteins rather than protein identification. The goals of these two chapters seek to identify and characterize a variety of post-translational modifications (Chapter IV) as well as assess the importance of these modifications with respect to protein function (Chapter V).

REFERENCES

(1) Karas, M.; Ingendoh, A.; Bahr, U.; Hillenkamp, F. *Biomed. Environ. Mass Spectrom.* **1989**, *18*, 841.

(2) Fenn J. B.; Mann, M.; Meng. C. K.; Wong, S. F.; Whitehouse *Mass Spectrom. Rev.* **1990**, *9*, 37.

(3) Remy, R.; Ambard-Bretteville, F.; *Methods Enzymol.* **1987**, *148,* 623.

(4) Ames, G.; Nikaido, K. *Biochemistry* **1976**, *15*, 616.

(5) Adessi, C.; Miege, C.; Albrieux, C.; Rabilloud, T. *Electrophorsis* **1997**, *18*, 127.

(6) Butt, A.; Davison, M.; Smith, G.; Young, J.; Gaskell, S.; Oliver, R.; Beynon, J.; *Proteomics* **2001**, *1*, 42.

(7) Opiteck, G.; Ramirez, S.; Jorgenson, J.; Moseley, M. *Anal. Biochem.* **1998**, *258*, 349.

(8) Dell, A.; *Science* **2001**, *291*, 2351.

(9) Ehring, H.; Karas, M.; Hillenkamp, F. *Org. Mass Spectrom.* **1992**, *27*, 472.

(10) Mann, M.; Hendrickson, R.; Pandey, A. *Annu. Rev. Biochem.* **2001**, *70*, 437.

CHAPTER II

LITERATURE REVIEW

Mass Spectrometry

The fundamentals of modern mass spectrometry were established more than a century ago by Sir Joseph John Thomson. These original contributions were accomplished just prior to the 20th century utilizing a device comprised of the basic components found in modern day mass spectrometers. Early experiments eventually lead to his discovery of the electron. This finding, along with others completed at the Cavendish Laboratories, were the earliest examples of the power of mass spectrometry.[1]

Mass spectrometry developed in the following decades to become an "interdisciplinary research methodology, impacting virtually every area of science, from physics through chemistry and biology".[2] Because all molecules contain mass, mass spectrometry has the inherent characteristic of being universally applicable to all materials.[2] To analyze a material by mass spectrometry certain fundamental aspects must be overcome: First, volatilization of the material since mass spectrometry is a gas phase technique and second, formation of charge on the molecule being analyzed. Recent advances in ionization methods have made it possible for most materials to be both volatilized and ionized, making mass spectrometry widely applicable. Mass analysis by mass spectrometry is accomplished by detecting the mass-to-charge ratio of an analyte. Basically there are two types of ions detected; molecular ions and fragment ions. Molecular ions refer to the detection of the intact analyte molecule plus some charge associated with it (typically a proton for biomolecular ions). Fragment ions refer to the detection of some portion of the analyte ion where charge retention lies with the detected fragment. Nonetheless, in all cases it is the mass of the ion divided by its charge state that is recorded by the mass spectrometer. With many ionization methods, such as electron ionization, laser desorption, or chemical ionization, z (charge) is typically one. Modern ionization methods such as electrospray ionization can, and often does, impart more than a single charge to the analyte ion. For example, a peptide with a mass of 1000 containing one proton will have a m/z (as recorded by the mass spectrometer) of 1001 Daltons (Da). This same peptide with 2 charges will be detected by the mass spectrometer at a m/z of 501 Da.

The vast majority of mass spectrometric applications in the early and mid- 20th century were limited to small volatile materials.[1] However, with the advent of modern ionization methods, mass spectrometry has enjoyed vast success in areas such as synthetic polymers[3], large and small bio-polymers[4], whole proteomes[5], whole cells,[6] and others which encompass the realm of large non-volatile molecules. This relatively broad applicability coupled with exceptional sensitivity, high resolution, and large mass ranges easily accounts for the widespread use of mass spectrometry in industrial, clinical, and academic settings. Fueling these fields is unparalled innovation in instrument design coupled with creative applications.[2] Among the most notable of these innovations are the development of matrix assisted laser desorption ionization (MALDI) and electrospray ionization (ESI).

Matrix Assisted Laser Desorption Ionization (MALDI)

MALDI (Figure II-I) was introduced in 1988 as a result of the pioneering work of Karas[7], HillenKamp[7], and Tanaka[8]. Characteristic to MALDI, as opposed to other laser desorption methods, is the use of a matrix to aid in volatilization and ionization. It is theorized that the matrix acts as an energy buffer between the laser and the analyte.[9]

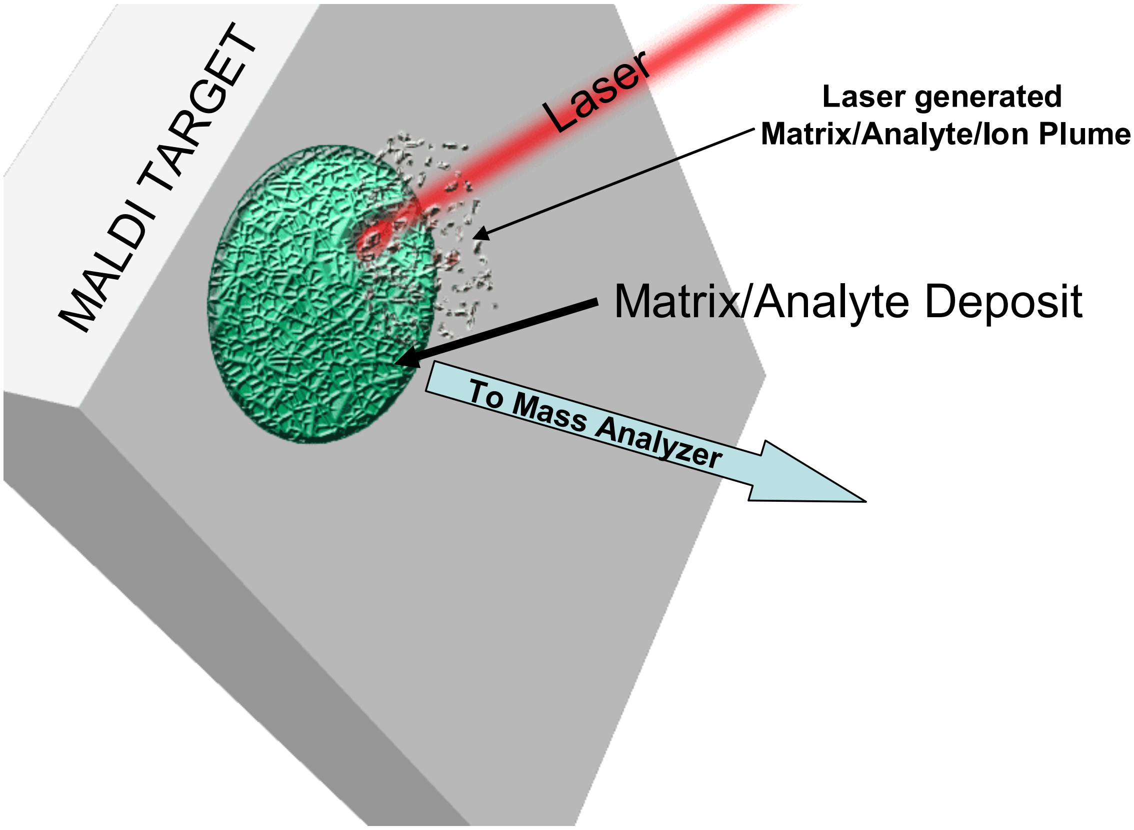


Figure II-I. The Matrix Assisted Laser Desorption Ionization (MALDI) Process.

Although the exact mechanism for ionization is under investigation, it is believed that laser energy absorbed by the matrix is transferred, partially, to the analyte. In addition to buffering laser energy, the matrix is also believed to be a source of protons essential during the ionization phase of ion generation.[7,10] Another important, less recognized, role of the matrix is to minimize aggregation of the analyte by distributing it within the crystal lattice formed by the matrix.[11]

The formation of the analyte ion is ultimately the result of a solvent, matrix, probe surface, and laser.[9] In simplest terms, the analyte and matrix are dissolved, independently, in a suitable solvent, typically methanol, acetonitrile, water, or a mixture of these.[9] Selection of matrix is not only dictated by the material to be analyzed but also by the laser used to generate the ions. Of importance is the absorbance wavelength of the matrix. It is critical that the matrix absorb the laser energy in order for the MALDI process to occur. Some common matrices used are sinapinic acid (SA), 2,5 dihydroxybenzoic acid (DHB), and α-cyano-4-hydroxy-trans-cinnamic acid (HCCA).

Once the matrix and analyte have been dissolved they are either co-deposited, or deposited independently, onto the surface of a probe. Typical ratios of matrix:analyte are 100:1 or greater. Probe surfaces are typically stainless steel but various articles have demonstrated the feasibility of other materials as suitable surfaces.[12] Once deposition is complete the sample is dried at atmosphere or under vacuum resulting in a crystal deposit on the surface of the probe. The probe is then placed into an evacuated source where it is irradiated with a pulsed laser. The most common laser employed in MALDI is the pulsed nitrogen laser (λ=337 nm). Recently IR and other UV lasers have also found use in MALDI mass spectrometry.

The energy emitted by the laser is absorbed by the matrix and transferred to the materials present within the crystal structure of the matrix. This localized energy-absorption results in the expulsion of a plume containing gaseous matrix and analyte molecules/ions. The mechanism for the formation of ions is largely unknown. Currently there is much deliberation as to whether ions are actually formed in the gas phase or the solid phase.[13] Nonetheless ionization is the result of the association of a proton or alkali-metal in the case of positive ion formation or loss of a proton during negative ion formation.

Due to the soft ionization characteristics of MALDI and its ability to transfer non-volatile materials from the solid phase to the gas phase, the technique has enjoyed success in a variety of fields. In recent years MALDI has become an increasingly popular tool for the analysis of biological systems. Researchers have successfully demonstrated the use of MALDI-MS for the analysis of both hydrophilic and hydrophobic proteins.[14-16] This is partially due to the ability of MALDI to handle moderate amounts of detergents and salts, allowing the researcher to bypass sample cleanup procedures which can be time consuming and laborious.[9] In addition to protein identification, various other aspects of proteins have been explored utilizing MALDI mass spectrometry, including the identification of post-translational modifications[17-19] and explorations of non-covalent protein interactions[20,21]. MALDI has also begun to show promise in the field of DNA analysis. Recent reports have demonstrated the use of MALDI-MS for the sequencing of short chain DNA as well as for the diagnosis of various genetic disorders such as cystic fibrosis and Huntington's disease.[22,23] There have also been a number of articles demonstrating the use of MALDI-MS for the analysis and identification of whole bacterial cells.[6, 24]

Electrospray Ionization (ESI)

In the mid- to late 1980s Fenn and co-workers published work demonstrating the use of a liquid flow system, under the influence of an electric field, to generate charged analyte molecules[25,26] from large non-volatile materials.[27,28] This work lead directly to the development of commercial electrospray ionization (ESI) based mass spectrometers. Fundamentally, ESI forms ions from a solution that is pumped through a small capillary held at ~ 2000 to 4000 volts. This potential is applied relative to a counter electrode (skimmer) in close proximity to the capillary (Figure II-II). The applied potential induces the formation of a Taylor cone (Figure II-III) due to the accumulation of like charges near the meniscus of the liquid at the exit of the fused silica capillary. Emitted from the Taylor cone is a spray composed of small droplets of micron diameter. These droplets traverse the area between the capillary tip and the entrance of the mass spectrometer, a region normally held at close to atmospheric pressure. Both a potential and a pressure gradient facilitate the movement of the droplets toward the entrance to the mass spectrometer.[4] As the ions traverse this region a counter flow of gas is applied, in addition to collisions with atmospheric gases, to help desolvate the droplets. As the solvents composing the droplet evaporate, the amount of charge remains constant resulting in an increase in overall charge density. Once the charge density becomes large enough, the columbic repulsions exceed the surface tension of the droplet (known as the

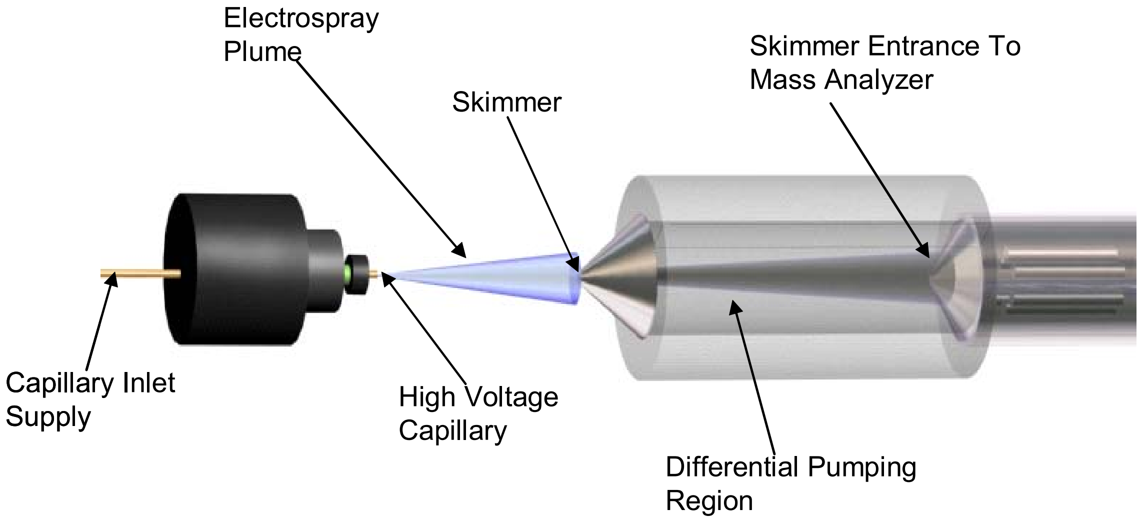


Figure II-II. Diagram of an electrospray ionization source.

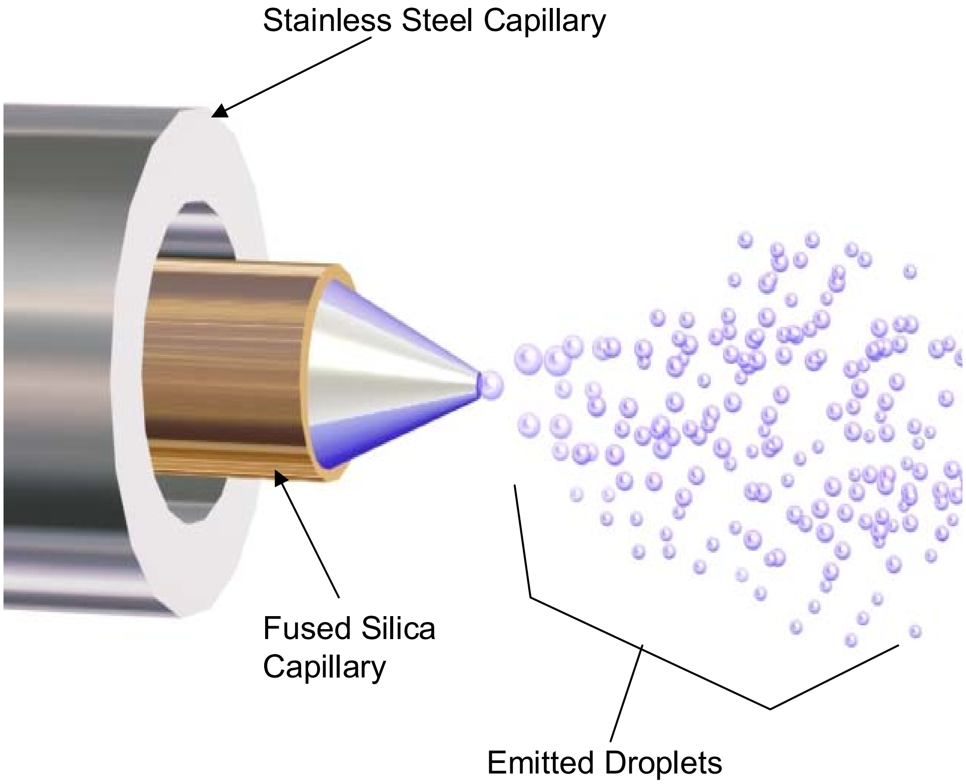


Figure II-III. Diagram illustrating the formation of a Taylor cone and the droplets emitted.

Rayleigh limit)[29] and the droplet breaks apart to form several smaller droplets, eventually leading to droplets with nanometer diameters.[30]

The exact mechanism for the formation of free gas phase ions, from the nanometer size droplets, is still not fully understood. Currently two mechanisms are used to describe the probable events that would lead to the formation of free ions. First, the 'charge residue model' describes the process of forming free ions through a series of desolvations/Rayleigh explosions until a single solvated analyte ion remains. Further desolvation of this single analyte ion results in the formation of a single gas phase ion[31] to be sampled by the mass spectrometer. Second, the 'ion evaporation model' describes a process of direct emission of gaseous ions from droplets with diameters less than ~10 nm whereby single ions are expelled from the surface of small droplets and sampled by the mass spectrometer.[32] Numerous publications have detailed the various aspects of the electrospray process.[30, 33-36]

Nonetheless, the soft method by which ions are formed results in little fragmentation of analyte during the ion production phase. Perhaps most unique to ESI is the formation of multiple charges on the analyte (addition of protons in positive ionization and abstraction of protons in negative ionization), see Figure II-IV. This characteristic imparts almost no upper limit to the mass of materials that can be analyzed by ESI-MS since the addition of multiple charges can effectively return a high molecular weight biopolymer, for example, into the mass range of many commercially available mass spectrometers.

Since the introduction of electrospray mass spectrometry the ES ionization source has undergone various design modifications. Shown in Figure II-V is one such design

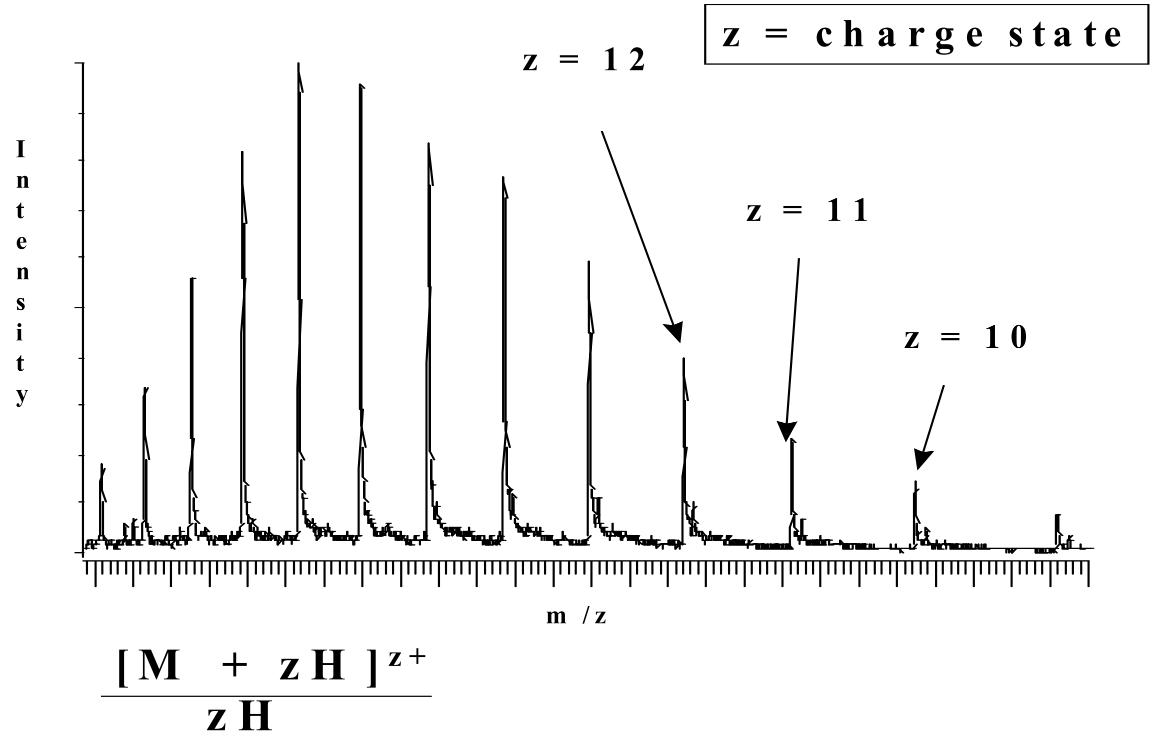


Figure II-IV. - ES mass spectrum of horse heart myoglobin, a naturally occurring protein. Each peak represents the same molecule but in a different charge state.

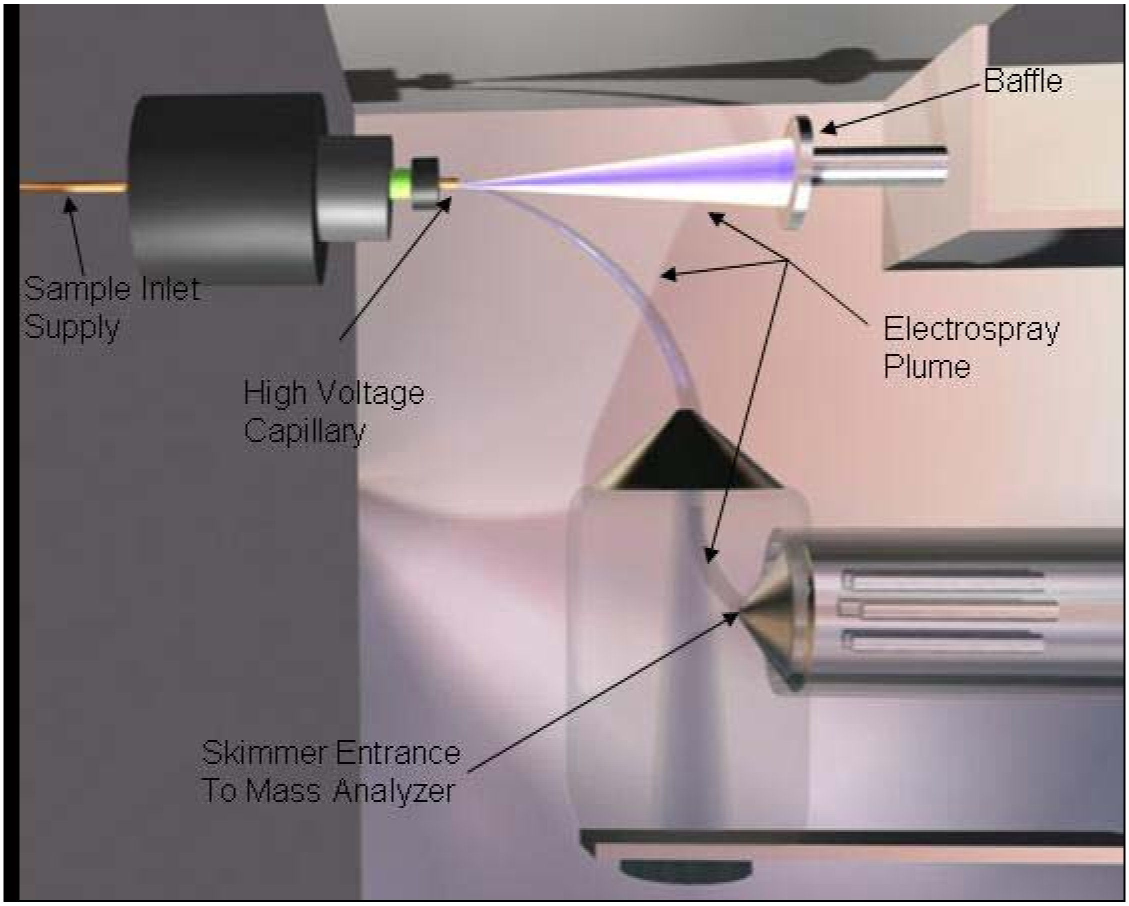


Figure II-V. Diagram of a Z-Spray^{TM} Electrospray Ionization Source.

modification developed by Micromass. The Z-Spray system helps to improve sensitivity and minimize issues associated with a variety of contaminants. Nonetheless it is the basic principles of electrospray that makes the Z-Spray system functional. As with the Z-spray system, modifications have primarily focused on circumventing issues with ES such as toleration of salts and detergents, coupling of ES sources with other established analytical methods suchs as capillary electrophoresis (CE) and high-performance liquid chromatography (HPLC), reducing flowrates and sample consumption, and increasing overall sensitivity.[9] Many of these modifications account for the relative success of electrospray ionization mass spectrometry. ESI-MS has been used successfully for the characterization of synthetic polymers,[3, 37-39] protein identification,[40-42] determination of post-translational modifications,[42-45] proteome analysis,[46,47] analysis of drug metabolites,[48] DNA sequencing and analysis,[49-51] small molecule analysis in a variety of media,[52-55] and many others.

While there are other ionization techniques such as fast atom bombardment (FAB),[56] laser desorption (LD),[57] chemical ionization (CI),[58] and electron ionization (EI),[59] these techniques lie outside the scope of this dissertation and will thus not be presented to the reader. To a large extent the aforementioned ionization methods have found limited to no success with high molecular weight/low volatility materials. As a result they have seen limited use in applications focused on in this research.

In addition to the variety of ionization methods available to the mass spectrometry community there also exists a variety of mass analyzers. Whereas the ionization method serves the purpose of converting solid, solution, or gas phase molecules to ions, the mass analyzer serves to separate the formed ions based on mass-to-charge (m/z). Even though

Table II-I. Figures of merit for common commercially available mass analyzers.[1]

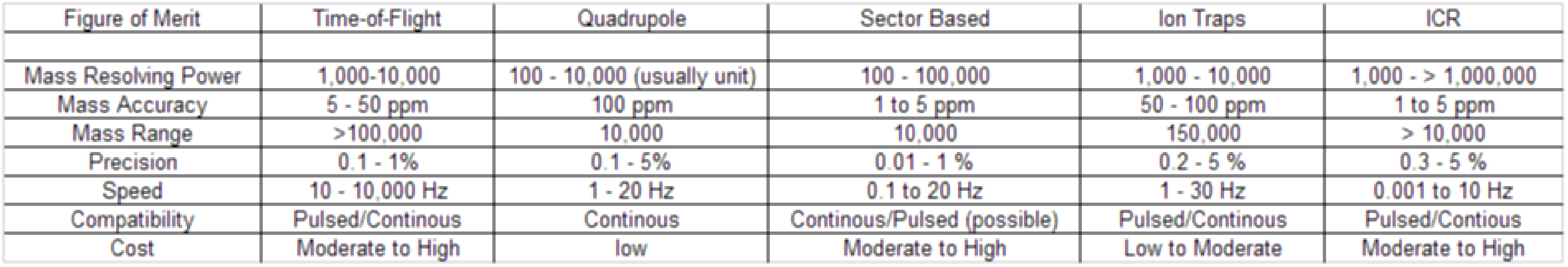


Figure of Merit	Time-of-Flight	Quadrupole	Sector Based	Ion Traps	ICR
Mass Resolving Power	1,000-10,000	100 - 10,000 (usually unit)	100 - 100,000	1,000 - 10,000	1,000 - > 1,000,000
Mass Accuracy	5 - 50 ppm	100 ppm	1 to 5 ppm	50 - 100 ppm	1 to 5 ppm
Mass Range	>100,000	10,000	10,000	150,000	> 10,000
Precision	0.1 - 1%	0.1 - 5%	0.01 - 1 %	0.2 - 5 %	0.3 - 5 %
Speed	10 - 10,000 Hz	1 - 20 Hz	0.1 to 20 Hz	1 - 30 Hz	0.001 to 10 Hz
Compatibility	Pulsed/Continous	Continous	Continous/Pulsed (possible)	Pulsed/Continous	Pulsed/Contious
Cost	Moderate to High	low	Moderate to High	Low to Moderate	Moderate to High

there are a variety of mass analyzers with which to measure ions based on their m/z, only a few of these analyzers have found wide spread use within the scientific community. Much of this can be attributed to the various figures of merit used to define the relative operational strengths and weakness of the different mass analyzers (table 1).[1] However, no single mass analyzer can be labeled superior in all aspects of research served by mass spectrometry. Rather, selection of a mass analyzer depends on the placement of importance, defined by use, with regards to the various figures of merit shown in table 1.[1]

"If is often said, and rightly so, that time-of-flight mass spectrometry has undergone a renaissance over the course of the past two decades. The technique has had a remarkable impact upon the practice of mass spectrometry over the last 20 years, and its "footprint" continues to grow."[1] Much of the growth of time-of-flight (TOF) analyzers can be linked to the development of MALDI. Since TOF analyzers are well suited for pulsed sources, such as MALDI, the marriage of the two is not surprising.

TOF analyzers are typically found in two configurations; linear (Figure II-VI) and reflectron (Figure II-VII). The simplest of the two configurations is the linear time-of-flight analyzer. Independent of the analyzer is the method by which ions are directed into the analyzer. Typically a voltage differential is applied between a repeller and extractor plate. This differential accelerates the ions along the axis of the flight tube (region of the TOF analyzer the ions traverse prior to detection). Not only does the voltage accelerate the ions into the flight tube (field free, low pressure region), but also imparts a kinetic energy (KE) which is integral to the separation theory that allows TOF analyzers to "discriminate" based on m/z.

