Structural Characterization of Protein Glycosylation Utilizing Fragmentation from Gas Phase Ion-Electron Reactions and Vibrational Excitation

by

Julie Taptim Adamson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
in The University of Michigan
2008

Doctoral Committee:

Assistant Professor Kristina I. Håkansson, Chair
Professor Daniel M. Burns
Professor Robert T. Kennedy
Professor David M. Lubman

To family

Acknowledgements

The work presented in this dissertation was directed under the guidance of my advisor, Professor Kristina Håkansson. Throughout my graduate career, she has continually provided me with valuable advice and support to help me achieve my goals. She approaches research with an enthusiasm, dedication, and intelligence that is an inspiration to all of her students.

I would like to thank all the members of my graduate committee, including Professor Daniel Burns, Professor David Lubman, and Professor Robert Kennedy for their valuable advice and suggestions during my oral examination, data meeting, and on my dissertation. I would also like to thank my collaborators, including Professor Irwin Goldstein, Professor Harry Winter, Professor Diane Simeone, and Stephanie Laurinec.

Financial support from the Searle Scholars Program, Eli Lilly & Company, a Dow Corning Assistant Professorship, an Eastman Chemical Company Summer Fellowship, a Rackham Predoctoral Fellowship, and the University of Michigan also needs to be acknowledged.

Current and former members of the Håkansson lab helped contribute to this dissertation, through their support and friendship. Special thanks go to my labmate Natasa Kalli, for enduring sitting next to me the past four years.

Of course, thanks must go to my friends and family. I would not be here today without the support of my parents. I would also like to thank Dr. Rick White, who encouraged me to pursue a PhD in analytical chemistry. I appreciate all of my friends who have supported me throughout this process, especially Scott C. Walker, for his advice and support.

Julie T. Adamson
April 28, 2008

Table of Contents

List of Figures

Figure

List of Schemes

Scheme

List of Appendices

Appendix

List of Abbreviations

AI-ECD	Activated ion electron capture dissociation
BIRD	Blackbody infrared radiative dissociation
CAD	Collision activated dissociation
ECD	Electron capture dissociation
EDD	Electron detachment dissociation
ESI	Electrospray ionization
ETD	Electron transfer dissociation
FT-ICR MS	Fourier transform ion cyclotron resonance mass spectrometry
GAG	Glycosaminoglycan
heCAD	High energy collision activated dissociation
HPLC	High performance liquid chromatography
IRMPD	Infrared multiphoton dissociation
LC/MS	Liquid chromatography/mass spectrometry
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
MS^n	Tandem mass spectrometry
MS/MS	Tandem mass spectrometry
m/z	Mass-to-charge ratio
NMR	Nuclear magnetic resonance
PTM	Post-translational modification
Q	Quadrupole
SID	Surface induced dissociation
S/N	Signal-to-noise ratio
SPE	Solid-phase extraction
SORI-CAD	Sustained off-resonance irradiation collision activated dissociation
T	Tesla, a unit of magnetic field

Abstract

Structural Characterization of Protein Glycosylation Utilizing Fragmentation from Gas Phase Ion-Electron Reactions and Vibrational Excitation

by

Julie Taptim Adamson

Chair: Kristina I. Håkansson

Glycosylation is one of the most common post-translational modifications found in eukaryotes. Glycan structural characterization is a difficult undertaking, because full characterization demands knowledge of saccharide linkage, branching, sequence, glycosylation location, heterogeneity, and occupancy. Mass spectrometry, with its sensitivity and tandem mass spectrometric capabilities, is an extremely valuable tool in this endeavor.

In this thesis, the utility of gas-phase ion electron reactions, including electron capture dissociation (ECD) and electron detachment dissociation (EDD), for glycosylation structural characterization is explored. Both ECD and EDD have been shown to be valuable techniques for the characterization of post-translationally modified peptides, including both phospho- and glycopeptides. However, further applications of

these techniques remain to be explored.

The combination of vibrational excitation (here infrared multiphoton dissociation (IRMPD)) and ion-electron based fragmentation (ECD) for *de novo* sequencing of a lectin with rare carbohydrate binding specificity is demonstrated. Despite the disadvantages associated with IRMPD and ECD, when used in conjunction these techniques proved to be a powerful tool for sequencing purposes. Over 75% of the protein was sequenced with the combination of vibrational excitation and ion-electron based fragmentation.

IRMPD and ECD are also utilized for the characterization of high-mannose type glycopeptides. This category of glycopeptides had not been previously examined with these techniques. IRMPD of high-mannose type glycopeptides is shown to often result in a mixture of glycan and peptide backbone cleavage, while ECD results in exclusively peptide cleavage (allowing for glycosylation localization).

Ion-electron based fragmentation for oligosaccharide structural characterization is investigated. ECD of metal–adducted oligosaccharides is shown to result in complementary structural information compared to IRMPD, and provide more cross-ring fragmentation than ECD of protonated species. Neutral and sialylated oligosaccharides are also examined with vibrational excitation and EDD. EDD often results in more cross-ring fragmentation compared to vibrational excitation, thus providing additional structural information.

Strategies for examining pancreatic cancer associated O-linked glycans are explored. A protocol is developed for examining oligosaccharides in conditioned media from pancreatic cancer cell lines.

Chapter 1

Introduction

1.1 Background

Since the introduction of electrospray ionization (ESI)[1,2] and matrix-assisted laser desorption/ionization (MALDI),[3,4] these "soft" ionization techniques have dramatically changed the field of mass spectrometry (MS). ESI and MALDI allow for the ionization of large nonvolatile molecules, while yielding little or no fragmentation. They are also applicable to a wide variety of biomacromolecules, ranging from peptides, proteins, oligonucleotides, oligosaccharides, polymers, to lipids. Currently, ESI and MALDI are the most commonly used ionization techniques for biomolecules.

The unique ability of ESI and MALDI to generate gas-phase ions from proteins and peptides dramatically changed the field of protein analysis when they were introduced in the late 1980s. Due to the sensitivity, resolution, and mass accuracy provided by mass spectrometry, along with the ability to generate protein and peptide ions from ESI and MALDI, mass spectrometry has become an essential tool in proteomics. MS-based proteomics is made possible by the availability of gene and genome databases, which permit the identification of proteins by matching peptide masses from a proteolytic digest or peptide fragmentation data to proteins

represented in available databases.[5-12] Peptide fragmentation is accomplished through tandem mass spectrometry (MS/MS or MS^n) and provides primary structure information; MS/MS techniques are described in detail in Section 1.3.3. In addition to protein identification and primary structure determination, mass spectrometry is also playing an increased role in characterizing protein higher order structure and non-covalent protein complexes, and also applications in imaging.[13-15]

One major challenge in proteomics is the characterization of post-translational modifications (PTMs). The proteome is highly dependent on the physiological conditions encountered by the cell and is by no means a direct translation of gene sequences to protein sequences, but instead is highly complex and dynamic. A protein often exists in several diverse states within a cell, largely due to different splice variants and post-translational modifications that can occur at multiple positions within the protein. PTMs are chemical modifications to proteins including proteolytic cleavage, addition of a functional group, or the formation of inter- and intra-peptidic linkages (including disulfide bridges).[16] PTMs on a protein may affect properties such as protein localization, stability, activity, turnover, and interactions with other proteins.[17] Common PTMs include phosphorylation, acetylation, methylation, and glycosylation. To date, over 300 varieties of PTMs are known and new ones are currently being discovered.[18]

Among PTMs, glycosylation is one of the most prevalent in eukaryotes, with a previous survey indicating that at least 50% of all proteins are glycosylated.[19] Glycoproteins have been found to play several key roles in biological systems including: cell-cell adhesion, cell-extra cellular membrane adhesion, folding and secretion, fertilization, glycoprotein targeting, and immune defense.[20,21] Glycan expression is

affected by cellular conditions, and it has often been shown that glycan structures change with the onset of cancer and inflammation. With cancer, glycosylation alterations may affect growth, differentiation, transformation, adhesion, metastasis, and immune surveillance of the tumor.[22] For several forms of cancer, such as ovarian and prostate, specific changes in glycosylation profiles indicate their utility as diagnostic tools.[23] Based on these changes, it may be possible to utilize the immune system to target cells based on abnormal glycosylation patterns. The development of glycan-based vaccines that target cancer-associated glycans is a rapidly developing field.

Unlike most other biomolecules, carbohydrate synthesis is not template-driven, an attribute which has encumbered biologists in their efforts to decipher the molecular details of carbohydrate synthesis and functionality. The attachment and processing of carbohydrates in eukaryotes occurs through complex pathways within the ER and Golgi apparatus, and are carried out by glycosyltransferases and glycosidases. The majority of glycosylation sites are found on plasma-membrane and secreted proteins, but simple glycosylated structures can also be observed on cytoplasmic and nucleoplasmic proteins.[24] There are two key types of protein glycosylation, which involve either covalent attachment of an oligosaccharide through the oxygen in serine or threonine (*O*-glycosylation) or through nitrogen in asparagine (*N*-glycosylation). With *N*-glycosylation, a common pentasaccharide core is attached to Asn in the consensus sequence Asn-X-Ser/Thr, where X may be any amino acid except Pro. There is no consensus sequence or single common saccharide core region for *O*-glycosylation, however the majority of eukaryotic O-linked glycans contain *N*-acetylgalactosamine (GalNAc) linked to serine or threonine. There are three categories of N-linked glycans,

which are defined by the saccharides attached to the common pentasaccharide core (see Figure 1.1). High-mannose glycans only contain mannose saccharides attached to the common core. With hybrid type structures, *N*-acetylglucosamine is attached to one core mannose saccharide. Both core mannose saccharides are linked to *N*-acetylglucosamines in the case of complex *N*-glycans. Complex *N*-glycans are the most common glycan type found in vertebrates.[20] Figure 1.1 illustrates representative structures of *N*-glycans, common *O*-glycan cores, and saccharide chemical structures.

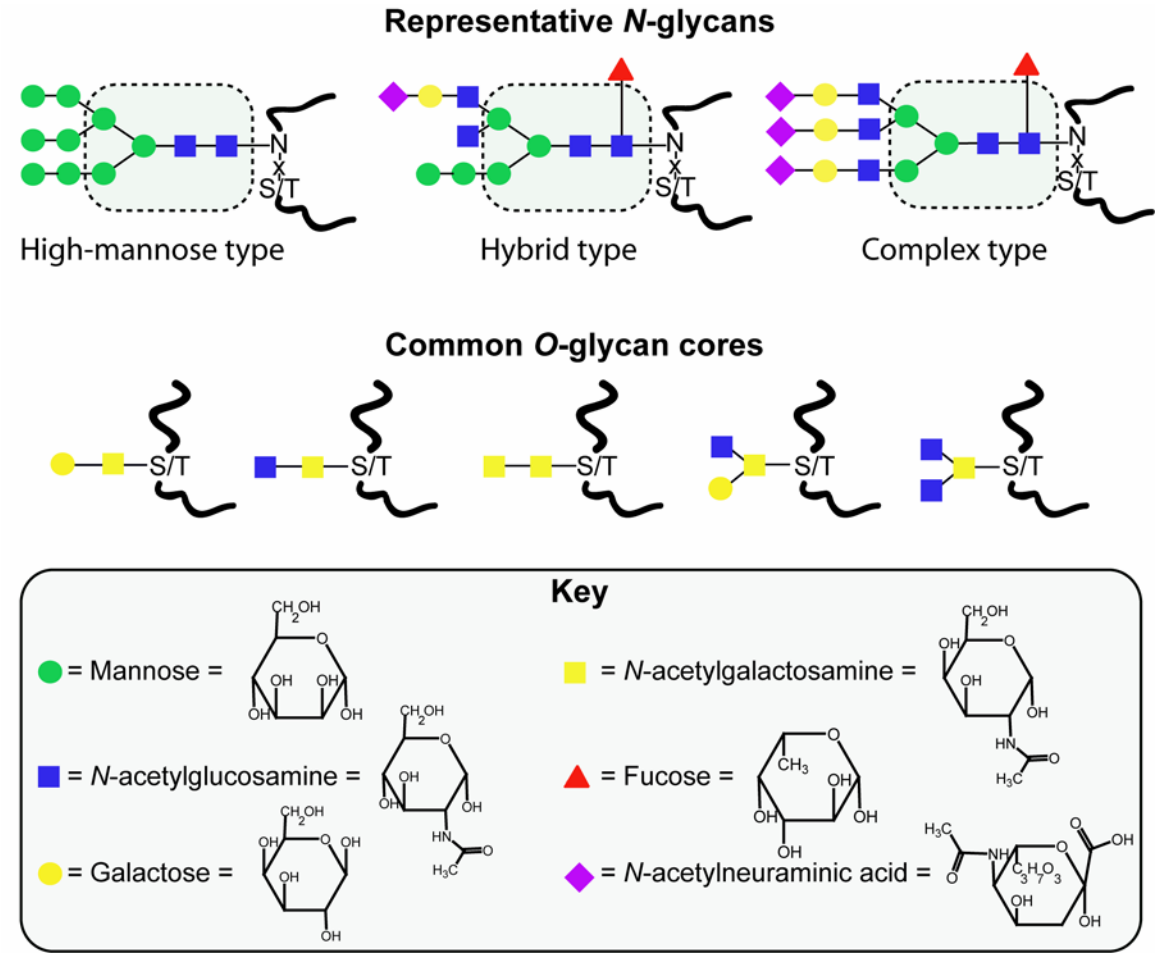


Figure 1.1. Representative *N*-glycans and common *O*-glycan cores. The common pentasaccharide N-linked core is indicated with dashed lines. The saccharides which are attached to this core define whether the glycan is high-mannose, hybrid, or complex type. The majority of O-linked structures have *N*-acetylgalactosamine linked to serine or threonine.

Carbohydrates, or glycans, are unlike linear biomolecules such as DNA, RNA, and proteins in that they can form complicated highly branched structures, containing multiple antennae, where saccharide units may be connected to each other through a variety of linkage types. The number of antennae for N-linked glycans typically ranges from two (biantennary) to four (tetraantennary),[21] whereas O-linked glycans tend to be less branched. Several different enzymatic reactions, as carried out by glycosyltransferases and glycosidases, are responsible for forming the carbohydrate portion of glycoproteins. These enzymatic reactions may not go to completion, which results in a wide array of glycoprotein forms (glycoforms). Consequently, glycosylation sites within a protein may not be fully occupied and may contain a wide range of glycans at an individual site. Due to their vast complexity, the determination of glycan profiles and glycan structures is a difficult undertaking, and no single technique is capable of providing a complete profile. Complete glycan structural characterization demands knowledge of saccharide linkage, branching, sequence, glycosylation location, heterogeneity, and occupancy.[25]

A plethora of techniques have been utilized for glycosylation characterization, and are discussed in Section 1.2. Mass spectrometry, often in combination with separation methodologies, is one of the most versatile techniques for characterizing glycans and glycoproteins due to its sensitivity and low sample consumption. Tandem mass spectrometry in particular, has proven to be a valuable tool for glycoprotein analysis and the characterization of glycopeptide and saccharide structures.[26,27] Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS), with its ultrahigh resolution

and mass accuracy is an especially powerful tool for glycosylation characterization. FT-ICR mass spectrometry is discussed in Section 1.3.

1.2 Current Methods for Glycosylation Characterization

As previously discussed, complete glycosylation characterization demands knowledge of saccharide linkage, branching, sequence, glycosylation location, heterogeneity, and occupancy. There are three main approaches for glycosylation characterization. One is to analyze the intact glycoprotein, following glycoprotein purification. Another common approach is investigation at the glycopeptide level, after enzymatic digestion of the intact glycoprotein. Alternatively, glycans can be removed from the protein and both species characterized separately.

The first step towards glycoprotein characterization is the initial isolation of the glycoprotein of interest. Glycoproteins may be found as either soluble or membrane-bound molecules, and thus isolation strategies will differ depending on where the protein is located in the cell. Both electrophoretic separations in gels and liquid chromatographic separations are commonly used for glycoprotein isolation. Liquid chromatographic separations include ion exchange, size exclusion, partition, hydrophobic interaction, dye ligand, and affinity chromatography.[28] Lectin affinity chromatography is an especially powerful technique. A wide variety of lectins from plant and animal sources, with different binding specificities, are now available. In addition, capillary electrophoresis has been reported in several instances for the separation of glycoprotein isoforms as well as glycopeptides, and oligosaccharides.[25,29]

To some extent, conventional MALDI-TOF-MS and ESI-MS are capable of characterizing intact glycoproteins.[25,28,30] Glycoproteins can often be resolved to their

individual glycoforms depending on the molecular weight and extent of glycosylation. Typically, intact glycoprotein analysis with mass spectrometry has been limited to smaller glycoproteins, less than 40 kDa.[28] In addition to mass spectrometry, X-ray crystallography is another powerful tool for whole protein structural analysis. However, X-ray crystallography is not usually amenable to glycoprotein characterization, because part or all of the glycan portion is not observed in the high-resolution electron density map. This is often due to glycan structure heterogeneity and the flexibility of the attached saccharides.

To determine glycosylation locations and heterogeneity, investigation of protein glycosylation at the glycopeptide level is another general approach. It is often difficult to detect glycopeptides in a complex protein digestion by mass spectrometry, partly due to the low sensitivity of detection of glycopeptides due to glycan heterogeneity. Also, acidic glycans may result in glycopeptide signal suppression, since positive ion mode mass spectrometry is typically used in proteomic applications. For these reasons, it is often necessary to separate glycopeptide and peptide mixtures with liquid chromatography prior to mass spectrometric analysis. A variety of strategies have also been reported for glycopeptide enrichment, and often rely upon lectin affinity chromatography. Alternatively, NMR spectroscopy has been utilized for glycopeptide characterization.[31] NMR analysis of glycopeptides is more difficult compared to NMR of oligosaccharides, and generally leads to less precise structural interpretation.[32]

A final strategy for glycosylation characterization is the chemical or enzymatic removal of glycans from the protein, and subsequent analysis of the glycan mixture. A disadvantage of this technique is that information regarding glycosylation location and

heterogeneity is lost. Chemically, both N- and O-linked glycans can be released from glycoproteins with hydrazine. There are several disadvantages to this procedure, including cleavage of peptide bonds and subsequent destruction of the protein, hydrolysis of saccharide acyl groups (requiring a reacetylation step), and incomplete removal of intermediates from the reducing terminus of the released glycans.[30] O-linked glycans are also frequently removed via β-elimination. A disadvantage of β-elimination is that the reducing terminus of the oligosaccharide is reduced, which typically prevents subsequent labeling (which may be needed for the attachment of a chromophore or other tag). N-linked glycans can also be released enzymatically, with several glycosidases commercially available. Unfortunately, no reliable enzymes are available for the enzymatic release of O-linked glycans.

Following glycan release, chromatographic and sometimes electromigration techniques may be utilized for distinguishing isomeric glycan structures (a feat which is difficult with MS). In particular, high-performance anion-exchange chromatography with pulsed-amperometric detection is an extremely powerful technique for oligosaccharide analysis, due to its separation of anomeric, structural, linkage and branch isomers. However, this form of chromatography is not easily coupled to MS due to its high salt requirements. Hydrophilic-interaction chromatography, reverse-phase chromatography, and graphitized carbon chromatography are also commonly used for glycan separation and are much more compatible with MS. With these forms of chromatography, derivatization of the glycan is often implemented in order to improve detection with the attachment of a chromophore. This derivatization often improves separation with reverse-phase LC. In addition to mass spectrometry, nuclear magnetic

resonance (NMR) spectroscopy is a customary technique for oligosaccharide structural characterization.[32] NMR provides unambiguous structural assignments, which is difficult with mass spectrometry. However, a chief disadvantage of NMR compared to mass spectrometry is sensitivity, which often makes NMR an unsuitable technique for oligosaccharide analysis in proteomic-type applications.

1.3 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

1.3.1 Overview

Fourier transform ion cyclotron resonance mass spectrometry was first introduced by Comisarow and Marshall in 1974.[33] There are several advantages for using FT-ICR MS for biomolecular structural characterization, including the capability of providing extremely high mass accuracy, on the order of low to sub-ppm levels. Another advantage offered by FT-ICR MS is its tremendous resolving power: For large biomolecules, resolving power in excess of 10^6 has been obtained. This high performance is partly due to the numerous cycles of cyclotron motion in the analyzer cell (resulting in long-lasting signal), and because cyclotron frequency is not affected by a spread in kinetic energy. Unit mass resolution, which allows for the determination of the charge state of a molecule, based upon the *m/z* spacing between isotopic peaks, has been achieved for proteins as large as 112 kDa with FT-ICR MS.[34]

There are several key features common to all FT-ICR instruments, including an ionization source, ultrahigh vacuum system, magnet, analyzer cell (ICR cell), and data acquisition system.[35] The most commonly used ionization technique combined with FT-ICR MS for biomolecule analysis is ESI.[36,37] One characteristic of ESI is its ability to generate multiply-charged ions, which is beneficial because FT-ICR MS resolving power

decreases linearly with increasing *m/z*.[38] Furthermore, mass resolving power, mass accuracy, and limits of detection are also enhanced in the low *m/z* region. Contrary to ESI, MALDI typically produces singly-charged ions, resulting in larger *m/z* values. MALDI can also be coupled to FT-ICR MS, and constitutes a powerful technique for the characterization of moderate size biomolecules (< 3-5 kDa).[39-45] An ultrahigh vacuum system is necessary in an FT-ICR mass spectrometer because ion detection needs to occur in a low pressure environment, typically in the region of 10^{-9} – 10^{-10} Torr.[35,38] Superconducting magnets of 3 to 15 Tesla are typically utilized in FT-ICR mass spectrometers. Increasing magnetic field strength offers several improved parameters, including mass resolving power, upper mass limit, mass accuracy, and signal-to-noise ratio.[46] Ion motion within the ICR cell, which is located within the magnet, is the governing principle behind FT-ICR MS and will be discussed shortly in the next section.

1.3.2 Operating Principles

Under the influence of a spatially homogenous magnetic field, $\mathbf{B} = -B_0\mathbf{z}$ where B_0 is magnetic field strength, an ion experiences a force that is perpendicular to both the direction of its velocity and the magnetic field:

$$\mathbf{F} = \text{mass} \bullet \mathbf{acceleration} = q\mathbf{v} \times \mathbf{B} \qquad 1\text{-}1$$

where q is charge and **v** is velocity. This force, called the Lorentz force, bends ions into a circular orbit which is perpendicular to the magnetic field axis. If v_{xy} denotes ion velocity in the *xy* plane (perpendicular to the magnetic field axis) and because angular acceleration is v^2_{xy}/r, equation 1-1 can be re-written as:

$$\frac{mv^2_{xy}}{r} = qv_{xy}B_0 \qquad 1\text{-}2$$

Here, m is ionic mass and r is the radius of the circular orbit. If these ions do not change speed (no collisions), equation 1-2 can be re-written as the basic cyclotron equation:

$$\omega = \frac{qB_0}{m} \text{ (S.I. units)} \qquad 1\text{-}3$$

where ω is angular velocity. By dividing equation 1-3 by 2π, an ion's cyclotron frequency, ν (in Hz), can be determined. The unique feature of an ion's cyclotron frequency is that it is independent of ion velocity and kinetic energy, eliminating the need for "focusing" ions which is often necessary in other types of mass spectrometers.[35,38]

The magnetic field in an FT-ICR instrument traps ions radially in the ICR cell, but does not prevent them from leaving the cell along the axial direction. Trapping plates are mounted perpendicular to the magnetic field with an applied potential. This configuration creates a potential well in the analyzer cell, causing ions to oscillate between the two endplates.

The initial ion cyclotron radius in the ICR cell is small (<1 mm) compared to the analyzer cell dimensions (on the order of 10 cm). Ion excitation is necessary in order to accelerate ions to a larger detectable cyclotron radius. Such acceleration is accomplished by applying differential rf voltages to a pair of opposing excitation plates in the analyzer cell (shown in Figure 1.2). Those ions which are in resonance with the frequencies of the applied rf voltage gain kinetic energy and spiral outwards to a larger cyclotron radius. Ions are excited coherently as a tight packet, commonly referred to as an ion cloud. The orbiting ion cloud induces an alternating current on the two opposing detection plates, which have parallel resistance and capacitance components and connects to a detection preamplifier.[38] The image current is measured as a time-domain signal and is composed of all the frequencies of the ion clouds in the analyzer cell. By applying a Fourier

transform to the time domain signal, the frequency components of the signal can be determined. The inherent advantages of ion detection in an FT- ICR MS are that it is non-destructive and a wide range of *m/z* ratios can be detected simultaneously.

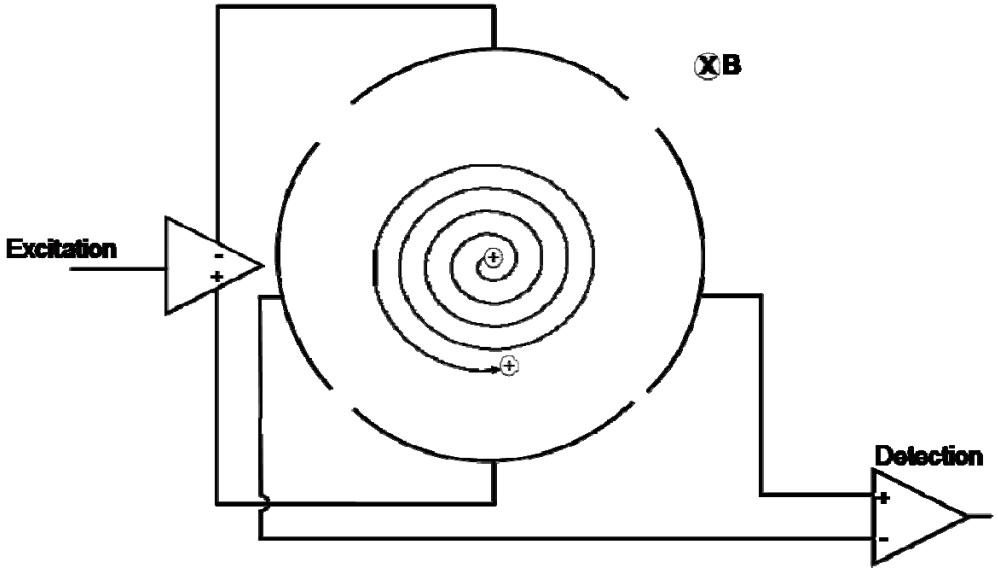


Figure 1.2. Ion excitation and detection in an ICR analyzer cell. A sinusoidal voltage is applied to a pair of opposing excitation plates. Ions in resonance with the applied frequencies spiral outwards to a larger cyclotron radius. An image current is induced in an opposing pair of detection plates, which are connected to an amplifier.

1.3.3 Tandem Mass Spectrometry

There are a wide variety of fragmentation techniques available in FT-ICR MS. These include sustained off-resonance irradiation collision activated dissociation (SORI-CAD),[47] infrared multiphoton dissociation (IRMPD),[48,49] blackbody infrared radiative dissociation (BIRD),[50,51] surface induced dissociation (SID),[52-54] electron capture dissociation (ECD),[55] and electron detachment dissociation (EDD).[56] CAD is the most common ion activation method in modern day mass spectrometers.[57-60] Beam-type CAD can be achieved with hybrid FT-ICR mass spectrometers, such as the one used in this

work. Fragmentation techniques used in this dissertation include CAD, IRMPD, ECD, and EDD (see Figure 1.3).

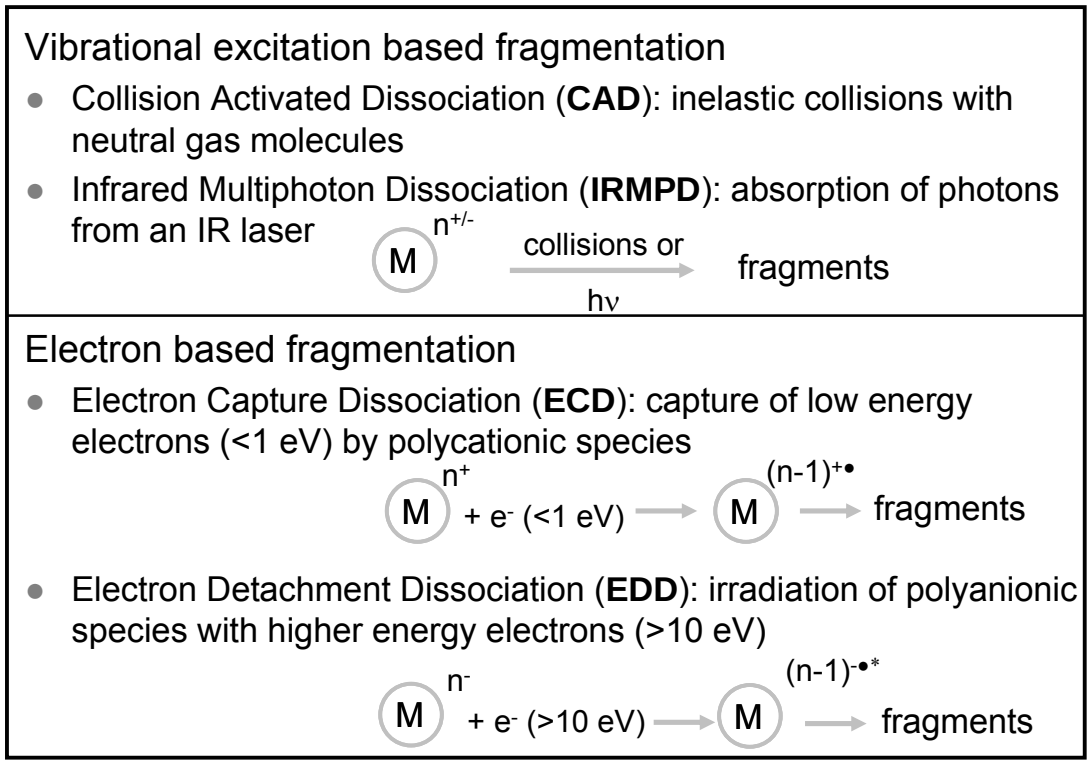


Figure 1.3. A summary of the fragmentation techniques used throughout this dissertation. CAD and IRMPD are vibrational excitation based fragmentation techniques, while ECD and EDD are electron based. "M" indicates the precursor ion. For vibrational excitation techniques, "n" can be any value and both cations and anions may be examined. For both ECD and EDD, "n" must be greater than two because electron capture and electron detachment will result in charge reduction of the precursor ion. Charge reduced radical species will undergo further fragmentation.

Vibrational Excitation Based Fragmentation. CAD, SORI-CAD, IRMPD, and BIRD can be categorized as vibrational excitation based fragmentation techniques. Fragmentation products from these "slow-heating" techniques form by the lowest energy pathways.[61] For peptides, this slow internal temperature increase results in preferential C-N bond cleavage, producing *b*- and *y*-type ions (see Figure 1.4).[62] The mobile proton model is the generally accepted mechanism describing peptide backbone cleavage.[57,63] For modified peptides, the application of vibrational excitation based fragmentation

techniques generally results in cleavage of the modification, such as glycosylation. Vibrational excitation techniques (typically CAD and IRMPD) are also frequently utilized for the structural characterization of oligosaccharides, and may result in several product ion types as illustrated in Figure 1.5. There are two types of cleavages observed for oligosaccharides; glycosidic cleavages correspond to cleavages between saccharide residues (B, C, Y, and Z-type ions) whereas cross-ring cleavages occur across the saccharide ring (A and X-type ions).

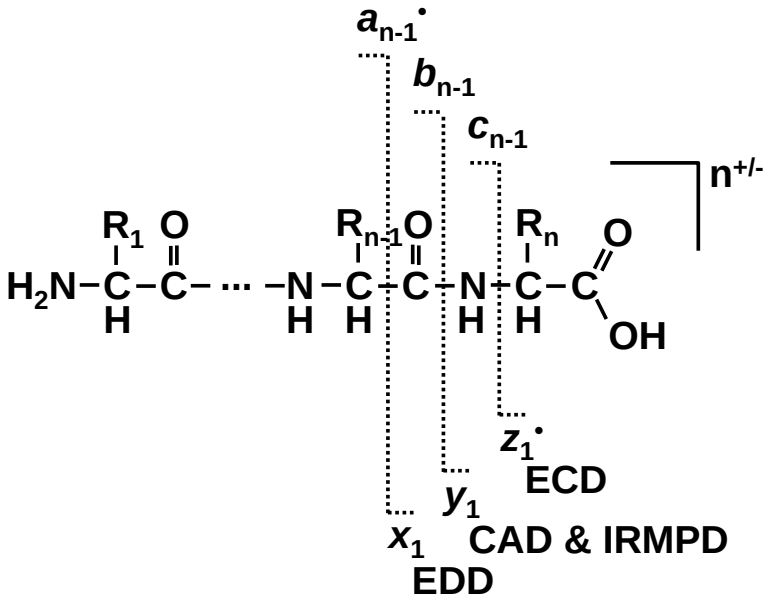


Figure 1.4. Peptide cleavages observed following vibrational excitation (CAD and IRMPD) and electron-based fragmentation techniques (ECD and EDD). Vibrational excitation results in preferential cleavage of the peptide C-N bond, yielding *b* and *y*-type ions. ECD of multiply charged peptide cations results in preferential cleavage of the N-C_α bond, yielding *c* and radical *z*-type ions. EDD of multiply charged peptide anions results in preferential cleavage of the C_α-C bond, yielding radial *a* and *x* type ions.

During CAD, ions are activated through inelastic collisions with neutral gas molecules, which results in a conversion of translational energy into internal energy (see Figure 1.3). Internal energy is redistributed throughout the ion, and if it exceeds its threshold energy dissociation occurs. SORI-CAD is similar to CAD, except during the

fragmentation process an excitation frequency is applied which is slightly off-resonance from the precursor ion's cyclotron frequency, resulting in an alternating increase and decrease in ion cyclotron radius as ions collide with inert gas molecules pulsed into the analyzer cell.[64,65] Another commonly used vibrational excitation technique is IRMPD, which also results in preferential C-N bond cleavage. IRMPD requires the absorption of an extensive number of IR photons before dissociation occurs (Figure 1.3). Typically, 10.6 μm CO_2 lasers are used for IRMPD. BIRD is a less commonly used fragmentation technique, and is rarely applied in proteomic applications with FT-ICR MS.

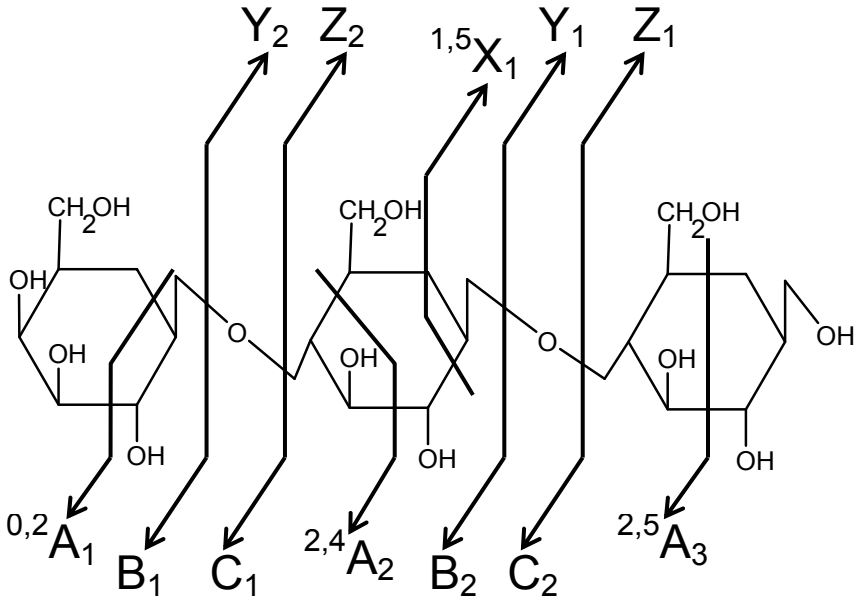


Figure 1.5. Oligosaccharide fragmentation nomenclature.[66] Glycosidic cleavages occur between saccharides (B, C, Y, and Z ions) while cross-ring cleavages are due to the cleavage of two bonds (A and X ions)

Electron Based Fragmentation. Electron capture dissociation and electron detachment dissociation are relatively recent ion activation techniques, and rely upon ion-electron reactions to induce fragmentation. The application of ion-electron reactions to biomolecules typically results in different fragmentation patterns compared to vibrational excitation and thus these techniques are complementary to most other MS/MS techniques.

Electron capture and detachment triggers gas-phase radical ion chemistry, however the precise mechanisms of these ion-electron reactions have yet to be fully elucidated, and are still widely debated.[56,67-74]

ECD was first applied in 1998 by Zubarev and co-workers for the dissociation of multiply charged protein cations.[55,75-77] This technique is based upon the dissociative recombination of polycationic molecules with low-energy electrons, typically less than 1 eV (see Figure 1.3). Unlike traditional fragmentation techniques, which generally cleave the peptide amide bond to yield *b* and *y* fragments, ECD is dominated by N-C_α bond cleavage yielding predominantly *c* and radical *z*-type ions (see Figure 1.4).[77] ECD cleavage tends to be more random than vibrational excitation techniques, and more extensive sequence coverage is usually achieved for peptides and proteins. When applied to modified peptides, ECD tends to result in extensive backbone cleavage without the loss of modifications, such as phosphorylation[78,79] and glycosylation.[80-83] The combination of vibrational excitation and ECD thus provides complementary structural information for modified peptides.

Electron detachment dissociation was introduced in 2001 as a fragmentation technique for polyanions.[56] In EDD, polyanions are irradiated with >10 eV electrons, resulting in electron detachment and subsequent product ions (see Figure 1.3). EDD of peptides yields predominantly radical *a* ions and *x* ions due to C_α-C bond cleavage (see Figure 1.4). Fragment ions generated by EDD have been shown to retain modifications such as sulfation and phosphorylation, which indicates that EDD also shows the potential for peptide sequencing and for mapping PTMs.[56,84,85]

Conversely, one chief disadvantage of ECD and EDD compared to other ion activation techniques is that fragmentation efficiency is relatively low, with typically less than 30% of precursor ions being converted to product ions.[86] As a consequence, a high signal-to-noise (S/N) ratio is required for precursor ions. In addition, ECD and EDD are difficult to implement in mass spectrometers other than FT-ICR instruments because of the trapping requirements; ions and electrons must be trapped simultaneously for dissociation to occur. Despite these disadvantages, the applications of ECD and EDD for biomolecular structural characterization are steadily increasing. In addition to PTM characterization, ECD has been utilized for *de novo* sequencing of peptides,[87] top-down protein analysis,[88,89] characterizing protein-ligand non-covalent interactions,[90] and oligosaccharide structural characterization.[91] EDD has also been applied to oligonucleotides,[92,93] gangliosides,[94] and glycosaminoglycans (GAGs).[95,96]

1.3.4. Experimental Setup

A schematic diagram of the 7T ESI quadrupole-FT-ICR mass spectrometer (Bruker Daltonics, Billerica, MA) used in this work is shown in Figure 1.6. The instrument is equipped with either a regular or nano-electrospray ion source, and a dual ion funnel inlet. The ion funnels consist of a series of ring electrodes with applied rf and dc electric potentials. Smith *et. al.* first introduced an ESI ion funnel in 1998 as a means to more effectively focus and transmit ions from regions of higher pressure to lower pressure.[97] Our instrument is also equipped with two hexapoles for external ion accumulation of electrosprayed ions.[98] The quadrupole and second hexapole allow for mass-selective ion accumulation in addition to external CAD. The dynamic range and sensitivity of an FT-ICR mass spectrometer is generally enhanced when external ion

selection is utilized prior to ion accumulation in the ICR cell.[99] Furthermore, ESI is a continuous ion source, whereas FT-ICR MS is a pulsed detector. External ion accumulation is utilized to keep the duty cycle high. To transport ions into the ICR cell and overcome the magnetic mirror effect, the instrument contains a series of ion transfer optics. The ICR cell in this instrument is cylindrical, which is a relatively common geometry. Located behind the ICR cell is a hollow cathode (HeatWave, Watsonville, CA) and 10.6 μm CO_2 laser (Synrad, Mukilteo, WA), which are used to fragment ions by ECD, EDD, and IRMPD.

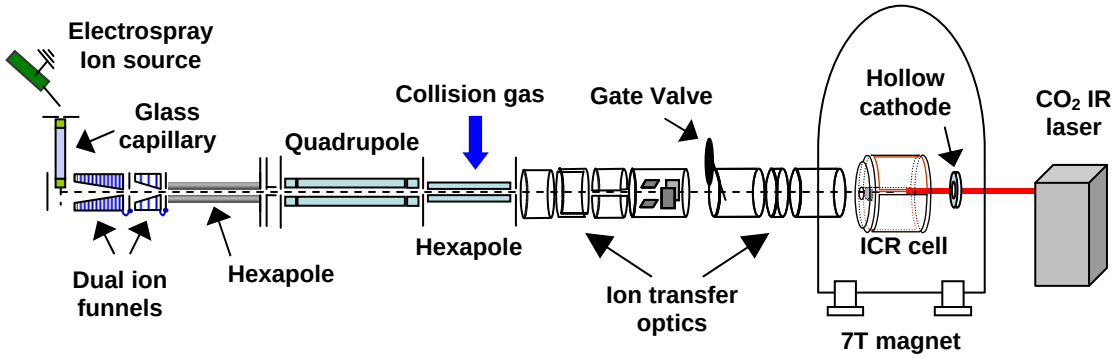


Figure 1.6. A schematic diagram of a 7 T Q-FT-ICR mass spectrometer used throughout this work. This instrument is equipped with an electrospray source, dual ion funnels, a quadrupole for mass selective accumulation, hexapole for external CAD, hollow cathode for gas phase ion-electron reactions, and a 10.6 μm CO_2 laser for IRMPD.

1.4. Dissertation Overview

The research presented in this dissertation focuses on the utilization of ion-electron based fragmentation techniques and vibrational excitation for glycosylation characterization with an FT-ICR mass spectrometer. As discussed in Section 1.3.3, both ECD and EDD have shown utility for the characterization of post-translationally modified peptides, including both phosphorylation and glycosylation. However, further applications of these techniques remain to be explored. In the following chapters, the

utility of ion-electron reactions are investigated for *de novo* sequencing of a carbohydrate binding protein (Chapter 2), characterization of high-mannose type glycopeptides (Chapter 3), and the structural characterization of oligosaccharides (Chapters 4 and 5). Chapter 6 focuses on strategies for the analysis of pancreatic cancer associated O-linked glycans.

Chapter 2 explores the combination of vibrational excitation and ECD for *de novo* sequencing of a lectin from the mushroom species *Lyophyllum decastes*. This work was a collaborative project with Professor Irwin Goldstein in the Department of Biological Chemistry at the University of Michigan The combination of ECD and vibrational excitation has previously been shown to be an attractive tool for *de novo* sequencing of proteins.[87,100] However, unlike previous efforts which have utilized the combination of ECD and CAD for sequencing purposes, the genome of this organism has not been sequenced. At the time this work was conducted, the combination of ECD and IRMPD (or CAD) had not been employed previously to *de novo* sequence an entire protein.

Chapter 3 explores the application of IRMPD and ECD for the structural characterization of high-mannose type glycopeptides. While ECD had been applied to the structural analysis of xylose type[83,101] and complex type[82] *N*-glycopeptides and *O*-glycosylated peptides,[80,82] it had not been applied towards the characterization of high-mannose type *N*-glycopeptides. Similarly, neither had the utility of IRMPD for the structural characterization of such glycopeptides been explored. This work has been published in the *Journal of Proteome Research* (2006, volume 5, page 493).

Chapter 4 describes the combination of metal adduction and ECD for oligosaccharide structural analysis. ECD had been previously applied to protonated

chitooligosaccharides, and yielded primarily glycosidic cleavages (similar to vibrational excitation).[91] Is has been previously shown that oligosaccharides ionized with alkali, alkaline earth, and transition metals often fragment to yield more cross-ring cleavages compared to their protonated counterparts with CAD and IRMPD. ECD fragmentation patterns of oligosaccharides ionized with various cationizing agents were examined, to determine whether the combination of ECD and metal ion adduction is a viable analytical tool for oligosaccharide structural characterization. Following ECD and IRMPD, fragmentation patterns were compared to determine whether ECD provides complementary information. This work has been published in *Analytical Chemistry* (2007, volume 79, page 2901).

Chapter 5 focuses on the combination of EDD and vibrational excitation for the analysis of neutral and sialylated oligosaccharides. EDD has, to our knowledge, not previously been applied to a wide variety of oligosaccharides, including neutral and sialylated glycans, as well as branched oligosaccharides. In Chapter 5, the fragmentation patterns of neutral and sialylated oligosaccharides following EDD are examined and compared to those obtained from IRMPD and CAD. This work has been published in the *Journal of the American Society for Mass Spectrometry* (2007, volume 18, page 2162).

Chapter 6 explores strategies for the analysis of pancreatic cancer associated O-linked glycans. This work was a collaborative project with Professor Diane Simeone in the Department of Surgery at the University of Michigan. This project focuses on the investigation of *O*-glycans from glycoproteins shed into growth media by pancreatic cancer cells. Lebrilla and co-workers presented a similar approach for ovarian cancer cell lines with a MALDI FT-ICR mass spectrometer.[102] However, ESI has a lower tolerance

to salt contaminants and glycan ionization is not as efficient compared to MALDI. Chapter 6 investigates various approaches for sample preparation with a model glycoprotein, fetuin. A protocol was designed and applied for the examination of oligosaccharides from various cell lines.

In addition, a summary of all results in this dissertation are presented in Chapter 7, along with future directions. The appendix includes data of glycopeptide anion fragmentation with EDD and IRMPD as well as EDD of chloride adducted oligosaccharides.