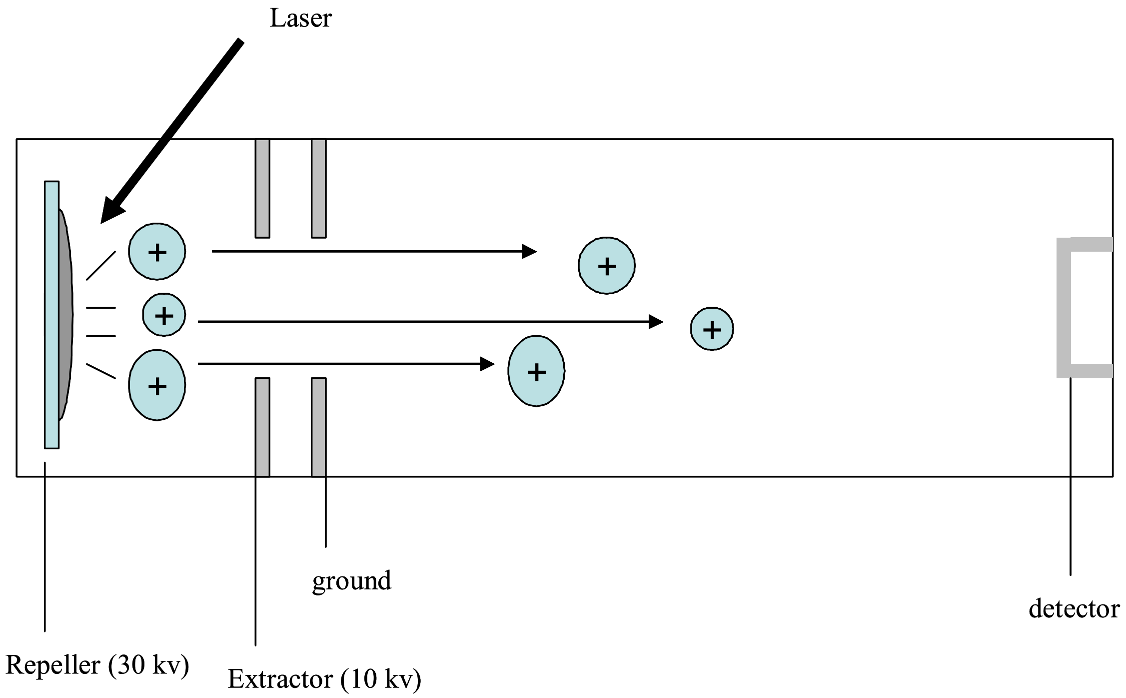


Figure II-VI. Schematic of a linear time-of-flight mass analyzer.

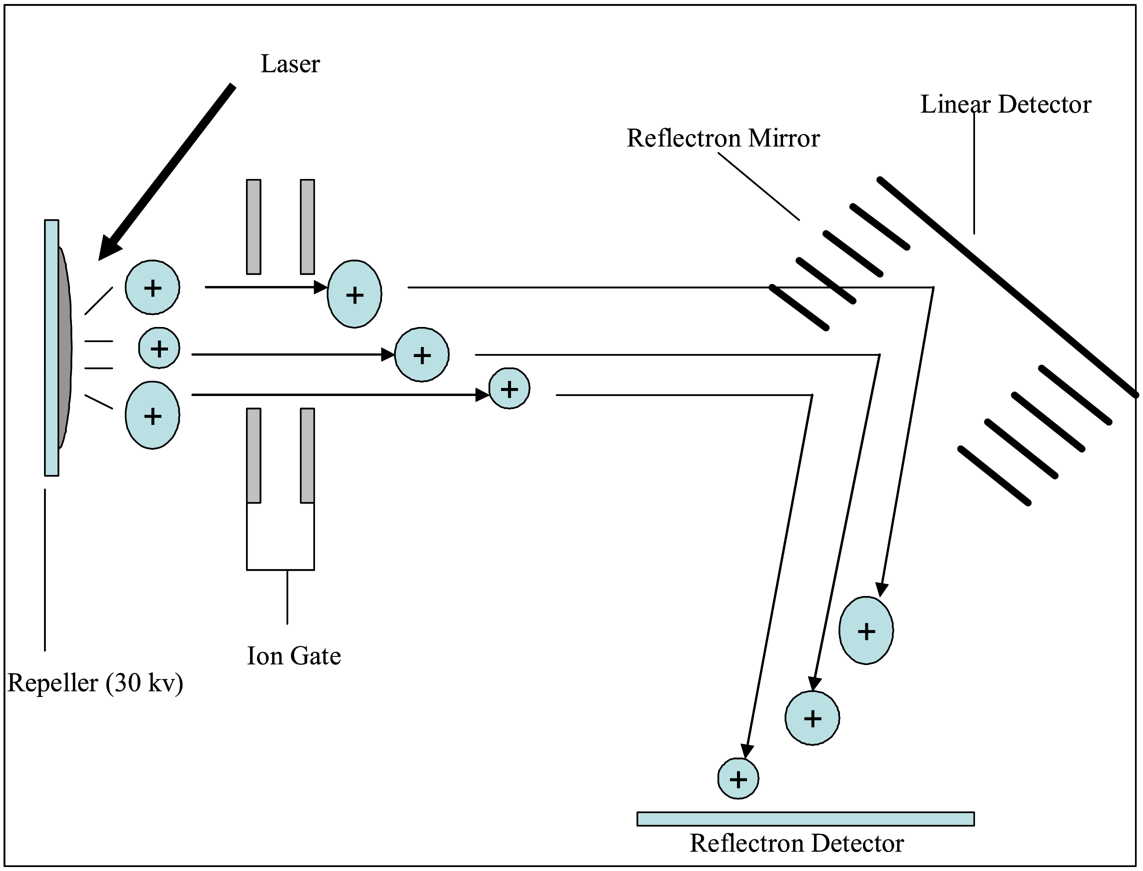


Figure II-VII. Schematic diagram of a reflectron time-of-flight analyzer.

$$KE = 1/2mv^2$$

$$t = \left(\frac{m}{2KE}\right)^{1/2} D$$

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

Figure II-VIII. Equations relating mass, kinetic energy, velocity, and time-of-flight

Essentially, the kinetic energy of the ions can be related to velocity and mass where the velocity is directly proportional to the square-root of the kinetic energy and inversely proportional to the square-root of the mass-to-charge (Figure II-VIII). In simplest terms, the time required for an ion to reach the detector is mass dependent (the basic premise on which TOF analyzers work). Since all ions have the same approximate kinetic energy (20 to 30 keV), their velocities will vary as a function of mass and thus the time required to traverse the flight tube will differ as a function of mass. In theory all ions of the same mass, starting simultaneously at a given point, will arrive simultaneously at the detector generating a single discrete detector response.

In reality the process by which the ions are generated and moved into the analyzer yields a population of ions containing both kinetic energy and spatial distributions. Spatial distributions, simply defined, relate to positions within the source occupied by ions. Ideally all ions should occupy a position where individual distances from the source to the detector are equal. Kinetic energy distributions arise from ionization events and the spatial distribution of ion packets within the instrument source. Both spatial and kinetic energy distributions translate to small velocity variations within a given population of ions with like mass. Ultimately these velocity variations lead to isobaric ions being detected at slightly different times and thus appearing to have slightly different masses (non-isobaric). This type of phenomenon degrades the resolution of the TOF mass analyzer since the width of the peak for a given mass is dependent upon the velocity distribution of that mass.[60]

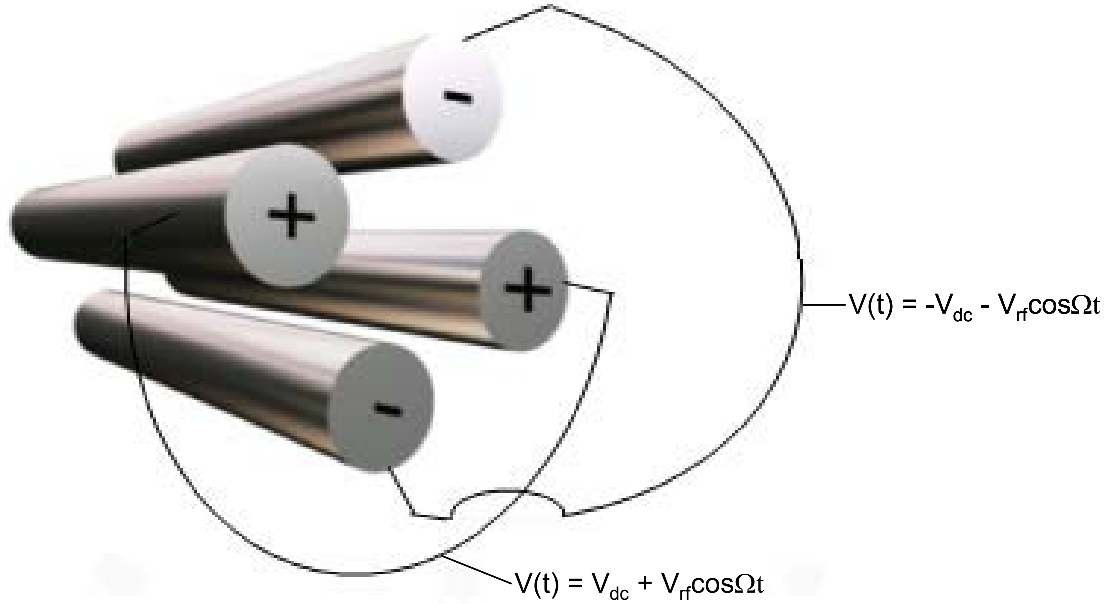
Several techniques have been explored to help reduce energy and spatial distributions within ion populations. The two most widely accepted methods are delayed extraction and reflectron based TOF analyzers, or a combination of both methods. Delayed extraction[60] refers to the absence of an extraction potential (normally used to accelerate ions into the flight tube) for a brief amount of time, typically several hundred nanoseconds. During this delay, the ions generated during ionization are allowed to spatially distribute themselves according to the amount of kinetic energy imparted as a result of the ionization process. Ions with greater energy are allowed to move further into the source. Conversely, lower energy ions move a shorter distance into the source. When the accelerating voltage is applied (after the short delay) it is done such that a potential-gradient is produced within the source. The gradient places the higher energy ions into a lower potential region whereas lower energy ions are exposed to stronger regions of the gradient (higher potential). The result is that ions exposed to stronger potential regions of the gradient are imparted slightly higher kinetic energies than those in the weaker regions of the gradient. This applied distribution of energies results in the spatial focusing of like masses as the ions traverse the flight tube. Thus, isobaric ions with differing kinetic energies, present as a function of ionization, arrive at the detector 'simultaneously', improving the overall resolution of the TOF analyzer.[60]

Reflectron based TOF analyzers also correct for kinetic energy spreads by re-focusing the ions. These analyzers employ an ion mirror commonly referred to as a reflectron (figure II-VII). The ion mirror is composed of an electric field that serves to reverse, or change, the direction of the flight of ions as they traverse the flight tube. In doing so the ion mirror functions to focus ions with the same mass to charge but differing kinetic energies (velocities). Focusing is accomplished by effectively increasing the flight path of ions with greater velocities relative to ions with lower velocities. Since ions with greater velocities will penetrate further into the electric field than ions with lower velocities, higher velocity ions 'see' a longer flight path relative to the detector. The depth of penetration into the electric field serves to refocus ions with like m/z spatially, thus improving the overall resolution.

Due to the nature by which molecules are mass-measured in TOF analyzers there is 'essentially' no upper mass limit to the mass of molecules that can be applied to TOF analyzers. The rapid rate with which data can be accumulated, coupled with high sensitivity, has placed TOF analyzers among the most widely used within the mass spectrometry community. Additionally, hybrid and ToF-ToF based systems have expanded the application range of TOF analyzers. For example, Quad-TOF and TOF-TOF mass spectrometers have become the instruments of choice within the proteomics community due to the wealth of information obtainable when analyzing proteomes using these systems.[1]

The quadrupole mass filter (figure II-IX) is one of the most widely used mass filters for low resolution mass spectrometry experiments.[1] For many years the quadrupole mass filter enjoyed success as a standalone mass filter found in a variety of mass spectrometers. In recent years, the quadrupole has been successfully employed in applications coupling the quadrupole with other mass analyzers (in multi-stage mass spectrometers such as the Q-TOF (quadrupole-time of flight) mass spectrometer). This success can be linked to the high performance, ease of use, and low cost associated with quadrupole based mass spectrometers.

Basically, quadrupole mass filters consist of four rods (either circular or hyperbolic) arranged in an opposing two rod configuration (figure II-IX). The opposing



$$\mathbf{m/z = [k'/\Omega^2 r^2]V}$$

Figure II-IX. Schematic diagram of a quadrupole mass filter.

sets of rods have alternating current (ac/rf (radio frequency)) and dc (direct current) potentials applied to them. The voltages are applied to the rods such that one set of opposing rods receive a positive voltage while the remaining set receive a negative voltage.[61] As with all mass spectrometers, ions are discriminated based on their mass to charge ratio (m/z). Essentially this discrimination is accomplished by the de-stabilization of the flight path of the ions as a function of the applied dc/rf voltages on the quadrupoles (ion filtering). The fundamentals of the filtering effect associated with quadrupoles can be explained by examination of the voltages applied to opposing sets of rods within the quadrupole mass analyzer. Each set of quadrupoles serve to filter ions in, essentially, two planes (xz and yz). In the absence of the dc potential, the ac potential causes the ions to oscillate by converging ions toward the center of the quadrupoles during the positive half of the ac cycle and diverging the ions during the negative half of the ac cycle. Application of a positive or negative dc potential also serves to influence the trajectory of ions. For the positive dc potential, ions will be influenced (primarily by the dc potential) if its m/z is large or its oscillation frequency is high. Alternatively, the application of a negative dc potential serves to destabilize all positive ions by drawing them to the positively charged pair of quadrupoles. For smaller ions this movement may be offset by the oscillation of the ion imposed by the ac potential.[61, 62]

Combining the high and low pass filter characteristics of the two sets of rods produces a ‘bandpass’ filter which acts to filter all ions except those within a small range of masses (figure II-X). In this way there exists a mutual stability region shared by the two sets of rods (the two filters) where only ions with a given range of m/z range can successfully pass. Ions without the proper m/z have un-stable trajectories and collide

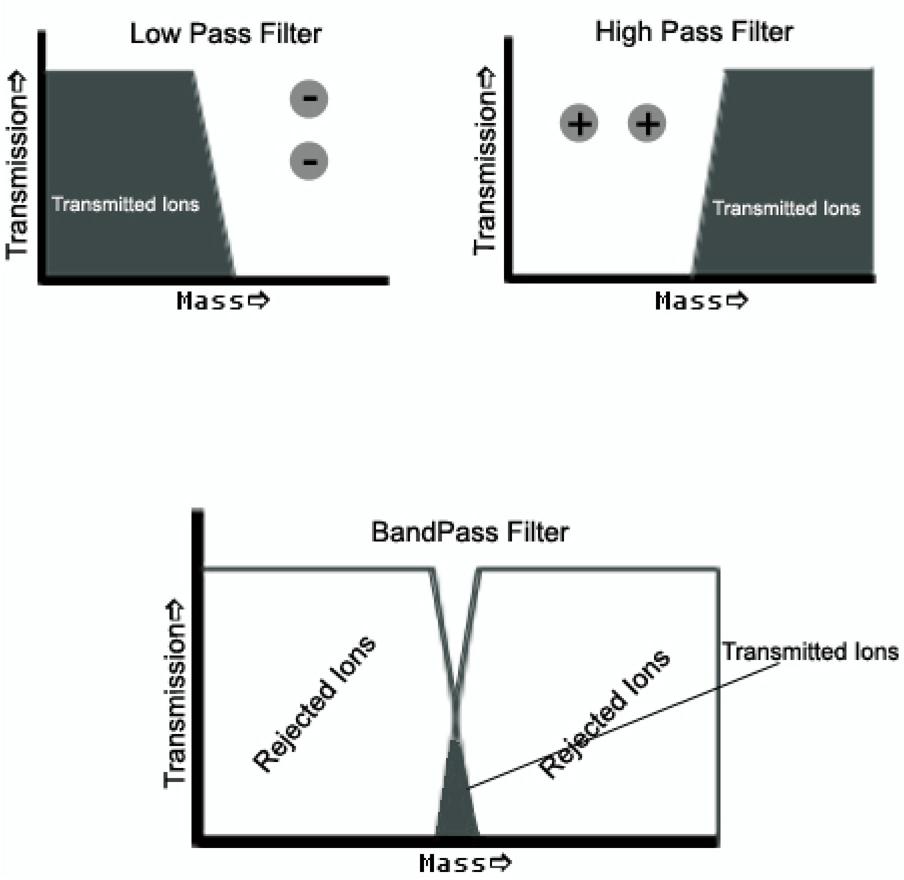


Figure II-X. Formation of bandpass filter by combining the filtering properties of each set of rods.

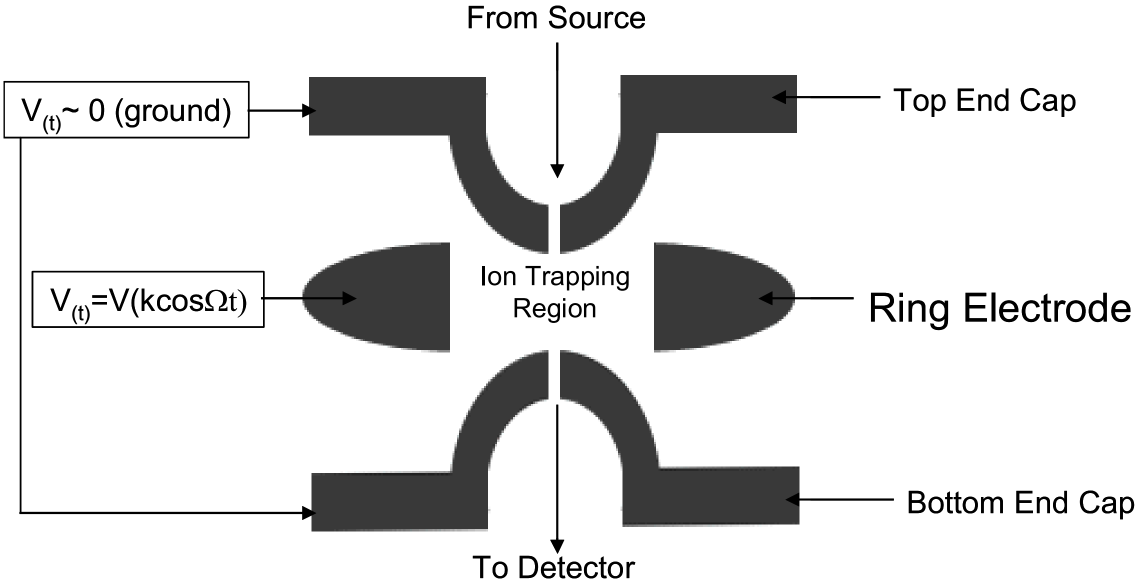
with the rods rather than exiting the mass filter. The mass to charge of the ions that can successfully traverse the mass filter are proportional to the voltage applied to the rods as demonstrated in figures II-IX and II-X. Thus by ramping the voltages applied to the rods, the mass filter is able to 'scan' a range of masses, effectively filtering all masses with the exception of those that are considered stable at the currently applied voltages. It should be noted that the ratio of the applied dc to rf voltages is held constant during the scanning process for any given resolution. Typically quadrupoles are operated with 'unit mass resolution'. This means that ions that differ by 1 da are fully resolved from each other at any given mass (capable of being filtered by the quadrupoles). It is possible to increase or decrease the resolving capabilities of quadrupoles. Increasing resolving power has the adverse effect of decreasing the sensitivity by decreasing the % transmission of ions through the quads. Alternatively de-tuning the quads (decreasing the resolving power) improves sensitivity but decreases mass accuracy and resolution.[61]

It is also possible to operate quadrupoles in such a way as to contain or transmit all ions regardless of their mass to charge. This type of operation is achieved by applying an rf-only field. Operation in rf-only mode can be extremely useful in situations where no mass filtering is needed. Essentially, rf-only quadrupoles act as lens systems or ion pipes. One common use of rf-only quadrupoles is as collision cells within tandem mass spectrometers, although in recent years rf-only quads have been replaced with rf-only octapole systems for this function. Other uses of rf-only quadrupoles include pre- and post- filters which help to focus the ions prior to and/or after mass analysis.[1]

Regardless of the extensive use found by quadrupoles, there has been little change in the figures of merit that are characteristic of the quadrupole mass filter. Until recently quadrupoles found use in applications associated with low and relatively low mass molecules.[1] Typically, the mass range of a quadrupole was 10 to 2000 (maximum of 4000). This placed limits on the mass of the analyte to be studied and often precluded the analysis of many biological molecules. With the development of electrospray ionization, quadrupole mass filters now enjoy success in a variety of biological mass analysis experiments. This is primarily due to the multiple charging (a result of electrospray) effectively bringing the m/z of many large bio-molecules into the mass range of quadrupole mass filters. The low cost, compact design, ease of use, modest linear dynamic range, etc. have made quadrupoles one of the most common mass analyzers/filters in use within the mass spectrometry community.

Figure II-XI shows a diagram of an ion trap mass analyzer. The theory of ion-traps[63, 64] are similar to that of quadrupoles in that rf-voltages are applied to produce stable ion trajectories which form the fundamental basis for mass discrimination. Unlike the quadrupole, however, ion traps do not act to filter a continuous beam of ions. Rather ion-traps serve to 'trap' a discrete quantity of ions and destabilize the trajectory of the ion population as a function of mass. Destabilization is achieved by the addition of an rf voltage that serves to influence the stability of ions based on mass-to-charge. Differing from quadrupoles, destabilization does not 'filter' the ion rather it ejects the ion from the ion- trap into the detector. Thus all ions are trapped within the mass analyzer and remain until ejected by the addition of a destabilizating rf-voltage. [65]

The design of the ion trap is fundamentally linked to quadrupoles in that ion traps can be thought of as a linear rod from a quadrupole bent into a doughnut shape. Above and below the doughnut shaped electrode are two hemi-spherical electrodes that serve to



$$m/z = [k'/\Omega^2 r^2]V$$

Figure II-XI. Schematic diagram of an ion trap analyzer.

support ion trapping orthogonal to the central ring electrode. By application of an rf only voltage to the ring electrodes, the ions within the trap experience an oscillating trajectory, related to their mass-to-charge, near the center of the ring electrode. This oscillation is damped, and further stabilized, by the addition of a gas, typically helium, within the mass analyzer. As the ring-electrode voltage is changed the stability of ions within the ring electrode also changes. Thus, mass scanning is achieved by varying the amplitude of the rf voltage that is imposed upon the ring electrode. Simultaneous with the ramping of the ring electrode-voltage, a small rf-voltage is applied to the end-caps. Once the end-cap frequency matches the resonance frequency of an ion trapped within the ring electrode, the ion will be ejected along the axis of the ring electrode. The process takes approximately 50 to 100 milliseconds (the scan time of an ion trap) to cover a mass range of a few thousand Daltons.[65-67]

Probably one of the most notable characteristics of the ion trap is the high sensitivity associated with this mass analyzer. The sensitivity of a mass analyzer can be related, essentially, to two attributes: transmission of the ions from the source to the detector and utilization of those ions sampled by the source. Ion accumulation directly affects the attribute associated with sensitivity. During the initial phase of ion trap mass analysis, ions are allowed to accumulate within the ring electrode. This is known as the ion accumulation time. Greater times allow for more ions to become trapped inside the ring electrode. Once trapping is complete the ions are then ejected for detection.[67] In an ideal situation all ions accumulated can be ejected and detected leading to an extremely efficient sampling of trapped ions. Ion traps excel in both transmission and ‘utilization of ions’, lending to the high sensitivity associated with them. In addition to high sensitivity, ion traps are generally low cost analyzers and have a small benchtop foot print (due to the compact nature of the analyzer itself). The variety of experiments (SIM, MS^n) that ion traps are capable of has also aided in their recent success within the mass spectrometry industry.[1]

Several other analyzers are found in use among the numerous applications within the field of mass spectrometry. Probably the most notable of these are the ICR analyzers associated with the rapidly growing field of Fourier transform mass spectrometry (FTMS)[64,]. Despite its recent success, FTMS has been an active area of research for over two decades.[1] Currently FTMS enjoys the highest mass accuracies and resolving power of any mass analyzer. As a result, there are many applications where FTMS is the analyzer of choice. Sector based instruments are also found in use within the mass spectrometry community, although to a lesser degree in recent years. Sector instruments are by far the most mature of the analyzers available.[1] Recent advancements in TOF and ion-trapping technologies have all but precluded sector-based instruments for the field of bio-analytical mass spectrometry. "However, in many inorganic and elemental applications, sector field mass spectrometry remains the gold standard for performance, particularly in the areas of dynamic range, abundance, sensitivity, and precision."[1]

Mass spectrometry's success dates back over 100 years. The variety of ionization methods and analyzers have lead directly to the vast acceptance of mass spectrometry as one of the most powerful and widely applicable analytical methods available today. The continued success of mass spectrometry will be driven by the need for greater sensitivity, information, and speed. As addressed earlier, it will likely not be a single analyzer or source that answers all these needs. Rather, advancements in a variety of fields will present the researcher with the needed tools to overcome the new challenges presented by the medical, biological, organic, and elemental fields which are served by mass spectrometry today.

REFERENCES

(1) McLuckey, S. A.; Wells, M. J. *Chem.Rev.* **2001**, *101*, 571-606.

(2) Bierbaum, V. M. *Chem.Rev.* **2001**, *101*, 209-210.

(3) Prebyl, B; Johnson, J. D.; Tuinman, A. A.; Zhou S.; Cook, K. D. *J. Am. Soc. for Mass Spectrom*., **2002**, *13*, 921-927.

(4) Jonsson, A. P. *Cellular and Molecular Life Sciences* **2001**, *58*, 868-884.

(5) Mann, M.; Hendrickson, R. C.; Pandey, A. *Annu.Rev.Biochem.* **2001**, *70*, 437-473.

(6) Fenselau, C.; Demirev, P.A. *Mass Spectrom. Rev.* **2001,** *20*, 157-171.

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

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

(9) Stump, M. J.; et al. *Applied Spectroscopy Reviews* **2002**, *37*, 275-303.

(10) Ehring, H.; Karas, M.; Hillenkamp, F. *Org. Mass Spetrom.* **1992**, *27*, 472-480.

(11) Strupat, K.; Bridson, N.J. *J. Phys. D: Appl. Phys.* **1993,** *26*, 442-447.

(12) Yao, J.; Scott, J.R.; et al. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 805-813.

(13) Knochenmuss, R.; Dubois, F.; et al. *Rapid Comm. In Mass Spectrom.* **1996**, *10*, 871-877.

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

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

(16) Bird, G. H.; Lajmi, A. R.; Shin, J. A. *Anal. Chem.* **2002**, *74*, 219-225.

(17) Talbo, G. H.; Suckua, D.; Malkoski, M.; Reynolds, E. C. *Pept.* **2001**, *22*, 1093-1098.

(18) Mirgorodskaya, E.; Krogh, T. N.; Roepstorff, P. *Methods Mol. Biol.* **2000**, *146*, 273-292.

(19) Goletz, S.; Thiede, B.; Hanisch, F. G.; Schultz, M.; Peter-Katalinic, J. *Glycobiology* **1997**, *7*, 881-896.

(20) Farmer, T. B.; Caprioli, R. M. *J. Mass Spectrom.* **1998**, *33,* 697-704.

(21) Friess, S. D.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 810-818.

(22) Koomen, J. M.; Russel, W. K.; Tichy, S. E.; Russel, D. H. *J. Mass Spectrom.* **2002**, *37*, 357-371.

(23) Garcia, B. A.; Heaney, P. J.; Tang, K. *Anal. Chem*. **2002**, *74*, 2083-2091.

(24) Lay, J. O. *Mass Spectrom. Rev.* **2001**, *20*, 172-194.

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

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

(27) Meng, C. K.; Mann, M.; Fenn, J. B. *Z. Phys. D.* **1988**, *10*, 361-368.

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

(29) Rayleigh, Lord. *Philos. Mag*. **1882**, *14*, 184-186.

(30) On the Mechanism of Electrospray mass spectrometry. In *Electropray Ionization Mass Spectrometry: Fundamentals, Instrumentation, and Applications*; Cole, R. B. ed. Wiley, New York, 1997, pp 3-63.

(31) Dole, M.; Mach, L. L.; Hines, R. L.; et al. *J. Chem. Phys.* **1968**, *49*, 2240-2247.

(32) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.

(33) Fenn, J. B. *J. Am. Soc. Mass Spectrom*. **1993**, *4*, 524-535.

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

(35) Cole, R. B. *J. Mass Spectrom.* **2000**, *35,* 763-772.

(36) Kebarle, J. *J. Mass Spetrom.* **2000**, *35*, 804-817.

(37) McEwen, C. N.; Peacock, P.M. *Anal. Chem.* **2002**, *74*, 743-2748.

(38) Yalcin, T.; Gabryelski, W.; Li, L. |*Anal. Chem.* **2000**, *72*, 3847-3852.

(39) Hunt, S. M.; Sheil, M. M.; Belov, M. *Anal. Chem.* **1998**, *70*, 1812-1822.

(40) Keough, T.; et al. *Electrophoresis* **2000**, *21*, 2252-2265.

(41) Figeys, D.; van Oostveen, I.; Ducret, A. *Anal. Chem.* **1996**, *68*, 1822-1828.

(42) Shen, Y.; Zhao, R.; Berger, S. J. *Anal. Chem.* **2002**, *74*, 4235-4249.

(43) Artice about o-linked glycans and capping with amines.....

(44) Zappacosta, F.; Huddelston, M. J.; Karcher, R. L. *Anal. Chem.* **2002**, *74*, 1658-1664.

(45) Reid, G. E.; Stephenson, J. L.,; McLuckey, S. A. *Anal. Chem.* **2002**, *74*, 577-583.

(46) Meng, F.; Cargile, B. J.; Patrie, S. M. *Anal. Chem.* **2002**, *74*, 2923-2929.

(47) Shen, Y.; Tolic, N.; Zhao, R. *Anal. Chem.* **2001**, *73*, 3011-3021.

(48) Clauwaert, K. M.; Van Bocxlaer, J. F.; Lambert, W. E. *Anal. Chem.* **1998**, *70*, 2336-2344.

(49) Oberacher, H.; Wellenzohn, B.; Huber, C. G. *Anal. Chem.* **2002**, *74*, 211-218.

(50) Little, D. P.; Aaserud, D. J.; Valaskovic, G. A. *Journal of the American Chemical Society* **1996**, *118*, 9352-9359.

(51) Zhang, S.; Van Pelt, C. K.; Schultz, G. A. *Anal. Chem.* **2001**, *73*, 2117-2125.

(52) LeRiche, T. *Rapid Commun Mass Spectrom.* **2001**, *15*, 608-614.

(53) Kassel, D. B. *Chemical Reviews* **2001**, *101*, 255-267.