1.5 References

1. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.
2. Aleksandrov, M. L.; Gall, L. M.; Krasnov, M. V.; Nikolaiev, V. I.; Shkurov, V. A. *Dokl. Akad. Nauk. SSSR* **1984**, *277*, 379-383.
3. Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.
4. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T. *Rapid Commun. Mass Spectrom.* **1988**, *2*, 151-153.
5. Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5011-5015.
6. Mann, M.; Hojrup, P.; Roepstorff, P. *Biol. Mass Spectrom.* **1993**, *22*, 338-345.
7. Pappin, D. J. C.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, *3*, 327-332.
8. James, P.; Quadroni, M.; Carafoli, E.; Gonnet, G. *Biochem. Biophys. Res. Commun.* **1993**, *195*, 58-64.
9. Yates, J. R.; Speicher, S.; Griffin, P. R.; Hunkapiller, T. *Anal. Biochem.* **1993**, *214*, 397-408.
10. Mann, M.; Wilm, M. *Anal. Chem.* **1994**, *66*, 4390-4399.
11. Mørtz, E.; O'Connor, P.; Roepstorff, P.; Kelleher, N. L.; Wood, T. D.; McLafferty, F. W.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8264-8267.
12. Shevchenko, A.; Jensen, O. N.; Podtelejnikov, A. V.; Sagliocco, F.; Wilm, M.; Vorm, O.; Mortensen, P.; Shevchenko, A.; Boucherie, H.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14440-14445.
13. Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269-295.
14. Loo, J. A. *Mass Spectrom. Rev.* **1997**, *16*, 1-23.
15. Pacholski, M. L.; Winograd, N. *Chem. Rev.* **1999**, *99*, 2977-3005.
16. Appel, R. D.; Bairoch, A. *Proteomics* **2004**, *4*, 1525-1526.
17. Mann, M.; Jensen, O. N. *Nature Biotechnol.* **2003**, *21*, 255-261.
18. Jensen, O. N. *Curr. Opin. Chem. Biol.* **2004**, 33-41.
19. Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Acta* **1999**, *1473*, 4-8.
20. Varki, A. *Glycobiol.* **1993**, 97-130.
21. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
22. Choudhury, A.; Moniaux, N.; Ulrich, A. B.; Schmied, B. M.; Standop, J.; Pour, P. M.; Gendler, S. J.; Hollingsworth, M. A.; Aubert, J. P.; Batra, S. K. *Br. J. Cancer* **2004**, *90*, 657-664.
23. Dube, D. H.; Bertozzi, C. R. *Nat. Rev. Drug Discov.* **2005**, *4*, 477-488.
24. Wopereis, S.; Lefeber, D. J.; Morava, E.; Wevers, R. A. *Clin. Chem.* **2006**, *52*, 574-600.
25. Mechref, Y.; Novotny, M. V. *Chem. Rev.* **2002**, *102*, 321-369.
26. Harvey, D. J. *Mass Spectrom. Rev.* **1999**, *18*, 349-451.
27. Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161-227.
28. Morelle, W.; Canis, K.; Chirat, F.; Faid, V.; Michalski, J. C. *Proteomics* **2006**, *6*, 3993-4015.
29. Zamfir, A.; Peter-Katalinic, J. *Electrophoresis* **2004**, *25*, 1949-1963.
30. Harvey, D. J. *Proteomics* **2001**, *1*, 311-328.
31. O'Connor, S. E.; Imperiali, B. *Chem. & Biol.* **1996**, *3*, 803-812.

32. Wormald, M. R.; Petrescu, A. J.; Pao, Y. L.; Glithero, A.; Elliott, T.; Dwek, R. A. *Chem. Rev.* **2002**, *102*, 371-386.
33. Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *25*, 282-283.
34. Kelleher, N. L.; Senko, M. W.; Siegel, M. M.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 380-383.
35. Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.
36. Hendrickson, C. L.; Emmett, M. R. *Annu. Rev. Phys. Chem.* **1999**, *50*, 517-536.
37. Lorenz, S. A.; Maziarz, E. P. I.; Wood, T. D. *Appl. Spectrosc.* **1999**, *53*, 18A-36A.
38. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.
39. Sampson, J. S.; Hawkridge, A. M.; Muddiman, D. C. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, in press.
40. Castro, J. A.; Koster, C.; Wilkins, C. L. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 239-241.
41. Dunphy, J. C.; Busch, K. L.; Hettich, R. L.; Buchanan, M. V. *Anal. Chem.* **1993**, *65*, 1329-1335.
42. McIver, R. T., Jr.; Li, y.; Hunter, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 4801-4805.
43. Solouki, T.; Marto, J. A.; White, F. M.; Guan, S.; Marshall, A. G. *Anal. Chem.* **1995**, *67*, 4139-4144.
44. Brock, A.; Horn, D. M.; Peters, E. C.; Shaw, C. M.; Ericson, C.; Phung, Q. T.; Salomon, A. R. *Anal. Chem.* **2003**, *75*, 3419-3428.
45. An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B. *Anal. Chem.* **2003**, *75*, 5628-5637.
46. Marshall, A. G.; Guan, S. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1819-1823.
47. Gauthier, J. W.; Trautman, T. R.; Jacobson, D. B. *Anal. Chim. Acta* **1991**, *246*, 211-225.
48. Woodlin, R. L.; Bomse, D. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250.
49. Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
50. Price, W. D.; Schnier, P. D.; Williams, E. R. *Anal. Chem.* **1996**, *68*, 859-866.
51. Dunbar, R. C.; McMahon, T. B. *Science* **1998**, *279*, 194-197.
52. Mabud, M. A.; Dekrey, M. J.; Cooks, R. G. *Int. J. Mass Spectrom. Ion Processes* **1985**, *67*, 285-294.
53. Ijames, C. F.; Wilkins, C. L. *Anal. Chem.* **1990**, *62*, 1295-1298.
54. Williams, E. R.; Henry, K. D.; McLafferty, F. W.; Shabanowitz, J.; Hunt, D. F. *J. Am. Soc. Mass Spectrom.* **1990**, *1*, 413-416.
55. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
56. Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
57. Biemann, K.; Martin, S. A. *Mass Spectrom. Rev.* **1987**, *6*, 1-76.
58. McLafferty, F. W. *Tandem Mass Spectrometry*; Wiley: New York, 1983.
59. McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599-614.

60. Sleno, L.; Volmer, D. A. *J. Mass Spectrom.* **2004**, *39*, 1091-1112.
61. McLuckey, S. A.; Goeringer, D. E. *J. Mass Spectrom.* **1997**, *35*, 461-474.
62. Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601-601.
63. Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.
64. Heck, A. J. R.; de Koning, L. J.; Pinske, F. A.; Nibbering, N. M. M. *Rapid Commun. Mass Spectrom.* **1991**, *5*, 406-414.
65. Mirgorodskaya, E.; O'Connor, P. B.; Costello, C. E. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 318-324.
66. Domon, B.; Costello, C. E. *Glycoconj. J.* **1988**, *5*, 397-409.
67. Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.
68. Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. Mass Spectrom.* **2002**, *8*, 337-349.
69. Turecek, F. *J. Am. Chem. Soc.* **2003**, *125*, 5954-5963.
70. Syrstad, E. A.; Stephens, D. D.; Turecek, F. *J. Phys. Chem. A* **2003**, *107*, 115-126.
71. Turecek, F.; Syrstad, E. A. *J. Am. Chem. Soc.* **2003**, *125*, 3353-3369.
72. Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949-8958.
73. Breuker, K.; Oh, H. B.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14011-14016.
74. Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208-224.
75. Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201-222.
76. Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57-77.
77. Zubarev, R. A. *Curr. Opin. Biotechnol.* **2004**, *15*, 12-16.
78. Stensballe, A.; Norregaard-Jensen, O.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.
79. Shi, S. D.-H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19-22.
80. Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.
81. Mormann, M.; Macek, B.; de Peredo, A. G.; Hofsteenge, J.; Peter-Katalinic, J. *Int. J. Mass Spectrom.* **2004**, *234*, 11-21.
82. Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 2355-2361.
83. Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.
84. Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803-1812.
85. Kweon, H. K.; Hakansson, K. *J. Proteome Res.* **2008**, *7*, 749-755.
86. Hakansson, K.; Cooper, H. J.; Hudgins, R. R.; Nilsson, C. L. *Curr. Org. Chem.* **2003**, *7*, 1503-1525.
87. Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Zubarev, R. A. *J. Proteome Res.* **2005**, *4*, 2348-2354.

88. Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J. H.; Begley, T. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 672-678.
89. Demirev, P. A.; Ramirez, J.; Fenselau, C. *Anal. Chem.* **2001**, *73*, 5725-5731.
90. Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 14432-14433.
91. Budnik, B. A.; Haselmann, K. F.; Elkin, Y. N.; Gorbach, V. I.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 5994-6001.
92. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1882.
93. Mo, J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.
94. McFarland, M. A.; Marshall, A. G.; Hendrickson, C. L.; Nilsson, C. L.; Fredman, P.; Mansson, J. E. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 752-762.
95. Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. *J. Am. Soc. Mass Spectrom.* **2006**, *18*, 234-244.
96. Wolff, J. J.; Chi, L.; Linhardt, R. J.; Amster, I. J. *Anal. Chem.* **2007**, *79*, 2015-2022.
97. Shaffer, S. A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D. *Anal. Chem.* **1998**, *70*, 4111-4119.
98. Senko, M. W.; Hendrickson, C. L.; Emmett, M. R.; Shi, S. D.-H.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 970-976.
99. Belov, M. E.; Nikolaev, E. N.; Anderson, G. A.; Udseth, H. R.; Conrads, T. P.; Veenstra, T. D.; Masselon, C. D.; Gorshkov, M. V.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 253-261.
100. Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10313-10317.
101. Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
102. An, H. J.; Miyamoto, S.; Lancaster, K. S.; Kirmiz, C.; Li, B.; Lam, K. S.; Leiserowitz, G. S.; Lebrilla, C. B. *J. Proteome Res.* **2006**, *5*, 1626-1635.

Chapter 2

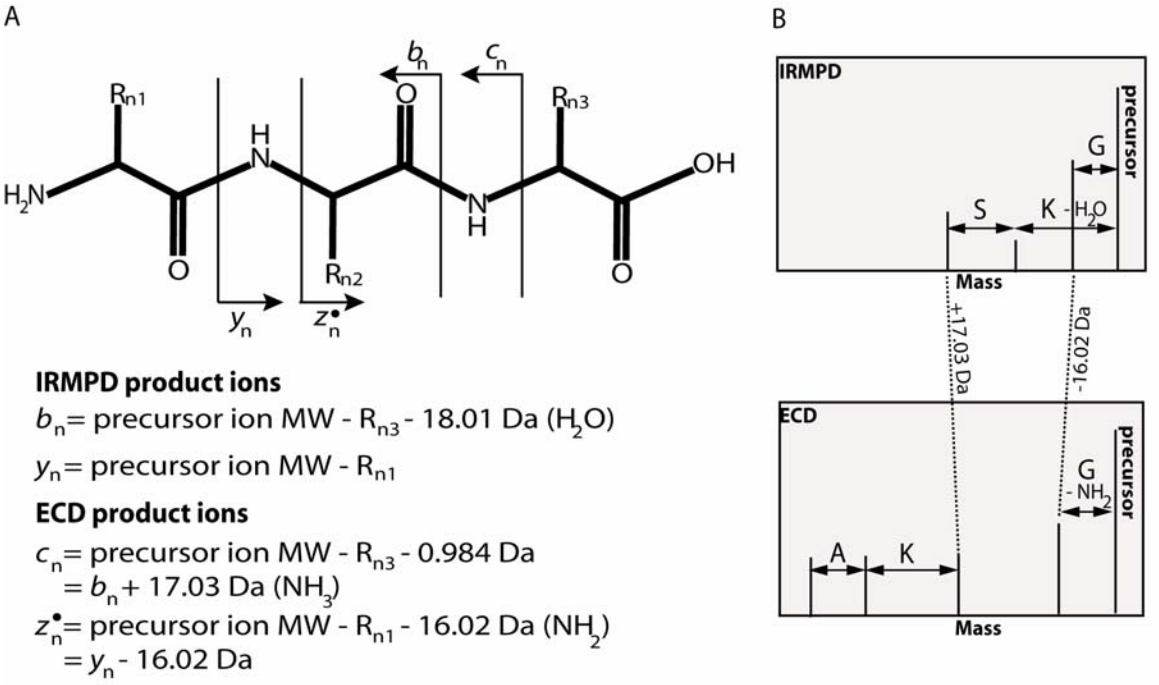
De Novo Sequencing of a Mushroom Lectin by Combined Infrared Multiphoton Dissociation and Electron Capture Dissociation

2.1 Introduction

Mass spectrometry is a vital tool in proteomic analysis for the characterization of peptides and proteins.[1-3] It can be used to identify unknown proteins by matching peptide masses from a proteolytic digest or peptide fragmentation data to proteins represented in a genomic database.[4-11] However, there are several factors which complicate this approach, including modifications after transcription (such as alternative splicing), post-translational modifications, unknown mutations, and sequence errors in the database. This method also fails if the genome of the organism from which the protein is derived has not been sequenced. In such cases, protein *de novo* sequencing becomes necessary.

Protein *de novo* sequencing aims to reconstruct a protein sequence exclusively using fragmentation data from peptide (bottom-up approach) or intact protein (top-down approach) tandem mass spectra (MS/MS spectra). One of the most commonly used techniques for peptide fragmentation is collision activated dissociation (CAD).[12-14] CAD is a vibrational excitation fragmentation technique that results in preferential C-N bond cleavage of peptides, producing *b*- and *y*-type ions (see Scheme 1).[15] Another commonly

used vibrational excitation technique is infrared multiphoton dissociation (IRMPD),[16,17] which also results in preferential C-N bond cleavage. Vibrational excitation of peptide cations shows a strong preference for cleavage on the N-terminal side of proline and also



Scheme 2.1. a) Peptide fragmentation nomenclature. IRMPD preferentially cleaves C-N bonds, resulting in *b* and *y*-type ions whereas ECD preferentially cleaves N-C_α bonds, resulting in *c* and radical *z*-type ions. Characteristic differences in masses between *b* and *c*-type, and *y* and *z*-type ions (the "golden complementary pair" rules) help distinguish N-terminal from C-terminal product ions. N- and C-terminal amino acids are assigned based upon characteristic mass differences between the precursor ion mass and a product ion mass corresponding to cleavage adjacent to the N- or C- terminus. For example, a *c*-type cleavage would appear as the loss of the C-terminal amino acid and an additional 0.984 Da. Potential *c*-type cleavages adjacent to the C-terminus are indicated with quotation marks around amino acid assignments. b) An example of how "golden pairs" can be utilized for *de novo* sequencing. Here, the combination of IRMPD and ECD and the appearance of two "golden pairs" sequences this peptide as GAKSK.

the C-terminal side of aspartic acid.[18,19] Due to preferential cleavage sites, a major difficulty facing *de novo* sequencing with CAD or IRMPD is incomplete backbone fragmentation, i.e., lack of cleavage between all amino acid residues. Furthermore, vibrational excitation of peptides often results in multiple fragmentation sites, generating

product ions which do not necessarily contain the N- or C- terminus of the peptide. Such fragments often complicate mass spectral interpretation. Although several peptide derivatization techniques have been developed in order to improve mass spectral quality,[20] a disadvantage is that they generally increase sample requirements. Due to several drawbacks associated with vibrational excitation based fragmentation, a technique which provides complementary sequence information is desired for *de novo* sequencing of proteins.

Electron capture dissociation (ECD) was introduced in 1998 as a tandem mass spectrometric fragmentation technique for peptide polycations.[21] ECD is based upon the dissociative recombination of protonated polypeptide molecules with low-energy electrons (<0.2 eV).[21-24] This technique has been shown to be complementary to vibrational excitation and a valuable tool for proteomic applications. Unlike CAD or IRMPD, ECD shows preferential cleavage of N-C_α bonds, resulting in *c*- and radical $z^\bullet$-type product ions (see Scheme 1).[21] ECD has also been shown to retain labile modifications such as glycosylation and phosphorylation and shows preferential cleavage of disulfide bonds.[22,23] Unlike vibrational excitation techniques, cleavage on the N-terminal side of proline is generally absent following ECD due to the cyclic structure of proline. ECD data can be utilized for determining whether ions observed from vibrational excitation are *b* or *y*- type ions, based on characteristic mass differences between *b* and *c*, and *y* and radical *z* type ions (see Scheme 1), so called "golden pairs".[25] A recent comprehensive study on 15,000 pairs of ECD and CAD mass spectra showed that the cleavage sites in ECD and CAD are truly complementary,[26] attesting that the combination of ECD and vibrational excitation is a powerful tool for peptide

characterization. However, one disadvantage of ECD is its lower fragmentation efficiency compared to vibrational excitation.[27,28] In addition, product ions observed from ECD can lose or gain an additional hydrogen, which can complicate spectral interpretation.[29,30] However, the complementary information that ECD provides makes it an attractive method to be used in conjunction with vibrational excitation for *de novo* sequencing. Recently, Savitski et al. developed a *de novo* sequencing algorithm, which uses the combination of ECD and CAD fragmentation. They demonstrated a proteomics-grade *de novo* sequencing approach which had the same level of efficiency and reliability as conventional database-identification strategies, but which alleviated the problems associated with using only CAD for sequencing.[31] In order to achieve this advantage, high mass accuracy measurements from a Fourier transform or similar performing instrument are required. In a recent bottom-up effort, combined ECD and CAD in a hybrid linear ion trap-Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer were utilized to *de novo* sequence cytochrome c_4 from the organism *Thiocapsa roseopersicina* whose genome has not yet been sequenced.[32] However, in that work, golden pair product ions were not utilized and hydrogen rearrangement, which frequently occurs in ECD, was not discussed.[29,30] Also, because a gene sequence is unavailable, it is difficult to evaluate potential protein sequence errors although the authors mention, e.g., oxidation and deamidation during sample handling.

The high mass accuracy and combined ECD/IRMPD capabilities of an FT-ICR mass spectrometer are utilized here for *de novo* sequencing of a 10 kDa lectin from the mushroom *Lyophyllum decastes*. This protein is a recently discovered α-galactosyl binding lectin from this organism.[33] Relatively few lectins with α-galactosyl specificity

have been described but such proteins constitute valuable reagents for glycobiological experiments. The genome of *L. decastes* has not been sequenced. Our goal was to utilize the combination of ECD and IRMPD to re-construct the amino acid sequence of the *L. decastes* lectin. To our knowledge, the combination of ECD and IRMPD has not been employed previously to *de novo* sequence an unknown protein. In addition to the amino acid sequence, the molecular weight and disulfide bond content of this protein were also examined. Furthermore, we compare the MS-derived sequence with that obtained from concurrent biochemical experiments, including Edman sequencing and, ultimately, cloning and gene sequencing. Challenges of MS-based analysis are also discussed.

2.2 Experimental

2.2.1 Sample Preparation

Lectin isolation and purification were performed as previously described.[33] Samples (1-2 mg) of a lectin fraction from melibiose-sepharose affinity chromatography were dialyzed against distilled water and lyophilized prior to mass spectrometric analysis. Automated Edman sequencing and matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) mass analysis were performed by the Protein Structure Facility at the University of Michigan. Reverse transcription polymerase chain reaction (RT-PCR) was performed from mRNA with oligonucleotide primers based on peptide sequences determined by Edman sequencing that allowed for the lowest amount of degeneracy. 3' rapid amplification of cDNA ends (RACE) was employed to obtain the protein sequence located 3' to the amino acids sequenced by RT-PCR. Details of RNA isolation and cDNA synthesis have been previously described.[33]

For disulfide bond determination, 0.25 nmoles of protein was treated with 10 mM dithiothreitol for 15 minutes at 100°C, diluted to 2 μM in 50% acetonitrile/water with 0.1% formic acid, and immediately analyzed by FT-ICR mass spectrometry. A 1 μM solution of untreated protein was also examined. The monoisotopic mass of the protein was determined from the 8+ and 7+ charge states, using external calibration with reference masses from Hewlett-Packard ESI-MS tuning mixture (part number G2421A). Average molecular weights were determined with MIDAS HD Helper.[34] Lectin peptides for MS/MS analysis were generated with a variety of endoproteases, including trypsin (Promega, Madison, WI), chymotrypsin (Sigma, St. Louis, MI), and GluC (Roche, Indianapolis, IN). Between 0.125 - 1 nmoles of protein was reduced in 10 mM dithiothreitol (Sigma) at 95-100°C for approximately 15 minutes. Samples were next treated with 45 mM iodoacetamide (Sigma) at room temperature in darkness for approximately 45 minutes. Ammonium bicarbonate (Fisher, Fair Lawn, NJ) was added to all digestion solutions, to a final concentration of 50 mM. Chymotrypsin samples also had 10 mM $CaCl_2$ (Sigma) present. The enzyme:protein ratio was between 1:30-1:50 for overnight digestions. For a short trypsin digestion, the protein was digested for 15 minutes at an enzyme:protein ratio of 1:50. For a short chymotrypsin digestion, the protein was digested for 2 hours at an enzyme:protein ratio of 1:15. All samples were desalted with either C18 ziptips (Millipore, Billerica, MA) or PepClean C-18 spin columns (Pierce, Rockford, IL). After desalting, overnight digestion samples were diluted into a solution of 50% acetonitrile (Fisher)/water with 0.1% formic acid (Acros Organics, Morris Plains, NJ) to a final concentration of 5-6 μM prior to mass spectrometric analysis. Short digestion samples were diluted to a concentration of 1 μM.

2.2.2 FT-ICR Mass Spectrometry

All experiments were performed with an actively shielded 7 T FT-ICR mass spectrometer with a quadrupole front-end (APEX-Q, Bruker Daltonics), as previously described.[35] Samples were electrosprayed via either an Apollo electrospray ionization (ESI) source at 60 μL/hour, or with a static nanoelectrospray source using PicoTip GlassTips (New Objective, Woburn, MA). Following a brief (0.05 ms) ion accumulation in the first hexapole, ions were mass selectively accumulated in the second hexapole for 2 - 4 s. Ions were then transferred through high-voltage ion optics and captured with dynamic trapping in an Infinity ICR cell.[36] Only ions with at least two charges were further examined with IRMPD and ECD. Precursor ions were further isolated in the ICR cell by correlated harmonic excitation fields (CHEF).[37] IRMPD was performed with a vertically mounted 25-W, 10.6-μm, CO_2 laser (Synrad, Mukilteo, WA). An indirectly heated hollow dispenser cathode (HeatWave, Watsonville, CA) was used to perform ECD.[38] The cathode heating current was kept constant at 1.8 A and the lens electrode located immediately in front of the cathode was kept at 1-2 V. For activated-ion ECD (aiECD), ions were heated with a brief (55 ms) IR laser shot prior to ECD. Product ions between *m/z* 200-2000 were detected.

2.2.3 Data Analysis

All mass spectra were acquired with XMASS software (Bruker Daltonics) with 256k data points and summed over 20-80 scans. Data processing was performed with MIDAS software.[39] Data were zero filled once and Hanning apodized. IRMPD and ECD spectra were manually charge-deconvoluted to their neutral state based on product ion isotopic distributions. Low abundance product ions, typically with *m*/*z* values below 800,

often appeared without their heavier isotopes. In these instances, these ions were assumed to be singly charged. This trend has been repeatedly observed following IRMPD and ECD of doubly and triply protonated peptides in this instrument.

From the list of neutral masses, mass gaps between product ions were manually assigned to a potential amino acid residue (with an initial mass error tolerance of ± 0.1 Da). Product ions due to water loss (18.01 Da) were assigned in IRMPD spectra and not considered when assigning "golden pairs". Following ECD, hydrogen addition/abstraction often occurs,[29,30] therefore deviations of ± 1 Da were also considered for ECD data. From these criteria, several potential amino acid assignments could often be made for each mass gap. However, confident amino acid assignments were made when IRMPD and ECD data agreed (so called "golden pairs" were observed). Amino acid assignments made from IRMPD data had a mass error typically less then 0.005 Da while from ECD the error was usually less than 0.01 Da. Larger mass errors observed following ECD were typically due to smaller signal-to-noise ratios and because of space-charge effects caused by the introduction of electrons into the ICR cell. To distinguish N-terminal and C-terminal fragments, IRMPD and ECD data were examined for "golden pairs".[25] When IRMPD and ECD cleavage occurs between the same amino acid residues, the mass difference between a *b* ion and a *c* ion is - 17.03 Da while the difference between a *y* ion and a radical *z* ion is 16.02 Da (see Scheme 1). Deviations of ± 1 Da were also considered in assigning "golden" pairs, to account for *c* and *z* product ions with additional hydrogen abstraction/addition. N- and C-terminal amino acids are assigned based upon characteristic mass differences between the precursor ion and a product ion corresponding to cleavage adjacent to the N- or C- terminus (see Scheme 2.1).

2.3 Results and Discussion

FT-ICR mass spectrometric characterization of the novel α-galactosyl binding lectin occurred concurrently with other biochemical characterization. Automated Edman sequencing determined that the N-terminal sequence of the lectin was ACWKANSCPGSAFESKDRLRLxFALLYxRYN (where x is undetermined). This information, along with knowledge from MALDI-TOF of the intact molecular weight (10,279 Da ± 0.01%), was the only structural information available at the onset of the investigation presented here, which focuses on *de novo* sequencing of this protein with mass spectrometry. Several proteolytic enzymes, including trypsin, GluC, and chymotrypsin, were utilized to generate a variety of multiply charged peptides. Examples of IRMPD and ECD spectra from various proteolytic digestions are presented below to demonstrate how the combination of these two techniques can be utilized for the purpose of sequencing unknowns. A comparison of the final mass spectrometric results versus those obtained through gene sequencing is also provided.

With positive ion mode ESI FT-ICR MS, the monoisotopic mass of the untreated protein was observed as 10,270.1 Da (Figure 2.1a). Only one form of the protein was detected, indicating that this lectin does not contain abundant heterogeneous forms. Following treatment with DTT the average molecular weight of the protein increased by 5.9 Da (Figure 2.1b), indicating the presence of three disulfide bonds. Ion abundances in Figure 1b were lower than those in 1a, likely due to the presence of DTT in the solution. For *de novo* sequencing of this protein, all peptide samples were reduced and alkylated prior to mass spectrometric analysis. Therefore, information regarding which cysteine residues are involved in a particular disulfide bond could not be determined. Although a

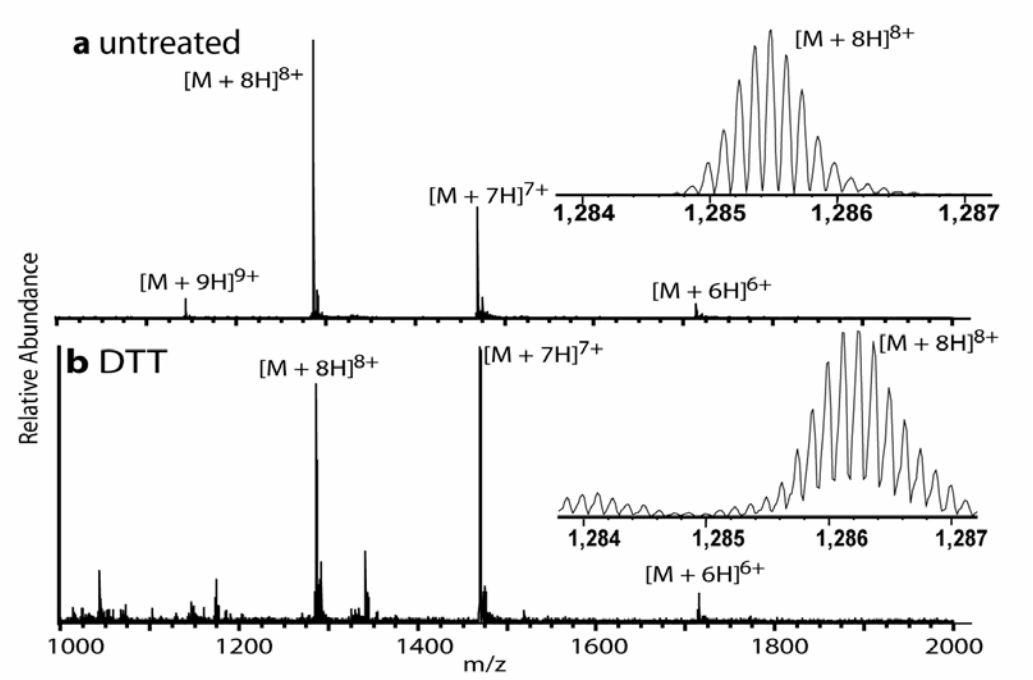


Figure 2.1. a) Positive ion mode FT-ICR mass spectrum (5 scans) of a lectin from the mushroom *Lyophyllum decastes*. b) Positive ion mode FT-ICR mass spectrum (30 scans) of the lectin treated with dithiothreitol for 15 minutes. The average molecular weight was determined from the 7+ and 8+ charge states. A mass increase of 5.9 Da was observed, indicating the presence of three disulfide bonds. Insets show a zoomed-in view of the 8+ charge state before and after DTT treatment.

top-down characterization of the untreated protein has the potential to indicate disulfide bond positions, both IRMPD and ECD of the intact untreated protein resulted in relatively poor fragmentation. Trypsin digestion of the native protein was also performed, in order to obtain disulfide linked peptides which could then be subjected to MS/MS. However, no disulfide linked peptides were observed with this strategy. Multiple chemical or enzymatic cleavage reagents may be able to produce disulfide-containing peptides,[40] however this approach was not explored.

2.3.1 Trypsin Analysis

Following an overnight trypsin digestion, several peptides were selected for further MS/MS characterization. Results for one of these peptides are shown in Figure

2.2, wherein the combination of IRMPD and ECD resulted in extensive sequence information. This Figure shows the deconvoluted FT-ICR MS/MS spectra of a doubly charged tryptic peptide with a mass of 1253.5 Da. Arrows indicate potential single amino acid differences between fragments (based on mass) and values in parentheses indicate the corresponding mass error in Daltons. For example, the label E (0.002) in Figure 2.2a signifies that the mass gap between the indicated product ions is 129.044 Da (0.002 Da heavier than glutamic acid). Filled squares indicate product ions that are likely due to secondary water loss in IRMPD. Neutral molecule losses, such as water (18.011 Da) and ammonium (17.027 Da), are commonly observed in vibrational excitation fragmentation. Generally, these neutral losses complicate spectral interpretation. Several potential short amino acid sequences are indicated in Figure 2.2a, including the sequences EFA (or AFE) and PCS (or SCP). The IRMPD spectrum does not confirm the directionality of these sequences or how they are arranged within the peptide. However, the product ion with a mass of 821 Da (labeled as cleavage next to proline) is one of the most abundant product ions observed following IRMPD and can therefore be hypothesized to correspond to cleavage on the N-terminal side of proline, given that such cleavage is preferred following vibrational excitation. This product ion, and all others in the same series, are therefore potentially *y*-type ions. In Figure 2.2a, there are two sets of product ions which indicate the sequence FA (or AF). One set is due to *b* or *y* type cleavage, while the other set is likely due to internal fragmentation.

Figure 2.2b shows the ECD spectrum of the same peptide as examined in Figure 2.2a. This spectrum is more difficult to interpret compared to the IRMPD spectrum

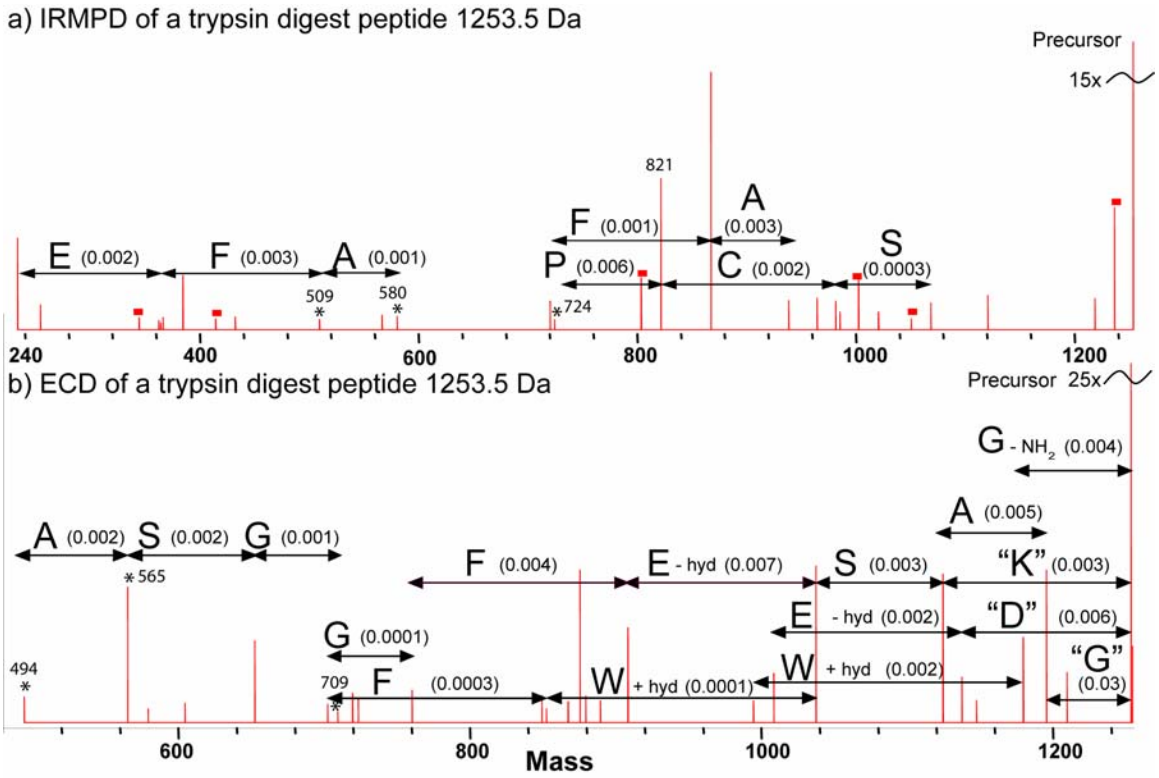


Figure 2.2. Deconvoluted a) IRMPD FT-ICR MS/MS spectrum (40 scans, 1.5 s, 10 W) and b) ECD FT-ICR MS/MS spectrum (80 scans, 75 ms, - 0.5 bias voltage) of a doubly protonated peptide from trypsin digestion (1253.5 Daltons). Arrows indicate a potential amino acid assignment, and values in parentheses indicate the mass error (Daltons). Squares indicate peaks which are likely due to water loss. +/- hyd indicates gain/loss of a hydrogen. Quotations around an amino acid assignment indicate a potential *c*-type ion (the mass gap is equal to the mass of the indicated amino acid and 0.984 Da, see Scheme 2.1). "Golden pairs" are indicated with asterisks. Specific product ions discussed in the text are indicated with mass labels.

because hydrogen addition/abstraction, which often occurs in ECD,[29,30] must be considered. Product ions with potential hydrogen addition and abstraction are indicated by the labels "- hyd" and "+ hyd.". Although *c*- and *z*-type cleavages are prevalent following ECD, *a*- and *y*-type ions may also be observed.[22,23] The possibility of all these product ions to be either even or odd electron species increases the number of potential amino acid assignments. In addition, several reports have demonstrated the occurrence of amino acid side chain losses in ECD of peptides.[41-44] Despite these difficulties, several

product ions observed in Figure 2.2b are complementary to those from IRMPD (Figure 2.2a). The ECD product ion with a mass of 494 Da (labeled as cleavage next to alanine) is 15.01 Da lighter than the IRMPD product ion at mass 509, which was also assigned as cleavage adjacent to an alanine residue. This mass difference indicates a "golden pair", consisting of a *y*-type ion and an even-electron *z* ion ("golden pairs" are denoted with asterisks in Figure 2.2). The ECD product ions with a mass of 565 Da (labeled as cleavage next to serine) and 709 Da (cleavage adjacent to glycine) also have corresponding *y*-type product ions in Figure 2.2a. With these three sets of "golden pairs", a large portion of the peptide, SCPGSAFE, can be confidently sequenced. Assuming this sequence is correct, Figure 2.2a indicates that the total mass of un-sequenced amino acids on the C-terminal end of the peptide is 215.125 Da. Meanwhile, the ECD spectrum indicates that FESK is a potential amino acid sequence for the C-terminal end. The potential loss of lysine from the precursor ion is denoted as "K" in Figure 2.2b, because *c*-type cleavage adjacent to the C-terminus appears as the loss of the amino acid residue and an additional 0.984 Da (as shown in Scheme 2.1). Given that this peptide was generated from a trypsin digestion, the presence of a C-terminal lysine or arginine would be highly expected. The combined mass of serine and lysine is 215.127 Da and, because none of the other potential C-terminal amino acid assignments (E"D" or W"G") in Figure 2.2b corresponds to this mass, the combined data imply the amino acid sequence SCPGSAFESK. This sequence corresponds to 1068.5 Da of this peptide, which is 85% of the total mass (1253.5 Da). Thus, 85% of the peptide was confidently sequenced from combined IRMPD and ECD. Although golden pairs were not observed within the C-terminal portion of the peptide, a potential C-terminal amino acid sequence was deduced

from ECD fragmentation information. However, even with the combination of IRMPD and ECD, the N-terminal amino acids of this peptide could not be sequenced.

2.3.2 GluC Analysis

Following an overnight GluC digestion, five peptides were selected and examined with IRMPD and ECD. The amino acid sequence of one of the GluC peptides was found to partially overlap with the previously examined trypsin digest peptide (Figure 2.2). Figure 2.3a and 2.3b show deconvoluted IRMPD and ECD spectra of a doubly charged peptide with a mass of 1583.6 Da generated from a GluC digest. The IRMPD spectrum contains the sequence SC (or CS), which was also identified in Figure 2.2. Several other potential amino acid assignments are also indicated in this spectrum, such as the presence of glutamic acid and valine, but these were later identified as either due to internal fragmentation (for glutamic acid), or as product ions due to multiple neutral molecule losses (valine). Similar to Figure 2.2b, the ECD spectrum in Figure 2.3b is more difficult to interpret than the corresponding IRMPD spectrum. A comparison of the IRMPD and ECD spectra for the GluC digest peptide indicates two "golden pairs" consisting of two pairs of *b*- and *c*-type cleavages. These ions confirm the presence of alanine, with the product ion at mass 1306 Da in Figure 2.3b and at mass 1289 in Figure 2.3a. Figure 2.3b indicates the C-terminal sequence GSAFE, which is also present in the sequence determined from Figure 2.2. From a GluC digestion of the protein, the terminal amino acid of this peptide is expected to be glutamic acid. The mass difference between the product ion at mass 834 Da (labeled as cleavage adjacent to serine) and the product at 1092 Da (due to cleavage next to glycine) is 258.07. Referring to the sequence SCPGSAFESK determined from Figure 2.2, this mass gap would correspond to the

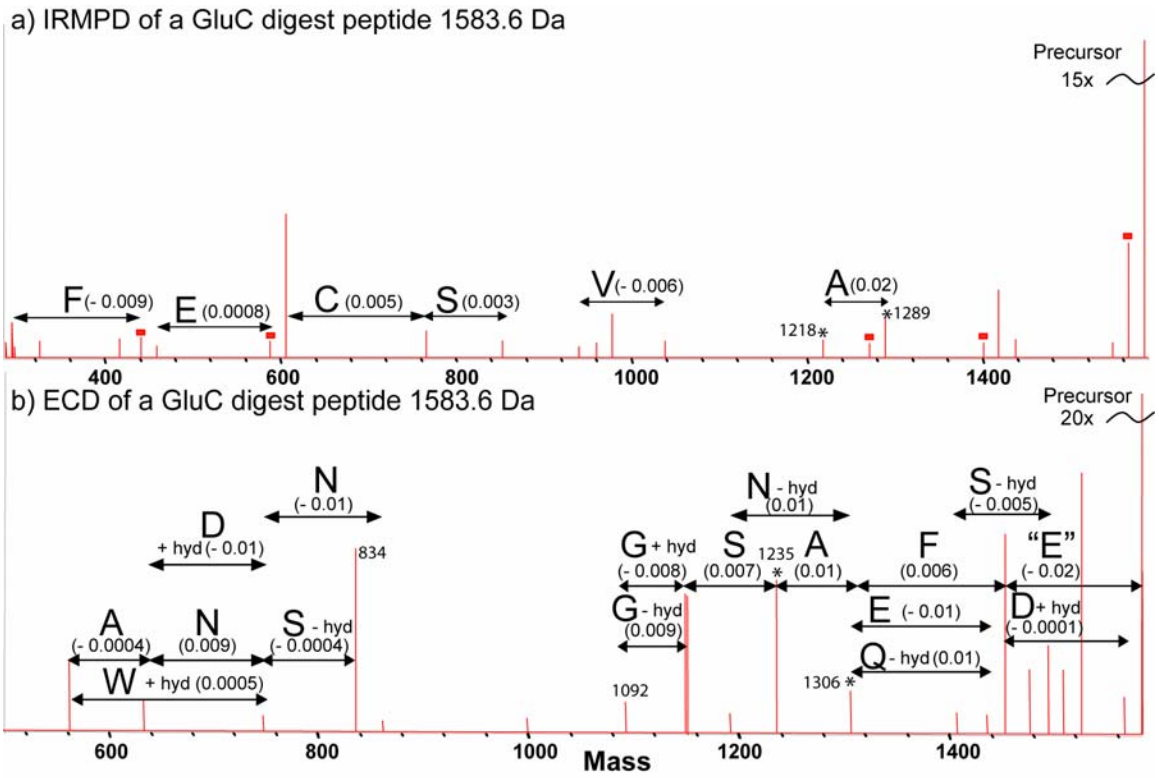


Figure 2.3. Deconvoluted a) IRMPD FT-ICR MS/MS spectrum (40 scans, 0.25 s, 10 W) and b) ECD FT-ICR MS/MS spectrum (80 scans, 60 ms, - 0.5 bias voltage) of a doubly protonated peptide from a GluC digestion (1583.6 Da). This peptide overlaps with the peptide examined in Figure 2.2. Squares indicate peaks which are likely due to water loss. +/- hyd indicates gain/loss of a hydrogen. Quotations around an amino acid assignment indicate a potential *c*-type ion (the mass gap is equal to the mass of the indicated amino acid and 0.984 Da, see Scheme 2.1). "Golden pairs" are indicated with asterisks. Specific product ions discussed in the text are indicated with mass labels.

combined mass of proline and an alkylated cystine residue along with hydrogen abstraction. Figure 2.3b also fills in the missing amino acids not sequenced from MS/MS data in Figure 2.2, and indicates the sequence ANSCPGSAFE. This sequence corresponds to 66% of the total mass of this GluC digest peptide, with a remaining 545.2 Da still un-sequenced at the N-terminal end. However, N-terminal Edman sequencing of the protein provided the sequence ACWKANSCPGSAFESKDRLRLxFALLYxRYN, which contains the ANSCPGSAFE partial sequence from Figure 2.3 preceded by ACWK, which indeed corresponds to the missing 545.2 Da. This example highlights

that, in many cases, IRMPD and ECD of a single peptide do not always allow for complete sequencing. This outcome is due to incomplete fragmentation and because of complications arising from several potential amino acid assignments with similar mass errors. When "golden pairs" are observed, there is higher confidence in amino acid assignments. However, in Figure 2.2 only three "golden pairs" were observed while only two were observed in Figure 2.3. To provide missing amino acid assignments, MS/MS of overlapping peptides (such as the ones shown in Figures 2.2 and 2.3) was often necessary.

2.3.3 Chymotrypsin Analysis

Following an overnight chymotrypsin digestion, sixteen peptides were selected and examined with IRMPD and ECD. Unlike trypsin and GluC, chymotrypsin hydrolyzes peptide bonds on the C-terminal side of numerous amino acids, including tyrosine, phenylalanine, tryptophan, leucine and, to some extent, methionine, isoleucine, serine, threonine, valine, histidine, glycine and alanine.[45] MS/MS revealed that the majority of peptides examined from a chymotrypsin digest had overlapping sequences, which is expected given the nature of chymotrypsin. Figure 2.4 shows deconvoluted IRMPD and ECD spectra of a doubly protonated chymotrypsin digest peptide with a mass of 2107.9 Da. IRMPD of this peptide resulted in extensive fragmentation (Figure 2.4a). The product ion observed in Figure 2.4a at mass 1943 corresponds to loss of a phenylalanine residue and water from the precursor ion, which indicates that it is a *b*-type ion (as shown in Scheme 2.1). A C-terminal phenylalanine is consistent with the cleavage tendencies of chymotrypsin. Several additional adjacent *b*-type product ions are also observed, indicating the C-terminal sequence ETEAQCI(L)NTF. Isoleucine and

leucine have the same monoisotopic mass, and therefore cannot be distinguished with IRMPD. In the lower mass region of Figure 2.4a, another series of product ions provides the sequence I(L)CQAETE. These ions overlap with the previously mentioned *b* ion series. The latter series is presumably a *y* ion series, indicated by the reversal in sequence (since amino acids are now lost from the N-terminus as opposed to the C-terminus for *b*-type ions). This assignment can be further verified by adding corresponding *b*-type ion masses to these *y*-type ion masses, to observe a total mass of 2107.9 Da (the mass of the intact peptide).

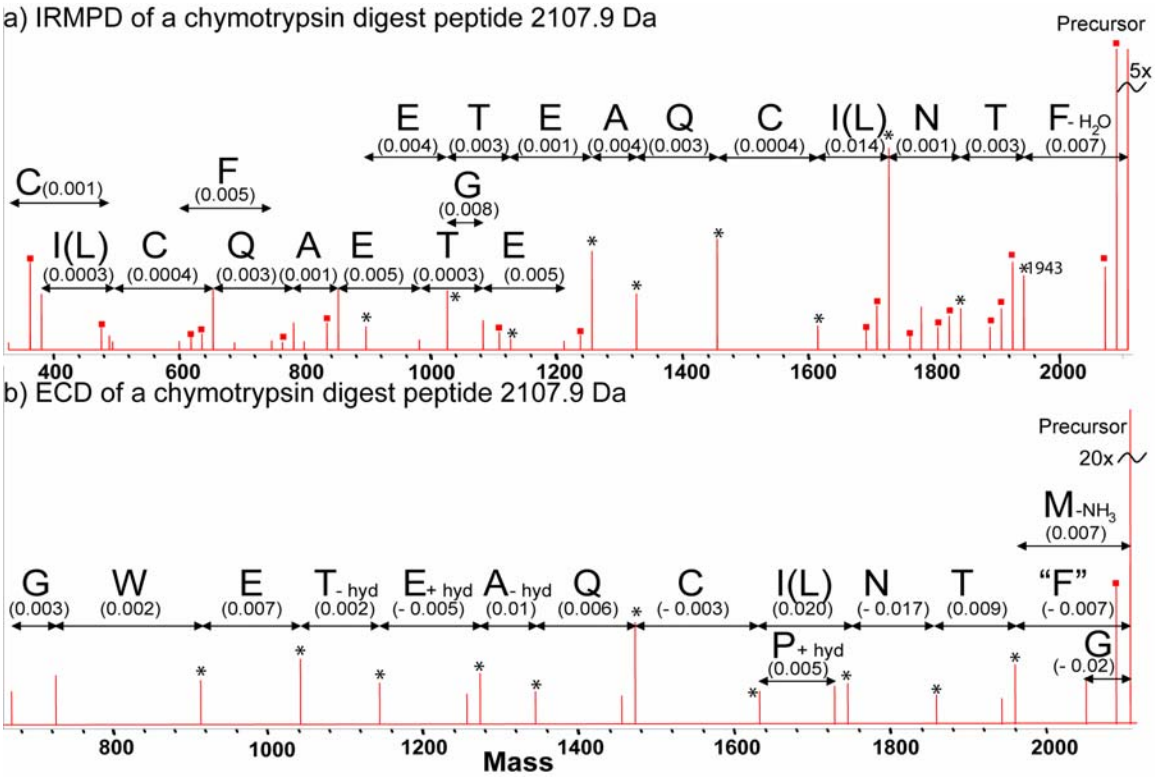


Figure 2.4. Deconvoluted a) IRMPD FT-ICR MS/MS spectrum (30 scans, 0.75 s, 10 W) and b) aiECD FT-ICR MS/MS spectrum (60 scans, 75 ms, - 0.5 bias voltage, preceded by 55 ms 10 W laser pulse) of a doubly protonated peptide from a chymotrypsin digest (2107.9 Da). Squares indicate peaks which are likely due to water loss. +/- hyd indicates gain/loss of a hydrogen. Quotations around an amino acid assignment indicate a potential *c*-type ion (the mass gap is equal to the mass of the indicated amino acid and 0.984 Da, see Scheme 2.1). "Golden pairs" are indicated with asterisks. Specific product ions discussed in the text are indicated with mass labels.