(54) Griffey, R. H.; SannesLowery, K. A.; Drader, J. J.; Mohan, V.; Swayze, V. V.; Hofstadtler, S. A. *J. of the Am. Chem. Soc.* **2000**, *122*, 9933-9938.

(55) Naylor, S.; Benson, L. M.; Tomlinson, A. J. *J. Chromatogr. A* **1996**, *735*, 415-438.

(56) Wang, Y.; Shi, L.; Song, F. R.; Liu, Z. Q.; Liu, S. Y. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 279-284.

(57) Barber, M.; Bordoli, R. S.; Sedgwick, R. D.; Tetler, L. W. *Org. Mass Spectrom.* **1981**, *16*, 265-273.

(58) Harrison, A.G. *Chemical Ionization Mass Spectrometry* 2nd Ed., CRC Press, Boca Raton, Fl., 1992.

(59) Harkey, M.; Henderson, G.; Zhou, C. *J. Analytic Toxicology* **1991**, *15*, 260-266.

(60) Vestal, M. L.; Juhasz, P.; Martin, S. A. *Rapid Comm. Mass Spectrom.* **1995**, *9*, 1044-1050.

(61) Miller, P. E.; Denton, M. B. *J. Chem. Educ.* **1986**, *7*, 617-622.

(62) Ying, J.; Douglas, D. J. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 649-655.

(63) March, R. E. *Rapid Commun. Mass Spectrom.* **1998**, *12*, 1543-1544.

(64) Burlingame, A. L.; Boyd, R. K.; Gaskell, S. J. *Anal. Chem.* **1998**, *70*, 467R-716R.

(65) Schwartz, S. A.; Jardine, I. *Methods. Enzymol.* **1996**, *270*, 552-586.

(66) McLukey, S. A.; Van Berkel, G. J.; Goeringer, D. E.; Glische, G. L. *Anal. Chem.* **1994**, *66*, 737A-743A.

(67) Jonscher, K. R.; Yaters, J. R. III *Anal. Biochem.* **1997**, *244*, 1-15.

(68) Buchanan, M. V.; Hettich, R. L. *Anal. Chem.* **1996**, *65,* 245A-259A.

CHAPTER III

IDENTIFICATION OF ALLERGENIC PROTEINS ISOLATED FROM *PENICILLIUM CHRYSOGENUM*[1]

---

[1] Johnson, J.D., Orlando, R.C. To be submitted to *J Allergy Clin Immunology*.

ABSTRACT

Allergic reactions can derive from a variety of in-home and outdoor sources. With regard to in-home allergens, little information has been obtained on fungal sources as potential stimulants of allergic reactions. A few publications have presented data indicating fungal species as sources for a variety of allergenic responses including allergic rhinitis, extrinsic allergic alveolitis, and sinusitis. However, little information exists with regards to the components within the fungal source that act as 'triggering agents' in the allergic response process. Presented here are experiments utilizing *Penicillium Chrysogenum*, a common indoor fungus, to determine proteinaceous components integral to the production of immuno-globulin E (IgE), an antibody associated with allergic response. Protein identification is often dependent upon searchable databases with which sequence information or peptide mass fingerprint data can be utilized to determine the identification of isolated or bulk proteins. Homology based protein identification against established fungal databases was used to tentatively identify protein(s) determined to be 'allergenic' based on a series of antibody specific interactions/isolation followed by liquid chromatography-tandem mass spectrometry.

INTRODUCTION

Reports from the Centers for Disease Control and Prevention state that asthma rates have nearly doubled in the US since 1980.[1] Childhood asthma has shown the most dramatic increases. Abnormal immunological (allergic) reaction is the primary cause of most reported asthma cases. This type of 'reaction' arises from one of three primary modes: inhalation of a bio-aerosol, ingestion, or dermal contact. For any of the modes to initiate an allergic response there must be an 'allergen' present. Some well characterized allergens are proteinecous materials derived from pets, cockroaches, and dust mites. Thus, most asthma cases are the result of exposure to an allergen within the atopic (genetically predisposed) individual's environment. In addition to pets, cockroaches, and dust mites, a less characterized but believed to be equally important source of allergens is indoor fungi.[2, 3] Few reports have been published utilizing human challenges for the study of fungal allergenic pathogens, and little information is known as to the identity of fungal allergenic proteins.

In addition to the potential problems fungi pose with regards to in-home allergens, fungi have been confirmed as a source for allergic rhinitis, extrinsic allergic alveolitis,[4, 5] and sinusitis.[6] Based on current lifestyle ~ 90% of a human's life is spent indoors. Thus, indoor contaminants could potentially be a greater issue than exposure to outdoor allergens. Identification of indoor fungal allergens allows for the quantification of exposure to specific fungi, which in turn, allows for a correlation of fungal exposure to the onset of allergies. The immediate goal arising from quantification studies would be to determine exposure limits in order to develop methods or practices to help reduce exposure below critical levels.

Since little information is known as to the identities of fungal allergens, standardized fungal proteinaceous allergens do not exist. This complicates the study of fungal proteins with respect to their role in human health.[5] In the past, crude fungal extracts have been prepared in a variety of ways including freeze-dried liquid cultures from artificial media. However, batch to batch variability within fungal extracts is often greater than variability observed among different strains, species, and even genera.[7]

The goal of this research was to identify a model process by which specific proteins found within fungal protein-extracts could be associated with allergenic responses. Additional work would focus on the batch purification of said proteins to be utilized as suitable standards for further studies including the production of fungal species-specific reagents (therapeutics).

The hallmark of allergenic disease is the up-regulation (abnormally high serum levels) of IgE's. It is known that IgE's are a direct consequence of exposure to allergenic agents. By utilizing a combination of mass spectrometry and immuno-chemical techniques it was possible to identify target proteins associated with allergenic reactions. Gel Electrophoresis and immunochemical techniques provided a way of isolating those proteins associated with up-regulation of IgE's. Mass spectrometry was then utilized to tentatively identify the protein through micro-sequencing. One common species of indoor air fungi was used, initially, in the development of the isolation/identification protocol: *Penicillium chrysogenum.*

EXPERIMENTAL

Hyperimmunization and Serum Extractraction

15 Balb-C mice were inoculated with 25 mg of a whole protein extract of *Penicillium Chrysogenum* suspended in 0.2 ml of 1.3% alhydrogel, by intra-peritoneal (IP) injection. This was followed by two boost injections (15 mg of extract in 0.2 ml Hank's balanced salt solution) at three week intervals. Blood was collected (cardiac bleed) five days after the final inoculation. Each mouse yielded ~ 300 μl of whole blood. Serum was separated by centrifugation at 12,000 rpm in a microfuge for 10 minutes at 4 ^{o}C.

Bradford Assay

Protein extracts of *Penicillium Chrysogenum* were used as received with the exception of spin filtration over a 0.22 μm filter (Millipore, Bellireca, MA, USA). Prior to use a micro-Bradford analysis was performed to determine total protein content. Samples were analyzed in triplicate as follows. A protein concentration calibration plot was generated using a series of BSA standards at 0.1 mg/ml, 0.2 mg/ml, 0.3 mg/ml, 0.4 mg/ml, 0.5 mg/ml, 1.0 mg/ml, and 10 mg/ml. A blank, composed of distilled water, was also used. *Penicillium Chrysogenum* extracts were prepared in 1:10, 1:5, and 1:1 dilutions utilizing distilled water as the diluent. For all reactions 10 μl of either extract or standard was mixed with 200 μl of 1X Bradford reagent (Biorad, Bradford, PA, USA). Reactions were carried out in a microtiter plate. The microtiter Plate was read using a Biorad 1022 Plate reader (Biorad, Bradford, PA, USA) at 480 nm. Total protein content was determined to be ~ 8 mg/ml (protein extract sample 1) and ~ 80 μg/ml (protein extract sample 2).

IgE Generation/Assay -

IgE content was measured utilizing a modified Enzyme Linked Immuno-Sorbent Assay (ELISA). Due to the high levels of other antibodies present (primarily IgG's) detection of IgE's was facilitated by utilizing a capture antibody (purified rat anti-mouse IgE monoclonal antibody, BD Pharmingen, San Diego, CA, USA). Each well of a microtiter plate (Imullon 2, Dynatech, Chantilly,VA, USA) was coated with 100 μL of purified rat anti-mouse IgE diluted to a 1:250 ratio in coating buffer (phosphate buffered saline (PBS), pH = 7.2). The plate was covered and incubated at 4 ^{o}C overnight. Next, the plate was then aspirated and washed 3X with 200 μL of wash solution (0.05% Tween-20 in PBS, 1 minute per wash). Next, 300 μL of blocking solution (1% BSA in PBS) was added to each well and the plate was covered and incubated for 1 hour at room temperature. The plate was then aspirated and washed 3X, at 1 min intervals, with wash buffer. Next, standards and serum samples (unknowns) were added to assigned wells on the microtiter plate as follows: 100 μL of either standard or serum (diluted in assay diluent (1% BSA in PBS) was added to the plate. Serum samples were diluted to a 1:10, 1:100, and 1:1000 ratio. Standards (Purified mouse IgE, BD Pharmingen, San Diego, CA, USA) were prepared at the following dilutions – 0 mg/ml (assay diluent alone, serves as blank), 1.2 ng/ml, 3.6 ng/ml, 6.3 ng/ml, 12.5 ng/ml, 25 ng/ml, 50 ng/ml, and 100 ng/ml. The plate was allowed to incubate at room temperature for 1 hour, washed 5X with wash buffer, then 100 μL of working detector (1:1 biotinylated Rat anti-mouse

IgE (BD Pharmingen, San Diego, CA, USA) and Avidin HRP Conjugate (BD Pharmingen, San Diego, CA, USA) diluted to 2 μg/ml in blocking solution was added to each well. The plate was covered and allowed to incubate at room temperature for 2 hours, aspirated and washed 7X (1 minute intervals) with wash buffer. Plates were allowed to soak at least 1 minute between each was step. 100 μL of tetramethylbenzidine (TMB) substrate solution (CalBiochem, La Jolla, CA, USA) was added to each well, and the plates were incubated in the dark at room temperature for 30 minutes. Absorbances were read at 450 nm.

1D Gel Electrophoresis

Prior to preparation of samples for 1D gel electrophoresis the extract was passed over a .22 μM filter (Millipore, Bradford, PA, USA). 12.5 μl of fungal protein extract was mixed with 12.5 μl Laemmli buffer (Biorad, Hercules, CA, USA). This solution was loaded into two 15% tris-HCl (Biorad, Hercules, CA, USA) gels for 1 dimensional gel electrophoretic analysis. Gels were run utilizing a Biorad Protean II 7 cm system (Biorad, Hercules, CA, USA) at 300 Volts/30 mAmps (constant) until the bromophenol blue band was within 0.5 cm of the gel bottom. Gels, for 1d whole protein detection, were developed as follows: 1) washed in 50% ethanol twice at 15 minute intervals. 2) pretreated, for 1 minute, in Sodium Thiosulfate (Sigma, St. Louis, MO, USA). 3) rinsed 3X for 20 seconds per rinse with DI water. 4) impregnated with silver nitrate (2g/L dissolved in .028% formaldehyde). (Sigma, St. Louis, MO, USA). 5) rinsed 2X for 20 seconds per rinse with DI water. 6) allowed to develop (Sodium Carbonate (60g/L), 0.5

ml 37% formaldehyde, sodium thiosulfate (4 mg/L)) until bands were visible. 7) Developing was stopped by the addition of 50% methanol containing 12% acetic acid.

Western Analysis

Orthogonal electroelution was carried out utilizing a Protean II Western Kit (Biorad, Bradford, PA). Proteins were electro-eluted (trans-migration) from the 1d gels onto nitro-cellulose at 45 Volts/200mAmps for 1 hour. After migration the cellulose membrane was removed and immediately placed in 1% BSA (Sigma, St. Louis, MO, USA) for 30 minutes. Next, the membrane was rinsed for 30 seconds in ultrapure water, followed by 2X rinsing in 0.1% tween-20 tris buffered saline (TTBS) at pH 7.5 for 1 minute intervals. 45 ml of primary antibody solution (1 ml of mouse serum added to 45 ml of 1% BSA containing 0.05% Tween-20) was added and allowed to incubate at 4 ^{o}C overnight. The membrane was then washed 3X with TTBS at 3 minute intervals. Secondary antibody, alkaline phosphatase conjugated rat anti-mouse IgE, (BD Pharmingen, San Diego, CA, USA) was added and incubated at room temperature for 1 hour. After incubation the membrane was washed 3X at 3 minute intervals with TTBS followed by 3 washings at 3 minute intervals with tris-buffered saline (pH = 7.5). Finally, developing solution (600 μL BCIP (Sigma, St. Louis, MO, USA) + 600 μL NBT (Sigma, St. Louis, MO, USA), diluted in 60 ml of detection buffer, was added and the membrane was allowed to incubate until bands appeared. Typical developing times were on the order of 5 to 6 hours.

In-Gel Digestion

Gel bands determined to contain proteins (correlation of western membrane with 1D gel separated protein extract) with allergenic properties were excised and in-gel digested with trypsin using a ProGest autodigester (Genomics Solutions, Ann Arbor, MI, USA). Gel bands were first excised and cut into 1 x 1 mm pieces. Note –solutions are discarded between each step except where noted. Gel pieces were washed with 200 µl of HPLC water for 10 minutes. Next, 200 µl of freshly prepared 50 mM sodium thiosulfate was added and gels pieces were incubated at room temperature for 15 minutes. This was repeated one time. Gel pieces were then washed 2X for 10 minutes with 30 mM Potassium Ferricyanide. Next, 250 µl of HPLC water was added and gel pieces were allowed to incubate for 10 minutes followed by a 10 minute wash with 200 µl of 25mM ammonium bicarbonate. Gel pieces were then dehydrated 2X for 10 minutes with 200 µl 25mM ammonium bicarbonate/50% acetonitrile, at which point the gel pieces shrink and become white in color. Gel slices were then re-hydrated in 30 µl of freshly prepared dithiothreitol (in 25 mM ammonium bicarbonate) and allowed to incubate for 30 minutes at 56 ^{o}C. Next, 15 µl of freshly prepared 55 mM iodoacetamide (in 25 mM ammonium bicarbonate)was added and allowed to incubate in the dark at room temperature for 30 minutes. The reaction was stopped by the addition of 200 µl of 25 mM ammonium bicarbonate. Gel pieces were again dehydrated with 200 µl of 25 mM ammonium bicarbonate in 50% acetonitrile. Supernatant was discarded and gel pieces were allowed to dry for 10 minutes. Gel pieces were then rehydrated at room temperature with 25 µl of sequencing grade trypsin (10 ng/ml (Promega, Madison, WI, USA)) in 25 mM ammonium bicarbonate pH 8.0 for 15 minutes (solution was not discarded prior to next step). 50 µl of ammonium bicarbonate, pH 8.0, was added. Gel pieces were allowed to incubate for 18 hours at 37 ^{o}C. Gel pieces were then extracted using 40 µl of distilled water containing 0.1% trifluroacetic acid (solution was not discarded). Two additional extractions were performed using 45 µl of 5% trifluroacetic acid/70% acetonitrile for 20 minutes (solutions were not discarded). The resulting 'extraction' solution was transferred to a total recovery vial (Waters, Milford, PA, USA) and dried in a vacuum centrifuge. Dried peptides were reconstituted in 5 µl of HPLC water prior to LC-MS analysis.

MASS SPECTOMETRY (LC-MS/MS)

All experiments were carried out utilizing a Q-Tof II (Micromass, Manchester, UK) mass spectrometer operating in auto-switching mode. Data were collected in positive ion MCA mode. 'Auto Switching' was utilized to obtain product ion data for ions pre-selected in MS1. Instrument conditions were as follows - Cone Voltage (CV) = +40V, Capillary Voltage = +2800V, flowrate = 180 nL/min (defined by LC), collision gas = argon, gas cell pressure = 15 psi.

A Waters Capillary LC was coupled to the instrument for sample separation and clean-up prior to entry into the mass spectrometer. Peptides collected from in-gel digestion were injected (5 µL) and separated on a 15 cm x 75 µM id C-18 column (LC Packings, San Francisco, CA, USA). LC conditions were as follows: a linear gradient was employed where %B (90 % acetonitrile, 0.1% formic acid) was ramped from 10% to 70 % over 95 minutes relative to mobile phase component A (0.1 % formic Acid). Those ions with sufficient signal-to-noise were subjected to collision induced dissociation and the resulting MS/MS spectrum collected. Data obtained for MS experiments were uploaded to Mascot (www.matrixsciences.com) and searched against fungal databases containing MS/MS peptide information.

RESULTS AND DISCUSSION

Bradford analysis of whole protein extracts

Whole protein extracts were subjected to Bradford assays in order to quantitate total protein content. This step provides no information as to the allergenic properties of the proteins found within the extract, rather it simply provides a measure of the success of the protein extraction, with higher protein concentrations indicating more successful extraction of proteins from cellular materials. Shown in Figure III-I is a representative plot from the Bradford analysis of protein extract sample 2. For experiments presented in this chapter protein extract 1 was utilized due to the greater concentration of proteins based on the Bradford assay. Experiments utilizing protein extract sample 2 yielded results commensurate with the decreased protein content (including lack of protein detected in 1d gel experiments and subsequent western transfer assays).

Enzyme linked immuno-sorbent assay of hyperimmunized mice

ELISA analysis of serum samples was done to determine total IgE content. Two serum 'pools' were obtained from separate inoculations of two groups of 15 mice. IgE content within the first serum pool (SP1) was determined to be ~80 to 100 ng/ml (Figure III-II). IgE content within the second serum pool was determined to be ~ 5 µg/ml (Figure III-III). The second pool (SP2) differed from the first pool in that individual serum

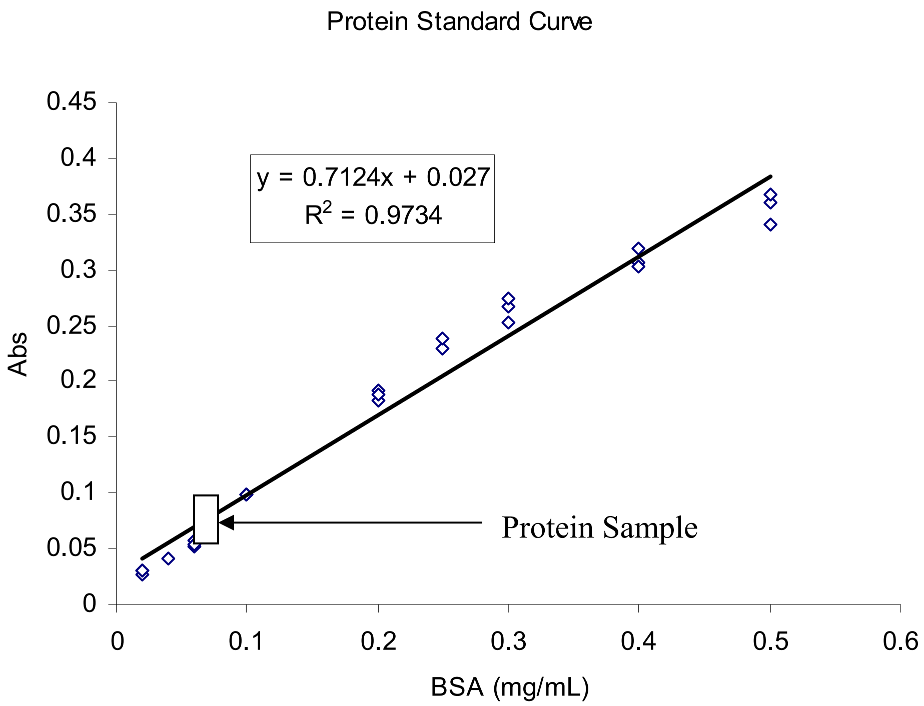


Figure III-I – Bradford Assay for Total Protein Content (fungal extract)

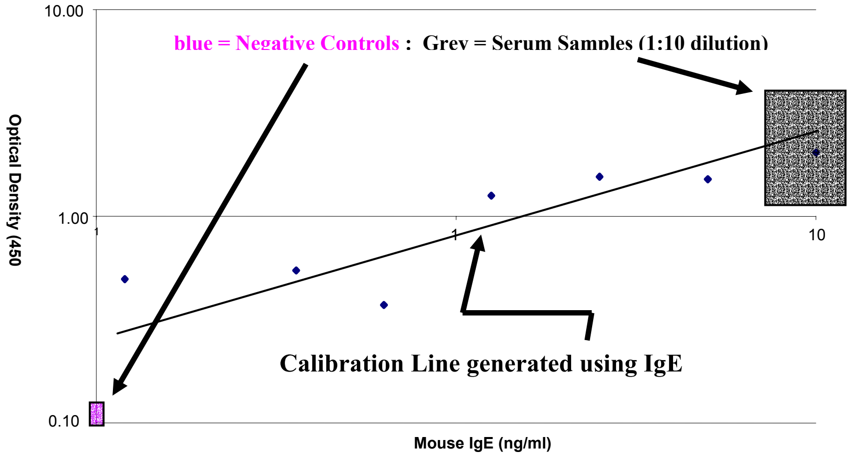


Figure III-II – ELISA determination of IgE content for SP1.

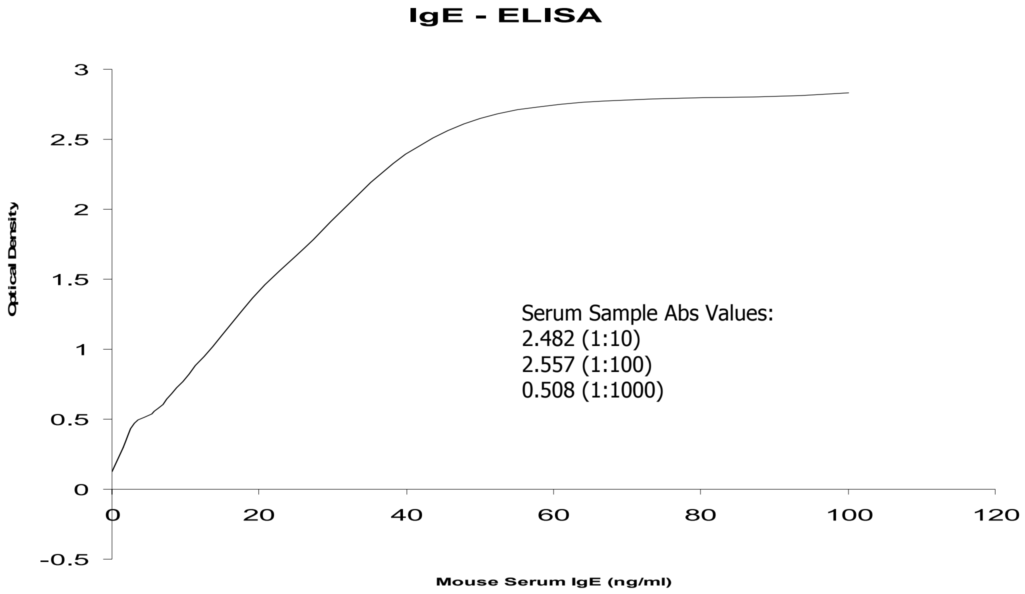


Figure III-III – ELISA determination of IgE content for SP2.

fractions (from individual mice) were pooled prior to ELISA analysis. With SP1 the serum samples were kept separate (i.e. the serum isolated from a given mouse was not mixed with other serum samples from other mice). Based on ELISA data it was determined that pooling the fractions would result in a more uniform representation of IgE content and reduce run-to-run variability as a result of differing IgE content within serum samples. It is speculated that IgE content within mouse serum will fluctuate depending on a mouse's 'sensitivity' to allergens, much like humans have differing degrees of sensitivity to allergen toxicity.

Western analysis/western transfer

Once IgE content had been confirmed Western Transfer (WT) experiments were employed to determine if up-regulation of IgEs could be attributed to proteins within the fungal extracts. Utilization of the mouse serum as the primary antibody would also allow for specific identification of proteinaceous components within the fungal extract causing up-regulation of IgE antibodies.

'1D Gel Analysis' (see figure III-IV) demonstrated the separation of the proteinaceous contents within the fungal extract based on molecular size. Subsequent 'western transfer' of the 1D separated proteins (see figure III-IV) allowed for the antibody specific detection of certain (isolated) proteins from the 1D analysis. The western transfer orthogonally elutes the proteins from within the gel matrix onto a piece of nitrocellulose. Once proteins were immobilized onto the nitrocellulose membrane it became possible to treat the proteins with a solution of antibodies in order to test for antigenic function within the population of immobilized proteins. For these experiments

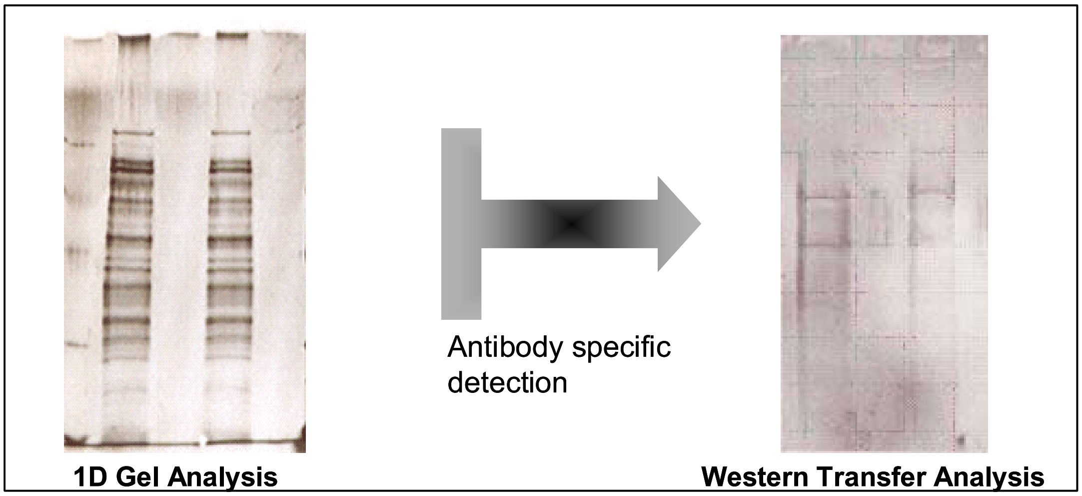


Figure III-IV - Western Transfer Identification of Allergenic Protein Components within Mouse Serum.

the serum isolated from the mice served to supply the antibodies with which the immobolized fungal proteins were assayed against. The results of the WT experiments clearly indicated that some components of the fungal extract were antigenic toward the antibodies found within the serum of mice hyperimmunized with the fungal extract, indicating that the mice were in fact sensitive to select components of the fungal extract in a way indicative of allergenic models.

Not only did the Western Transfer serve as proof that some IgE's within the serum are specific for components within the extract, but it also served as a map to correlate which bands present within the 1D gel corresponded to 'allergenic proteins'. By using this correlation map allergenic proteins contained within bands were in-gel digested and analyzed by LC-MS/MS.

In-gel digestion/Protein Identification

Using the Western Map (Figure III-IV) a band was excised from the 1d gel (Figure III-IV) corresponding to a fungal protein. This protein was then in-gel digested and analyzed by LC-MS/LC-MS/MS. Data accumulated over a mass range of 400 to 2000 daltons were combined to form a total ion chromatogram (Figure III-V). MS/MS data were then processed using the Masslynx software to output a .pkl file. This .pkl file contains all peaks, for a given MS/MS experiment, centroided and voided of all isotope peaks. The output data file was then uploaded to Mascot (www.matrixscience.com) and searched against online databases (fungal specific) for primary amino-acid sequence matches. Since there exists little genome information for *P. Chrysogenum*, protein identification was largely homology based. Thus, proteins were identified using only

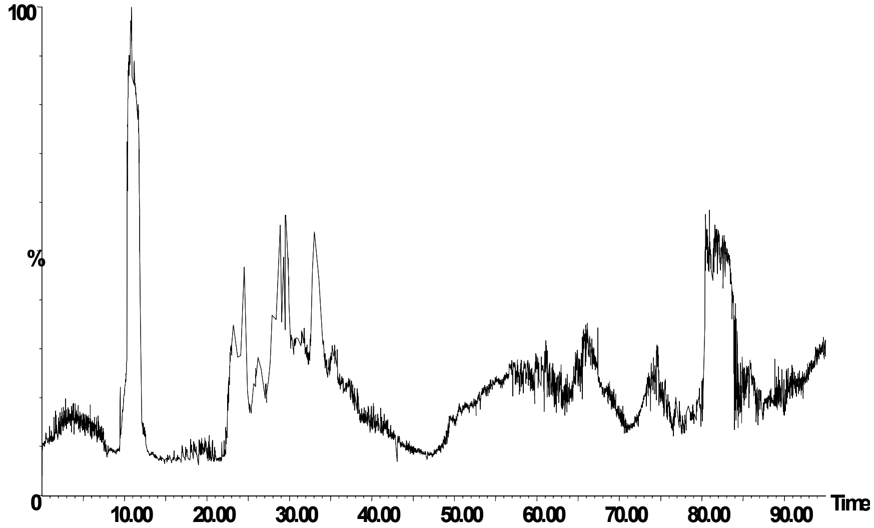


Figure III-V – Total ion chromatogram from representative in-gel digestion one peptide since the likelihood of detecting multiple homologous peptide regions was low.

Shown in Figure III-VI is the corresponding MS/MS spectrum of a peptide that matched a protein within the Mascot database. This process identified a protein, present within the band, with a high degree of homology for the MOT2 (MOT; modulator of transcription) protein (MW = ~67,000) from *Candida Albicans*. *Candida Albicans* is a fungus that is responsible for the most frequently occuring fungal infections in humans.[8] It has also been strongly linked to allergens resulting in bronchial asthma.[9]

While other peptides, from the same gel band, are readily detectable in the LC-MS and LC-MS/MS experiments none corresponded to the same MOT2 protein. As stated previously this could be due to the lack of other 'homologous regions' between the *C. Albicans* protein and the *P. Chrysogenum* protein. A further factor to consider when doing homology based searches is the likelihood that a protein may contain no regions of significant homology with those present within a database(s). Shown in Figure III-VII is data obtained from the in-gel digestion of another band suspected of containing an 'allergenic' protein based on the Western data described earlier. Although the MS/MS data is sufficient to obtain sequence information no significant match was returned. This is likely a result of little or no homology shared between the protein from the gel-band and one in the online database. As databases become more extensive the likelihood of such non-matched hits will decrease, the confidence of homology based hits should increase, and species specific protein identifications should become more common.