Figure 2.4b shows the ECD spectrum of the same peptide as examined in Figure 2.4a. To efficiently fragment this peptide, activated ion ECD was necessary, presumably due to intramolecular non-covalent interactions, which are more common for larger peptides and prevent product ions from separating in ECD.[46-48] For aiECD, ions were heated with an IR laser pulse before they underwent electron capture. This activation was to aid in the destabilization of gas-phase secondary structure.[46] Following ECD, *c*-type ions were identified along with several *b*- and *y*-type ions. The latter products were assumed to be due to vibrational excitation from the laser pulse. Unlike previous IRMPD and ECD spectra, several abundant "golden pairs" are observed. A total of ten *b*/*c* type "golden pairs" are observed, consisting of 10 *b*-type ions observed following IRMPD and 10 corresponding *c*-type ions observed following ECD. No *y*/*z* type "golden pairs" were identified. These 10 "golden pairs" confirm the sequence ETEAQCI(L)NTF, as determined from the IRMPD spectrum in Figure 2.4a. ECD also indicates that residues GW are preceding this sequence. For this peptide, the combination of IRMPD and ECD allowed 69% of the total peptide mass to be sequenced. However, neither fragmentation technique provided information regarding the amino acid sequence at the N-terminal end of the peptide. In addition, neither IRMPD nor ECD can distinguish isoleucine from leucine. Zubarev and co-workers have demonstrated that hot electron capture dissociation (heECD) may result in secondary fragmentation of radical *z*-type ions, generating residue specific *w*-type ions which distinguish isoleucine from leucine. However, in these experiments heECD was not examined.

As a final example, Figure 2.5 shows the deconvoluted IRMPD and ECD spectra of a doubly charged peptide from a chymotrypsin digest with a mass of 1517.7 Da. The

deconvoluted IRMPD spectrum of this peptide has several potential amino acid assignments and is relatively difficult to interpret (see Figure 2.5a). Unlike the other IRMPD spectra shown above, this spectrum likely contains extensive internal fragments and neutral molecule losses, which complicate spectral interpretation. Between masses 698-1055 Da, the IRMPD spectrum indicates that this peptide contains the sequence YPP (or PPY). The presence of multiple proline residues within this peptide may result in extensive internal fragmentation.[49-51] The amino acids adjacent to this sequence are unclear; however the sequences YPPKG (or GKPPY) and YPPGK (or KGPPY) are both possible assignments (direction of peptides is unknown). Product ions in the low mass region (the ion series indicating the partial sequence PKG (or GKP)) support the sequence YPPKG. Similar to the other ECD spectra shown above, several potential amino acid assignments can be made from the ECD spectrum in Figure 2.5b. This spectrum suggests that the C-terminal sequence of the peptide is FGY, indicating a peptide with at least one chymotrypsin missed cleavage. The sequence preceding FGY could not be confidently sequenced here, due to several possible assignments with similar mass error, and lack of overlap with the IRMPD data. A comparison of Figures 2.5a and b shows the presence of only two "golden pairs", consisting of one *b*/*c* pair and one *y*/*z* pair. Overall, there is very little overlap between IRMPD and ECD data and due to the complicated nature of both the spectra no additional sequence information could be obtained. However, IRMPD and ECD spectra of other chymotrypsin digest peptides which were found to overlap with this peptide confirmed the partial sequence KPPY, and indicated the longer sequence GKPPYKGQFGY (data not shown). Overall, Figure 2.5 demonstrates that, in some cases, the combination of IRMPD and ECD provided very

little sequence information due to a lack of extensive *b* and *y*-type product ions following IRMPD and poor overlap between IRMPD and ECD data. Thus, it was necessary to perform MS/MS of several additional overlapping peptides in order to sequence this stretch of amino acids.

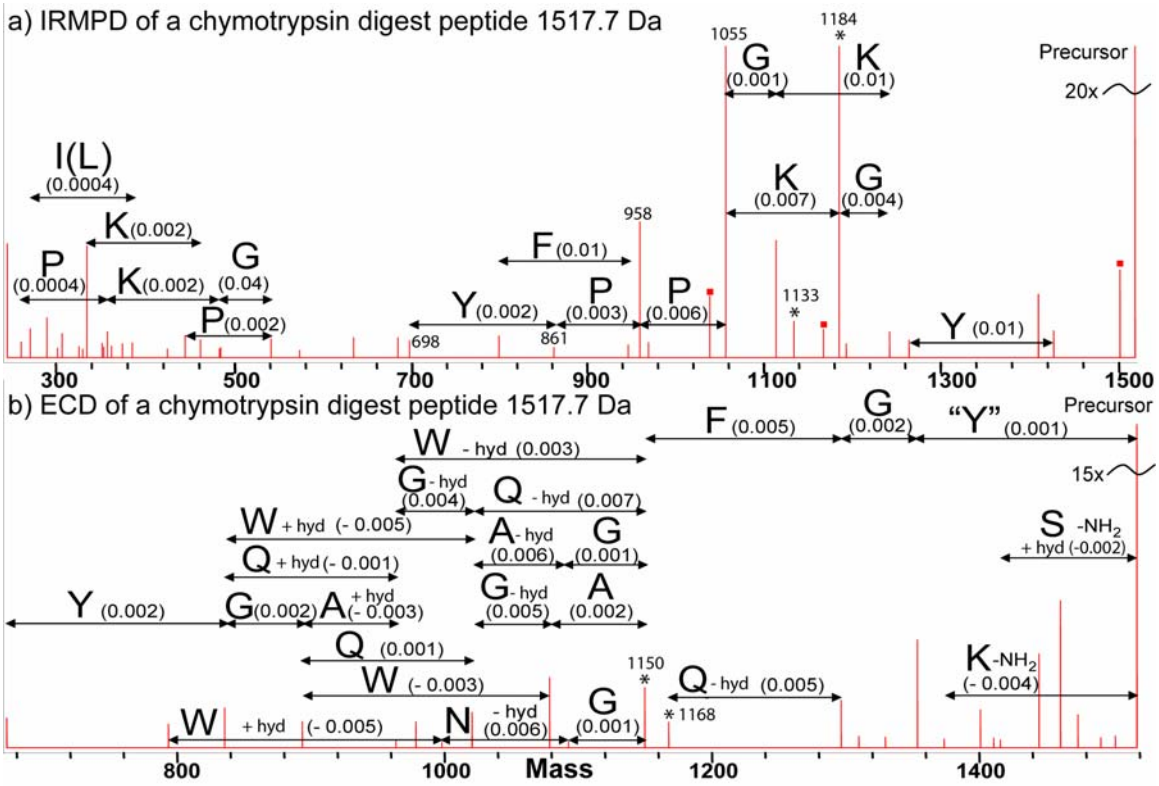


Figure 2.5. Deconvoluted a) IRMPD FT-ICR MS/MS spectrum (24 scans, 0.09 s, 10 W) and b) ECD FT-ICR MS/MS spectrum (80 scans, 35 ms, - 0.2 bias voltage) of a doubly protonated peptide from a chymotrypsin digest (1517.7). Squares indicate peaks which are likely due to water loss. +/- hyd indicates gain/loss of a hydrogen. Quotations around an amino acid assignment indicate a potential *c*-type ion (the mass gap is equal to the mass of the indicated amino acid and 0.984 Da, see Scheme 2.1). "Golden pairs" are indicated with asterisks. Specific product ions discussed in the text are indicated with mass labels.

2.3.4 Comparison of MS/MS Derived Sequence to Biochemical Gene Sequencing

A total of 28 peptides from overnight trypsin, GluC, and chymotrypsin digests were examined with IRMPD and ECD. Although large portions of each peptide could frequently be sequenced from a combination of fragmentation techniques (see Figures

2.2-2.5), only in some cases did these peptide sequences overlap. To arrange these peptide sequences within the protein, larger peptides from limited trypsin and chymotrypsin digestions were also examined. Several peptides, ranging in size from 2,100 Da to 6,100 Da, were fragmented with IRMPD and ECD (data not shown). Although MS/MS of these larger peptides could have been used directly for sequencing, their relatively large size resulted in fragmentation which was insufficient for *de novo* sequencing purposes (for both IRMPD and ECD). A summary of the amino acid sequence of the lectin derived from MS/MS data is shown in Figure 2.6a. The combination of IRMPD and ECD revealed the sequences of several large portions of the protein. However, several sections could not be sequenced with tandem mass spectrometry, and are underlined in Figure 2.6a. These areas could not be sequenced due to poor fragmentation. In some examples, the mass gaps can be quantified from tandem mass spectra of peptides from limited trypsin and chymotrypsin digests.

As previously discussed, a mass spectrometric characterization of this protein occurred concurrently with gene sequencing. Following the completion of *de novo* sequencing with tandem mass spectrometry, the protein sequence was also obtained with RT-PCR and 3' RACE. The peptide sequence determined from automated Edman sequencing was used to generate forward and reverse primer combinations for RT-PCR. This analysis identified the first 57 amino acids of the protein. 3'RACE was employed to obtain the protein sequence 3' to these 57 amino acids. The complete protein sequence obtained through these techniques is shown in Figure 2.6b, along with a comparison of the sequence obtained from tandem mass spectrometry. The protein, as sequenced with RT-PCR and 3'RACE, has a monoisotopic mass of 10,275.8 Da, which, with the

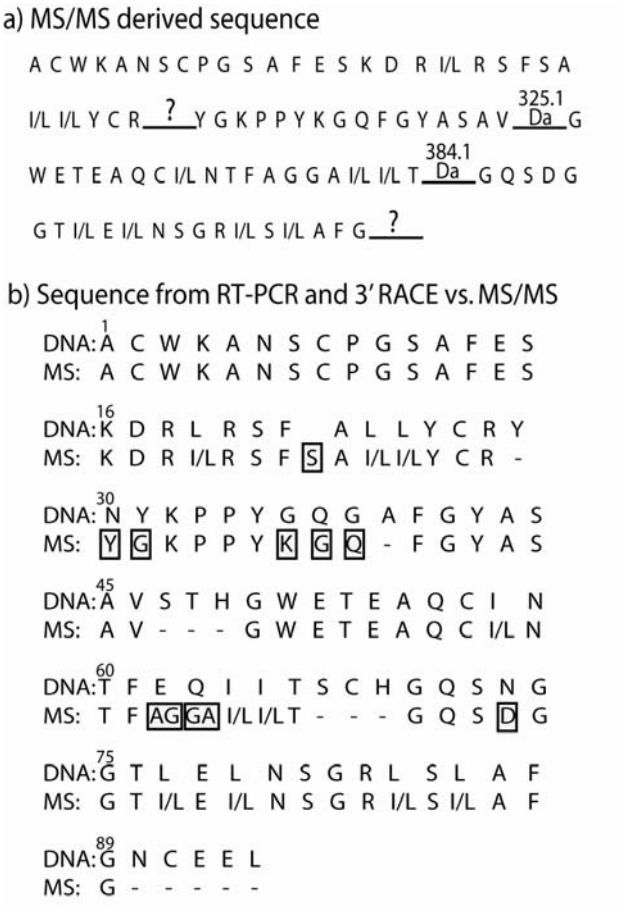


Figure 2.6. a) MS/MS derived amino acid sequence of a lectin from *Lyophyllum decastes*. Several portions of the protein could be sequenced using a combination of IRMPD and ECD. Unsequenced portions are underlined in the Figure, and when the mass of the missing gap is known, these have been indicated. b) A comparison of the sequence derived from MS/MS versus that obtained through RT-PCR and 3' RACE. Squares around an amino acid indicate incorrectly sequenced residues (by MS/MS), while dashes indicate residues which were not sequenced with MS/MS.

presence of three disulfide bonds would be 10269.8 Da. This value is within 0.3 Da of the value measured with ESI FT-ICR MS. Figure 2.6b indicates that the combination of IRMPD and ECD correctly sequenced 73 out of 94 amino acids, which corresponds to 77% of the total mass of the protein. Of the amino acids not sequenced correctly, 9 were incorrectly sequenced with mass spectrometry (indicated in Figure 2.6b with boxes) and 13 could not be sequenced at all due to insufficient fragmentation information (indicated in Figure 2.6b with dashes). Figure 2.6a indicates several gaps in the amino acid sequence which could not be determined with tandem mass spectrometry. The masses of two of these gaps are 325.1 Da and 384.1 Da. RT-PCR and 3' RACE indicated that these gaps correspond to STH (325.1 Da) and SCH (384.1 Da). In Figure 2.6b, Residue 73 was

incorrectly sequenced with MS/MS as aspartic acid instead of asparagine. This discrepancy is likely an artifactual deamidation that was introduced during enzymatic digestion. MS/MS also incorrectly sequenced residue 63 as GA instead of Q. The combined mass of glycine and alanine is the same as that of glutamine. MS/MS data indicated a glycine residue in this position, which may have been due to internal fragmentation, or a contaminant peak. MS/MS data also incorrectly sequenced a serine residue after phenylalanine at residue 22. In this example, ECD indicated the peptide sequence FSAI(L)(L)YCR, but IRMPD data could not confirm this sequence. Residues 36-38 were also incorrectly sequenced from ECD data, possibly due to artifact or contaminant mass spectral peaks.

2.4 Conclusions

The results presented here demonstrate that the combination of IRMPD and ECD is a valuable tool for *de novo* protein sequencing, and can be especially useful for proteins from organisms with unsequenced genomes. While IRMPD spectra were complicated by internal fragmentation and neutral molecule losses, ECD spectra were complicated by hydrogen addition and abstraction, multiple product ion types (*a*, *c*, *y*, and *z*-type), and low fragmentation efficiency. Despite these disadvantages, when used in conjunction these techniques proved to be a powerful tool for sequencing purposes. In these experiments, the overlap between IRMPD and ECD data, indicated by so called "golden pairs", was used to sequence peptides generated from trypsin, chymotrypsin, and GluC digestions. In instances where very few "golden pairs" were observed, fragmentation information from either IRMPD or ECD could be used for sequencing.

Several enzymes were used in order to generate overlapping peptides, which was found to be vital for *de novo* sequencing. Although 100% of the amino acid sequence of this protein was not sequenced, several improvements can be made to make complete sequencing feasible. Utilizing a wider variety of proteolytic enzymes and increasing the total number of peptides examined would generate additional MS/MS data, and fill in the blanks where amino acid sequences were unassigned due to insufficient fragmentation data. In addition, automated data interpretation would improve sequencing efficacy.

2.5 References

1. Aebersold, R.; Mann, M. *Nature* **2003**, *13*, 198-207.
2. Mann, M.; Jensen, O. N. *Nature Biotechnol.* **2003**, *21*, 255-261.
3. Pandey, A.; Mann, M. *Nature* **2000**, *405*, 837-846.
4. Henzel, W. J.; Billeci, T. M.; Stults, J. T.; Wong, S. C.; Grimley, C.; Watanabe, C. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5011-5015.
5. Mann, M.; Hojrup, P.; Roepstorff, P. *Biol. Mass Spectrom.* **1993**, *22*, 338-345.
6. Pappin, D. J. C.; Hojrup, P.; Bleasby, A. J. *Curr. Biol.* **1993**, *3*, 327-332.
7. James, P.; Quadroni, M.; Carafoli, E.; Gonnet, G. *Biochem. Biophys. Res. Commun.* **1993**, *195*, 58-64.
8. Yates, J. R.; Speicher, S.; Griffin, P. R.; Hunkapiller, T. *Anal. Biochem.* **1993**, *214*, 397-408.
9. Mann, M.; Wilm, M. *Anal. Chem.* **1994**, *66*, 4390-4399.
10. Mørtz, E.; O'Connor, P.; Roepstorff, P.; Kelleher, N. L.; Wood, T. D.; McLafferty, F. W.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8264-8267.
11. Shevchenko, A.; Jensen, O. N.; Podtelejnikov, A. V.; Sagliocco, F.; Wilm, M.; Vorm, O.; Mortensen, P.; Shevchenko, A.; Boucherie, H.; Mann, M. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 14440-14445.
12. McLafferty, F. W. *Tandem Mass Spectrometry*; Wiley: New York, 1983.
13. Biemann, K.; Martin, S. A. *Mass Spectrom. Rev.* **1987**, *6*, 1-76.
14. McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599-614.
15. Roepstorff, P.; Fohlman, J. *Biomed. Mass Spectrom.* **1984**, *11*, 601-601.
16. Woodlin, R. L.; Bomse, D. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250.
17. Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
18. Yu, W.; Vath, J. E.; Huberty, M. C.; Martin, S. A. *Anal. Chem.* **1993**, 3015-3023.
19. Huang, Y.; Triscari, J. M.; Pasa-Tolic, L.; Anderson, G. A.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *J. Am. Chem. Soc.* **2004**, 3034-3035.
20. Roth, K. D.; Huang, Z. H.; Sadagopan, N.; Watson, J. T. *Mass Spectrom. Rev.* **1998**, *17*, 255-274.
21. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
22. Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201-222.
23. Zubarev, R. A. *Mass Spectrom. Rev.* **2003**, *22*, 57-77.
24. Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208-224.
25. Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 10313-10317.
26. Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *Angew. Chem. Int. Ed.* **2006**, *45*, 5301-5303.
27. Hakansson, K.; Cooper, H. J.; Hudgins, R. R.; Nilsson, C. L. *Curr. Org. Chem.* **2003**, *7*, 1503-1525.
28. Zubarev, R. *Exp. Rev. Proteomics* **2006**, *3*, 251-261.

29. Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2006**, *18*, 113-120.
30. O'Connor, P. B.; Lin, C.; Cournoyer, J. J.; Pittman, J. L.; Belyayev, M.; Budnik, B. A. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 576-585.
31. Savitski, M. M.; Nielsen, M. L.; Kjeldsen, F.; Zubarev, R. A. *J. Proteome Res.* **2005**, *4*, 2348-2354.
32. Branca, R. M. M.; Bodo, G.; Bagyinka, C.; Prokai, L. *J. Mass Spectrom.* **2007**, *42*, 1569-1582.
33. Goldstein, I. J.; Winter, H. C.; Aurandt, J.; Confer, L.; Adamson, J. T.; Hakansson, K.; Remmer, H. *Arch. Biochem. Biophys.* **2007**, *467*, 268-274.
34. Blakney, G. T.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Proc. 50th ASMS Conference on Mass Spectrometry and Allied Topics*, Orlando, FL, June 2-6 2002; CD-ROM.
35. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2004**, *77*, 1876-1882.
36. Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26*, 514-518.
37. Heck, A. J. R.; Derrick, P. J. *Anal. Chem.* **1997**, *69*, 3603-3607.
38. Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.
39. Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.
40. Smith, D. L.; Zhou, Z. *Methods Enzymol.* **1990**, *193*, 374-389.
41. Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241-249.
42. Fung, Y. M. E.; Chan, T. W. D. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1523-1535.
43. Chalkley, R. J.; Brinkworth, C. S.; Burlingame, A. L. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1271-1274.
44. Savitski, M. M.; Nielsen, M. L.; Zubarev, R. *Anal. Chem.* **2007**, *79*, 2296-2302.
45. Burrell, M. M., Ed. *Enzymes of Molecular Biology, Vol. 16*; Humana Press: Totowa, NH, 1993.
46. Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.
47. Sze, S. K.; Ge, Y.; Oh, H.; McLafferty, F. W. *Anal. Chem.* **2003**, *75*, 1599-1603.
48. Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
49. Schwartz, B. L.; Bursey, M. M. *Biol. Mass Spectrom.* **1992**, *21*, 92-96.
50. Tang, X. J.; Thibault, P.; Boyd, R. K. *Anal. Chem.* **1993**, *65*, 2824-2834.
51. Leymarie, N.; Berg, E. A.; McComb, M. E.; O'Connor, P. B.; Grogan, J.; Oppenheim, F. G.; Costello, C. E. *Anal. Chem.* **2002**, *74*, 4124-4132.

Chapter 3

Infrared Multiphoton Dissociation and Electron Capture Dissociation of High-Mannose Type Glycopeptides

3.1 Introduction

A protein can exist in several diverse states within a cell, largely due to different splice variants and the great number of post-translational modifications (PTMs) that can occur at multiple positions within the protein. Glycoproteins are a highly diverse class of biomolecules that have been found to play several key roles in biological systems including: cell-cell adhesion, cell-extra cellular membrane adhesion, folding and secretion, fertilization, glycoprotein targeting, and immune defense.[1,2] Abnormal glycosylation patterns have been linked to several disease states, such as protein misfolding in neurodegenerative diseases,[3,4] susceptibility to infection,[5] evasion of the immune system by cancer cells,[6,7] and congenital disorders of glycosylation (CDGs).[8] Despite their frequency and importance to protein structure and function, analyses of carbohydrates have been underrepresented compared to nucleic acids and proteins.

The two key types of glycosylation involve either covalent attachment of an oligosaccharide through the oxygen in serine or threonine (*O*-glycosylation) or through nitrogen in asparagine (*N*-glycosylation). In cases of *N*-glycosylation, a common

pentasaccharide core is attached to Asn in the consensus sequence Asn-X-Ser/Thr, where X may be any amino acid except Pro. There are three main classes of *N*-glycosylation, distinguished by the saccharide units which extend beyond the common core. These include hybrid type, complex type, and high-mannose type glycans. Representative structures of these glycans are shown in Figure 1.1 in Chapter 1. Although typically not considered a key glycan category, xylose type glycans are distinguished by the attachment of xylose to the common pentasaccharide core and are found predominantly in plants.[9]

Mass spectrometry (MS) is a valuable tool for glycoprotein analysis, and tandem mass spectrometry (MS^n) is particularly effective in characterizing peptide and saccharide structures.[10-12] Fourier transform ion cyclotron resonance (FT-ICR) MS, with its ultrahigh resolution, mass accuracy, and multiple tandem mass spectrometry abilities, is a powerful tool for the determination and structural examination of several types of biological macromolecules.[13] The current procedure for mass spectrometry-based glycoprotein characterization often employs release of glycans through chemical or enzymatic means, followed by separate analyses of the protein and carbohydrate. Another approach is to investigate glycosylation on the glycopeptide level with tandem mass spectrometry. This strategy minimizes sample manipulation and allows mapping of glycan structures to specific sites.[14-16]

Electron capture dissociation (ECD)[17] and infrared multiphoton dissociation (IRMPD)[18,19] are powerful fragmentation techniques for FT-ICR MS characterization of glycoproteins. ECD is a rather recently introduced technique, but its expanding implementation has been a significant advance in the field of biomolecular structural

analysis. ECD involves the irradiation of multiply charged cationic analyte ions with low energy electrons (< 1 eV), generating charge reduced radical species and product ions.[20-23] It has been extensively shown that ECD of peptides results in preferential cleavage along the peptide backbone generating *c′*-type and radical *z*-type ions (Zubarev nomenclature)[24] while retaining labile modifications such as phosphorylation and glycosylation.[25-32] IRMPD of *N*-glycosylated peptides is understood to selectively cleave glycosidic bonds rather than the peptide backbone. This preference has been demonstrated for both xylose type and complex type glycopeptides.[29,31,33] Accordingly, the combination of ECD and IRMPD has been shown to provide complementary structural information regarding peptide and glycan structure.[29,31] In addition, the recent combination of electron transfer dissociation (ETD) and CAD in a quadrupole ion trap mass spectrometer has shown similar capabilities for a xylose type glycopeptide.[34]

While ECD has been applied to the structural analysis of xylose type[29,31] and complex type[30] *N*-glycopeptides and *O*-glycosylated peptides,[26,30,35] it has not been widely applied towards the characterization of high-mannose type *N*-glycopeptides. Similarly, neither has the utility of IRMPD for the structural characterization of such glycopeptides been extensively explored.[36] Ribonuclease B, which has been characterized with MS,[37-40] X-ray crystallography,[41] and NMR,[42] was chosen as a model glycoprotein containing high-mannose type glycans. With this model, we employ the combination of ECD and IRMPD for the structural characterization of high-mannose type glycopeptides. The utility of ECD and IRMPD is examined for glycopeptides ranging in size from four to forty amino acids, the latter being, to our knowledge, the largest *N*-glycosylated peptide characterized through these fragmentation techniques to date.