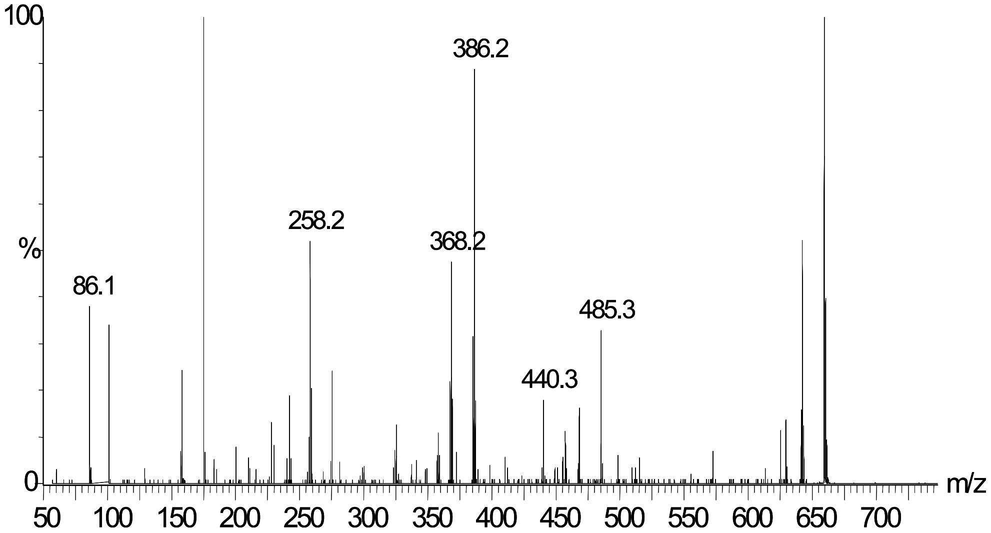


Figure III-VI – MS/MS spectrum of Mascot Identified Peptide.

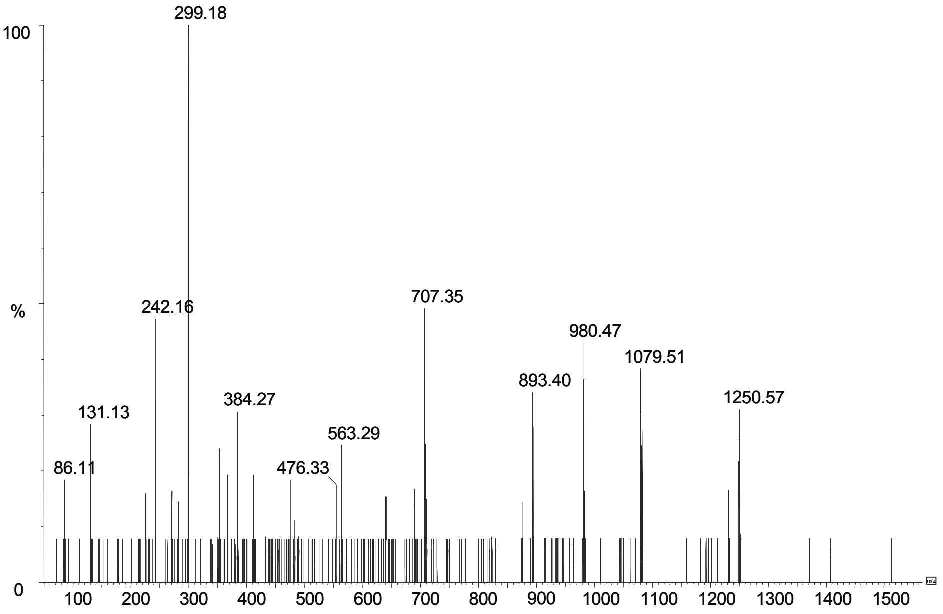


Figure III-VII – MS/MS data from ingel digest of an additional protein.

CONCLUSIONS

Mice were successfully hyperimmunized resulting in the formation of antibodies in response to the foreign fungal proteins. Western analysis combined with 1D gel electrophoresis was utilized to identify proteins associated with up-regulation of IgE's resulting from hyper-immunization with a *P. Chrysogenum* protein extract. This provided data for the first time that clearly indicated proteins isolated from *P. Chrysogenum* are likely to induce reactions indicative of allergenic models.

Since little genome information was available for *P. chrysogenum*, protein identification was homology based. While several homology based hits have been obtained only one yielded extensive homology, desired for a high-confidence identification. This could be for several reasons. First homology based hits depend on the detection of a peptide that shares a similar sequence region with peptides from an organism other than itself. While there exists the possibility that such an event may occur it is as likely, if not more so, that homologous regions will not be detected within the generated peptides. Additionally, peptide coverage from gels are lower than from solutions making the detection of homologous peptides even less likely. However, more efficient digestions and detection of peptides will increase the likelihood of successful identification of proteins utilizing homology based protein searching. Second, there could be errors in the in-silico generation of a sequence from MS/MS data. As an alternative de Novo sequencing was also employed to more accurately discern peptide sequences prior to database searching, especially in situations where no homology 'hits' were obtained. Comparisons of *de Novo* and *in Silico* sequences showed that computer generated data were accurate for those peptides searched.

Given the protein extract contains a high number of unique proteins, misidentification of proteins as the result of multiple proteins occupying a single band is also an issue. Two dimensional gel electrophoresis provides far superior separation of proteins, greatly reducing the likelihood that multiple proteins could occupy a single spot/band. Also, due to the nature by which proteins are separated in 2D analysis the proteins are more 'concentrated' yielding the likelihood of detecting additional proteins associated with allergenic response.

Thus, future work might center on the generation of 2D 'western' maps as opposed to 1d maps. Subsequent in-gel digestion and MS/MS analysis should again be utilized to identify proteins associated with IgE up-regulation.

ACKNOWLEDGEMENTS

We graciously thank the Environmental Protection Agency (Grant number R-82833701-0) for financial support of this project. This work was presented in preliminary form at the 43rd ASMS conference on Mass Spectrometry and Allied Topics, June 2001, Chicago, Illinois and at the 44th ASMS conference on Mass Spectrometry and Allied Topics, June 2002, Orlando, Florida.

REFERENCES

(1) Centers for Disease Control and Prevention. 1996. Asthma surveillance programs in public health departments – United States. *MMWR* **45**: 802-804.

(2) Salvaggio, J.; Aukurst, L. *J Allergy Clin Immunology* **1981,** *68*, 327–346.

(3) Miller, J.D. *Atmosphere Environ*. 1992, *26a*, 2163-2172.

(4) Flannigan, B.; Miller, J. 1994. Health implications of fungi in indoor environments: an overview. In: *Health Implications of Fungi in Indoor Environments*, pp. 3 – 28. Elsevier Science Publications, Amsterdam, The Netherlands.

(5) Verhoeff, A.P.; Burge, H. A. *Annals of Allergy, Asthma, and Immunology* **1997**, *78*, 544- 554.

(6) Ponikau, J. U.; Sherris D. A.; Kern E. B.; Homburger H. A.; Frigas E.; Gaffer, T. A.; Roberts, G. D. *Mayo Clin. Proc*. **1999**, *74*, 877–884

(7) Burge, H. A. *Immunol. Allergy Clin N Am.* **1989**, *9*, 307-319.

(8) http://www.pasteur.fr/actu/presse/com/dossiers/GBgenomics/GBgencandi.html

(9) http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=10938515&dopt=Abstract

CHAPTER IV

CHARACTERIZATION OF POST-TRANSLATIONAL MODIFICATIONS OF RECOMBINANT *PECTATE LYASE*[1]

---

[1] Johnson, J.D., Bergmann, C., Orlando, R. To be submitted to *J Am. Soc. Mass Spectrom*.

Abstract

Pectate lyase from *Aspergillus niger* was over-expressed homologously in its native host as well as heterologously in *Aspergillus nidullans*. This resulted in four recombinant proteins, two from each host. The MALDI MS data of the four proteins showed that their molecular masses were greater than the molecular masses calculated from the amino acid sequence. Two of the proteins, one from each host, had molecular masses ~1200 Da greater than expected while the other two proteins had molecular masses ~2400 Da greater than expected. Two explanations for the observed discrepancy in molecular masses are glycosylation and phosphorylation. A variety of techniques, including lectin analysis, solid phase extraction, liquid chromatography, site specific tagging, enhanced chemiluminescence, protease digestions, glycosidase digestions, and mass spectrometry were utilized to determine the presence and location of all post-translational modifications. From these data evidence for the presence of a single N-linked glycan was found in two of the four recombinant proteins, accounting for part of the observed molecular mass discrepancies. The remainder of the additional mass was determined to be a result of a series of smaller O-linked glycans.

INTRODUCTION

Pectate lyases catalyze the cleavage of pectate, the de-esterified form of pectin, through a beta-elimination reaction. Pectate lyase belongs to a class of enzymes known as pectin degrading enzymes and are produced by phytopathic fungi as part of the pathogenic attack orchestrated to degrade pectin within higher plant cell walls. Pectin primarily consists of homopolymeric partially methylated poly-$\alpha(1\rightarrow4)$-D-galacturonic acid residues, known as the "smooth" regions. There also exists "hairy" regions of alternating α-$(1\rightarrow2)$-L-rhamnosyl-α-$(1\rightarrow4)$-D-galacturonan which typically contains branches of L-arabinose and D-galactose. To a lesser degree pectins may also have rhamanogalacturonan sidechains containing D-xylose, D-apiose, 3-deoxy-D-manno-2-octulosonic acid, D-glucuronic acid, L-fucose, and 3-deoxy-D-lyxo-2-heptulosonic acid attached to poly-α-$(1\rightarrow4)$-D-galacturonic acid regions.[1, 2]

The ability of enzymes to degrade plant cell walls is of interest in a number of industrial areas. Primarily, the pectin degrading enzymes have found extensive use in the food industry where they aid in the release of color in juices and wines, clarification of juices, controlling the consistency of purees, among others.[1,3] This extensive use in industry has necessitated the development of easy, fast, and robust over-expression protocols to aid in the demand for large quantities of pectin and pectate degrading enzymes.[4] The development of over-expression systems often requires changes in a number of conditions such as host organism, vector, and growth media in order to optimize the production of the over-expressed protein. Variations in over-expression conditions have been shown to result in changes in not only the amount of protein produced but also the degree of glycosylation present on the recombinant protein as well as other protein properties.[5,6] As a result the product of any over-expression needs to be thoroughly characterized in order to determine if the recombinant protein is identical to its native form. In many instances the thorough characterization associated with the validation of over-expressed enzymes can also provide additional insight into the mechanisms associated with enzymolysis; for example the mode-of-action of PDEs toward pectin polysaccharides.[7]

The complete characterization of a protein encompasses far more than elucidation of the primary amino acid sequence. Information about the types and locations of all post-translational modifications must also be obtained. Of particular interest is glycosylation because it may affect the properties of the protein by alterations in protein stability, increased resistance to proteases, or changes in enzymatic activity/mode-of-action.[7,8,9] A number of the pectin degrading enzymes isolated from fungal sources have been shown to contain glycosylation.[7,10,11, 12] Since the role of the carbohydrate group is largely unknown it is important to discern that any variations in glycosylation between the native and recombinant forms does not lead to variations in enzyme properties associated with the over-expressed protein(s). N- and O-linked glycosylation are the two major types of glycosylation encountered in fungal proteins.

N-linked glycans are bound to the protein through a common triad of amino acids known as the N-linked consensus sequence. More specifically the glycan is attached through the nitrogen of an asparagine residue within the triad. N-linked glycans also share a common core structure of $Man_3GlcNac_2$ but may contain additional monosaccharides on the core structure (forming bisected carbohydrates for example). Also, N-linked carbohydrates are classified into one of three possible families; high mannose, complex, and hybrid structures. O-linked glyans have fewer common characterisitics (such as core structure) and lack the requirement of a consensus sequence for attachment to the protein backbone. These oligosaccharides are linked through the oxygen of a serine or threonine residue.

The study of glyoproteins is often associated with the need for high sensitivity and the ability to characterize structurally complex materials. In years past the amount of material needed to successfully characterize a glycoprotein was often in the milligram to gram range.[20] A common tool used today for the characterization of glycoproteins is mass spectrometry.[13-20] Both MALDI and electrospray based mass spectrometry have enjoyed success in the field of glycoprotein analysis. Additionally mass spectrometry, in combination with a variety of glycosidases (enzymes that cleave glycosidic bonds), has been used to characterize both carbohydrates attached to proteins/peptides as well as free carbohydrates. Several groups have reported the digestion of glycopeptides and glycoproteins on the surface of a MALDI probe utilizing a variety of endo- and exo-glycosidases.[21-24] The use of glycosidases aids in discerning the types of glycosidic linkages both to the peptide backbone and between sugar residues, the types of sugar residues present within the glycan, and the anomeric configuration at each residue.[25]

Unfortunately there exists a far more limited array of O-glycanases amenable to similar techniques that have become standards in the realm of N-linked glycan characterization. This makes the release and characterization of O-linked glycans considerably more challenging. As a result chemical cleavage mthods have become the only alternative for O-linked release and characterization.[20] Nonetheless chemical methods lack the specificity and ease of use associated with the glycosidases used to study N-linked structures.

Alkaline β-elimination is the most commonly used chemical method to release O-linked glycans. However, the use of strong alkaline systems have a number of problems associated with them. First, side reactions such as peeling can be a serious issue depending on the goal of the experiment being performed.[26] Also, the production of salts often associated with the neutralization of the alkaline releasing agent can be a serious issue depending on the analytical technique used to study the released glycan or the product protein.[20]

We present here the use of a variety of analytical techniques, including mass spectrometry, solid phase extraction, enhanced chemiluminescence, and anion exchange chromatography for the localization and structural characterization of the post translational modifications located on four isoforms of recombinant pectate lyase over-expressed in *Aspergillus Niger* and *Aspergillus Nidulans*. The four glycoproteins were first analyzed by MALDI-TOF to assess the protein molecular masses for comparison to the molecular mass based on the primary amino acid sequence (figure V-I). Analysis of the primary sequence for pectate lyase also revealed the presence of a single consensus sequence which could be occupied by an N-linked glycan. Next, all proteins were digested and analyzed by LC-MS and LC-MS/MS to obtain sequence information and possibly identify peptides containing carbohydrates. One method used to identify glycoproteins utilized LC-MS/MS where MS/MS data were analyzed for the presence of characteristic carbohydrate fragment ions generated as a result of in-cell collisional activation. Further characterization of attached glycans was achieved by MS/MS analysis

```
ATVSDAAFGY    ASLNGGTTGG   AGGTTTTVSS   YAAFTSAVSG
DDAKVVYVDG    TIKQTADQVK    IGSNTSIIGK     DANAILEGFG
VLVKEKENVI     IRNLGVSKVL    ADNGDAIGVQ   YSNNVWIDHC
DVSSDRDHDK   DYYDGLIDIT     HGSDYVTVSN   TFIHDHWKAS
LVGHSDSNED    EDSGHLTVTY   ANNYWYNVNS  RAPSFRFGTG
HVYNSYYLDV    SDGINTRDGA  QLLVESNQFV    DSKKALYSTD
DGYAVSNDND   FGDSENTAEE  GTLTSMPYDY    TLLGSANVKA
AVVGTAGQTL    TF
```

Figure V-I. Amino acid sequence of *Aspergillus Niger* pectate Lyase. N-linked consensus sequence is shown in blue.

of glycopeptides, site specific labeling followed by in-gel separation and detection, site-specific amine labeling followed by MS/MS characterization to determine location of site specific modifications, and finally anion exchange chromatography to determined the identities of attached glycans.

EXPERIMENTAL

MALDI-TOF analysis of recombinant isoforms

Data were accumulated on a Hewlett Packard G2030A (Palo Alto, CA) MALDI-TOF mass spectrometer. The instrument was operated in positive ion mode at a pressure of ~ 1 x 10^{-6} Torr, 28 kV accelerating voltage and 7 kV extractor voltage. Samples were mixed 1:1 with sinapinic acid (10 mg/ml; 70% aqueous acetonitrile (Sigma, St. Louis, MO, USA) containing 0.01% trifluroacetic acid (Sigma-Aldrich, St. Louis, MO, USA)) followed by vacuum drying. Samples were desorbed and ionized using UV radiation emitted from a nitrogen laser at an output wavelength of 337 nm.

Liquid chromatography-mass spectrometry

All experiments were carried out utilizing a Q-Tof II (Micromass, Manchester, UK) mass spectrometer operating in auto-switching mode. Data were collected in positive ion MCA mode. ‘Auto Switching’ was utilized to obtain product ion data for ions pre-selected in MS1. Instrument conditions were as follows - Cone Voltage (CV) = +41V, Capillary Voltage = +2800V, flowrate = 180 nL/min (defined by LC).

A Waters Capillary LC was coupled to the instrument for sample separation and clean-up prior to entry into the mass spectrometer. Peptide solutions were injected (5 µL) and separated on a 15 cm x 75 µM id C-18 column (LC Packings, San Francisco, CA, USA). LC conditions were as follows: a linear gradient was employed where %B (90 % acetonitrile, .1% formic acid) was ramped from 10% to 70 % over 220 minutes relative to mobile phase component A (.1 % formic Acid). Those ions with sufficient signal-to-noise were subjected to collision induced dissociation and the resulting MS/MS spectrum collected. Collison cell pressure was maintained at 15 psi with argon as the collision gas. Collison energies were optimized within the auto-switching setup based on the mass and charge of the precursor ion being collisionally activated.

Alkaline β-Elemination of recombinant proteins

50 µl of each recombinant protein was mixed with 150 µl of 0.1 M sodium hydroxide. Samples were incubated overnight at room temperature. Next, 150 µl of 0.1 M HCl was added to neutralize the sodium hydroxide. Prior to MALDI-MS analysis a small aliquot of the protein solution was extracted using a C-4 Ziptip (Millipore, Bellareca, MA, USA) according to the manufacture's instructions.

Trypsin digests of recombinant isoforms

50 µg of each recombinant protein was digested using sequencing grade trypsin (Promega, Madison, WI, USA), using a ratio of 1:50 (w/w) enzyme to substrate in 100 mM ammonium bicarbonate (Sigma, St. Louis, MO, USA), pH 7.8, for 18 hours at 37 ^{o}C with gentle shaking.

ZipTip$_{MC}$ extraction of trypsin digested recombinant proteins

Trypsin digested proteins were extracted utilizing a ZipTip$_{MC}$ (Millipore, Bellireca, MA, USA) pipette tip according to instructions provider by the manufacturer.

Lectin analysis

The recombinant proteins were screened against a series of lectins as a rapid method to determine information about the structures of carbohydrates that may be present on the proteins. The protocol included in the DIG Glycan Differentiation (Roche Diagnostics, Mannheim, Germany) kit was modified to an ELISA style experiment. A Immulon 2 microtiter plate (Dynatech, Chantilly,VA, USA) was coated (20 wells; 5 wells per recombinant pectate lyase isoform) with 100 µL of each of the four recombinant proteins and allowed to incubate at room temperature for 2 hours. Four control proteins, carboxypeptidase Y, transferrin, fetuin, and asialofetuin were also added to individual wells and allowed to incubate simultaneously with the recombinant proteins. Next, the microtiter plate was washed 5 times at 5 minute intervals with Tris buffered saline (TBS, 0.05 M Tris-HCl, 0.15 M NaCl, pH 7.5). The plate was then aspirated to remove wash solution. Blocking solution (supplied in DIG Glycan Differentiation Kit) was added (200 µl) to each well and the plate was allowed to incubate for 1 hour at room temperature. The plate was then washed 5X at three minute intervals with Tris-buffered saline. The plate was then washed with Buffer 1 (supplied with the kit: TBS containing 1 mM $MgCl_2$, 1 mM $MnCl_2$, 1 mM $CaCl_2$, pH 7.5). Lectin solutions (GNA, SNA, DSA – 10 µl of lectin + 10 ml Buffer 1; MAA - 50 µl lectin + 10 ml of Buffer 1; PNA – 100 µl lectin + 10 ml of Buffer 1) were then added to microtiter wells (100 µl) such that each recombinant isoform was screened with each of the five lectins present in the kit (one lectin per well, 5 lectins per protein). Lectins were allowed to incubate in the microtiter plate for 1 hour at room temperature. The plate was then washed 5X at 5 minute intervals with TBS. Next, 100 µl of anti-digoxigenin-AP was added to each well and allowed to incubate for 1 hour. Plate was then aspirated and washed 5X at 5 minute intervals with TBS. Wells were allowed to develop (15 to 30 minutes). Reaction was confirmed by a yellow color change (clear to yellow).

Phosphate specific enzyme linked immunosorbent assay (ELISA)

Phosphorylation content was assayed utilizing a modified ELISA. A microtiter plate (Nunc Immuno plate (Imullon 2), Nunc, San Diego, CA, USA) was coated (selected wells) with 100 µL of each of the four recombinant proteins (10 µg in tris-buffered saline((TBS), 0.05 M Tris-HCl, 0.15 M NaCl, pH 7.5). The plate was covered and incubated at room temperature for 2 hours then aspirated and washed 3X with 200 µL of wash solution (0.05% Tween-20 in PBS, 1 minute per wash). Next, 300 µL of blocking solution (1% BSA in TBS) was added to each well and the plate was covered and incubated for 1 hour at room temperature. The plate was then aspirated and washed 3X, at 1 min intervals, with wash buffer. Next, antibodies (mouse AP conjugated mouse anti-phosphoserine/threonine and AP conjugated mouse anti-phosphotyrosine (Zymed, San Francisco, CA, USA) diluted 1:1000 in TBS) were added (200 µL) and allowed to incubate for 1 hour. The plate was then aspirated and washed 5X with 300 µL of wash solution (0.05% Tween-20 in TBS, 3 minutes per wash). 100 µL paranitrophenylphosphine (PNPP), 1 mg/ml, (CalBiochem, La Jolla, CA, USA) was added to each well. Detection of yellow color change indicated a positive reaction with antibody.

Carbohydrate labeling for enhanced chemi-luminescence detection

10 µg of each recombinant protein was lyophilized to dryness followed by reconstitution in 100 mM sodium acetate (Sigma, St. Louis, MO, USA). 10 µl of 30 mM sodium metaperiodate (Amersham Biosciences, Piscataway, NJ, USA) was added and the solution was allowed to incubate for 20 minutes (note – addition of metaperiodate and incubation must be done in a dark area). Next, 10 µl 20 mM sodium meta-bisulphite was added and the solution was incubated for 5 minutes. Finally, 5 µl of biotin hydrazide was added and the solution was incubated for 60 minutes.

1D gel electrophoresis

1 µl of protein solution was mixed with 9 µl of Laemmli buffer (Biorad, Hercules, CA, USA). 10 µl of protein/Laemmli solution was loaded into a 15% tris-HCl (Biorad, Hercules, CA, USA) gel for 1 dimensional gel electrophoresis analysis. Prior to loading of samples onto gels, the samples were passed over a .22 µM filter (Millipore, Bradford, PA, USA). Gels were run utilizing a Biorad Protean II 7 cm system (Biorad, Hercules, CA, USA) at 300 Volts/30 mAmps (constant) until bromophenol blue band was within 0.5 cm of the gel bottom.

Western transfer/ECL analysis

PVDF membranes (Biorad, Hercules, CA, USA) were immersed in 100% HPLC grade methanol (Aldrich, St. Louis, MO, USA) until the membrane became translucent. The membrane was then transferred to trans-blotting buffer (25 mM Tris at pH 8.3 (Biorad, Hercules, CA, USA), 192 mM glycine (Biorad, Hercules, CA, USA), 20% methanol) just prior to use where it was allowed to incubate for 3 minutes. Care was taken to insure that membrane was not allowed to dry at any point prior to completion of experiment(s). Orthogonal electroelution was carried out utilizing a Protean II Western Kit (Biorad, Bradford, PA, USA). Proteins were electro-eluted (trans-migration) from the 1d gels onto the PVDF membrane(s) at 45 Volts/200mAmps for 1 hour. After trans-migration the membrane was removed and rinsed 3X at 5 minute intervals with distilled water. The membrane was then incubated in ECL (enhanced chemi-luminescence) glycoprotein blocking agent (Amersham Biosciences, Piscataway, NJ, USA), overnight at 2 to 8 ^{o}C. Next, the membrane was briefly rinsed 3X at 10 minute intervals in phosphate buffered saline (PBS). The membrane was then incubated in Streptavidin-HRP (horse radish peroxidase) solution (supplied in ECL Glycoprotein Detection Module (Amersham Biosciences, Piscataway, NJ, USA) for 30 minutes followed by three rinses (10 minute intervals) with PBS. Following the protocol supplied in the ECL Reagents Kit (Amersham Biosciences, Piscataway, NJ, USA) the two supplied ECL reagents were mixed 1:1. Approximately 2 ml of this mixture was distributed evenly over the PVDF membrane and allowed to incubate for 1 minute. The solution was then poured from the surface of the PVDF membrane and the membrane was sealed inside a small plastic bag. Note – the remainder of the procedures outlined in this protocol must be carried out in a darkroom due to the extreme sensitivity of autoradiography paper to light. Once radiography paper has been 'fixed' it is safe to expose the film to light sources. The sealed PVDF membrane was placed in an autoradiography cassette. A sheet of Hyperfilm ECL (Amersham Biosciences, Piscataway, NJ, USA) was place directly above the sealed PVDF membrane and the cassette was closed. The film was exposed to the PVDF membrane for a period of 15 to 60 seconds. Next the autoradiography film was removed and placed in developing solution (1:2 dilution of Kodak GBX Developer (Kodak, Rochester, NY, USA)) for ~ 1 to 2 minutes. After the appearance of bands the autoradiography sheet was transferred to a fixing solution (1:5 dilution of Kodak GBX fixer (Kodak, Rochester, NY, USA) for 3 minutes followed by washing with water.

Ethylamine Beta-elimination/labeling

50 μg of each recombinant protein/peptide digest was lyophilized to dryness and reconstituted in 100 μl of 70% ethylamine (Aldrich, St. Louis, MO,USA). The solution was allowed to incubate for 18 hours at 50 ^{o}C. Amine-labeled peptides to be analyzed by LC-MS were injected directly with no sample cleanup. Amine labeled proteins to be analyzed by MALDI-TOF were dried in a vacuum centrifuge and reconstituted in HPLC grade water prior to preparation for MALDI-MS analysis.

Anion Exchange Chromography – monosaccharide composition analysis

Neutral and amino sugars were released from portions (100 μg) of each recombinant protein with 2 N trifluroacetic acid at 100 ^{o}C for 4 hours. After hydrolysis samples were dried using a vacuum centrifuge followed by re-suspension in 200 μl of

HPLC water. Alternatively, proteins and standards were treated with 1M sodium borohydride and 0.05 M sodium hydroxide overnight at 45 °C. This resulted in release of glycans with no hydrolysis of the individual sugar residues. Calibration standards (fucose, glucose, galactose, N-acetyl glucosamine, and N-acetyl galactosamine) were hydrolyzed in parallel with the samples. Samples (10 µl) and standards (0.5, 1.0, and 2.0 nmol) were analyzed using a Dionex DX500 system (SunnyVale, CA, USA) equipped with a GP40 gradient pump, ED40 electrochemical detector, Thermoseparations (Mountain View, CA, USA) AS3500 autosampler, and a Dionex Carbopac PA10 (4 x 250 mm) analytical column with an Amino Trap and Borate trap column. The gradient program used eluents A (water) and B (100 mM sodium hydroxide). For trifluroacetic acid hydrolysis experiments the following gradient was used: from 0 to 25 minutes mobile phase A was held at 92%, from 25 to 27 minutes mobile phase B was ramped linearly from 8 to 100% and held at 100% from 27 to 37 minutes. Mobile phase B was then ramped linearly back to 8% from 37 to 39 minutes and held at 8% from 39 to 50 minutes. Flowrate was maintained at 0.250 ml/min throughout the experiment. For the sodium borohydride/sodium hydroxide hydrolysis experiments the following gradient was used: Mobile phase A was maintained at 90% from 0 to 4.80 minutes, from 4.80 to 15.90 minutes mobile phase B was ramped linearly from 10 to 30 % and held at 30 % until 26 minutes, from 26 to 27.10 minutes mobile phase B was ramped linearly from 30 to 10 % and held at 10 % until 40 minutes. Flowrate was maintained at 0.400 ml/min throughout the experiment.

RESULTS AND DISCUSSION

Pectate lyase overexpressed homologously in *Apergillus Niger* produced two isoforms (referred to as ply I and ply II). Pectate lyase heterologously overexpressed in *Aspergillus Nidullans* produced two isoforms (referred to as pal I and pal II). Due to the extensive similarities between ply I and pal I the data presented for these two recombinant proteins will often be collectively referred to as isoform I. Similarly, data for ply II and pal II will often be collectively referred to as isoform II. The four recombinant isoforms (ply I, ply II, pal I, and pal II), see figure V-II, were studied to determine if differences existed with regards to the types of post-translational modifications present. The MALDI mass spectra taken for the four recombinant isoforms revealed that all four isoforms showed masses in excess of the mass based on the amino acid sequence (31115.6 Da), see figure IV-III. Both pal II and ply II showed masses ~ 1200 Da greater than the calculated mass whereas the isoform I recombinant proteins (ply I and pal I) showed masses ~ 2400 Da greater than the calculated mass.