3.2 Experimental

3.2.1 Ribonuclease B Preparation

Approximately 1.25 nmoles of ribonuclease B (Sigma, St Louis, MI) in 100 mM ammonium bicarbonate (NH_4HCO_3, Fisher Scientific, Fair Lawn, NJ) was mixed with 10 µL of 10 mM dithiotreitol (DTT, Sigma) and incubated at 56°C for one hour. After cooling to room temperature (RT), 10 µL of 100 mM iodoacetamide (Sigma) was added followed by incubation in the dark at RT for one hour. Another 10 µL of 10 mM DTT was added to the solution at RT and, after 45 minutes, either 12.5 ng/µL of trypsin (Princeton Separations, Adelphia, NJ) or GluC (Roche, Indianapolis, IN) in 25 mM NH_4HCO_3 was added (modified procedure[43]). For a tryptic digestion, 100 µL of trypsin was used and the digestion proceeded for 15 hours at 37 °C. For GluC digests, between 15-50 µL of enzyme was used and digestion proceeded at 25°C for 15-18 hours. In addition, a limited tryptic digest was prepared by adding 14 µL of trypsin and allowing the reaction to proceed for five minutes at 2 °C. Following both trypsin and GluC digestion, all solutions were acidified to 0.1% formic acid.

For the overnight trypsin digest, concanavalin A tips (Glygen Corporation, Columbia, MD) were used to selectively bind glycopeptides. Loading buffer consisted of 100 mM sodium acetate (CH_3COONa, Sigma) and 200 mM NaCl (Fisher), washing solution consisted of 50 mM CH_3COONa and 200 mM NaCl, and elution was performed in 150 mM glucose (Fluka, Milwaukee, WI), 25 mM CH_3COONa, and 100 mM NaCl. Following enrichment with concanavalin A tips, the glycopeptide sample was desalted using a C18 reverse-phase micro-column, Ziptip (Millipore, Billerica, MA). For GluC digests and the short tryptic digest, samples were desalted following the addition of

formic acid. All samples were diluted to $5x10^{-6}$ M with electrospray solvent consisting of 1:1 methanol and water with 2% acetic acid.

3.2.2 Lectin Preparation

Approximately 150 μmoles of a lectin from *Erythrina cristagalli* (Sigma) in 50 μL of water was mixed with 20 μL of 12.5 ng/μL of trypsin in 25 mM NH_4HCO_3. Digestion proceeded for 15 hours at 37 °C. Desalting conditions were used as described above. The sample was diluted to $1x10^{-6}$ M with electrospray solvent.

3.2.3 FT-ICR Mass Spectrometry

All experiments were performed with an actively shielded 7 T FT-ICR mass spectrometer with a quadrupole front-end (APEX-Q, Bruker Daltonics, Billerica, MA), as previously described.[44] An indirectly heated hollow dispenser cathode[45] was used to perform ECD. IRMPD was performed with a vertically mounted 25-W, 10.6-μm, CO_2 laser (Synrad, Mukilteo, WA).

Peptide solutions were infused via an external Apollo electrospray ion source at a flow rate of 60 μL/h with the assistance of N_2 nebulizing gas. Ions were accumulated in the first hexapole for 0.05 s, transferred through a mass-selective quadrupole, and mass-selectively accumulated (2-10 *m/z* isolation window) in the second hexapole for 1-4 s. Ions were then transferred through high-voltage ion optics and captured by dynamic trapping (IRMPD experiments) or gas-assisted dynamic trapping (ECD experiments) with argon as the collision gas in an Infinity ICR cell.[46] The experimental sequence up to the ICR cell fill was looped 1-5 times to achieve maximum precursor ion signal. The cathode heating current was 1.8 A, and during the ECD event the cathode voltage was pulsed for 15 ms to - 0.25 V to generate low energy electrons. A lens electrode located

immediately in front of the cathode was kept at 1 V to improve electron injection. IRMPD was performed at 30% laser power with firing times ranging from 45-100 ms. All mass spectra were acquired with XMASS software (Bruker Daltonics) with 256 data points from *m/z* 200-2500 and summed over 30-50 scans. Data processing was performed with MIDAS software[47,48] and exported to Microsoft Excel for internal frequency-to-mass calibration with a two-term calibration equation.[49] All IRMPD product ions were within 10 ppm error whereas ECD product ions were assigned within 20 ppm. The larger error allowed in ECD was due to smaller signal to noise ratios, the pulsing of gas into the cell, and because a two-point calibration does not properly account for space-charge effects caused by the introduction of electrons into the ICR cell.

3.3 Results and Discussion

Ribonuclease B is a protein of 124 amino acids, with the heterogeneous glycan structure $GlcNAc_2Man_{5-9}$ (GlcNAc=N-acetylglucosamine, Man=mannose) attached to Asn34. NMR experiments have determined the structures of the several high-mannose type glycans found on Asn34, as well as their relative abundances.[42] With such extensive characterization, this glycoprotein is a suitable model system to examine high-mannose type glycopeptides. When glycan structures are shown throughout this Chapter, the most abundant glycan species are indicated in cases where multiple isomers exist. All saccharide units, excluding xylose (indicated by a rhombus), are represented according to the nomenclature designated by the Consortium for Functional Glycomics (CFG).[50] In this nomenclature scheme, green circles indicate mannose, blue squares indicate N-acetylglucosamine, and red triangles indicate fucose.

3.3.1 IRMPD and ECD of a Trypsin Digest Glycopeptide

The electrospray ionization (ESI) FT-ICR mass spectrum from a non-separated tryptic digest of ribonuclease B showed that glycopeptide ion abundances were very low compared to non-glycosylated peptides (data not shown). This behavior is a general occurrence for glycoprotein digests, due to the spread of glycopeptide ion signals over a population of several peptides with varying glycan structures. For this reason, chromatographic separation is often necessary to enrich glycopeptides prior to further analysis. Lectin affinity chromatography is a commonly used strategy for purification and fractionation of glycoproteins, glycopeptides, and oligosaccharides.[51-53] This technique is based on the high affinity of lectins towards carbohydrate structures. For example, concanavalin A is a lectin with broad carbohydrate specificity and was used here to enrich glycopeptides from the ribonuclease B tryptic digest.

Figure 3.1 shows the IRMPD mass spectrum from a doubly protonated *N*-glycosylated high-mannose type glycopeptide observed following concanavalin A treatment. The peptide sequence and predicted glycan structure are also shown in Figure 3.1, along with observed cleavage sites (indicated by red dotted lines). Most product ions observed correspond to the precursor ion, $[M + 2H]^{2+}$, with one or more monosaccharide losses and with charge retention on the reducing end of the glycopeptide. Also observed are several singly protonated, dehydrated sugar ions. All glycosidic bonds are cleaved, yielding extensive information regarding glycan composition. The sequential loss of seven 162 Da masses confirms the known mannose content of this glycopeptide. Similar to previous IRMPD spectra of xylose type and complex type glycopeptides, no peptide backbone cleavage is observed. These results re-emphasize the utility of IRMPD for

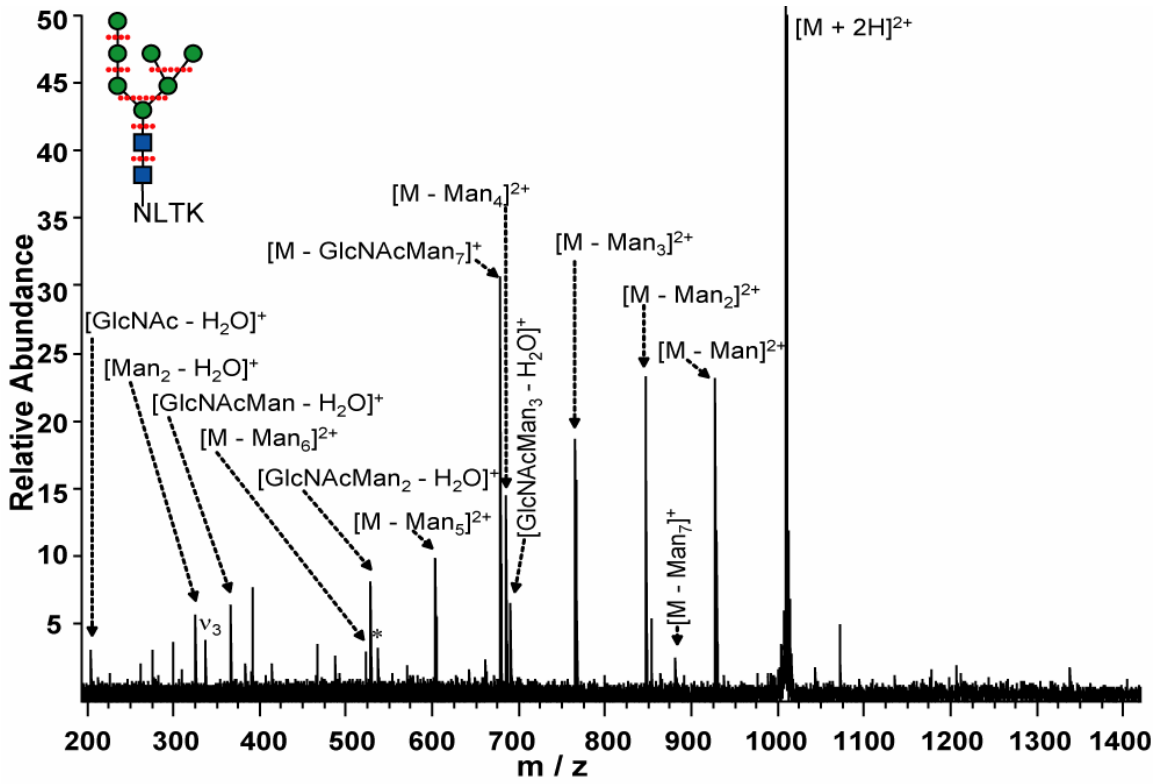


Figure 3.1. IRMPD FT-ICR (50 scans, 45 ms irradiation at 7.5 W laser power) tandem mass spectrum of a ribonuclease B glycopeptide from an overnight trypsin digestion. The precursor ion, denoted as $[M + 2H]^{2+}$, consists of amino acids 34-37 with a high mannose glycan, $GlcNAc_2Man_7$, attached at Asn34. Extensive cleavage within the glycan is observed (indicated by red dotted lines); however, no cleavage is seen between Asn-GlcNAc. Similarly to previous IRMPD of xylose type and complex type glycopeptides, no backbone fragmentation is observed. (ν_3=harmonic peak, *=noise, ■=GlcNAc, ●=Man).

glycopeptide structural characterization, this technique being particularly useful for uncharacterized glycopeptides due to selective glycan cleavage.

Analogous to previous characterizations of xylose type and complex type glycans, the ECD spectrum of the same high-mannose glycopeptide resulted in only peptide backbone cleavage (data not shown). No cleavage occurred within the glycan but, because of the small size of the peptide, only one *c′*-type ion (containing the glycans) was detected. However, as with previous analyses, the fragmentation observed in IRMPD

and ECD was strictly complementary.

3.3.2 IRMPD of a GluC Digest Glycopeptide

To investigate the utility of the IRMPD/ECD approach for the characterization of larger high-mannose type glycopeptides, we performed a GluC digest of ribonuclease B. In the mass spectrum obtained following direct infusion of this digest (without glycopeptide enrichment), several glycopeptides in the 5900-6600 Da range were readily observed. One of those was selected for further analysis with IRMPD and ECD. To demonstrate that these two fragmentation techniques are applicable to a variety of high-mannose glycopeptides, a glycopeptide with a different glycan structure, $GlcNAc_2Man_6$, was chosen.

Figure 3.2 shows the IRMPD mass spectrum of this high-mannose type glycopeptide consisting of amino acids 10-49 with the glycan $GlcNAc_2Man_6$ attached at Asn34. The precursor ion is indicated by $[M + 7H]^{7+}$. Unlike the IRMPD mass spectrum of the smaller tryptic glycopeptide, very limited glycan cleavage is observed. The only ions that correspond to glycosidic cleavage are losses of one and two mannoses from the precursor ion. In contrast to previous IRMPD mass spectra of glycopeptides, peptide backbone cleavage is readily observed, including five *b* type ions and one *y′* type ion (Zubarev nomenclature).[24] This effect is not due to a longer laser irradiation time used here (100 ms), versus the shorter pulse (45 ms) used for the smaller trypsin digest glycopeptide. A longer irradiation time was necessary for fragmentation due to the increased number of vibrational degrees of freedom of this larger glycopeptide. Irradiation times ranging from 60-90 ms still resulted in peptide backbone cleavage, and irradiation for less then 60 ms resulted in the loss of one mannose, with no other

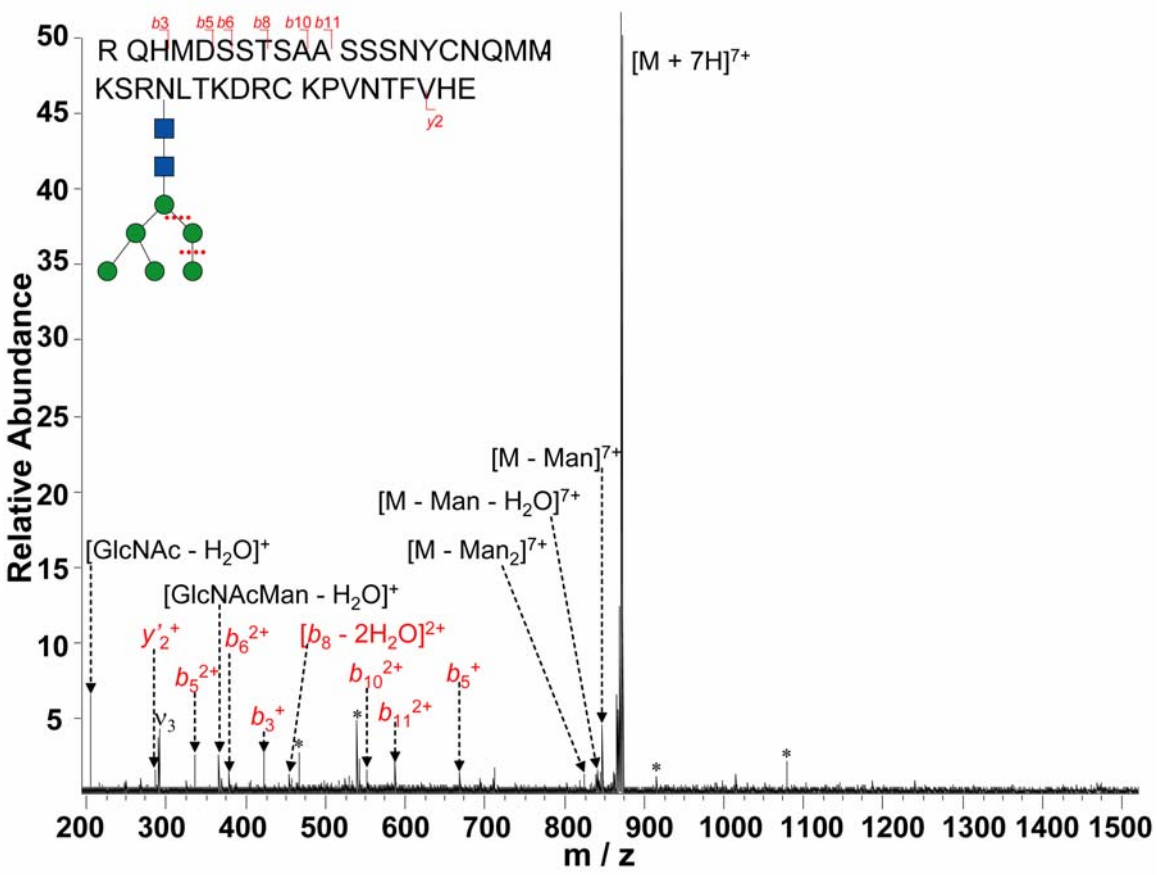


Figure 3.2. IRMPD FT-ICR (30 scans, 100 ms irradiation at 7.5 W laser power) tandem mass spectrum of a ribonuclease B glycopeptide from a GluC digestion. The precursor ion, denoted as $[M + 7H]^{7+}$, consists of amino acids 10-49 with a high-mannose glycan, $GlcNAc_2Man_6$, attached at Asn34. Very limited glycan cleavage is observed (the most likely glycosidic cleavages are indicated by red dotted lines). Contrary to what has been shown previously for IRMPD of glycopeptides, several *b* and *y*′ type product ions are also detected (indicated by red solid lines). (ν_3=harmonic peak, *=noise, ■=GlcNAc, ●=Man).

cleavages observed (data not shown). These experiments demonstrate that the fragmentation pattern observed was not a result of "harsh" experimental parameters. Increasing the irradiation time to 150 ms resulted in no additional glycan cleavages, indicating that incomplete glycosidic bond cleavage was not due to insufficient energy input (data not shown).

It is also interesting to note here that *b* and *y*′ ion abundances are similar to those corresponding to glycan cleavage, indicating that the degree of backbone cleavage and glycan cleavage is comparable. This mixture of glycan and peptide backbone cleavage greatly complicates spectral interpretation, and would render structural characterization of an unknown glycopeptide difficult. These results are in disagreement with IRMPD data observed for xylose type and complex type glycopeptides, which showed that IRMPD selectively induces dissociation of glycosidic bonds in N-linked glycans and leaves the peptide backbone intact.[29,31,33]

Caprioli and co-workers have shown that CAD of high-mannose type glycopeptides results in minor peptide backbone cleavage in addition to extensive glycosidic cleavage.[54] In one particular example, poor glycopeptide fragmentation was explained due to ion suppression effects and or as a result of increased glycopeptide mass preventing efficient dissociation and/or detection of product ions. However, Zaia has reasoned that the tendency towards peptide backbone cleavage is due to the lack of a labile HexNAc (*N*-acetylhexosamine) residue in the antennae region of the sugar.[11] Cleavage to the reducing side of HexNAc is favored due to the proximity of a charged group to the glycosidic oxygen. Other labile saccharides include sialic acid and fucose.[11,55,56]

McLafferty and co-workers found that CAD of large (>6 kDa) high-mannose type glycopeptides from immunoglobulin constructs resulted in product ions formed from a mixture of peptide backbone cleavage and glycosidic cleavage.[57] In that example complete glycan sequencing was still feasible, contrary to the poor glycosidic cleavage observed here in IRMPD of a GluC ribonuclease B glycopeptide. The tendency towards

peptide backbone cleavage over glycan cleavage was also observed in ion trap CAD of whole ribonuclease B.[58] Here, the authors proposed that for the whole glycosylated protein, the N-linked sugar was "inert" to fragmentation due to the competition from facile amide bond cleavage predominantly N-terminal to Pro and C-terminal to Asp and Lys. This competition was able to "protect" glycosidic bonds from cleavage. To determine whether the lack of glycan cleavage for a ribonuclease B GluC glycopeptide was due to size effects, a peptide of smaller length, but larger than the previous tryptic glycopeptide, was also examined (see Section 3.3.4).

3.3.3 ECD of a GluC Digest Glycopeptide

Figure 3.3 shows the ECD mass spectrum of the same ribonuclease B GluC digest glycopeptide that was examined with IRMPD. To our knowledge, this GluC proteolytic peptide is the largest glycopeptide that has been subjected to ECD. Extensive fragmentation was observed, corresponding mostly to c' and $z^{\bullet}$ type ions. Of the 38 peptide backbone bonds available for fragmentation, cleavage was observed at 32 sites. The N-terminal side of proline was not considered because cleavage at this site is generally not observed in ECD due to proline's cyclic structure.[17] As seen in previous ECD experiments,[26,59,60] z ions were a mixture of both radical ($z^{\bullet}$) and even-electron (z, z') species. These even-electron z ions have been proposed to be formed by direct hydrogen atom capture during the ECD process or by hydrogen rearrangement.[59] In addition, two y' ions ($y'_{12}{}^{2+}$, $y'_{18}{}^{3+}$) and two radical a ($a_{22}{}^{3+\bullet}$, $a_{28}{}^{4+\bullet}$) ions were detected. Formation of such ions is a minor fragmentation pattern in ECD, although it has been seen as a predominant pathway for non-standard peptide-like structures.[61] As expected, no glycosidic cleavages were observed following ECD.

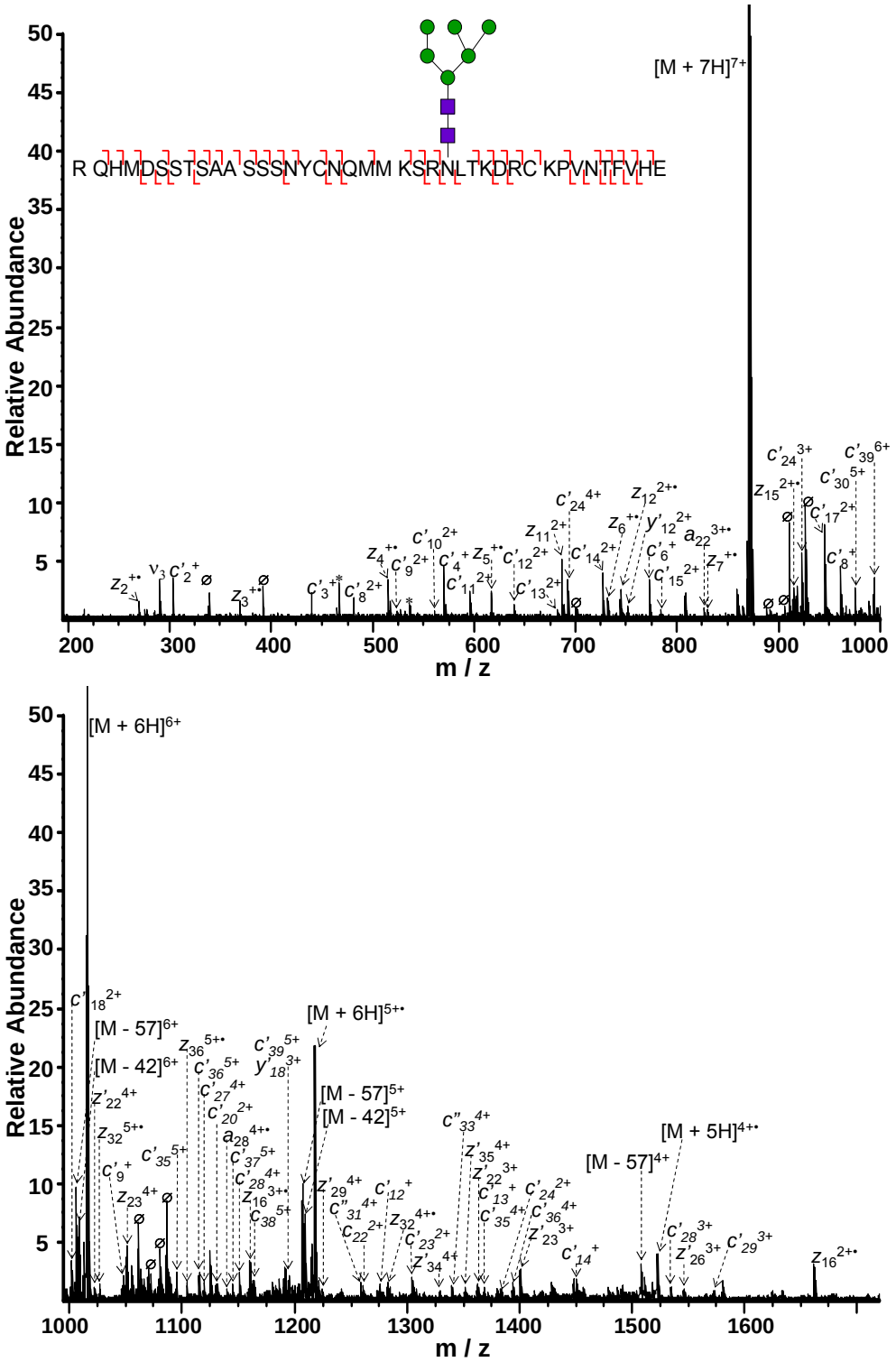


Figure 3.3. ECD FT-ICR (40 scans, 15 ms irradiation, - 0.25 V bias voltage) tandem mass spectrum of a ribonuclease B glycopeptide from a GluC digestion. The precursor ion, denoted as $[M + 7H]^{7+}$, consists of amino acids 10-49 with a high-mannose glycan, $GlcNAc_2Man_6$, attached at Asn34. Most product ions correspond to c' and $z^{\bullet}$ type ions, however some $a^{\bullet}$ and y' type ions are also observed. (ν_3=harmonic peak, *=noise, ø=due to quadrupole isolation, ■=GlcNAc, ●=Man).