Post-translational modifications often account for mass discrepancies between calculated and measured molecular masses. Larges mass discrepancies are often associated with glycosylation. Other types of post-translational modifications such as acetylation and phosphorylation result in mass increases smaller than 100 Da per site. In order to achieve the mass discrepancies observed between the calculated mass and the measured masses of the four recombinant proteins, multiple sites would have to modified if glycosylation were not the primary posttranslational modification. N- and O-linked glycosylation represent the two types of glycosylation found on glycoproteins. N-linked glycosylation requires the presence of an N-linked consensus sequence. This is a three

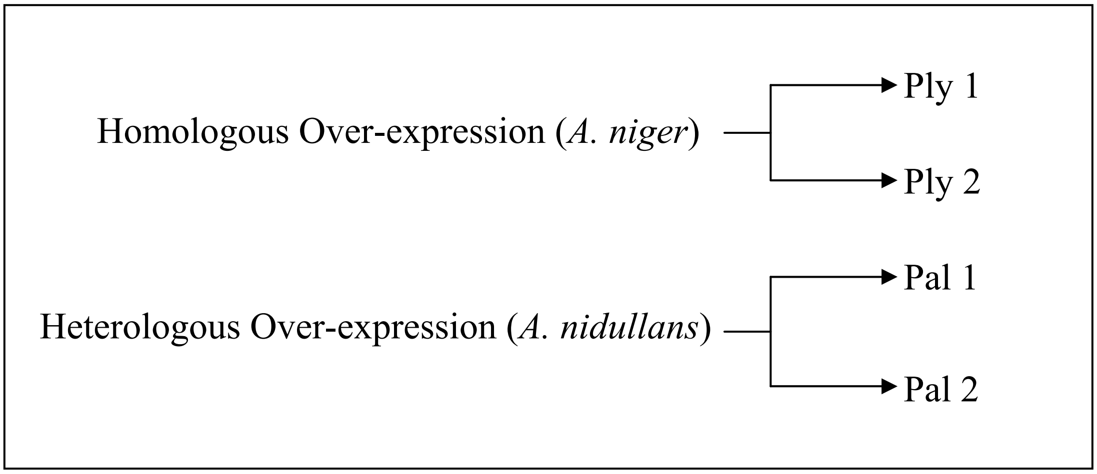
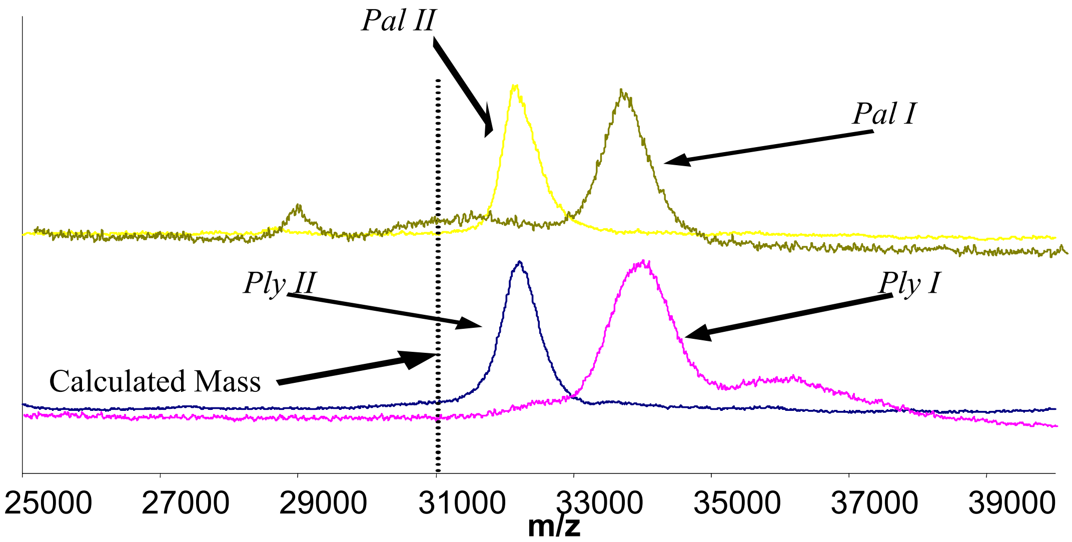


Figure IV-II. Homologous and heterologous overexpression of the *Aspergillus Niger* pectate lyase.



Calculated Mass (Primary Structure only) – 31115.6 Da

Measured Masses

A. Nidullans (Pal I) - 33,500 Da

A. Nidullans (Pal II) - 32,300 Da

A. Niger (Ply I) - 33,500 Da

A. Niger (Ply II) – 32,300 Da

Figure IV-III. MALDI-TOF Data demonstrating the greater mass observed with all recombinant proteins when compared to the mass based on the amino acid sequence.

amino acid triad composed of Asn-X-Ser/Thr, where X is any amino acid residue besides proline[11]. O-linked glycosylation can occur at any serine or threonine. Inspection of the amino acid sequence for pectate lyase, figure IV-I, shows the presence of one such N-linked consensus sequence indicating a potential site for N-linked glycan attachment.

As a method to screen for a variety of glycosylations the four recombinant proteins were tested for their ability to bind lectins. Lectins are proteins that bind carbohydrates with varying degrees of specificity. Thus the binding of a lectin to a particular protein indicates the presence of a carbohydrate (glycosylation) with structural features specific for the binding lectin. A series of lectins were utilized to probe for glycosylation within the four recombinant proteins. Of the lectins tested only a single lectin, *Galanthus nivalis* agglutinin (GNA), 'reacted' with two of the four recombinant proteins, ply 1 and pal 1. GNA recognizes terminal mannose and is suitable for the identification of high-mannose N-linked glycans. Thus, two of the four recombinant proteins, both isoform I proteins, are likely to contain a high mannose carbohydrate attached at the N-linked consensus sequence. Since neither of the isoform II recombinant proteins (ply II and pal II) reacted with any of the lectins, including GNA, the mass difference observed between the isoform I and isoform II recombinant proteins might be explained by the presence of an N-linked high mannose structure present on the isoform I proteins but not the isoform II proteins.

Liquid chromatography-mass spectrometry (LC-MS) was used to further characterize any existing N-linked glycans which might be present on the four recombinant proteins. Prior to LC-MS analysis the four recombinant proteins were digested with trypsin. Trypsin was chosen because it would isolate the N-linked consensus sequence to within a small peptide. The resulting total ion chromatograms for two of the four recombinant proteins (generated from summing all ion signals eluting from the LC as a function of time) are presented in figure IV-IV. By summing all the ions eluting over the mass window from ~ 80 to 180 minutes a mass spectrum was generated containing all ions detected over that 100 minute period, see figure IV-V. Analysis of the resulting mass spectra showed a large number of similar ions as would be expected should the two proteins share the same primary amino acid sequence. However, for the ply II sample, figure IV-V(a), the presence of a peptide at m/z 989.6 Da indicated that the consensus sequence was not occupied by a glycan since a mass of 989.6 Da corresponds to the molecular mass of the peptide based on its amino acid sequence alone. Analysis of figure IV-V(b), the mass spectrum for ply I, showed that no peptide was detected. This could indicate that the peptide contains an N-linked glycan since the presence of a carbohydrate would change the mass at which the peptide is detected.

As a rapid way to screen the large amount of data presented in the TIC, mass extractions were used to identify elution times for components that showed fragment ions characteristic of glycopeptides. For example m/z 204 Da, corresponding to the HexNAc oxonium ion, was extracted against the four TICs generated from the LC-MS experiments. The results indicated an intense elution of ions at ~ 81 minutes in ply II and pal II samples but not with either of the isoform I samples. By summing all ions eluting within the window of time (~75 to 85 minutes) that encompassed the differing peaks between the two TICs shown, a mass spectrum was generated containing all the ions detected by the mass spectrometer for the time window selected (75 to 85 minutes). This data is presented in figure IV-VI. Analysis of the resulting summed mass spectrum for

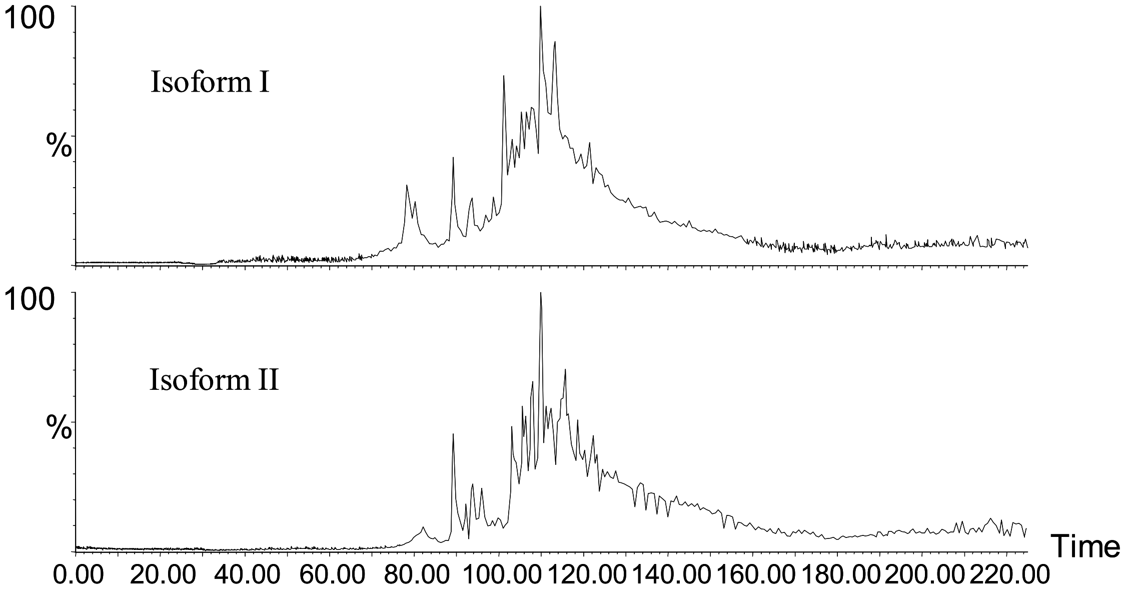
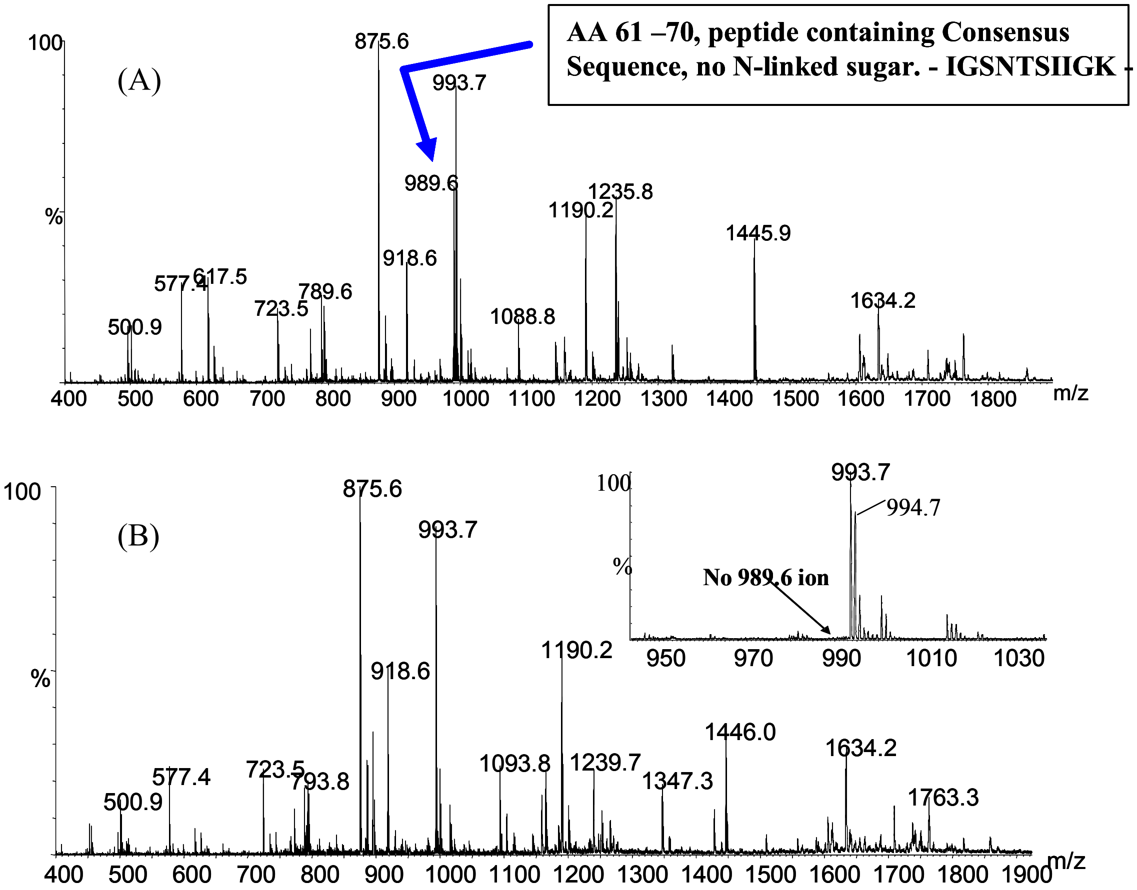


Figure IV-IV. LC-MS data for ply I and ply II. Note the increased intensity of the peak eluting at ~ 81 minutes.


AA 61 –70, peptide containing Consensus Sequence, no N-linked sugar. - IGSNTSIIGK -

No 989.6 ion

Figure V-V. Mass spectra showing the (a) presence of the peptide contain the consensus sequence in ply II and (b) the lack of the same peptide in ply I.

the isoform I recombinant proteins showed a series of peaks from m/z ~900 Da to 1700 Da. This same series is absent in the ply II and pal II (isoform II) recombinant proteins, the proteins where the N-linked consensus sequence is present but no carbohydrate was detected.

Analysis of the observed mass distribution detected in the isoform I proteins showed a peak spacing of 81 Da (2+). This corresponded to the residue mass of a hexose. Glycosylation is not a ‘fixed mass’ post translational modification. Rather, “glycoproteins usually exist as a complex mixture of glycosylated variants (glycoforms)”[27]. These glycoforms gave rise to mass distributions where sequential peaks were separated by the mass of the repeat residue(s). Recalling the results of the lectin experiments, both isoform I proteins screened positive against *Galanthus Nivalis* agglutinin, a lectin specific for high mannose structures. In agreement with those results was the detection of the mass distribution within the isoform I proteins (162 Da = residue mass of a hexose; mannose is a hexose). Thus, heterogeneity resulting from N-linked high mannose (common in yeast and fungus) glycoforms would present itself in the form of a mass distribution with 162 Da differences between sequential peaks. To better discern the structures of the ions detected in the mass distribution, several of the peaks were characterized using MS/MS. The results of one such experiment are presented in figure IV-VII. Examination of the selected ion showed a 0.5 Da spacing within the isotope distribution indicating a +2 charge state (z). Thus the mass of the glycoform being analyzed was 2692.6 Da (z*(measured m/z) – z). Since the mass of the free peptide was 989.6 Da the approximate mass of the attached carbohydrate was ~ 1703 Da. MS/MS analysis of the 1347.3 Da glycoform, figure IV-VII(B), yielded a number of

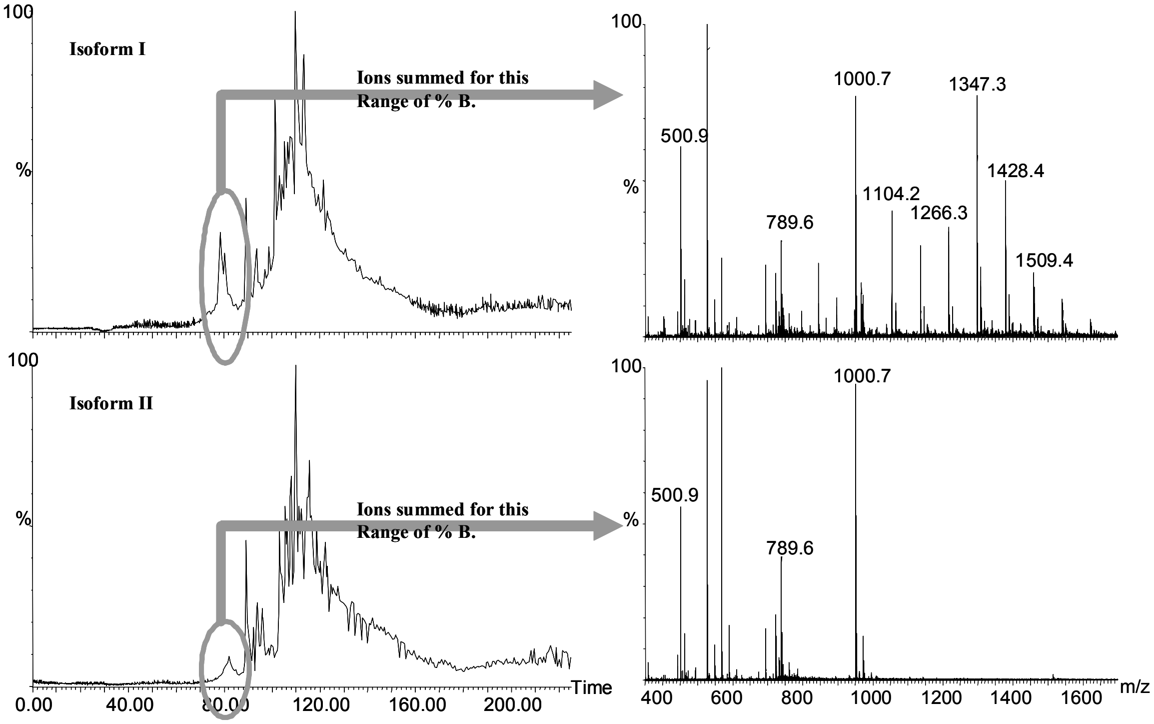


Figure IV-VI. Detection of glycoforms in Isoform I samples. Lack of glycoforms in Isoform II samples.

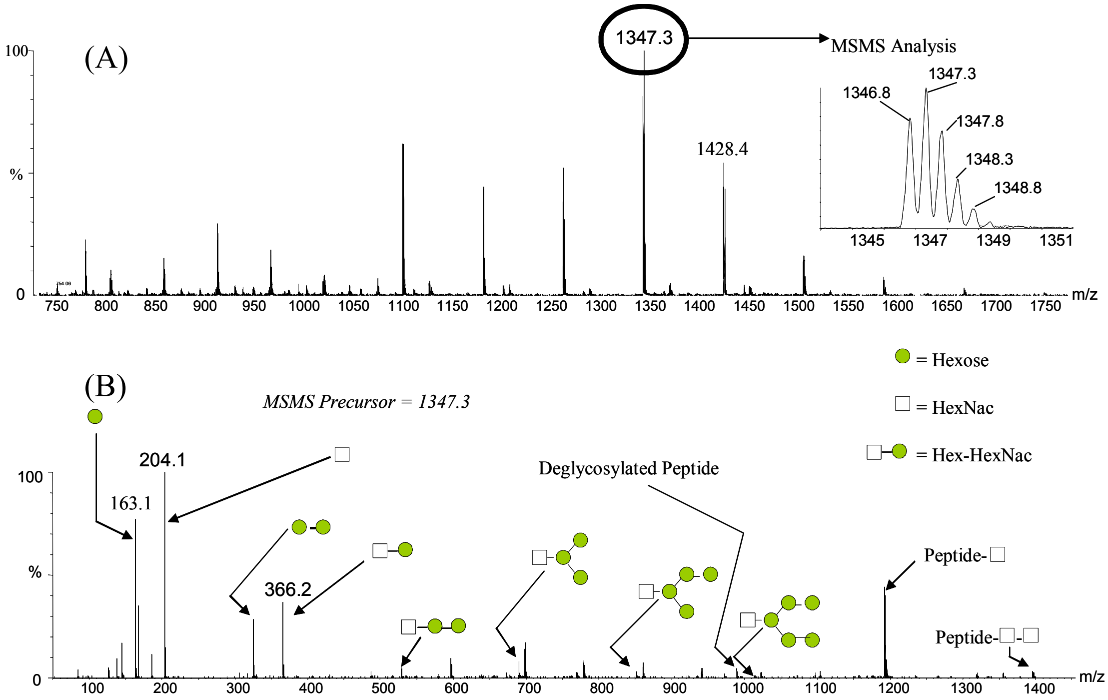


Figure IV-VII. Structural elucidation of selected glycoforms from ply I recombinant protein. Precursor m/z = 1347.3. (A) MS spectrum. (B) MS/MS spectrum.

fragment ions. Both the free hexose residue and the HexNAc oxonium ions were readily detected. The presence of the HexNAc oxonium (204 Da) ion supports that the glycan is attached through a nitrogen linkage. The presence of the hexose ion (163 Da) supports N-linked glycosylation as well since the core structure of N-linked glycans ($Man_3GlcNac_2$) contains mannose, a hexose sugar. Further support for the N-linked glycan is found at m/z 1193.7 and 1397.8. These two ions correspond to the peptide plus one and two HexNAc residues respectively.

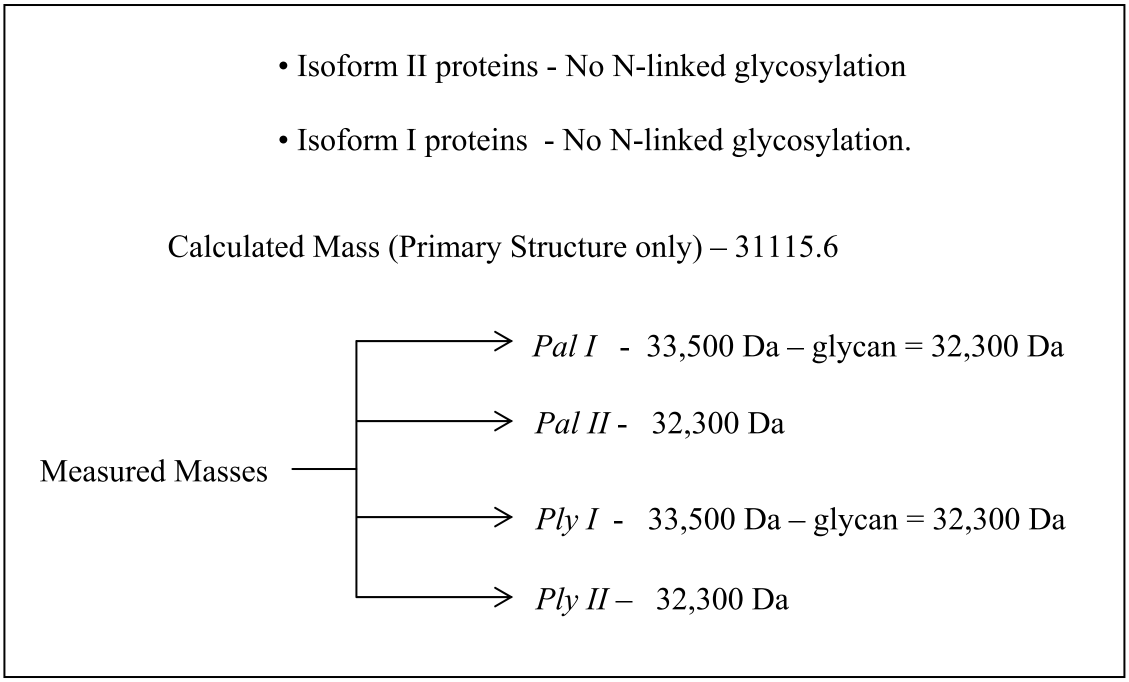
In addition to the N-linked core elements discussed above additional ions were observed representing the exterior makeup of the sugar (i.e. those residues attached to the antenna of the core structure). These are shown between the mass range of ~ 540 and 1000 Da in figure IV-VII(B). Although some stereochemistry is indicated in the structures shown in figure IV-VII(B) it is not possible to deduce the full structure from a mass spectrum alone. For example, the structure shown at ~1014 could easily have one mannose transferred from one antenna to the other. While structurally different they would yield the same m/z thus being indistinguishable from a mass spectrometry standpoint. Also, using the term mannose indicates a level of structural information not obtainable from the experiments presented here. Mannose is a hexose as is galactose and glucose and all 16 hexoses are stereoisomers and are thus indistinguishable using mass spectrometry. However, as stated earlier, knowledge of N-linked structural features allows for the assignment of structures to fragment ions found in MS/MS data. This information coupled with the results of the lectin screen allow for the assigning of structures that would not be possible using mass spectrometry alone. Although neither the GlcNAc-GlcNAc-Man nor the GlcNAc-GlcNAc structures are detected, an ion was detected at the mass of the deglycosylated peptide mass (989.6 Da) further strengthening previous results indicating that the distribution of glycovariants detected derive from an N-linked structure attached to the single identified consensus sequence.

Additional glycoforms were subjected to MS/MS analysis and found to yield data very similar to that presented. Most notably was the presence of the same fundamental peaks corresponding to core residues as well as terminal residues associated with structure outside of the core region. In most situations the same ions shown in figure V-VII(B) were readily detected in the other glycoforms. Higher molecular mass glycoforms (molecular mass in excess of 1600 Da) yielded the least favorable data. This could be due to two reasons: 1) increased size results in more efficient distribution of energy, imparted through collisions within the gas cell, leading to less efficient fragmentation. 2) The higher molecular mass ions had considerably weaker ion intensities making MS/MS analysis more difficult.

To further strengthen the argument in favor of an N-linked structure, ply I and pal I were subjected to glycosidase treatment with PNGase F. PNGase F is an exo-glycosidase that completely removes the N-linked sugar from a protein/peptide. MALDI-TOF analysis of the PNGase F treated proteins showed an ~1200 Da loss in mass. These PNGase F treated proteins were then digested with trypsin and analyzed by LC-MS. Prior to PNGase F treatment there was no detectable ion at m/z 989.6 (peptide containing the N-linked consensus sequence). After PNGase treatment an ion at 989.6 Da was readily detectable. These results also indicate that the N-linked consensus sequence for ply I and pal I do contain an N-linked structure.

Thus far the isoform I proteins have been shown to contain an N-linked glycan with an average molecular mass of ~ 1200 to 1500 Da. Taking into account the presence of the N-linked glycan there still exists an ~ 1200 Da mass discrepancy with the isoform I proteins and the calculated mass (see figure IV-VIII). Also, since there exists no N-linked glycosylation within the isoform II recombinant proteins the 1200 Da mass difference detected by MALDI-TOF is, as of yet, un-explained. This information is summarized in figure IV-VIII. In the absence of N-linked glycosylation two other types of post-translational modifications could account for the un-explained 1200 Da mass. Phosphorylation is a PTM where phosphates are attached to serines or threonines. Phosphorylation accounts for an 80 Da increase per phosphate. Alternatively O-glycosylation could account for all or part of the observed mass discrepancy. As with phosphorylation, O-linked glycans can attach to, essentially, any serine or threonine.

Two experiments were utilized to test for the presence of phosphates as a source for some, or all, of the 1200 Da observed mass difference. Immobilized metal affinity technology has been utilized for some time to selectively remove phosphopeptides from various proteolytic digests. Extraction of tryptic digests with IMAC (immobilized metal affinity chromatography) bedded ziptips resulted in the detection of a few peptides as shown in figure V-IX. However, comparison of the masses detected with masses for tryptic fragments indicated that ions detected contained no phosphates. For example, as is illustrated in figure V-IX, an ion at m/z 1189.9 Da was detected. Since this is a doubly charged ion the molecular mass of the corresponding peptide would be 2377.8 Da which corresponds to the molecular mass of a tryptic peptide spanning amino acids 197 to 217. This does not necessarily mean that the detected peptide was not phosphorylated



- Isoform II proteins - No N-linked glycosylation
- Isoform I proteins - No N-linked glycosylation.

Calculated Mass (Primary Structure only) – 31115.6

Measured Masses

Pal I - 33,500 Da – glycan = 32,300 Da

Pal II - 32,300 Da

Ply I - 33,500 Da – glycan = 32,300 Da

Ply II – 32,300 Da

Figure IV-VIII. Mass discrepancies observed after N-linked glycosylation was accounted for.

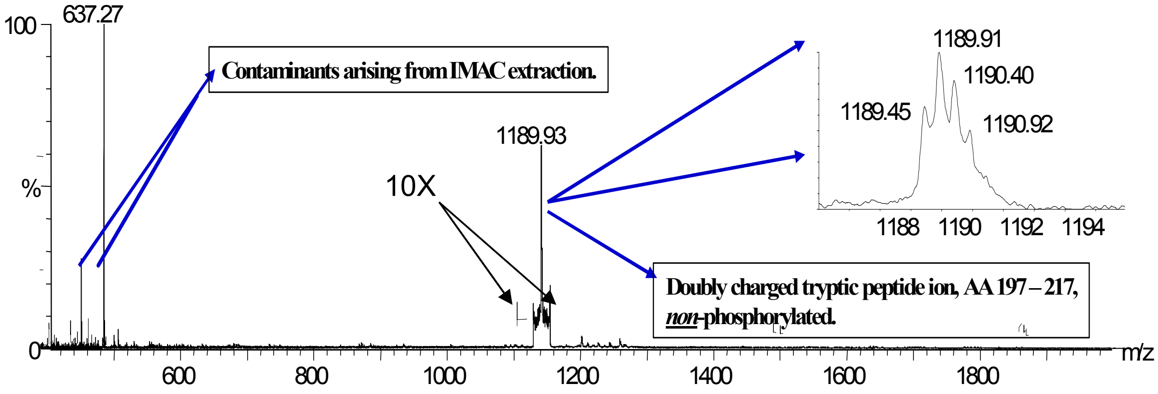


Figure IV-IX. MALDI Mass spectrum of IMAC extracted components from ply I tryptic digest.

prior to the MS experiment. Phosphopeptides are notoriously fragile and tend to undergo loss of phosphates even under the mildest of MS experiments. Thus, additional experiments needed to be completed in order to determine if, in fact, the protein was phosphorylated.

As an alternative to mass spectrometry based detection of phosphopeptides, an enzyme linked immunosorbent assay (ELISA) was developed to assay the four recombinant proteins for phosphorylation. Two phosphoprotein specific antibodies were used in the ELISA of the four proteins. In all cases no evidence was found indicating that the proteins contained phosphorylation. Given these results it is likely that the IMAC extracted peptides were not phosphorylated prior to MS analysis. Rather, IMAC resins, while highly specific for phosphopeptides, are also slightly specific for peptides containg acidic amino acids as was the case for all peptides detected in the IMAC extractions. The relative weakness of the peptides detected also supports the likelihood that those peptides extracted were acidic rather than phosphorylated (due to the lessened selectivity of the IMAC resin toward acidic peptides as opposed to phosphopeptides).

O-glycosylation was the next post-translational modification studied as a possible source of the higher-than-expected molecular masses presented by all four recombinant proteins. Unlike N-glycosylation, O-glycosylation has no requirement for a consensus sequence at the site of glycan attachment. Rather, there need be only a serine or threonine in order for an O-linked glycan to be attached. Analysis of the amino acid sequence for pectate lyase showed 29 serines and 26 threonines resulting in 55 potential sites for O-glycosylation. This results in nearly all peptides, generated from trypsin digestion, as having the capacity to contain O-linked glycans.

Rather than look for specific glycopeptides containing O-linked glycans, initially, β-elimination was used to chemically cleave any O-linked glycans attached to the intact recombinant proteins. A unique observation associated with the four recombinant proteins was the time required to release the O-linked glycans under β-elimination conditions. Control glycoproteins typically showed O-linked glycan release after only a few hours under β-elimination conditions. In the case of the recombinant pectate lyases 18 hour incubations resulted in detectable release of O-linked glycans but the reaction had not gone to completion as is demonstrated in figure IV-X. Nonetheless a detectable shift in the mass of the proteins was obvious in all four cases. In the case of the isoform I proteins (those containing the N-linked glycan) there was a detectable 600 Da loss after β-elimination. However, with the isoform II proteins a 1200 Da loss was detected after β-elimination. As a test to determine if the N-linked glycan was possibly protecting some of the O-linked glycans against β-elimination, in the isoform I proteins, the PNGase F released ply I protein was subjected to β-elimination. As with the other isoform I β-elimination results there was a detectable 600 Da shift. This could indicate that the isoform II proteins are differentially O-glycosylated compared to the isoform I proteins as was the case with N-linked glycosylation between the isoform I and II recombinant proteins.

Extraction of LC-MS data against other common carbohydrate fragment ions yielded no additional peptides, above those found for the N-linked glycopeptide, likely to contain carbohydrate groups. Alternatively, the individual LC-MS data files were scanned manually for the presence of peptides whose masses were an incremental value greater than the mass of a peptide not detected in the LC-MS experiment. For example,

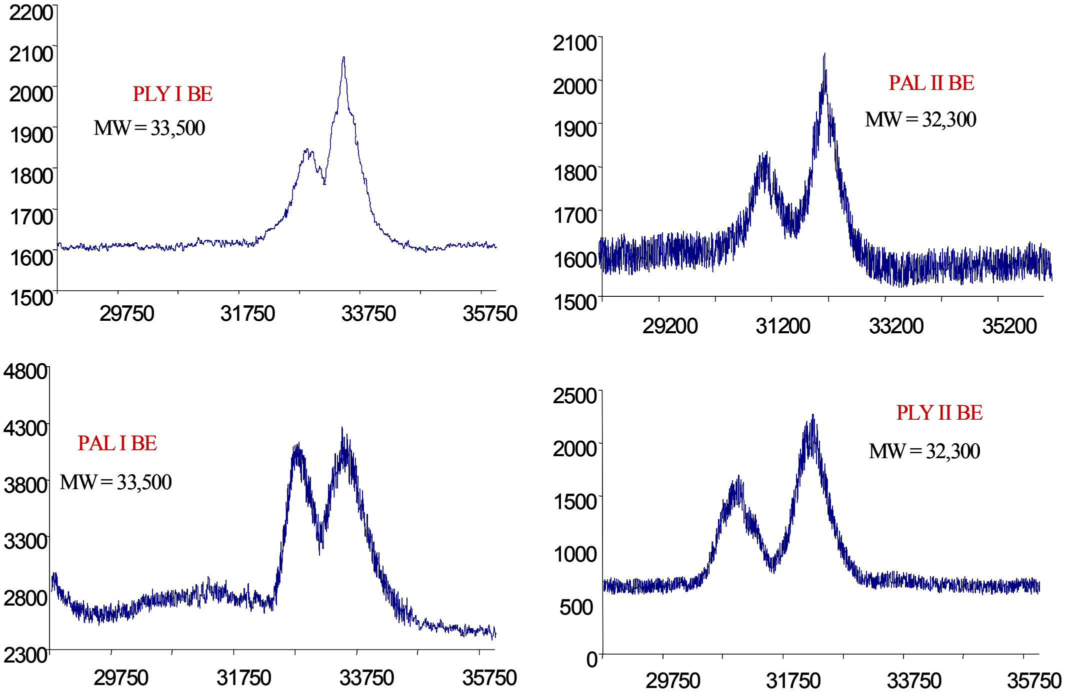


Figure IV-X. MALDI-TOF data demonstrating the mass loss associated with β-elimination of O-linked glycans from the recombinant pectate lyase proteins.

mannose is a common initiating sugar in O-linked carbohydrates attached to proteins. A peptide with molecular mass 1200 might not correspond to an expected tryptic peptide based on the m/z at which the ion is detected. This could be due to the presence of a post-translational modification inflating the measured m/z. If the peptide does contain a single mannose residue the expected molecular mass of the tryptic peptide is '1200 – 162 Da' or 1038 Da. Applying this to the LC-MS data for the four recombinant proteins showed no evidence supporting O-glycosylation.

Alternatives to mass spectrometry were pursued to better assess the presence of O-linked glycosylation since lack of detectable O-glycopeptides does not necessarily mean O-glycosylation was not present. One such alternative utilized periodate chemistry to selectively target all carbohydrates present on the four recombinant proteins. Basically the vicinal diol group of a sugar residue is oxidized to a di-aldehyde and then selectively labeled using biotin hydrazide. Once the recombinant proteins had been selectively biotinylated at the carbohydrate groups, they were separated from the residual biotinylation reagents by 1D gel electrophoresis, transferred to a PVDF membrane, then immunochemically detected utilizing horseradish peroxidase (HRP) conjugated streptavidin in the presence of luminol. HRP catalyzes the breakdown of cyclic diacylhydrazides such as luminol. Immediately following oxidation of luminol the molecule is left in an excited state. Decay from the excited state to ground state results in the emission of light which is detected by a sheet of autoradiography film.

The results of these experiments are shown in figure IV-XI. Five of the lanes contained non-biotin labeled proteins. For those lanes containing non-biotin labeled proteins no periodate oxidation was done, and the biotin hydrazide was added prior to 1D gel separation. This was done to insure that unreacted biotin hydrazide was being removed from the samples via gel electrophoresis. The presence of unreacted biotin hydrazide on the membranes would result in false positives. The remainder of the lanes contained proteins that were subjected to periodate oxidation and subsequent biotin labeling. The first lane contained transferrin, a known glycoprotein, the positive control. In all cases involving the periodate/biotin reacted recombinant proteins, detected bands indicated that all proteins were in fact glycosylated. Based on data presented earlier, the ply I and pal I proteins were already known to be glycosylated and acted as additional positive controls. Likewise, ply II and pal II proteins were known to *not* contain an N-linked glycan, thus any detection using the periodate-biotin system would have to be due to the presence of O-glycans. As shown in figure IV-XI both recombinant proteins are detected indicating that they do in fact contain O-linked glycans. Again, to test the feasibility that the isoform I proteins were also O-glycosylated the PNGase released forms of ply I and pal I were subjected to periodate oxidation/biotin labeling and re-analyzed using the ECL detection system. In both cases the proteins were detected indicating that O-linked glycans are present. Thus, these experiments further strengthen the results of the β-elimination studies indicating O-glycosylation as the reason for part of the mass inflation observed with the isoform I recombinant proteins and possibly all the mass inflation associated with the isoform II recombinant proteins. However, neither the β-elimination or the periodate/biotin experiments provided information as to the types or locations of the O-glycans present.

To locate the O-glycans, a modified β-elimination was performed using ethylamine as the base. Ethylamine acts to eliminate the glycan then adds back across the

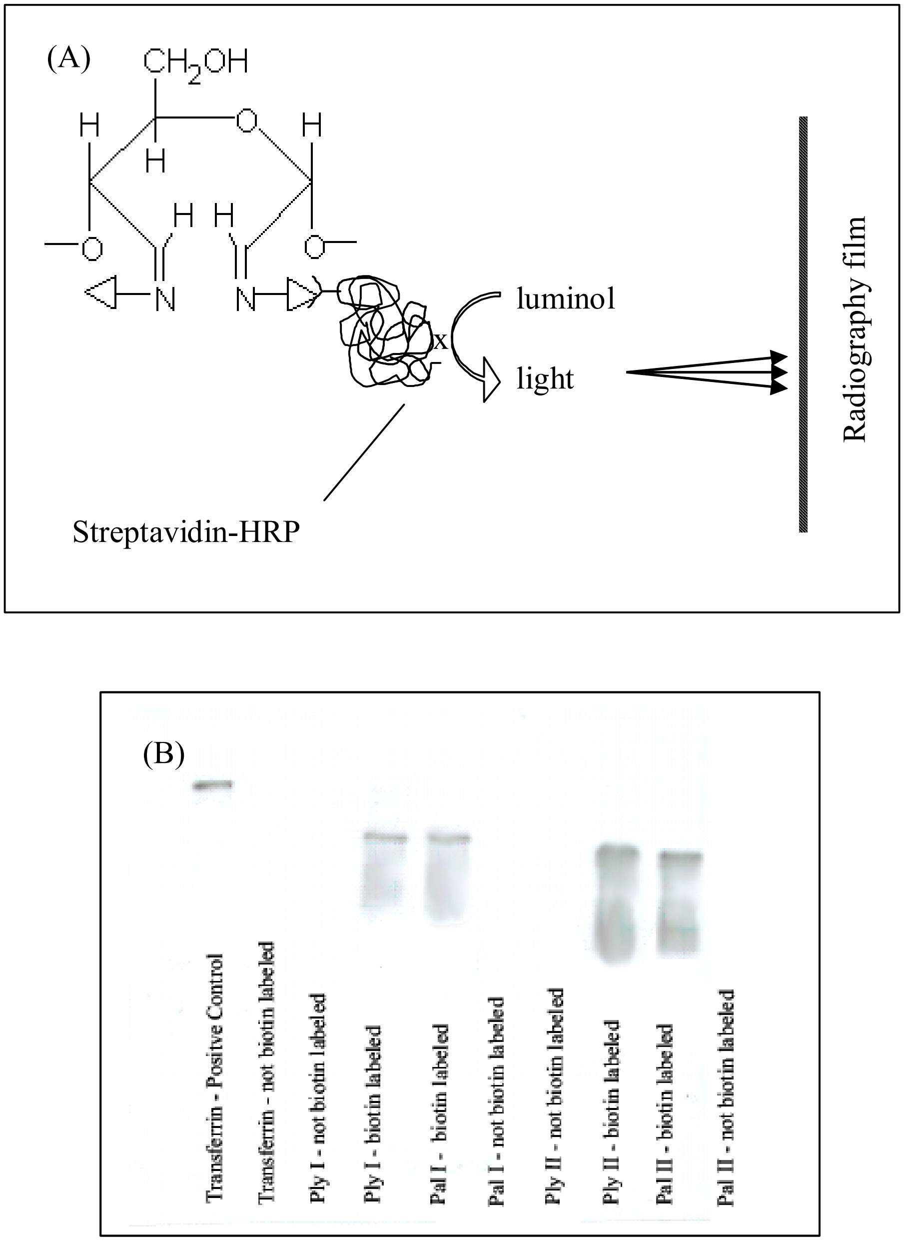


Figure IV-XI. (A) Illustration showing the ECL detection scheme. (B) Image of the autoradiography sheet resulting from the ECL glycoprotein analysis of the four recombinant proteins.

unsaturated bond, tagging the site where the glycan was eliminated. A subtle advantage in using ethylamine is derived from the secondary amine that results at the site where the glycan was attached. Secondary amines are good bases (i.e. they are good proton acceptors) which aids in the ionization of the labeled peptide, thus increasing the sensitivity of mass spectrometry toward these labeled peptides. The results of the ethylamine labeling of the ply I tryptic digest are shown in figure IV-XII. Comparison of figure IV-XII(B) with IV-XII(C) showed that prior to ethylamine labeling only a few weak peaks were detected (802.5 Da, 811.6 Da, etc.). After ethylamine labeling a series of 3 new peaks appeared (802.6 Da, 811.9 Da, 820.5 Da). The isotopic distributions of these peaks indicated they were triply charged thus the actual mass separation was 27 Da, the mass of an ethylamine label. The mass of the peptide after subtraction of the ethylamine label(s) was ~ 2378 Da. This corresponded to a tryptic peptide (AA 197 to 217) containing two serines and two threonines (4 possible sites for O-glycosylation with 3 sites occupied). Likewise similar results were obtained for a peptide at 889.3 Da containing two ethylamine labels. From these data five sites were identified as potentially containing O-linked glycans.

In both situations the corresponding non-labeled peptides were detected (weakly) in the LC-MS experiments involving only the tryptic digested proteins. Since most of the peptides from a tryptic digest of pectate lyase contained at least one serine or threonine, and only a few peptides were detected containing ethylamine labels, the likelihood of non-specific labeling of free serines or threonines seems minimal. The possibility exists that those peptides containing O-linked sugars are undergoing in-source de-glycosylation resulting in the detection of apparently non-glycosylated peptides.

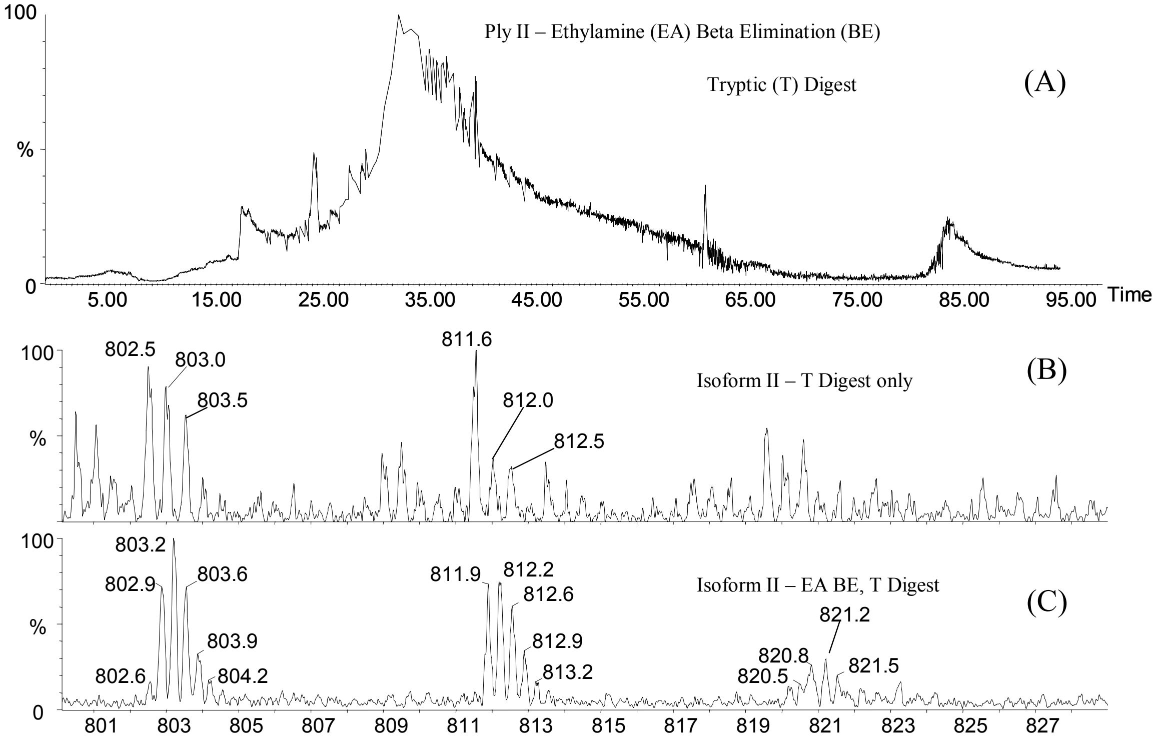


Figure IV-XII. (A) Total ion chromatogram from the ethylamine labeled tryptic digest of ply II. (B) Selected mass region from the tryptic digest of ply I (no ethylamine labeling). (C) Selected mass region showing the detection of ethylamine labeled peptides in the ply I ethylamine reacted tryptic digest.

While ethylamine labeling can provide information about sites occupied by O-glycans it does not provide structural information about the glycan. Anion exchange chromatography was utilized to obtain information about the O-linked structures. Prior to anion exchange chromatography potential glycans were hydrolyzed from the proteins utilizing two methods. Hydrolysis of sugars utilizing trifluoracetic acid results in the release of carbohydrates from the protein and subsequent hydrolysis of all glycosidic linkages. Thus any carbohydrates that existed as polysaccharides are converted to monosaccharides. Alternatively, release of sugars under sodium borohydride/sodium hydroxide results in no additional breakdown of the carbohydrates, i.e polysaccharides remain whole. However, with the use of the sodium borohydride/sodium hydroxide hydrolysis, terminal and free monosaccharides are reduced to form their alditol counterparts.

Analysis of the resulting trifluoracetic acid hydrolysis products showed a strong presence of mannose and GlcNAc for the isoform I proteins (figure IV-XIII). Since the isoform I protein contained an N-linked high mannose (a glycan composed of only GlcNAc and mannose) these two sugars were expected. In addition to mannose and GlcNAc, galactose and glucose were present. These could arise from the O-linked glycans suspected of occupying five potential sites (based on ethylamine results) within the recombinant protein. Isoform II proteins yielded similar results but sugars detected were much less intense compared to the isoform I recombinant proteins. Also, while GlcNAc was strongly detected in isoform I none was detected in isoform II, probably due to the absence of the N-linked glycan in the isoform II proteins. The primary sugars detected in the isoform II proteins were mannose and galactose. Mannose is known to

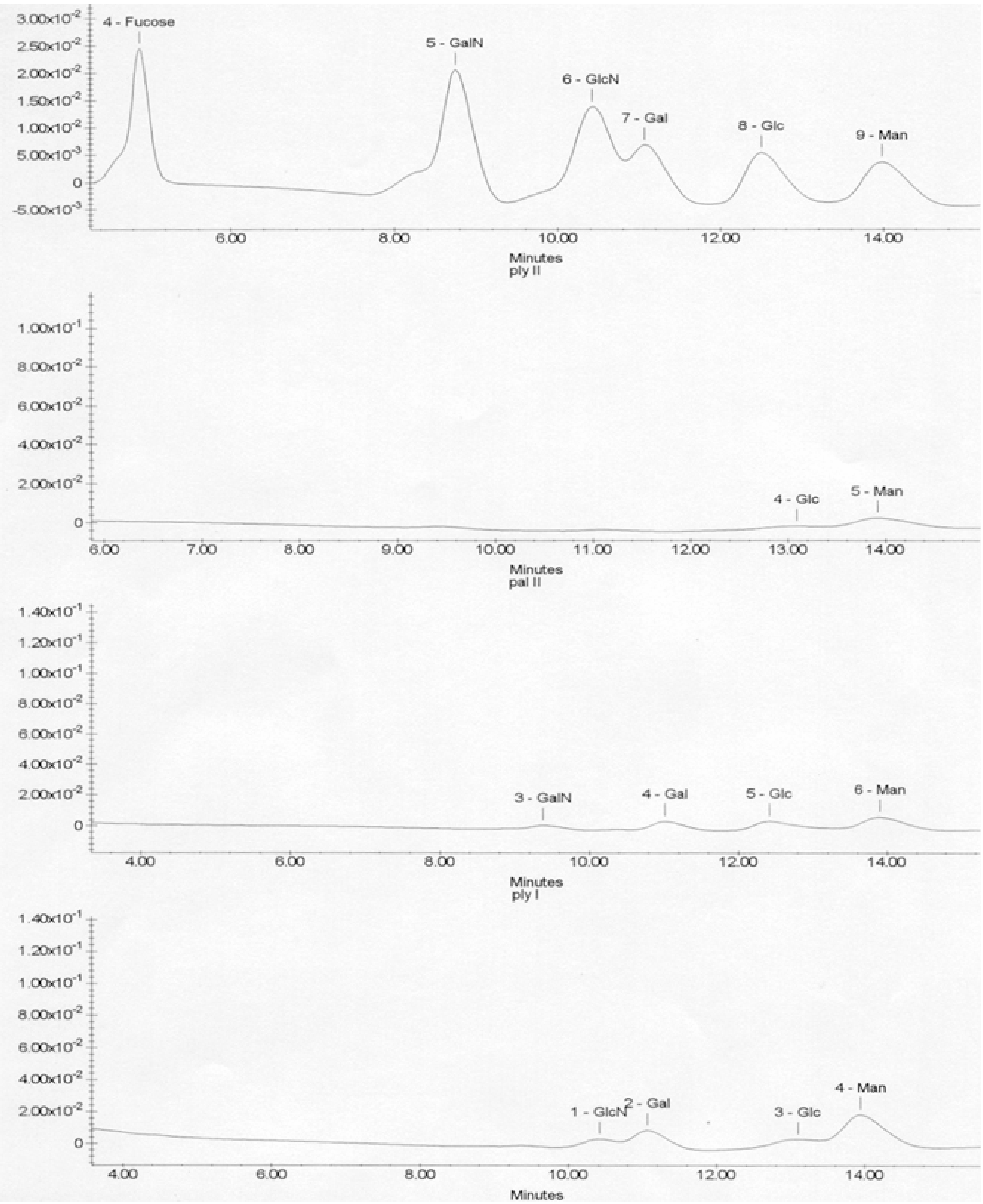


Figure IV-XIII. Anion exchange chromatogram for trifluoroacetic acid hydrolysis products. Note – top trace represents calibration standards.

serve as an O-linked glycan in some systems.[27] Additionally galactose could also serve as an O-linked glycan resulting in the presence of two unique O-linked sugars on the four recombinant proteins.

Analysis of the reaction products from the sodium borohydride/sodium hydroxide experiments showed results supporting the presence of mannose as a possible O-linked glycan (figure IV-XIV). For the three samples analyzed mannitol was readily detected. Additionally, no mannose was detected even in the Ply I sample indicating that side reactions such as peeling had not occurred. Thus indicating the sugars that were released were converted to their alditol forms, at least in the case of mannose.

Anion exchange chromatography has provided some insight as to the types of O-linked carbohydrates present and further confirmed the presence of the N-linked carbohydrate in the isoform I proteins and its absence in the isoform II proteins. To this end it appears that the O-linked sites are occupied by either mannose or possibly galactose, although the latter is considered doubtful due to the weak signal generated by galactose. However, the detection of a peak intermediate to fucose and galactose, in both the isoform I and isoform II recombinant proteins (trifluoracetic acid hydrolysis), has yet to be characterized. It is speculated that this peak could have resulted from the incomplete hydrolysis of some sugars thus leaving a di-saccharide which would elute earlier (as was observed) when compared to the monosaccharides detected.

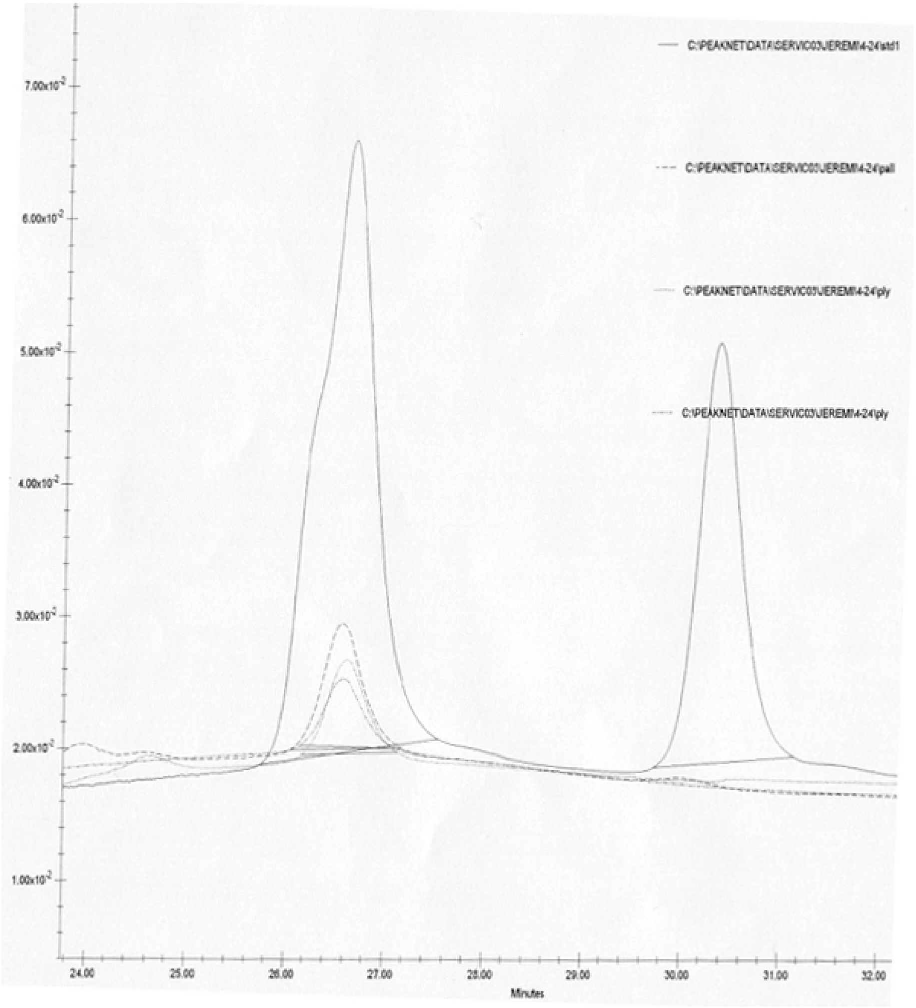


Figure IV-XIV. Anion exchange chromatogram of sodium borohydride/sodium hydroxide reaction products.

CONCLUSIONS

The mass differences detected by MALDI mass spectrometry can be ascribed to the presence of N-linked and O-linked glycans in the four recombinant pectate lyases studied. Part of the mass difference associated with the Isoform I recombinant proteins was due to the presence of an N-linked high mannose structure at residue 64. This was confirmed by lectin analysis, periodate oxidation chemistry, and anion exchange chromatography. Based on electrospray mass spectrometry experiments, utilizing a Q-TOF II, the molecular mass of the N-linked glycan was determined to be between 800 and 2000 Da - due to heterogeneity associated with N-linked carbohydrates. The average mass of the N-linked glycan was estimated to be ~ 1200 Da, the observed average mass difference between the isoform I and isoform II recombinant proteins (MALDI-TOF data). Since the isoform II proteins were shown to not contain an N-linked glycan the difference between the calculated and expected mass with these recombinant proteins must be ascribed to other post-translational modifications. Also, while an N-linked glycan accounts for 1200 Da of the additional mass observed for the two isoform I proteins, a 2400 Da mass difference was measured when compared to the mass based on the amino acid sequence. Thus, like the isoform I proteins, the isoform II proteins have additional mass associated with other types of post-translational modifications.

IMAC extractions and phosphate ELISA excluded phosphorylation as a source of the additional mass. However, based on results from periodate oxidation, ethylamine labeling, and anion exchange chromatography, O-linked glycosylation accounted for a portion of the additional mass observed in all four recombinant proteins. Ethylamine labeling yielded two tagged peptides (AA218-233 and AA197-217). Collectively these two peptides contain four serines and two threonines, or six potential sites for O-linked glycosylation. AA197-217 labeled with two ethylamines. AA218-213 labeled with three ethylamines. Assuming all five sites are occupied by a single carbohydrate residue (mannose) the combined molecular mass due to O-glycosylation is 810 Da, slightly lower than the expected 1200 Da. Alternatively, labeling excusively with GlcNAC would result in a total mass of 1020 Da. Again, this mass is slightly lower than the 1200 Da based on MALDI Data. However, should two or more of the sites be occupied with di-saccharides the mass due to O-linked glycosylation falls into better agreement with the 1200 Da mass difference observed (based on MALDI data).

ACKNOWLEDGEMENTS

This research was supported from grants from the National Science Foundation (grant # 9626835) and the National Institute of Health NCRR (P41RR05351). Part of this work was presented in preliminary form at the 42nd ASMS conference on Mass Spectrometry and Allied Topics, June 2000, Long Beach, California.

REFERENCES

(1) Http://www.sbu.ac.uk/water/hypec.html

(2) Warren, M. E.; et al. *Carbohydrate Research* **2002**, *337*, 803-812.

(3) Cooper, R. M.; in *Biochemical Plant Pathology*, pp. 101-135, Wiley, New York (1993).

(4) Archer, D. B.; Peberdy, J. F. *Crit. Rev. Biotechnol.* **1997**, *17*, 273-279.

(5) Kkitamoto, N.; Yoshino, S.; Ohmiya, K.; Tsukagoshi, N. *Appl. Environ. Microbiol.* **1999**, *65*, 20-24.

(6) Nehrke, K.; Hagen, F. K.; Tabak, L. A. *Glycobiology* **1998**, *8*, 367-381.

(7) Colangelo, J.; Licon, V.; Benen, J.; Visser, j.; Bergmann, C.; Orlando, R. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 1448-1453.

(8) Hounsell, E. F.; Davies, M. J.; Renouf, D. V. *Glycoconj. J.* **1996**, *13*, 19-25.

(9) Wyss, D. F.; Wagner, G. *Curr. Opin. Biotechnol.* **1996**, *7*, 409-418.

(10) Yang, Y.; Bergmann, C.; Bennan, J.; Orlando, R. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1257-1263.

(11) Colangelo, J.; Licon, V.; Benen, J.; Visser, j.; Bergmann, C.; Orlando, R. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 2382-2387.

(12) Mayans, O.; Scott, M.; Connerton, I.; Gravesen, T.; Benen, J.; Visser, J.; Pickersgill, R.; Jenkins, J. *Structure* **1991**, *5*, 677-685.

(13) Gert-Jan, W. M.; Alebeek, V.; Scherpenzeel, K. V.; Baldwin, G.; Schols, H. A.; Voragen, A. *Biochem. J.* **2003**, *372*, 211-218.

(14) Kester, H.; Benen, J.; Visser, J.; Warren, M.; Orlando, R.; Bergmann, C.; Magaud, D.; Anker, D.; Doutheau, A. *Biochem. J.* **2000**, *346*, 469–474.

(15) Warren, M. E.; Kester, H.; Benen, J.; Colangelo, J.; Visser, J.; Bergmann, C.; Orlando, R. *Carbohydrate Research* **2002**, *337*, 803-812.

(16) Harvey, D. *Proteomics* **2001**, *1*, 311-328.