The most abundant ion (Fig. 3.3, top) corresponds to the precursor ion, amino acids 10-49 with $GlcNAc_2Man_6$ attached at Asn34, indicated by $[M + 7H]^{7+}$. ECD product ions observed in Figure 3.3 are a result of electron capture induced dissociation of the precursor ion, $[M + 7H]^{7+}$, and charge stripped species resulting from quadrupole isolation of the precursor ion; $[M + 6H]^{6+}$, $[M + 5H]^{5+}$, and $[M + 4H]^{4+}$. The major product of electron capture dissociation is often $[M + nH]^{(n-1)+\bullet}$, which fragments to yield c' and $z^{\bullet}$ type ions. However, an examination of the isotopic distribution of the 6^+ species reveals that the even electron species (charge stripped species) is more dominant than the 6^+ radical species. An examination of the 5^+ and 4^+ species revealed that these were a mixture of even and odd electron species. Neutral losses were also quite dominant in the spectrum, labeled as $[M - 57]^{n+}$ and $[M - 42]^{n+}$. The former corresponds to the loss of C_2H_4ON (58.023 Da) from the charge reduced species (the charge reduced species is one hydrogen heavier then the neutral species, [M], thus the loss of 57 Da), which has been previously observed due to the elimination of a portion of the carboxyamidomethylated side chain of a cysteine residue.[32] The other prominent neutral loss is either from the loss of CH_3N_2 (43.029 Da)[62] or CH_3CO (43.018 Da)[63] from the charge reduced species. Both losses have been observed in ECD, although the latter was specifically attributed to the loss of an acetyl radical from a glycopeptide. These results demonstrate that ECD is a highly effective technique for the characterization of particularly large glycopeptides, resulting in almost complete sequence coverage and indicating the site of glycosylation for this high-mannose type glycopeptide.

3.3.4 IRMPD of a Short Trypsin Digest Glycopeptide

While the IRMPD spectrum of a relatively small glycopeptide from an over-night trypsin digest showed extensive glycosidic cleavage (Figure 3.1), that of a larger glycopeptide from a GluC digest showed little glycan cleavage and several product ions corresponding to peptide backbone cleavage (Figure 3.2). To determine whether the lack of glycan cleavage for the GluC glycopeptide was due to size effects, a peptide of smaller length, but larger than the previous tryptic glycopeptide, was examined. This peptide, consisting of amino acids 34-39 with the glycan $GlcNAc_2Man_5$ attached at Asn34, was produced by a short (5 min) trypsin digestion at 2 °C. Figure 3.4 shows the corresponding IRMPD spectrum of the doubly protonated $[M + 2H]^{2+}$ precursor ion. Similar to the IRMPD spectrum of the glycopeptide NLTK-$GlcNAc_2Man_7$ (Figure 3.1), extensive glycosidic cleavage is observed yielding all necessary information regarding glycan composition. However, highlighted in the spectrum is a *b*-type ion, corresponding to peptide backbone cleavage. This fragmentation pattern shows that IRMPD will likely result in a mixture of glycosidic and backbone cleavage for high-mannose type glycopeptides, regardless of their size. A resulting mixture of backbone and glycosidic cleavage is a major obstacle, because it can complicate spectral interpretation. This behavior is of particular concern in the current example, where the *b*-type ion exhibits both backbone cleavage and glycosidic cleavage. Thus, in cases where the glycan structure of a high-mannose type glycopeptide is unknown, IRMPD alone would likely be insufficient in determining the glycan composition. In such cases, alternative fragmentation techniques such as ECD would be necessary.

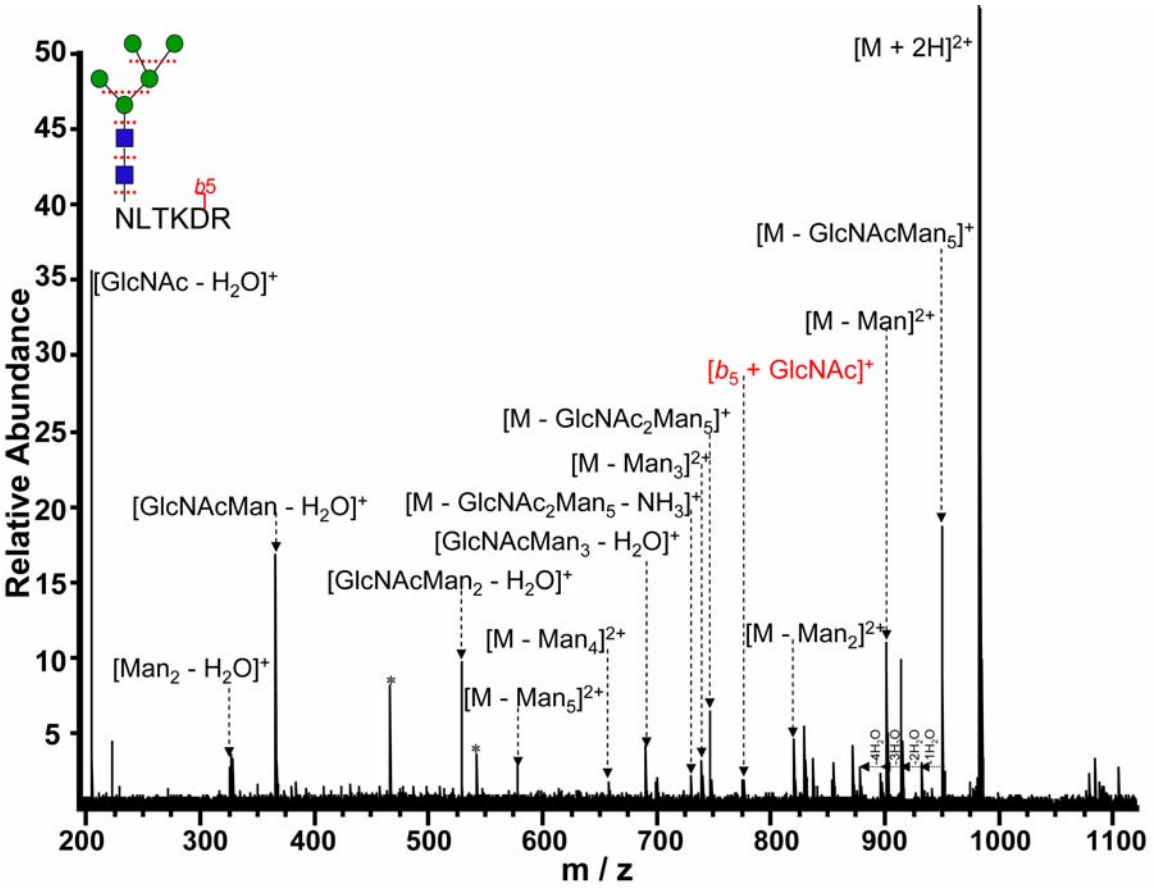


Figure 3.4. IRMPD FT-ICR (30 scans, 90 ms at 7.5 W laser power) tandem mass spectrum of a ribonuclease B glycopeptide from a five minute cold trypsin digestion. The precursor ion, denoted as $[M + 2H]^{2+}$, consists of amino acids 34-39 with a high-mannose glycan, $GlcNAc_2Man_5$, attached at Asn34. Similar to the IRMPD mass spectrum shown in Figure 3.1, extensive cleavage within the glycan is observed (indicated by red dotted lines). In addition, backbone cleavage is also observed in the form of a *b* type ion (indicated by red solid line). (ν_3=harmonic peak, *=noise, ø=due to quadrupole isolation, ■=GlcNAc, ●=Man).

3.3.5 IRMPD of a Xylose Type Glycopeptide

The results presented above show that for high-mannose type glycopeptides, peptide backbone cleavage competes with glycosidic bond cleavage when IRMPD is employed. This effect can be interpreted as a higher resistance to gas-phase cleavage by mannose-linking glycosidic bonds compared to peptide backbone bonds. To further investigate this occurrence, we re-examined the IRMPD behavior of a xylose type glycopeptide from a tryptic digest of *Erythrina cristagalli* lectin.[13] A similar xylose type

glycopeptide from *Erythrina corallodendron* had been previously characterized with IRMPD,[29,31] and those results are also used for comparison. Figure 3.5 shows the glycan structure for the xylose type glycopeptide from *Erythrina cristagalli*, and zoomed in regions corresponding to the loss of one and two monosaccharide residues. We expected that product ions corresponding to the loss of one mannose should be more abundant than those due to loss of xylose because mannose loss can occur through two combinations, i.e. via cleavage of two different glycosidic bonds. However, loss of mannose is not the most abundant pathway observed in the spectrum, demonstrating that mannose-linking glycosidic bonds are more resistant to cleavage compared to both xylose and fucose glycosidic bonds. These results are the same as those from an IRMPD characterization of an *Erythrina corallodendron* lectin glycopeptide.[29,31] As previously stated, fucose is considered a labile saccharide, consistent with our findings. However, our results also demonstrate lability of xylose, which has not been thoroughly discussed in the literature. An examination of product ions corresponding to two monosaccharide losses in Figure 3.5 shows similar results. Here, the combined loss of fucose and xylose is quite abundant, while the loss of two mannose residues is hardly observed. In previous *Erythrina corallodendron* IRMPD examinations of xylose type glycopeptides, the loss of two mannose residues was either not observed at all or also seen to a much lower degree.[29,31] Similarly, a CAD study of a xylose type glycopeptide did not report the loss of only two mannose residues.[34] These results are notable, because several current research efforts involve the gathering of statistical information that can be used to predict MS/MS behavior.[64-66] Several strategies are employed in order to aid the interpretation of mass spectra of carbohydrates. These strategies typically rely on carbohydrate MS/MS

libraries and/or algorithms to predict glycan structure based on experimental data.[67-76] No reliable strategy exists yet for the interpretation of glycopeptide fragmentation. Nonetheless, the unique behavior of mannose needs to be incorporated into such models if they are to be useful for glycopeptide structural characterization.

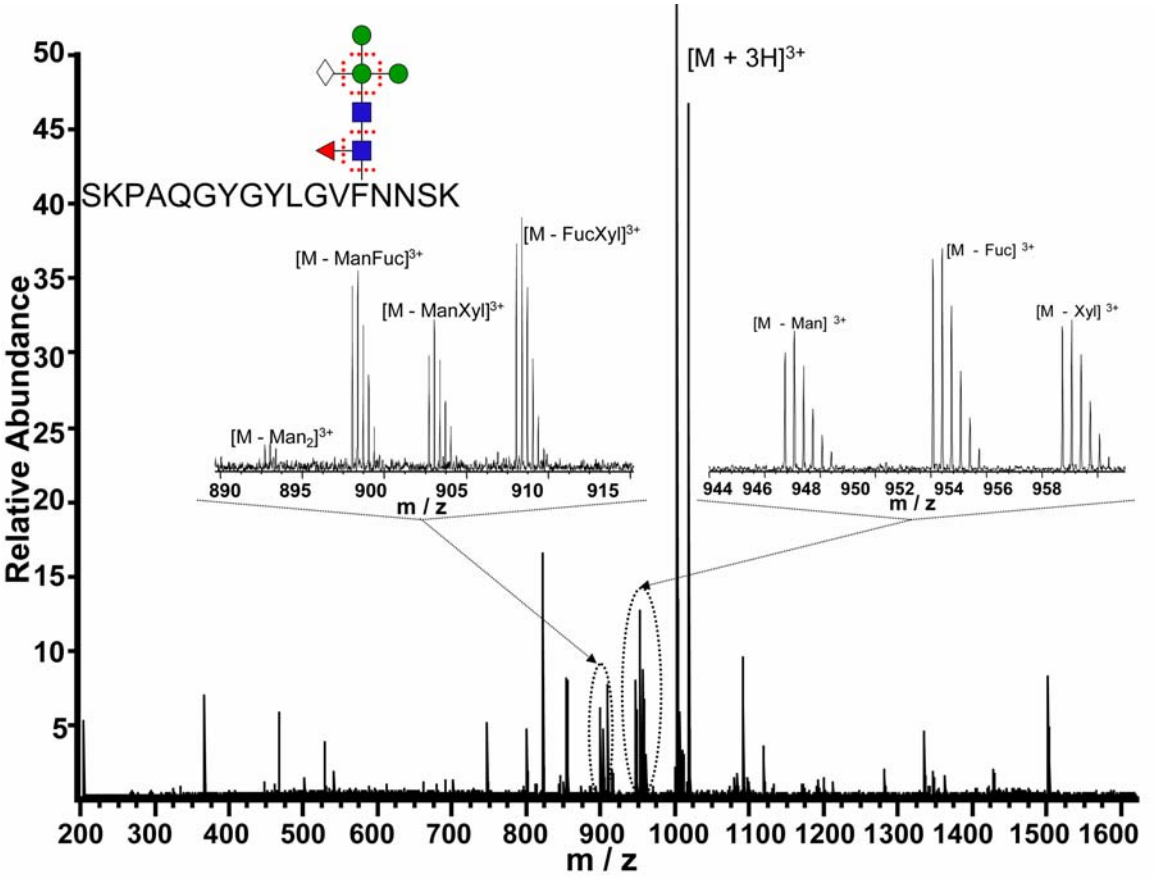


Figure 3.5. IRMPD FT-ICR (10 scans, 50 ms at 10 W laser power) tandem mass spectrum of a lectin glycopeptide from *Erythrina cristagalli*. The precursor ion, denoted as $[M + 3H]^{2+}$, consists of amino acids 100-116 with a xylose type glycan attached at Asn113. Similar to other IRMPD spectra of glycopeptides, extensive cleavage within the glycan is observed (indicated by red dotted lines). No backbone cleavage is seen in the spectrum. The zoomed in regions correspond to losses of one and two monosaccharides, all of which are the same charge state as the precursor ion. (■=GlcNAc, ●=Man, ◄=Fuc, ◊=Xyl).

3.4 Conclusions

We have found that high-mannose type glycopeptide fragmentation differs from xylose type and complex type glycopeptides in IRMPD. IRMPD analyses of high-mannose type glycopeptides from ribonuclease B demonstrated that peptide backbone cleavage competes with glycosidic cleavage, rather than selectively inducing glycosidic cleavage as has been previously observed. This observation is not due to size effects, as has been previously suggested, because it occurs for both small (~2000 Da) and large (~6100 Da) high-mannose type glycopeptides. As predicted, ECD of a high-mannose type glycopeptide results in extensive peptide backbone cleavage while none occurs within the glycan. Repeatedly, ECD has been shown to be a valuable technique for characterizing and locating post-translational modifications, and here we demonstrate its utility for high-mannose type glycoproteins. Furthermore, our results extend the mass range for the successful application of ECD towards glycopeptide characterization. Previous examinations of glycopeptides using IRMPD and ECD emphasized the effectiveness of this combination for glycopeptide structural characterization; the former technique was shown to selectively induce glycan cleavage while the latter cleaves the peptide backbone. However, caution must be used when applying IRMPD for the structural characterization of unknown glycopeptides, because spectral interpretation can be complicated by a mixture of glycan and peptide cleavage when examining high-mannose type glycopeptides. In addition, the unique behavior in IRMPD for high-mannose type glycopeptides needs to be incorporated in models predicting peptide fragmentation patterns for enhanced identification of unknowns.

3.5 References

1. Varki, A. *Glycobiol.* **1993**, 97-130.
2. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
3. Collinge, J.; Sidle, K. C. L.; Meads, J.; Ironside, J.; Hill, A. F. *Nature* **1996**, *383*, 685-690.
4. Prusiner, S. B. *Science* **1997**, *278*, 245-251.
5. Falk, P. G.; Bry, L.; Holgersson, J.; Gordon, J. I. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1515-1519.
6. Dennis, J. W.; Laferte, S.; Waghorne, C.; Breitman, M. L.; Kerbel, R. S. *Science* **1987**, *236*, 582-585.
7. Codington, J. F.; Haavik, S. *Glycobiol.* **1992**, *2*, 173-180.
8. Grünewalk, S.; Matthus, G.; Jaeken, J. *Pediatr. Res.* **2002**, *52*, 618-624.
9. Kamerling, J. P. *Pure Appl. Chem.* **1991**, *63*, 465-472.
10. Harvey, D. J. *Mass Spectrom. Rev.* **1999**, *18*, 349-451.
11. Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161-227.
12. Park, Y.; Lebrilla, C. B. *Mass Spectrom. Rev.* **2005**, *24*, 232-264.
13. Hakansson, K.; Cooper, H. J.; Hudgins, R. R.; Nilsson, C. L. *Curr. Org. Chem.* **2003**, *7*, 1503-1525.
14. Huddleston, M. J.; Bean, M. F.; Carr, S. A. *Anal. Chem.* **1993**, *65*, 877-884.
15. Wilm, M.; Neubauer, G.; Mann, M. *Anal. Chem.* **1996**, *68*, 527-533.
16. Küster, B.; Krogh, T. N.; Mortz, E.; Harvey, D. J. *Proteomics* **2001**, *1*, 350-361.
17. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
18. Woodlin, R. L.; Bomse, D. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250.
19. Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.
20. Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.
21. Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. Mass Spectrom.* **2002**, *8*, 337-349.
22. Syrstad, E. A.; Stephens, D. D.; Turecek, F. *J. Phys. Chem. A* **2003**, *107*, 115-126.
23. Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208-224.
24. Kjeldsen, F.; Haselmann, K.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Chem. Phys. Lett.* **2002**, *356*, 201-206.
25. Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal. Chem.* **1999**, *71*, 4250-4253.
26. Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.
27. Stensballe, A.; Norregaard-Jensen, O.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.
28. Shi, S. D.-H.; Hemling, M. E.; Carr, S. A.; Horn, D. M.; Lindh, I.; McLafferty, F. W. *Anal. Chem.* **2001**, *73*, 19-22.
29. Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.

30. Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 2355-2361.
31. Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
32. Mormann, M.; Macek, B.; de Peredo, A. G.; Hofsteenge, J.; Peter-Katalinic, J. *Int. J. Mass Spectrom.* **2004**, *234*, 11-21.
33. Hakansson, K.; Emmett, M. R.; Marshall, A. G.; Davidsson, P.; Nilsson, C. L. *J. Proteome Res.* **2003**, *2*, 581-588.
34. Hogan, J. M.; Pitteri, S. J.; Chrisman, P. A.; McLuckey, S. A. *J. Proteome Res.* **2005**, *4*, 628-632.
35. Renfrow, M. B.; Cooper, H. J.; Tomana, M.; Kulhavy, R.; Hiki, Y.; Toma, K.; Emmett, M. R.; Mestecky, J.; Marshall, A. G.; Novak, J. *J. Biol. Chem.* **2005**, *280*, 19136-19145.
36. Mormann, M.; Pohlentz, G.; Kölbl, S.; Peter-Katalinic, J. *Proc. 53rd ASMS Conference on Mass Spectrometry and Allied Topics*, San Antonio, Texas 2005; CD-ROM.
37. Mock, K. K.; Davey, M.; Cottrell, J. S. *Biochem. Biophys. Res. Commun.* **1991**, *177*, 644-651.
38. Rudd, P. M.; Scragg, I. G.; Coghill, E.; Dwek, R. A. *Glycoconj. J.* **1992**, *9*, 86-91.
39. Conboy, J. J.; Henion, J. D. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 804-814.
40. Camilleri, P.; Haskins, N. J.; Rudd, P. M.; Saunders, M. R. *Rapid Commun. Mass Spectrom.* **1993**, *7*, 332-335.
41. Williams, R. L.; Greene, S. M.; McPerson, A. *J. Biol. Chem.* **1987**, *262*, 16020-16031.
42. Fu, D.; Chen, L.; O'Neil, R. A. *Carbohydr. Res.* **1994**, *261*, 173-186.
43. Chen, S. S.; Deutsch, E. W.; Yi, E. C.; Li, X.; Goodlett, D. R.; Aebersold, R. *J. Proteome Res.* **2005**, *4*, 2174-2184.
44. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1882.
45. Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.
46. Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26*, 514-518.
47. Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.
48. Blakney, G. T.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *Proc. 50th ASMS Conference on Mass Spectrometry and Allied Topics*, Orlando, FL, June 2-6 2002; CD-ROM.
49. Ledford, E. B., Jr.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **1984**, *56*, 2744-2748.
50. Borman, S. In *Chem. Engin. News*, 2005; Vol. 83, pp 41-50.
51. Merkle, R. K.; Cummings, R. D. *Methods Enzymol.* **1987**, *138*, 232-259.
52. Carlsson, S. R. *Glycobiology: A Practical Approach*; Oxford University Press: Oxford, 1993.
53. Cummings, R. D. *Methods Enzymol.* **1994**, *230*, 66-86.
54. Nemeth, J. F.; Hochensang, G. P.; Marnett, L. J.; Caprioli, R. M. *Biochemistry* **2001**, *40*, 3109-3116.