(17) Wells, L.; Vosseller, K.; Cole, R.; Cronshaw, J.; Matunis, M. J.; Hart, G. W. *Mol. Cell. Proteomics* **2002**, *10,* 791-804.

(18) Harvey, D. J. *Int. Journal Mass Spectrom.* **2003**, *226*, 1-35.

(19) Nalivaeva, N. N; Turner, A. J. *Proteomics* **2001**, *1*, 735-747.

(20) Huang, Y.; Mechref, Y.; Novotny, M. V. *Anal. Chem.* **2001**, *73*, 6063-6069.

(21) Geyer, H.; Schmitt, S.; Wuhrer, M.; Geyer, R. *Anal. Chem.* **1999**, *71*, 476-482.

(22) Yang, Y.; Orlando, R. *Anal. Chem.* **1996**, *68*, 570-575.

(23) Kuster, B.; Naven, T.; Harvey, D. J. *J. Mass Spectrom.* **1996**, *31*, 1131-1140.

(24) Colangelo, J.; Orlando, R. *Anal. Chem.* **1999**, *71*, 1479-1486.

(25) Prime, S.; Dearlney, A. M.; Ventom, R.; Parekh, B.; Edge, C. J. *J. Chromatogr. A* **1996**, *720*, 263-271.

(26) White, C. C.; Kennedy, J. F. In *An Introduction to the Chemistry of Carbohydrates*, 3rd ed; Kennedy, J. F., Ed. Clarndon Press: Oxford, U.K., 1988, pp 42-67.

(27) Dell, A.; Morris, H. *Science* **2001**, *291*, 2351-2356.

CHAPTER V

EFFECTS OF N-GLYCOSYLATION ON THE SPECIFICITY OF PECTIN METHYLESTERASE FROM *ASPERGILLUS NIGER*[1]

---

[1] Johnson, J.D, Bergmann, C., Orlando, R. To be submitted to *European Journal of Mass Spectrometry*.

ABSTRACT

Pectin methylesterase is an enzyme secreted by fungi during pathogenic attack. Its primary role is to catalyze the hydrolysis of methyl esters located within the pectin portions of the cell wall. The role of N-linked glycosylation with respect to the mode of action and specificity of pectin methlyesterase is unknown. Pectin methlyesterase is a glycoprotein which contains three potential sites for N-linked glycosylation. Wildtype and mutant pectin methylesterases (PMEs) were prepared by homologous over-expression in *Aspergillus Niger*. Mutant PMEs were prepared by selectively converting key asparagine residues (associated with N-glycosylation) to glutamine. From these experiments six mutants (3 single mutants, 2 double mutants, and 1 triple mutant) were prepared in addition to the fully glycosylated wildtype form. By selectively removing N-linked glycans it is possible to ascertain the effects of such structural changes on the enzymatic properties of PME. The wildtype and mutant forms were then incubated in the presence of fully methylesterified hexa-galacturonate to ascertain the role, if any, of the N-linked glycans in the specifity, and to a lesser degree, activity of the respective PME.

INTRODUCTION

Pectin is a major constituent found in plant cell walls of most higher plants.[1] Pectin primarily consists of homopolymeric partially methylated poly-$\alpha(1\rightarrow4)$-D-galacturonic acid residues, known as the "smooth" regions. There also exists "hairy" regions of alternating α-$(1\rightarrow2)$-L-rhamnosyl-α-$(1\rightarrow4)$-D-galacturonan which typically contains branches of L-arabinose and D-galactose. To a lesser degree pectins may also have rhamanogalacturonan side chains containing D-xylose, D-apiose, 3-deoxy-D-manno-2-octulosonic acid, D-glucuronic acid, L-fucose, and 3-deoxy-D-lyxo-2-heptulosonic acid attached to poly-α-$(1\rightarrow4)$-D-galacturonic acid regions.[2] The backbone of pectic polysaccharides can be substituted. For example, galacturonic acid residues within the homogalacturonan (poly-$\alpha(1\rightarrow4)$-D-galacturonic acid) can be acetylated at the C-2 and C-3 hydroxy groups. Also, methyl-esterification is likely to take place at the C-6 position but can occur in either an entirely random fashion or in a manner in which entire blocks of galacturonic acid can be methyl-esterified.[3]

This complex pectic network serves as a defensive barrier for plant cells. In addition, the complex carbohydrate network serves as a "dynamic interface that participates directly in cellular responses to exogenous stimuli".[4] The direct role played by the cell wall is integral in the perception and response of the cell to incoming signals. When acted upon by phytopathogenic fungi the cell responds with various defense mechanisms designed to target the pathogenic attack of the fungus and limit cellular damage (various pathogenesis related proteins, for example). In response to the plant defense mechanisms many fungi have elaborate sets of enzymes which serve to degrade the pectin and cellular networks, allowing the fungus to absorb nutrient by-products from the cell wall as well as nutrients from the cytoplasm.

The pathogenic attack of fungi on a plant is often led with a set of enzymes that target the homogalacturonan regions of the cell wall pectin network. These enzymes include polygalacturonases, pectin esterases, and pectic lyases. Pectin methylesterase, a type of pectic esterase, functions to catalyze the desterification of methyl-ester groups within pectin. This initial and subtle step results in the formation of homogalacturonan with lowered degrees of esterification. This decreased esterification increases the susceptibility of the pectin network to the actions of polygalacturonases and pectate lyases.[5] In addition, small pectic fragments released during pathogenic attack act as signals that induce the expression of other pectolytic enzymes as well as the expression of genes that encode for distinct pectic isozymes.[4]

The ability of these enzymes to successfully degrade plant cell walls has been presented as both a serious problem (within the agricultural community) and a useful resource (within the food industry). The obvious destructive nature these fungal enzymes present to the agricultural community is easily demonstrated through lost crops. Alternatively, uses within the food industry continue to grow. Probably the most notable, and widely used from a commercial standpoint, are the PMEs from *Aspergillus Niger*. Industries such as food processing, fruit juice production, and wine producers spend hundreds of millions of dollars on pectinases annually. They are used to control viscosity in various food preparations, consistency and clarity in a variety of juices, and release of color in wines.[3]

Pectin methylesterases are found in a variety of micro-organisms, and are commonly found in economically important fungi such as *Aspergillus*[6] and *Botrytis*[7] or in bacteria such as *Erwinia*[8]. Nonetheless, limited information has been presented related to their structural characteristics and enzymatic mode-of-action. A few recent reports have demonstrated the successful study and partial characterization of some PMEs isolated from several yeasts.[9,10] Pectin methylesterases are typically medium sized enzymes with molecular masses of 25,000 to 50,000 daltons.[11] *A. niger* derived PME is a protein composed of 331 amino acids having a molecular mass of 33,927 Da. Most fungal PMEs demonstrate maximum activity at pH 4 to 6 whereas bacterial PMEs most often have an activity maximum between pH 6 to 8.[11] One other difference that manifests itself is the enzymatic mode of action. It is speculated that fungal PMEs de-esterify in a random fashion attacking the methyls in no specific order. Alternatively, plant and certain bacterial PMEs remove methyls in a blockwise fashion resulting in stretches of de-methylated GalA.[1,12] It has also been shown that PMEs are limited in that total de-methylation cannot be achieved. Rather, approximately 20-30% of the methyl-esters remain after extended incubation. While some of the aforementioned research has provided insight as to the mode of action and limitations of PMEs, little information is presented as to the role of glycosylyation in enzyme mode-of-action or specificity.

Recently, mass spectrometry has been utilized to provide insight into pectin-degrading enzymes. Kester and Warren recently reported on the mode-of-action of *A. niger* derived PME with respect to fully methylesterified oligogalacturonates as well as activity studies using a wildtype PME and a series of mutated PMEs.[1,3] Substrate specificity has been studied for a number of pectin degrading enzymes including polygalacturonases and pectin lyases.[1,3,11,13-15] MALDI mass spectrometry, in conjuction with high performance anion exchange chromatography, was used to study the products of the digestion of partially methyl-esterified pectin by endopolygalacturonase (EPG).[14] Korner, et al. used ion-trap mass spectrometry to characterize the products of a variety of enzyme/substrate reactions including polygalacturonase II, exopolygalacturonase and pectin lyase.[13, 15] To more thoroughly investigate the role of substrate size and type of esterification Gert-Jan, et al. investigated a variety of substrates and their susceptibility to de-esterification with *Aspergillus Niger* derived pectin methylesterase.[11]

The research described herein focuses on the characterization of products resulting from the incubation of wildtype PME as compared to mutated PMEs. The intent was to study substrate mode-of-action as a function of enzyme glycosylation. These experiments were carried out using a series of mutated pectin methylesterases derived from *Aspergillus Niger* as well as the fully glycosylated wildtype enzyme. As stated earlier the molecular mass of PME derived from *Aspergillus Niger* is 33,927 Da. Located within the primary amino acid sequence (Figure V-I) are three sequons (N-linked consensus sequences) available for potential N-linked glycan attachment. The three sequons initiate at the following residues; 95 (N95), 283 (N283), and 302 (N302). Mutation of the asparagines initiating the three sequons excludes glycosylation at the mutation site. For example, mutation of a single site resulted in a 'single mutant' variant – containing only two of the three N-linked glycans normally found on the wildtype enzyme. Mutation at two sites resulted in a 'double mutant' – containing only one of the three N-linked glycans normally found in the wildtype enzyme. Absence of glycans, as a result of mutation, was confirmed by MALDI mass spectrometry. The substrate utilized

MVKSILASVF FAATALAASR MTAPSGAIVV
AKSGGDYDTI SAAVDALSTT STETOTIFIE
EGSYDEQVYI PALSGKLIVY GQTEDTTTYT
SNLVN^{95}ITHAI ALADVDNDDE TATLRNYAEG
SAIYNLNIAN TCGQACHQAL AVSAYATEQG
YYACQFTGYQ DTLLAETGYQ VYAGTYIEGA
VDFIFGQHAR AWFHECDIRV LEGPSSASIT
ANGRSSESDD SYYVIHKSTV AADGNDVSS
GTYYLGRPWS QYARVCFQKT SMTDVINHLG
WTEWSTSTPN TEN^{283}VTFVEYG NTGAGSKGPR
AN^{302}FSSELTEP ISISTLLGSD WEDWVDTSYI
N

Protein MW = 33,927 Da

Figure V-I. Primary amino acid sequence of *Aspergillus niger*. N-linked consensus sequences are shown in blue. Superscript indicates corresponding residue number.

was a chemically synthesized fully methyl-esterified hexagalacturonate. The mode-of-action and activity of the enzymes were determined by characterizing the reaction products from a series of incubations using LC-MS and LC-MS/MS.

EXPERIMENTAL

Preparation of fully methyl-esterified hexagalacturonate

Hexagalacturonic acid (50 mg) was dissolved in 5 ml of methanol, 0.02 M HCl and allowed to incubate for 14 days at 4 ^{o}C. The esterification reaction was terminated by addition of Ag_2CO_3 (80 mg). The resulting AgCl precipitate was removed by centrifugation. Methanol was removed by evaporation under a stream of nitrogen. To remove partially methyl-esterified hexamers the sample was reconstituted in water and passed over a 4-ml AG 501-X8 mixed-bed resin column (Bio-Rad Laboratories, Richmond, CA, U.S.A) then eluted with water. Isolated fully methyl-esterified hexagalacturonate was then freeze dried. To determine the quantity of hexa-methyl-galacturonate, freeze dried samples were dissolved in water and the concentration determined by measuring uronate content.[16]

Enzyme Incubations and sample preparation for mass spectrometric analysis

Fully methyl-esterified hexagalacturonate was reconstituted in 50 mM sodium acetate, pH 5.0 followed by incubation with a selected PME (wildtype, triple mutant, etc.) for 0.5 and 24 hours at room temperature. Substrate concentration was maintained at 5 mM in 0.07 ml total reaction volume. After 30 minutes of incubation a 0.035 ml aliquot of reaction mixture was removed and treated as follows (after 24 hours the remaining

0.035 ml was treated as follows): To stop the reaction (enzymolysis) the solution was mixed with 0.2 ml Dowex 50W-X2 (H^+) beads in a Microcon 0.22 micron spin filter. 600 ml of water was added and the sample was gently agitated followed by incubation at room temperature for 15 minutes. The solution was separated from the beads by centrifugation at 10,000 g and subsequently freeze dried. Use of the cation exchange column 'selectively' removed the enzyme.

^{18}O-Exchange of the methyl-hexagalacturonate anomeric oxygen

Differentiation of the reducing and non-reducing ends of carbohydrates is required before sequencing is possible by tandem mass spectrometry. In order to sequence the products of the methyl hexagalacturonate/PME reaction the anomeric oxygen was exchanged in H_2O^{18}. The freeze dried reaction products (0.220 mg) were incubated in 95% H_2O^{18} (.220 ml) containing 0.5% (v/v) formic acid. The exchange was allowed to proceed for 24 hours at room temperature.

LC-MS and direct infusion mass spectrometry analysis

Direct Infusion MS and MSMS analysis

Mass Spectra were obtained using a Micromass Q-TOF II (Milford, MA, U.S.A) equipped with a microspray ion source. Data were accumulated under (+) ion mode utilizing a Cone Voltage (CV) of + 40V and a Capillary Voltage of + 2800V. Nitrogen was used for the nebulizing gas at a flow rate of 20 L/hr. Quadrupole resolution was optimized to limit the window of ions entering the collision cell to +/- 1 da. This was done to prevent the introduction of non-^{18}O labeled hexamers into the collision cell and subsequent analysis of those ^{16}O derived product ions. The sample was loaded into a 100 μL gas-tight syringe (Hamilton, Austin, Texas, U.S.A) and infused directly into the source of the mass spectrometer at a rate of 0.001 ml/min using a Harvard Apparatus model 22 syringe pump (South Natik, MA, U.S.A).

LC-MS and LC-MSMS analysis

Data were obtained using a Waters CapLC (Milford, PA, U.S.A) equipped with a LC-Packings C-18 Pepmap HPLC microbore (70 μm X 15 cm) column (San Francisco, CA, U.S.A) coupled to a Micromass Q-TOF II (Milford, MA, U.S.A) equipped with a microspray ion source. LC conditions were as follows: a linear gradient was employed where %B (100 % acetonitrile, 0.1% formic acid) was ramped from 10% to 70 % over 95 minutes relative to mobile phase component A (0.1 % formic Acid). Data were accumulated under (+) ion mode utilizing a Cone Voltage (CV) of + 40V and a Capillary Voltage of + 2800V. For MSMS analysis ions were pre-selected by inclusion in a parameter file which limited mass selection for MSMS analysis to m/z values defined by the user. Only ions whose m/z were the same as those derived from the enzymatic de-methylation were chosen for subsequent MSMS analysis. Argon was used for the collision gas (pressure = 15 psi). Collision energy was optimized based on the type of ion being fragmented in the collsion cell (H^+ adduct = 17 eV, Na^+ adduct = 55 eV). Nitrogen was used for the nebulizing gas at a flow rate of 20 L/hr. Quadrupole resolution was optimized to limit the window of ions entering the collision cell to +/- 1 da. This was done to prevent the introduction of non-^{18}O labeled hexamers into the collision cell and subsequent analysis of ^{16}O derived product ions.

MALDI Mass Spectrometry analysis

Data were accumulated on a Hewlett Packard G2030A (Palo Alto, CA) MALDI-TOF mass spectrometer. The instrument was operated in positive ion mode at a pressure of ~ 1 x 10^{-6} Torr, 28 kV accelerating voltage and 7 kV extractor voltage. Samples were mixed 1:1 with sinapinic acid (10 mg/ml; 70% aqueous acetonitrile (Sigma, St. Louis, MO, USA) containing 0.01% trifluroacetic acid (Sigma-Aldrich, St. Louis, MO, USA)) followed by vacuum drying. Samples were desorbed and ionized using UV radiation emitted from a nitrogen laser at an output wavelength of 337 nm.

RESULTS AND DISCUSSION

MALDI MS analysis of wildtype and mutant PME

In order to investigate the likelihood that site directed mutations had successfully de-glycosylated the protein, the predicted mass of the protein (based on the amino acid sequence, figure V-I) was compared to the measured molecular mass obtained by MALDI-MS (Table V-1). This type of experiment can only provide limited insight as to the presence and extent of N-linked glycosylation. The measured masses for the wildtype, single, and double mutant proteins were all substantially higher than the molecular mass based on the amino acid sequence. Mass shifts like those observed between the wildtype, single, and double mutants and the molecular mass as predicted by the amino acid sequence are typically due to the presence of post-translational modifications. More specifically large mass shifts, like those measured with these proteins (~ 1400 to 2600 daltons), are typically associated with N-linked glycosylation. This is due to the smaller measured mass increases associated with other types of post

Table V – 1. MALDI-MS measured molecular masses of Wildtype PME and the seven mutant forms.

Protein	Measured Molecular Weight (Da)	Mass Difference Measured – Predicted (Da)	Potential number of N-linked Glycans
Wildtype PME	**40512 +/- 400 Da**	**6585**	**3**
Single Mutants			
N95Q	**39029 +/- 390 Da**	**5102**	**2**
N283Q	**37993 +/- 380 Da**	**4066**	**2**
N302Q	**37826 +/- 380 Da**	**3899**	**2**
Double Mutants			
N95Q/N283Q	**36320 +/- 360 Da**	**2393**	**1**
N283Q/N302Q	**35311 +/- 350 Da**	**1384**	**1**
N95Q/N302Q	**36057 +/- 360 Da**	**2130**	**1**
Triple Mutant			
N95Q/N283Q/N302Q	**34055 +/- 340 Da**	**128**	**0**

translational modifications such as acetylation, phosphorylation, and in many instances O-linked glycoslylation. The average molecular mass of these N-linked glycans was measured to be ~ 2000 Daltons. Since all N-linked structures have a common core of $Man_3GlcNAC_2$ (molecular mass = 892 Da), and the observed shifts all exceed this value the likelihood that combinations of other types of PMTs accounting for the observed mass difference is decreased. When all three potential sites for N-linked glycosylation were mutated (triple mutant form of PME) the measured mass differed by ~ 128 Da from the predicted mass (based on amino acid sequence). The observed 128 Da difference is not large enough to be accounted for with N-linked glycosylation or O-linked glycosylation and also falls within the mass accuracy of the MALDI-TOF instrument used (typically 0.1% error). This would indicate that mutagenesis had successfully eliminated all potential sites for N-linked glycan attachment. Analysis of the mass differences (measured – predicted) for the double mutants is in agreement with the mass loss observed for the single mutants. For example, the measured mass for the N283Q/N302Q mutant is ~ 35311 Da. This particular PME had two sites mutated (no N-linked glycan attached) leaving a single site for the N-linked glycan at residue 95. The measured - predicted mass would leave a mass increase equal to the molecular mass of the glycan attached at residue 95 which was calculated to be 1384 Daltons. Likewise, the N95Q single mutant PME showed a mass deficit relative to the wildtype PME equal to the mass of the glycan attached at residue 95. This mass was calculated to be 1483 Da, in agreement with the mass based on the double mutant derived value of 1384 Da. As shown in table V-1 the differences in mass indicated that successful mutagenesis was likely achieved and based on the measured mass differences the changes observed are likely due to the removal of N-linked glycans. Also, the variations of the molecular masses observed by MALDI-MS indicated that structural variations between the attached N-linked glycans likely exist.

Effects of N-glycosylation of PME activity/specificity

Electrospray mass spectrometry was utilized to characterize the products formed from PME incubations with a selected substrate, fully methyl-esterified hexagalacturonate (figure V-II). To ascertain any differences in the rates of product formation as a result of mutagenesis, activities of the wildtype and mutants against fully methyl-esterified hexa-galacturonate were assessed at two time intervals, 0.5 hours and 24 hours. Rather than determine product progression as a function of time over shorter intervals the primary goal of this portion of the research was to simply ascertain if changes in glycosylation had profound effects on the short term and long term activity of the enzyme. This type of information is useful for industrial applications where the function of a PME may be to rapidly demethylate a target for further breakdown by additional enzymes.

Figure V-III shows the mass spectrum of the unreacted, O-18 exchanged hexagalacturonate (fully methyl-esterified). As is shown the primary species detected is the fully methyl esterified O^{18} form (m/z = 1183.2). Due to the incomplete exchange of the anomeric oxygen there was also a strong peak at 1181.2 resulting from the O^{16} form of the substrate. This data indicated an approximate 60 to 65% incorporation of O^{18}. To determine the effects of N-glycosylation on short and long term incubations of substrate with enzyme, aliquots of enzyme/substrate incubations were analyzed at 0.5 and 24

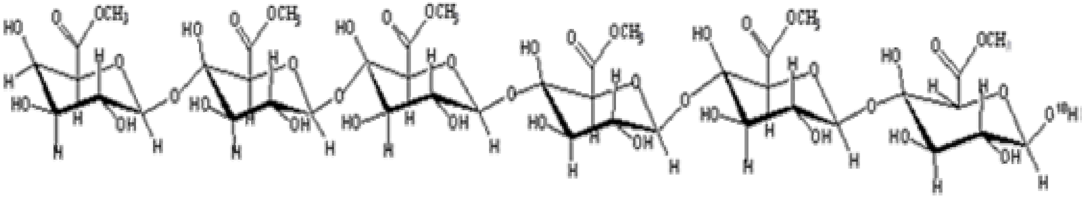


Figure V-II. O-18 exchanged fully methyl-esterified hexagalacturonate.

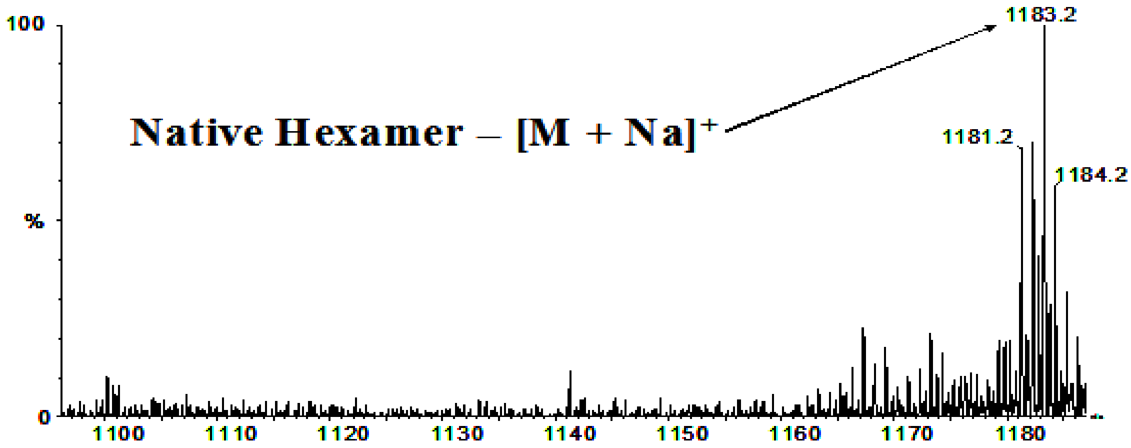


Figure V-III. Electrospray mass spectrum of fully methyl-esterified hexagalacturonate. Sodiated species.

hours. Figure V-IV shows the results of four such experiments utilizing the wildtype and triple mutant PME/substrate incubations.

Pectin methyl esterase serves to catalyze the de-methylation of certain methyl-ester groups. In fully methyl-esterified hexagalacturonate there are 6 methyl-ester groups available for enzymatic/catalytic de-methylation. The rate at which the enzyme can catalyze such events is defined as activity. Analysis of Figure V-IV shows that little, if any, variation was observed in the activity between the wildtype and triple mutant forms of PME. As PME de-methylates its substrate the change in mass is reflected by a 14 Dalton shift to lower mass. Thus, a substrate, such as fully methyl-esterified hexagalacturonate, with an initial mass of 1183.2 would be detected at 1169.2 once a single de-methylation had occurred. However, in the instance of fully methyl-esterified hexagalacturonate six sequential de-methylation events could give rise to a series of six ions whose mass differ by 14 Da (mass distribution) with each ion in the series representing an additional degree of de-methylation (table V-II).

Figure V-IV clearly demonstrates this sequential de-esterification through the presence of a mass distribution extending from mass 1183.2 (fully methyl-esterified form) to 1113.2 (loss of 5 methyl groups). At 30 minutes most of the de-esterified variants were represented, as demonstrated by the detection of ions with representative m/z values, with the most abundant form being the variant showing 3 de-methylation events detected at 1141.2. Alternatively the dominant form found in the two 24 hour incubations occured at 1113.2 Da corresponding to the variant with a single methyl-ester remaining (loss of 5 methyls). Comparison of the wildtype (figure V-IV(a)) and triple mutant (figure V-IV(b)) data for the 30 minute incubations clearly showed that little

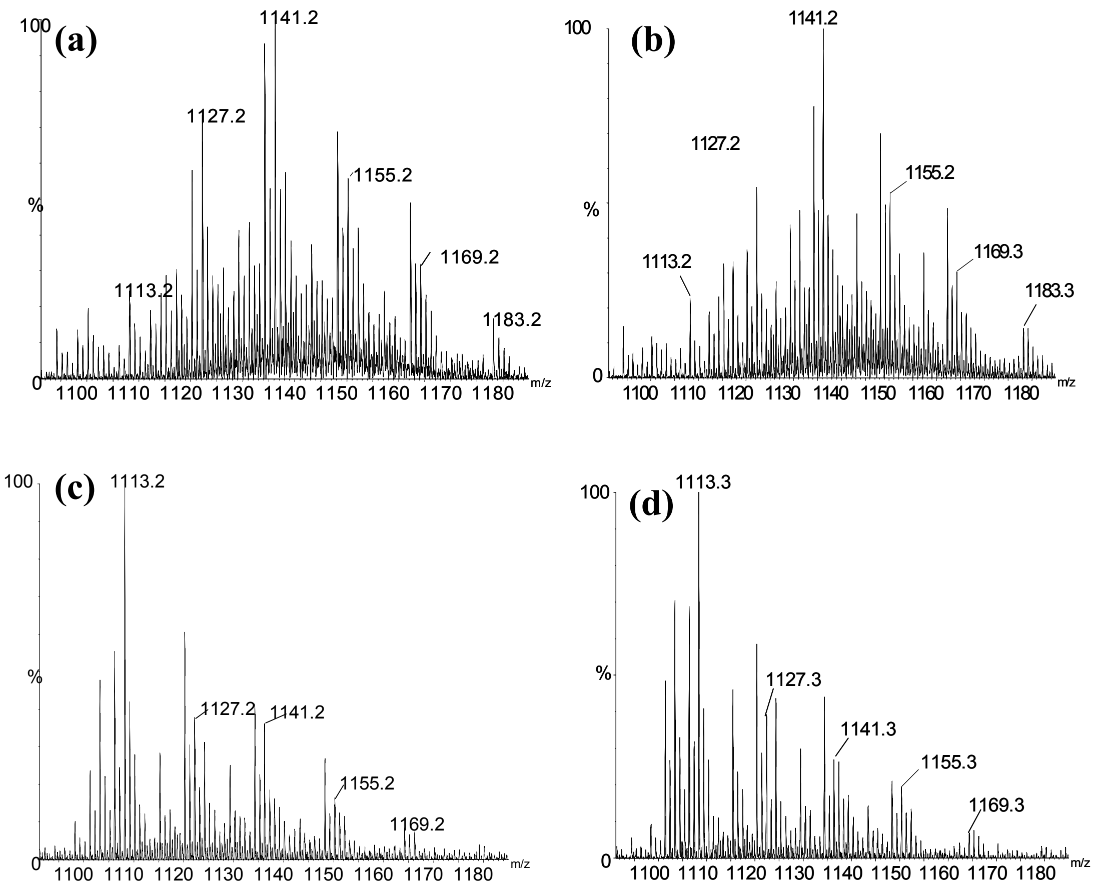


Figure V-IV. MS data showing the formation of de-methylated variants from incubation of wildtype and triple mutant pectin methylesterase. (a) Wildtype PME – 30 minute incubation. (b) Triple mutant PME – 30 minute incubation. (c) Wildtype PME – 24 hour incubation. (d) Triple mutant PME – 24 hour incubation.

Table V-II. Molecular masses of products formed by de-methylation of methlyl-esterified hexagalacturonate.

Compound	**Mass**	**Number of methyl-esters**	**Number of de-methylation events**
6Me-6Galacturonate	1183.2	6	0
5Me-6Galacturonate	1169.2	5	1
4Me-6Galacturonate	1155.2	4	2
3Me-6Galacturonate	1141.2	3	3
2Me-6Galacturonate	1127.2	2	4
1Me-6Galacturonate	1113.2	1	5
0Me-6Galacturonic Acid	1099.2	0	6

change, if any, was detected with regards to the short term activity of PME. This indicated that short term activity is not affected by the presence (wildtype) or absence (triple mutant) of the N-linked glycans. Similary longer incubation times yielded no differences when comparing data between the wildtype and triple mutants forms of PME. This indicated that N-glycosylation played no role in the activity of PME, at least within the time scales investigated.

Tandem mass spectrometric analysis of reaction products

Under typical tandem mass spectrometric conditions oligosaccharides dissociate in predictable patterns forming product ions that can be used to determine the structure of the precursor oligosaccharide. Product ions formed in tandem MS experiments are named based on the location of bond cleavage and charge retention. Ions where the charge is retained at the reducing end are termed ‘Y and Z’ ions. Alternatively, if charge is retained at the non-reducing terminus the resulting ion is termed either a ‘B’ or ‘C’ ion.[17] In simplest terms the Y and B fragment ions lack one oxygen compared to the Z and C ions. In quadrupole-TOF instruments, where collisions are energetically low and resonance times are typically short, dissociation occurs predominantly along the glycosidic linkage giving rise to mostly Y and B ions (figure V-V). In addition, as demonstrated in figure V-V, it is not uncommon to observe an ion series corresponding to water loss. In other types of mass spectrometric instruments where collision energies are substantially higher, or resonance times substantially longer, other product ions may be produced or even favored. For example, ion-trap mass spectrometers tend to also produce C, Z, and cross-ring cleavage product ions in addition to Y and B ions.[13]

The lack of intense C, Z, and cross-ring cleavage ions greatly simplify spectra associated with quadrupole-TOF and triple quadrupole based instruments.

To assess the mode-of-action of *Aspergillus Niger* derived pectin methylesterase on fully methyl-esterified hexagalacturonate each product formed from the incubation of enzyme with the substrate was subjected to MSMS analysis. Most of the ions analyzed by MSMS yielded a series of Y, B, and water-loss ions. For example, analysis of the mono-methyl-esterified hexagalacturonate MSMS spectrum, figure V-V, revealed a complete Y series (Y_6 thru Y_1 ions), complete B ion series ($B_6 – B_1$), and a complete 'Y – H_2O^{18}' series. This [figure V-V], being the simplest product spectrum to interpret due to the presence of a single methyl-ester, is best for demonstrating the basic premise of methyl-ester residue location based on MSMS data. By analysis of the mass differences between a given series it was possible to determine the location of the methyl-esterified residue within the oligosaccharide. The mass of a methyl-esterified galacturonate residue was 190 Da whereas the mass of a galacturonic acid residue (loss of the methyl-ester) was 176 Da. Analysis of the mass differences between sequential peaks ($Y_n – Y_{n-1}$) in a given series, for example the Y-ion series, should yield either a 190 Da difference or a 176 Da difference. As shown in figure V-VI the only 190 Da difference is observed between the Y_6 and Y_5 ions. All other differences corresponded to 176 Da. This indicated that the methyl-ester was always located on the non-reducing residue for the mono-methyl esterified hexagalacturonate product. All other residues were successfully converted to galacturonic acid based on the 176 mass difference observed between the corresponding ions. Conversely, similar calculations utilizing the B-ion series confirm the results obtained from the y-ion series. The peak at 213 corresponded to the sodium

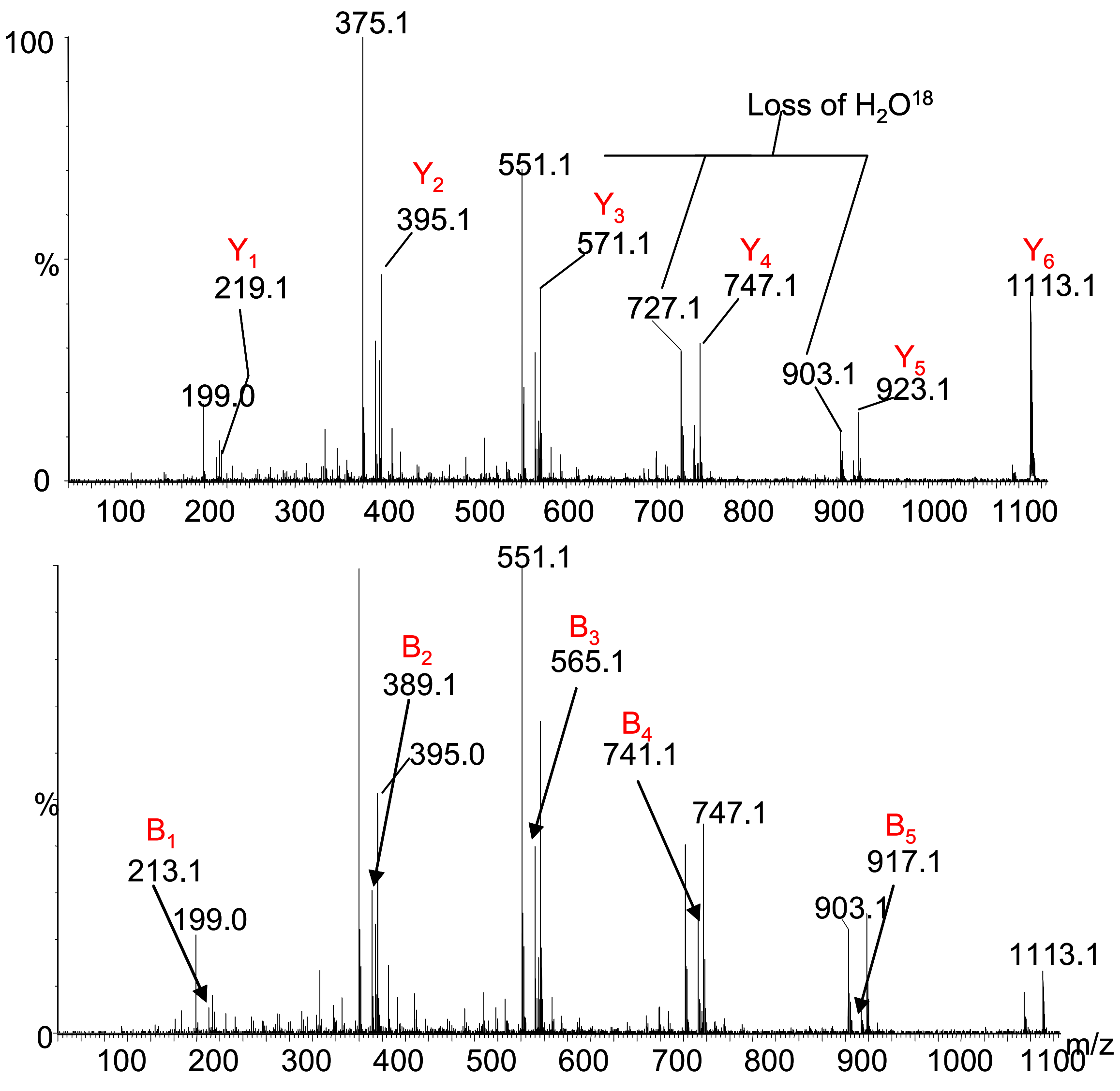


Figure V-V. Dissociation of an oligosaccharide in the collision cell of a quadrupole-TOF mass spectrometer. Y, B, and water-loss fragments correspond to the major product ions generated. Precursor ion = mono-methyl esterified hexagalacturonate, m/z 1113.1 Da.

adduct of the B_1 fragment ion containing a methyl-ester. As with the Y-ion series addition of 176 Da (galacturonic acid) follows sequentially for the B_2, B_3, B_4, and B_5 ions indicating a single methyl-ester at the non-reducing residue, see figure V-VII.

Assignment of methyl ester location, such as the one demonstrated above, is only possible when the reducing and non-reducing ends can be differentiated. Differentiation of the reducing and non-reducing ends is achieved by the incorporation of an ^{18}O at the anomeric oxygen (reducing end), see figure V-II. Thus, as illustrated in figure V-VII, Y fragment ions represent charge retention on the reducing end with residue loss (due to collisional activation) from the non-reducing end whereas B-ions represent charge retention on the non-reducing end but residue loss from the reducing end.

Comparison of the data from the MSMS analysis of the monomethyl hexa-galacturonate for all PMEs (wildtype and mutants) showed no difference in the location of the remaining methylester on the oligosaccharide. In all instances the single methyl-ester was always located on the non-reducing residue as indicated by both the detection of a B_1 ion with m/z of 213 Da and the detection of a 190 difference between the Y_6 and Y_5 ions. It has been previously reported that PME was unable to de-methylate the non-reducing residue of substrates with degrees-of-polymerization (DP) 3, 4 and 5 and suggested that PME is unable to completely de-esterify all methyl esters in the homogalacturonan part of pectin.[1,18] Data to this point was in agreement with these previous results and indicated that de-esterification of the non-reducing residue was also not possible on oligosaccharides with a DP of six.

Utilization of the same fundamental principles allowed for identification of the methyl-ester location(s) in other de-methylated products formed from the incubation of

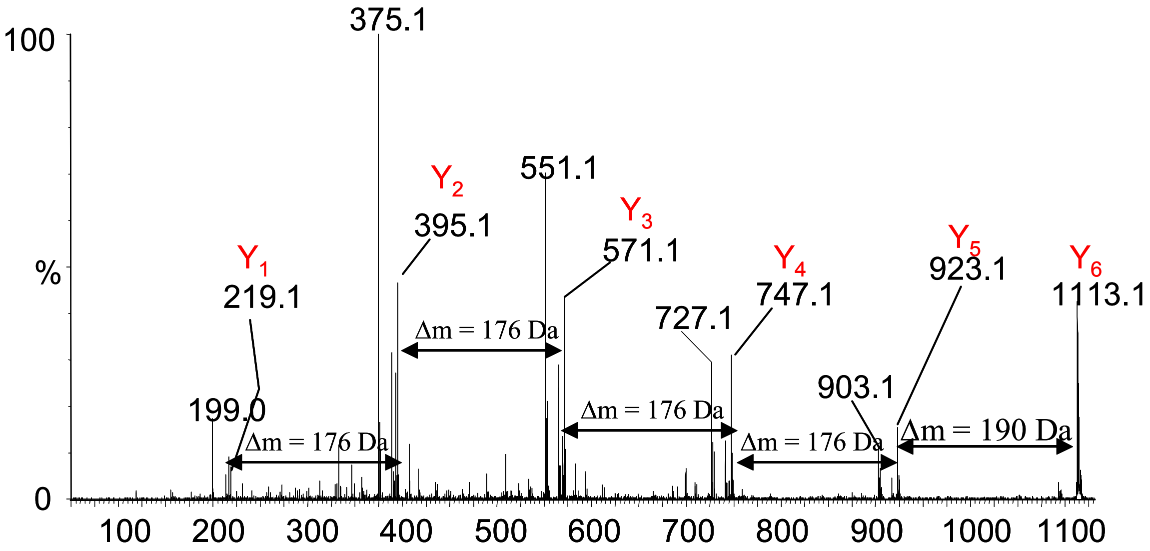


Figure V-VI. Localization of methyl-esterified residue based on mass differences within the Y-ion series for the mono-methyl hexagalacturonate product. Precursor ion = 1113.1 Da.

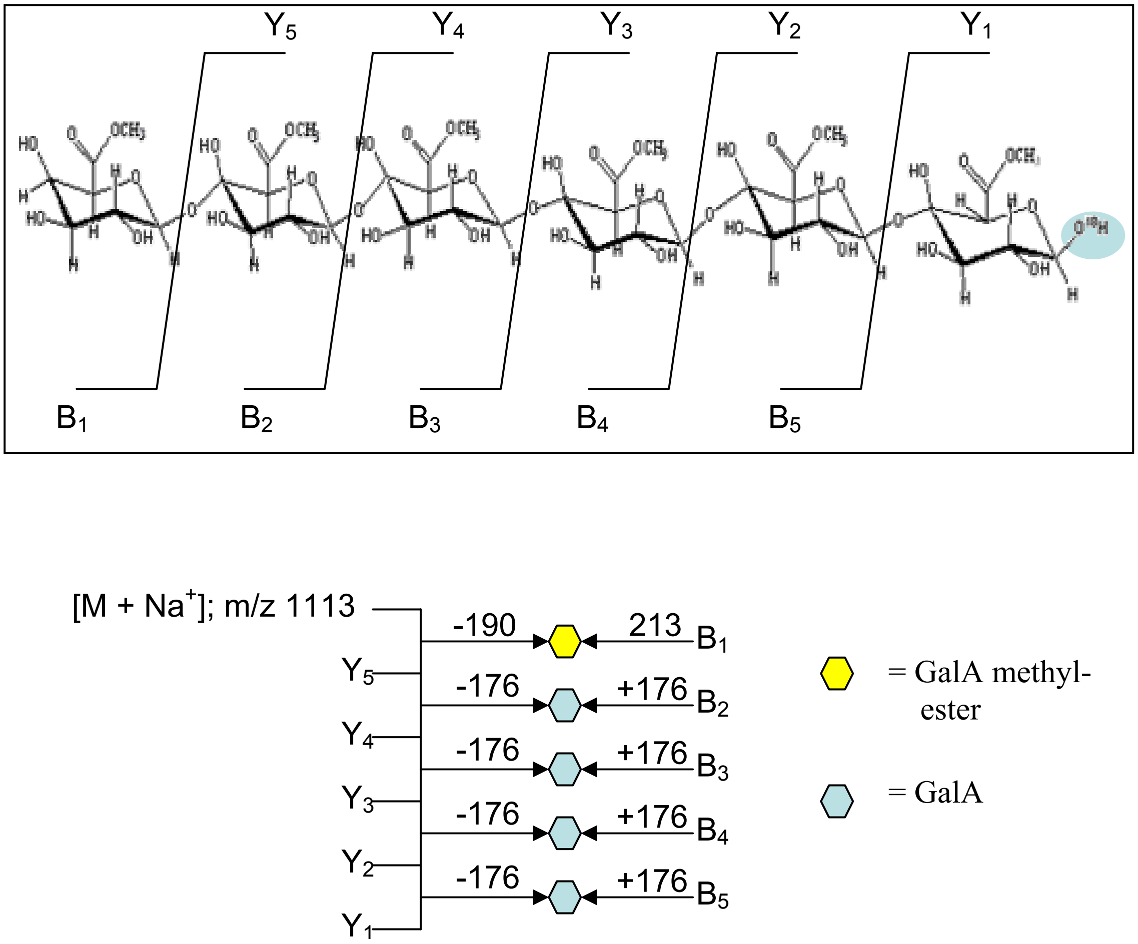


Figure V-VII. Illustration demonstrating the identification of the methyl-ester position based on analysis of mass difference between fragment ions produced by MSMS analysis for the mono-methyl hexagalacturonate product.

pectin methylesterase with fully methyl-esterified hexagalacturonate. Shown in figure V-VIII are the MSMS data for the di-methyl hexagalacturonate products from incubations with the wildtype and triple mutant forms of the PME. In agreement with the previous data was the detection of ions confirming the presence of one methyl-ester at the non-reducing residue (B_1 ion at m/z = 213 Da and a 'Y_6 - Y_5' value of 190 Da). Since one methyl ester must apparently remain on the non-reducing residue this leaves a single methyl-ester to occupy one of the remaining five residues.

Shown in Figure V-IX is an illustration of the masses expected based on specific losses required in order for the remaining methyl-ester to be located at each of the available positions. Comparison of the calculated masses from figure V-IX with the measured masses in figure V-VIII indicated that methyl ester loss was preferred from central regions of the oligosaccharide leaving terminal and near-terminal residues as the preferred sites for methyl-esters. Specifically the detection of 'Y_5 - Y_4', 'Y_3 – Y_2', and 'Y_2 – Y_1' mass shifts of 176 Da eluded that demethylation was favored on the central residues, and only occured at the terminal and near terminal residues when substrate had been largely de-methylated at the 3 and 4 positions. A 'Y_5 – Y_6' loss of 190 Da and the predominant detection of Y_1 at 233 Da indicated that the favored form of the reducing sugar was methyl galacturonate, again indicating that core demethylation is favored.

Utilizing the same strategies the remainder of the products from the PME digest of fully methyl-esterified hexagalacturonate were analyzed and structural information obtained as to the locations of methyl-esters within the respective oligosaccharide variants. In some situations it was not possible to locate all methyl-esters. This was particularly true of the tri-methyl variant. It is believed that the number of possible

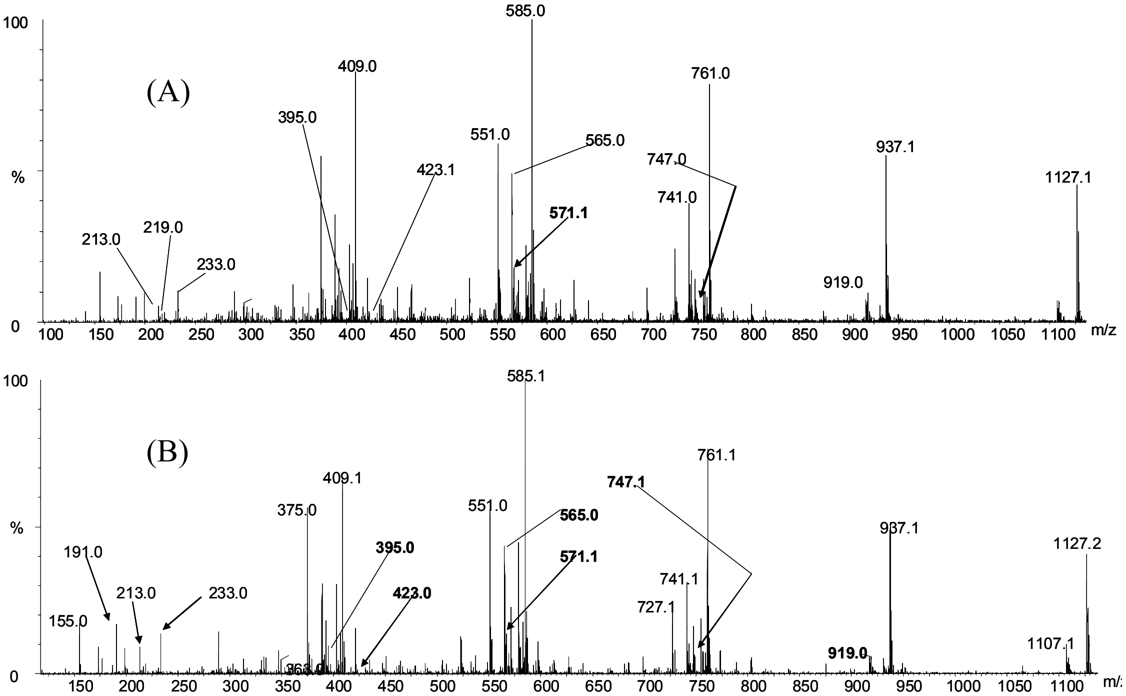
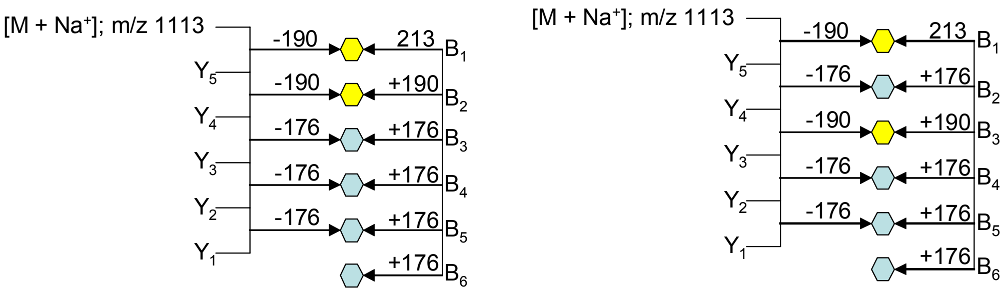
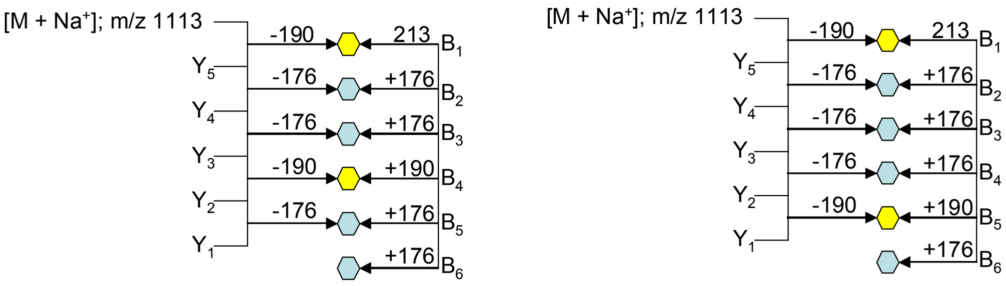


Figure V-VIII. MSMS data from PME generated di-methyl hexagalacturonate. Precursor m/z = ~1127.1 (A) Wildtype PME - 30 minute incubation time. (B) Triple mutant PME – 30 minute incubation time.



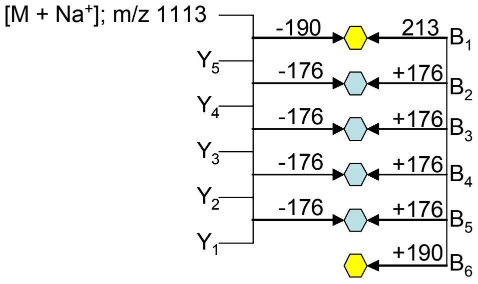
5,6 di-methyl

4,6 di-methyl



3,6 di-methyl

2,6 di-methyl



1,6 di-methyl

note: methyl location numbered from the reducing end.

⬢ = GalA methyl-ester

⬡ = GalA

Figure V-IX. Y and B ion masses used to determine the location of the 2 methyl-esters.

isoforms resulted in spectral dilution making the detection of some peaks impossible, under the experimental conditions used. Nonetheless it was clear within all variants that de-methylation never occurred at the non-reducing residue and that preferential de-glycosylation seemed to favor central residues as opposed to terminal and near-terminal residues. This latter observation was not as clearly discernable in the tri-, and to a lesser extent, the tetra-methyl variants due to lack of sufficient fragment ions need to readily determine all methyl-ester locations. Again, these data are in agreement with other data presented within this dissertation as well as previously reported similar experiments utilizing similar materials.

Of interest also was the comparison of the mode-of-action of the mutated proteins (exhibiting varying degrees of N-linked glycan absence) with the wildtype pectin protein. Specifically, does the presence of the N-linked glycans influence the mode-of-action of PME towards its substrate? Comparisons of MSMS data, e.g. figure V-VIII, between mutants and wildtype PMEs indicated that N-linked glycosylation played no role in the mode-of-action (MOA) of PME. In all cases the data obtained for the mutant enzymes were nearly identical to the data obtained for the wildtype. The absence of any significant differences in the triple mutant data when compared to the wildtype should be sufficient to indicate that N-linked glycosylation is irrelevant with regards to MOA. However, additional experiments were completed in order to verify that selective removal N-linked glycans would not produce a significant change in enzyme MOA. As was expected there were no observed differences.

CONCLUSIONS

MS analysis revealed that the reaction of substrate with each PME produced multiple compounds with varying degrees of esterification. At increased incubation times, the extent of de-esterification increased. However, variations in the m/z and intensities of ions detected within the mutant or wildtype data were not observed indicating that N-linked glycosylation did not affect enzyme activity, at least within the time frames investigated.

The products formed by PME action on fully methyl-esterified hexagalacturonate were further characterized by tandem mass spectrometry. Both Y and B ions were observed in the MS/MS spectra for each precursor ion selected from the MS experiments. The presence of these ions allowed for the identification of specific sites that had undergone de-methylation. From these data it was concluded that N-linked glycosylation did not play a role in the mode-of-action of PME toward fully methyl-esterified hexagalacturonate. However, the data indicated that preferential de-esterification of central residues was a common characteristic shared between wild-type and mutant forms as was the inability of mutant or wildtype PMEs to de-methylate the non-reducing residue.

ACKNOWLEGEMENTS

The authors are grateful for the financial support of this research provided by the National Institutes of Health (NIH grant number P41RRR05351) and the National Science Foundation (NSF grant MCB-0115132). This work was presented in preliminary form at the 45th ASMS conference on Mass Spectrometry and Allied Topics, June 2003, Montreal, Canada.

REFERENCES

(1) Kester, H.; et al. *Biochem. J.* **2000**, *346*, 469–474.

(2) Http://www.sbu.ac.uk/water/hypec.html

(3) Warren, M. E.; et al. *Carbohydrate Research* **2002**, *337*, 803-812.

(4) Wen, F.; Yanmin, Z.; Hawes, M. C. *The Plant Cell* **1999**, *11*, 1129-1140.

(5) Sakai, T.; Sakamato, T.; Hallaert, J.; Vandamme, E. J. *Adv. Appl. Microbiol.* **1993**, *39*, 213-250.

(6) Christgau, S.; Kofod, L. V.; Halkier, R.; Andersen, L. N.; Hockauf, M.; Dorreich, K.; Kauppinen, S. *Biochem. J.* **1996**, *319*, 705-712.

(7) Reignault, P.; Mercier, M.; Bompeix, G.; Boccara, M. *Microbiology* **1994**, *140,* 3249-3255.

(8) Plastow, G. S.; *Mol. Microbiol.* **1988**, *2*, 247-254.

(9) Gainvors, A.; Frezier, V.; Lemaresquier, H.; Lequart, C.; Aigle, M.; Belarbi, A. *Yeast* **1994**, *10*, 1311-1319.

(10) Nakagawa, T.; Miyaji, T.; Yurimoto, H.; Sakai, T.; Kato, N.; Tomizuka, N.; *Appl. Environ. Microbiol.* **2000**, *66*, 4253-4257.

(11) Gert-Jan, W. M.; Alebeek, V.; Scherpenzeel, K. V.; Baldwin, G.; Schols, H. A.; Voragen, A. *Biochem. J.* **2003**, *372*, 211-218.

(12) Grasdalen, H.; Andersen, A. K.; Larsen, B. *Carbohydr. Res.* **1996**, *289*, 105-114.

(13) Korner, R.; Limberg, G.; Mikkelsen, J. D.; Roepstorff, J. *Mass Spectrom.* **1998**, *33*, 836-842.

(14) Daas, P. J.; Arisz, P.W.; Schols, H. A.; De Ruiter, G. A.; Voragen, A. G. J. *Anal. Biochem.* **1998**, *257*, 195-202.

(15) Korner, R.; Limberg, G.; Christensen, T.; Mikkelsen, J. D.; Roepstorff, P. *Anal. Chem.* **1999**, *71*, 1421-1427.

(16) El Rayah Ahmed, A.; Labavitch, J. M. *J. Food Biochem.* **1977**, *1*, 361-365.

(17) Domon, B.; and Costello, C.E. *Glycoconjugate J.* **1998**, *5*, 397-409.

(18) Kohn, R.; Dongowski, G.; Bock, W. *Die Nahrung* **1985**, *29*, 75-85.

CHAPTER VI

CONCLUSIONS

(Chapter III) Mice were successfully hyperimmunized resulting in the formation of antibodies in response to the foreign fungal proteins. Western analysis combined with 1D gel electrophoresis was utilized to identify proteins associated with up-regulation of IgE's resulting from hyper-immunization with a *P. Chrysogenum* protein extract. This provided data for the first time that clearly indicated proteins isolated from *P. Chrysogenum* are likely to induce reactions indicative of allergenic models.

Since little genome information was available for *P. chrysogenum*, protein identification was homology based. While several homology based hits have been obtained only one yielded extensive homology, desired for a high-confidence identification. This could be for several reasons. First homology based hits depend on the detection of a peptide that shares a similar sequence region with peptides from an organism other than itself. While there exists the possibility that such an event may occur it is as likely, if not more so, that homologous regions will not be detected within the generated peptides. Additionally, peptide coverage from gels are lower than from solutions making the detection of homologous peptides even less likely. However, more efficient digestions and detection of peptides will increase the likelihood of successful identification of proteins utilizing homology based protein searching. Second, there could be errors in the in-silico generation of a sequence from MS/MS data. As an alternative de Novo sequencing was also employed to more accurately discern peptide sequences prior to database searching, especially in situations where no homology 'hits' were obtained. Comparisons of *de Novo* and *in Silico* sequences showed that computer generated data were accurate for those peptides searched.

Given the protein extract contains a high number of unique proteins, mis-identification of proteins as the result of multiple proteins occupying a single band is also an issue. Two dimensional gel electrophoresis provides far superior separation of proteins, greatly reducing the likelihood that multiple proteins could occupy a single spot/band. Also, due to the nature by which proteins are separated in 2D analysis the proteins are more 'concentrated' yielding the likelihood of detecting additional proteins associated with allergenic response.

Thus, future work might center on the generation of 2D 'western' maps as opposed to 1d maps. Subsequent in-gel digestion and MS/MS analysis should again be utilized to identify proteins associated with IgE up-regulation.

(Chapter IV) The mass differences detected by MALDI mass spectrometry can be ascribed to the presence of N-linked and O-linked glycans in the four recombinant pectate lyases studied. Part of the mass difference associated with the Isoform I recombinant proteins was due to the presence of an N-linked high mannose structure at residue 64. This was confirmed by lectin analysis, periodate oxidation chemistry, and anion exchange chromatography. Based on electrospray mass spectrometry experiments, utilizing a Q-TOF II, the molecular mass of the N-linked glycan was determined to be between 800 and 2000 Da - due to heterogeneity associated with N-linked carbohydrates. The average mass of the N-linked glycan was estimated to be ~ 1200 Da, the observed average mass difference between the isoform I and isoform II recombinant proteins (MALDI-TOF data). Since the isoform II proteins were shown to not contain an N-linked glycan the difference between the calculated and expected mass with these recombinant proteins must be ascribed to other post-translational modifications. Also, while an N-linked glycan accounts for 1200 Da of the additional mass observed for the two isoform I proteins, a 2400 Da mass difference was measured when compared to the mass based on the amino acid sequence. Thus, like the isoform I proteins, the isoform II proteins have additional mass associated with other types of post-translational modifications.

IMAC extractions and phosphate ELISA excluded phosphorylation as a source of the additional mass. However, based on results from periodate oxidation, ethylamine labeling, and anion exchange chromatography, O-linked glycosylation accounted for a portion of the additional mass observed in all four recombinant proteins. Ethylamine labeling yielded two tagged peptides (AA218-233 and AA197-217). Collectively these two peptides contain four serines and two threonines, or six potential sites for O-linked glycosylation. AA197-217 labeled with two ethylamines. AA218-213 labeled with three ethylamines. Assuming all five sites are occupied by a single carbohydrate residue (mannose) the combined molecular mass due to O-glycosylation is 810 Da, slightly lower than the expected 1200 Da. Alternatively, labeling excusively with GlcNAC would result in a total mass of 1020 Da. Again, this mass is slightly lower than the 1200 Da based on MALDI Data. However, should two or more of the sites be occupied with di-saccharides the mass due to O-linked glycosylation falls into better agreement with the 1200 Da mass difference observed (based on MALDI data).

(Chapter V) MS analysis revealed that the reaction of substrate with each PME produced multiple compounds with varying degrees of esterification. At increased incubation times, the extent of de-esterification increased. However, variations in the m/z and intensities of ions detected within the mutant or wildtype data were not observed indicating that N-linked glycosylation did not affect enzyme activity, at least within the time frames investigated.

The products formed by PME action on fully methyl-esterified hexagalacturonate were further characterized by tandem mass spectrometry. Both Y and B ions were observed in the MS/MS spectra for each precursor ion selected from the MS experiments. The presence of these ions allowed for the identification of specific sites that had undergone de-methylation. From these data it was concluded that N-linked glycosylation did not play a role in the mode-of-action of PME toward fully methyl-esterified hexagalacturonate. However, the data indicated that preferential de-esterification of central residues was a common characteristic shared between wild-type and mutant forms as was the inability of mutant or wildtype PMEs to de-methylate the non-reducing residue.