55. Harvey, D. J. *Proteomics* **2001**, *1*, 311-328.
56. Pfenninger, A.; Karas, M.; Finke, B.; Stahl, B.; Sawatzki, G. *J. Mass Spectrom.* **1999**, *34*, 98-104.
57. Fridriksson, E. K.; Beavil, A.; Holowka, D.; Gould, H. J.; Baird, B.; McLafferty, F. W. *Biochemistry* **2000**, *39*, 3369-3376.
58. Reid, G. E.; Stephenson, J. L.; McLuckey, S. A. *Anal. Chem.* **2002**, *74*, 577-583.
59. Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.
60. Hakansson, K.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2001**, *73*, 3605-3610.
61. Cooper, H. J.; Hudgins, R. R.; Marshall, A. G. *Int. J. Mass Spectrom.* **2004**, *234*, 23-35.
62. Cooper, H. J.; Hudgins, R. R.; Hakansson, K.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 241-249.
63. Mormann, M.; Paulsen, H.; Peter-Katalinic, J. *Eur. J. Mass Spectrom.* **2005**, in press.
64. Zhang, Z. Q. *Anal. Chem.* **2004**, *76*, 3908-3922.
65. Schutz, F.; Kapp, E. A.; Simpson, R. J.; Speed, T. P. *Biochem. Soc. Trans.* **2003**, *31*, 1479-1483.
66. Huang, Y. Y.; Triscari, J. M.; Tseng, G. C.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Anal. Chem.* **2005**, *77*, 5800-5813.
67. Tseng, K.; Hedrick, J. L.; Lebrilla, C. B. *Anal. Chem.* **1999**, *71*, 3747-3754.
68. Gaucher, S. P.; Morrow, J.; Leary, J. A. *Anal. Chem.* **2000**, *72*, 2331-2336.
69. Cooper, C. A.; Gasteiger, E.; Packer, N. H. *Proteomics* **2001**, *1*, 340-349.
70. Tseng, K.; Xie, Y.; Seeley, J.; Hedrick, J. L.; Lebrilla, C. B. *Glycoconj. J.* **2001**, *18*, 309-320.
71. Ethier, M.; Saba, J. A.; Spearman, M.; Krokhin, O.; Ens, W.; Standing, K. G.; Perreault, H. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1743-1754.
72. Leavell, M. D.; Leary, J. A.; Yamasaki, R. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 571-576.
73. Ethier, M.; Saba, J. A.; Spearman, M.; Krokhin, O.; Butler, M.; Ens, W.; Standing, K. G.; Perreault, H. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 2713-2720.
74. Joshi, H. J.; Harrison, M. J.; Schulz, B. L.; Cooper, C. A.; Packer, N. H.; Karlsson, N. G. *Proteomics* **2004**, *4*, 1650-1664.
75. Lohmann, K. K.; von der Lieth, C. W. *Nucleic Acids Res.* **2004**, *32*, 261-266.
76. Reinhold, V.; Ashline, D.; Lapadula, A.; Hanneman, A.; Hatcher, P. J.; Zhang, H.; Bullock, K. *Proc. 53rd ASMS Conference on Mass Spectrometry and Allied Topics*, San Antonio, Texas, June 5-9 2005; CD-ROM.

Chapter 4

Electron Capture Dissociation of Oligosaccharides Ionized with Alkali, Alkaline Earth, and Transition Metals

4.1 Introduction

Oligosaccharide-containing biomolecules, known as glycoconjugates, are highly diverse and prevalent in biological systems. The role of glycoconjugates in nature is extensive,[1] ranging from protein folding[2] to immune system response.[3] The multiple functions of glycoconjugates are largely due to the increased degree of complexity oligosaccharides impart to these biomolecules. Oligosaccharides may exist as several isomeric forms with diverse linkages and unlike other biomolecules can form highly branched structures. Complete structural characterization of oligosaccharides requires information regarding linkage, sequence, branching, and anomeric configuration (a term which refers to the configuration, α or β, of the glycosidic bond of a sugar). Structural characterization of the oligosaccharide portion of glycoconjugates is often accomplished by releasing the sugar from the biomolecule. It is frequently necessary to utilize a wide range of analytical methodologies in order to fully characterize these diverse and structurally complex molecules.

Mass spectrometry is an important tool for oligosaccharide characterization and offers high sensitivity and minimum sample requirements. Tandem mass spectrometry (MS^n) has been employed extensively for oligosaccharide structural analysis.[4-13] In particular, Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) offers several advantages for oligosaccharide analysis,[14] including high mass accuracy, ultrahigh resolution, and the availability of several tandem mass spectrometric techniques.

Oligosaccharides undergo two main types of fragmentation (shown in Figure 1.5 in Chapter 1). The predominant fragmentation pathway for protonated oligosaccharides is generally glycosidic cleavage, which occurs between monosaccharide units and provides information regarding saccharide sequence and branching. However, valuable information regarding sugar linkage can be gained from cross-ring cleavages, which are less prevalent than glycosidic cleavages. Several factors are known to affect the degree of oligosaccharide fragmentation and the extent of glycosidic versus cross-ring cleavage. These factors include variables such as the ionizing cation, the lifetime of the ion prior to detection and the energy deposited into the ion. Oligosaccharides ionized with alkali, alkaline earth, and transition metals often fragment to yield more cross-ring cleavages compared to their protonated counterparts.[15-27] These fragmentation patterns are highly dependent on the specific metal adduct chosen and also the degree of oligosaccharide branching.[22] In addition, fragmentation of deprotonated oligosaccharides may also result in additional cross-ring fragmentation.[10,28-30]

Conventionally, tandem mass spectrometry is accomplished with low energy collision activated dissociation (CAD), and results in mostly glycosidic cleavages.

However, several alternative fragmentation techniques have been utilized for oligosaccharides including high-energy CAD,[11,31-33] infrared multiphoton dissociation (IRMPD),[34-36] electron capture dissociation (ECD),[37] and 157 nm photodissociation.[38] High-energy CAD and 157 nm photodissociation of oligosaccharides results in more extensive cross-ring fragmentation. IRMPD and low-energy CAD are both low energy vibrational excitation techniques. However, Lebrilla and co-workers have shown that the fragmentation efficiency in IRMPD is greater than that in CAD for large oligosaccharides.[36] In addition, ECD has been applied to protonated chitooligosaccharides, and yielded primarily glycosidic cleavages corresponding to B and C-type ions. ECD is based on the dissociative recombination of polycationic molecules with low-energy electrons (<1 eV), generating charge reduced radical species and product ions.[39-44] The implementation of ECD in biomolecular structural analysis is rapidly expanding, and while it is known to be extremely valuable for post-translational modification analysis, including protein glycosylation,[45-50] its applicability towards oligosaccharide characterization has not been fully explored.

We have previously shown that divalent metal-adduction can greatly improve positive ion mode ionization efficiency of acidic sulfated peptides.[51] In addition, utilization of divalent metal adducts to positively ionize such acidic molecules facilitates formation of the doubly-charged species that are required for ECD. Furthermore, we and others have reported that ECD fragmentation can be drastically different as a function of metal charge carrier, thereby allowing directed bond cleavage.[51-54] Our current objective is to examine ECD fragmentation of oligosaccharides ionized with various cationizing agents, to determine whether the combination of ECD and metal ion adduction is a viable

analytical tool for oligosaccharide structural characterization. Following ECD and IRMPD, we compare fragmentation patterns to determine whether ECD provides complementary information. Cationizing agents were select Group I, Group II, and transition metal species, including Na^+, K^+, Ca^{2+}, Ba^{2+}, Mg^{2+}, Mn^{2+}, Co^{2+}, and Zn^{2+}. Several cationizing agents were chosen in order to determine which metal adducts, if any, provide the most extensive fragmentation (and thereby structural information) from ECD. Model oligosaccharides examined are maltoheptaose, para-lacto-*N*-hexaose, and an N-linked glycan from human α_1-acid glycoprotein. Both linear and branched oligosaccharides are investigated, to determine whether oligosaccharide branching affects ECD fragmentation.

4.2 Experimental

4.2.1 Sample Preparation

Maltoheptaose, para-lacto-*N* hexaose, and a complex type *N*-glycan from human α_1-acid glycoprotein were obtained from Sigma Chemical Co. (St. Louis, MO). All salts, including NaCl, KCl, $Ca(C_2H_3O_2)_2$, $MgBr_2$, $BaCl_2$, $CoBr_2$, $ZnBr_2$, and $McCl_2$ were also from Sigma. Samples were prepared by mixing oligosaccharides (5 or 2.5 μM final concentration) and salts (20 or 10 μM final concentration) in a water:methanol (1:1 v/v) solution. These concentrations are comparable to those used in similar analyses.[26,27] For protonated samples, oligosaccharides were prepared in 50% methanol and 2% acetic acid solution.

4.2.2 FT-ICR Mass Spectrometry

All experiments were performed with an actively shielded 7 T FT-ICR mass spectrometer with a quadrupole front-end (APEX-Q, Bruker Daltonics, Billerica, MA), as previously described.[55] An indirectly heated hollow dispenser cathode was used to perform ECD.[56] IRMPD was performed with a vertically mounted 25-W, 10.6-μm, CO_2 laser (Synrad, Mukilteo, WA). Samples were infused either via an external Apollo or Apollo II electrospray ion source at a flow rate of 60 μL/h with the assistance of N_2 nebulizing gas. Oligosaccharides adducted with either sodium, calcium, or potassium were examined using the Apollo II ion source, while the rest were analyzed with the original Apollo source. Due to the increased sensitivity with the Apollo II ion source, sample concentrations were reduced by a factor of two. Following ion accumulation in the first hexapole for 0.05 s, ions were mass selectively accumulated in the second hexapole for 1-7 s. Ions were then transferred through high-voltage ion optics and captured with dynamic trapping in an Infinity ICR cell.[57] The experimental sequence up to the ICR cell fill was looped 3 - 15 times to achieve maximum precursor ion signal. Due to contaminating salts present in all samples (all samples were used without further purification), several ion fills and longer accumulation times were often necessary in order to achieve abundant precursor ion signal of the metal-adducted species of interest. Precursor ions were further isolated in the ICR cell using correlated harmonic excitation fields (CHEF).[58,59] For ECD, the cathode heating current was kept constant at 1.8 A and the cathode voltage was pulsed during the ECD event to a bias voltage of (- 0.1)-(- 1.0) V for 30-175 ms to generate low energy electrons. A lens electrode located immediately in front of the cathode was kept at 1-2 V to improve electron injection. IRMPD was

performed with a laser power of 7.5-15 W and with firing times ranging from 0.1-1 s. For activated-ion ECD of the *N*-glycan, ions were heated before they underwent electron capture for 25-100 ms with a 7.5-15 W laser pulse. This activation was to aid in the destabilization of gas-phase secondary structure.[60]

4.2.3 Data Analysis

All mass spectra were acquired with XMASS software (Bruker Daltonics) with 256 data points from *m/z* 200-2500 and summed over 30-80 scans. Data processing was performed with MIDAS software.[61] Data were zero filled once, Hanning apodized, and exported to Microsoft Excel for internal frequency-to-mass calibration with a two-term calibration equation.[62] Product ion spectra were interpreted with the aid of the web application GlycoFragment (www.dkfz.de/spec/projekte/fragments/).[63] Product ions were not assigned unless they were at least 3x the noise level. Many unlabeled product ions in several spectra are often assignable, but due to high mass error were left unlabeled.

4.3 Results and Discussion

According to the Domon and Costello nomenclature[64] product ions containing the reducing end of an oligosaccharide are designated as X (cross-ring cleavage), Y and Z-type ions (glycosidic cleavage). Those fragments containing the non-reducing end are designated A (cross-ring), B and C-type ions (glycosidic). Subscript numerals denote where cleavage occurred, and superscript numerals denote between which bonds cross-ring cleavage occurred. In certain examples, when a product ion may correspond to

several possibilities, all possible ions are listed (alternative product ion assignments are indicated by parentheses).

4.3.1 Linear Oligosaccharide: Maltoheptaose

Due to the symmetric nature of maltoheptaose, a number of ions cannot be distinguished based on their m/z ratio (B and Z, C and Y, and several A and X). However, earlier studies have shown that product ions formed from cleavage on the non-reducing side of glycosidic oxygens (B and Y ions) are more abundant than those due to cleavage on the reducing side (C and Z ions).[4] For maltoheptaose ionized with divalent cations, Harvey observed several Y ions, but could not distinguish between B and Z ions.[27] Similarly, several of the A and X ions observed following IRMPD cannot be distinguished based on their m/z ratios, but CAD labeling studies have shown that cross-ring cleavage products begin at the reducing end of the oligosaccharide to generate A-type ions and proceed in a stepwise fashion towards the non-reducing end.[17,18] Commonly observed cross-ring cleavage fragments of metal-adducted oligosaccharides include ^{0,2}A and ^{2,4}A-type ions (low energy fragmentation techniques). Based on this information, assumptions can be made regarding which product ion types are observed following IRMPD. However, the precise ECD mechanism is yet to be fully elucidated,[40-43,65-68] and thus it is prudent to avoid making assumptions regarding which ion types are observed following ECD. Throughout our discussion, when a product ion may correspond to several possibilities, all possible ions are listed for both IRMPD and ECD.

The IRMPD (Figure 4.1a) and ECD (Figure 4.1b) spectra obtained from barium-adducted maltoheptaose are shown in Figure 4.1. It should be noted that barium is not typically used as a metal cationizing agent for oligosaccharides. Larger metal ions are

generally not favorable for oligosaccharide fragmentation because as metal ion size increases, there is a general tendency for poorer fragmentation.[19,22] This trend has been attributed to the inability of many metal complexes with large metal cations to undergo charge-induced fragmentation. For this reason, we were interested in examining whether barium would be an effective metal for catalyzing cross-ring fragmentation following ECD.

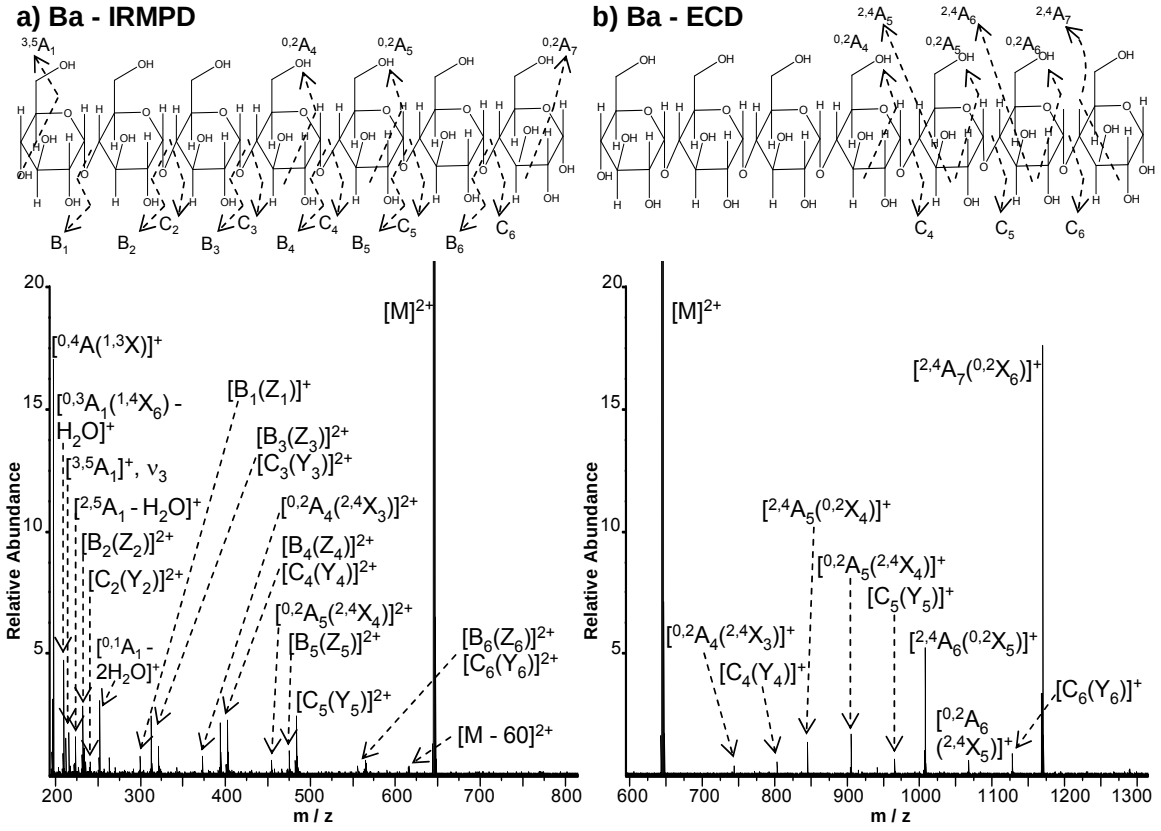


Figure 4.1. FT-ICR tandem mass spectra of Ba^{2+}-adducted maltoheptaose. a) IRMPD (30 scans, 120 ms irradiation with 10 W laser power) and b) ECD (40 scans, 100 ms with a bias voltage of (- 0.2) V). All ions in the spectra contain Ba. Specific cleavages are indicated in the oligosaccharide structure above each spectrum (product ions with neutral losses are not indicated in the upper figure). Due to the symmetry of the molecule, several product ions cannot be unambiguously assigned (indicated by parentheses in the labels in each spectra). In the upper oligosaccharide structure, only one cleavage site is indicated if several are possible (to aid readability).

In Figure 4.1a and 4.1b, all product ions retained barium. Singly-charged product ions observed following IRMPD are due to accompanying proton loss, which is likely from a hydroxyl group.[17] This proton loss would leave a structure containing an ion pair between the barium and an anionic oxygen,[27] resulting in a metal-adducted product ion with one less charge than the precursor ion and a complementary protonated product ion. These protonated product ions are not present in the maltoheptaose spectra, due to the low *m/z* excitation/detection being set to 200. In this example, IRMPD of barium-adducted maltoheptaose (Figure 4.1a) results in predominantly doubly-charged product ions. The preference for singly or doubly-charged product ions following IRMPD depends on the metal species, and reflects the second ionization energy of the coordinating metal.[27] IRMPD of barium-adducted maltohepatose results in extensive glycosidic cleavage, as well as several less abundant cross-ring cleavage fragments. In the IRMPD spectrum in Figure 4.1a, loss of 60 Da corresponds to loss of $C_2H_4O_2$, which is probably a ^{0,2}A cleavage in the reducing terminal saccharide.[23,27]

Following ECD of barium-adducted maltoheptaose (Figure 4.1b), all product ions are singly-charged even-electron ions, resulting from gain of an electron and subsequent loss of hydrogen. Comparing Figures 4.1a and 4.1b, IRMPD of barium-adducted maltoheptaose does not result in extensive cross-ring cleavage. However, in the ECD spectrum cross-ring cleavage is the dominant fragmentation pathway. In Figure 4.1b, only three glycosidic cleavage fragments are observed compared to six cross-ring fragments. It is interesting to note that glycosidic cleavage fragments from ECD correspond to the C(Y) ion type, but not the B(Z) type. In addition, a charge-reduced species, $[M + Ba]^{+\bullet}$, is not observed. ECD of protonated chitooligosaccharides,[37] and

polyethylene glycol molecules[69,70] also resulted in no apparent charge-reduced species. Following ECD of all metal-adducted forms of maltoheptaose, charge-reduced species were not observed, but in certain examples a proton-stripped species, $[M + metal - H]^+$, is present.

The ECD spectrum of $[M + 2Na]^{2+}$ is shown in Figure 4.2a. Here, the most abundant fragmentation channel is loss of a sodium ion. Similar to barium-adducted maltoheptaose, glycosidic cleavages observed following IRMPD of $[M + 2Na]^{2+}$ correspond to B(Z)-type ions, whereas ECD only results in C(Y)-type glycosidic cleavage. In Figure 4.2a, singly-charged ECD product ions are not due to loss of sodium but rather the gain of an electron and loss of hydrogen, as previously discussed. The product ion at *m/z* 1076 has lost an additional hydrogen.[71] Such additional hydrogen loss is observed occasionally following ECD (forming odd-electron species), but not following IRMPD. Similarly, electron induced dissociation of protonated chitooligosaccharides occasionally resulted in minor odd-electron species 1 Da lighter than the corresponding even-electron species.[37] ECD of the $[M + 2K]^{2+}$ species was also examined (see Figure 4.3). The results are similar to the $[M + 2Na]^{2+}$ species, with relatively poor fragmentation yield and a mixture of glycosidic and cross-ring cleavages.

Figure 4.2b shows the ECD spectrum of $[M + Mg]^{2+}$. Numerous cross-ring cleavage fragments corresponding to $^{0,2}A(^{2,4}X)$ and $^{2,4}A(^{0,2}X)$-type ions are observed, along with several glycosidic cleavage product ions. The ECD spectrum of $[M + Ca]^{2+}$ is shown in Figure 4.2c. The dominant fragmentation pathway following ECD is the generation of a $^{2,4}A_7(^{0,2}X_6)$ product ion, which was not observed following IRMPD. Compared to some of the other species examined, IRMPD of calcium-adducted

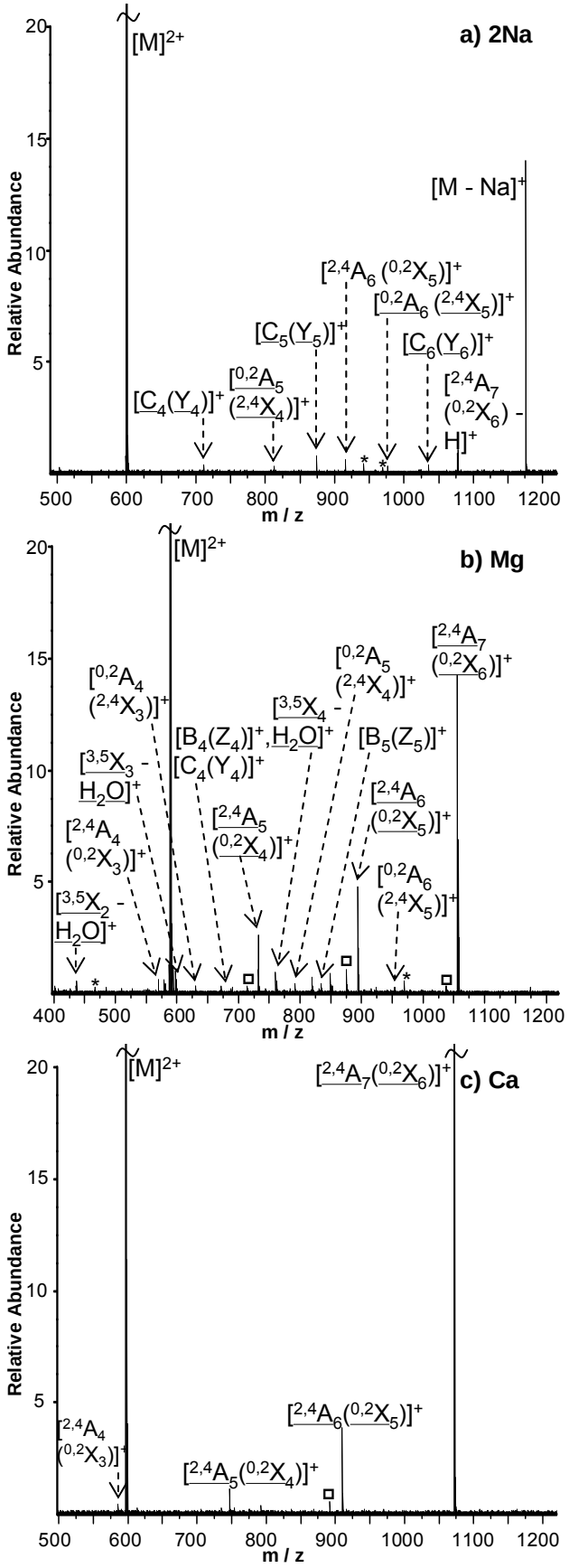


Figure 4.2. ECD FT-ICR tandem mass spectra of the a) 2 Na^+ adduct of maltoheptaose (60 scans, 50 ms with a bias voltage of (- 0.5) V) b) Mg^{2+} adduct of maltoheptaose (80 scans, 100 ms with a bias voltage of (- 0.2) V) and c) Ca^{2+} adduct of maltoheptaose (60 scans, 75 ms with a bias voltage of (- 0.2) V). * denotes electronic noise peaks. □ indicates water loss from an adjacent peak. Product ions are underlined if they were not observed following IRMPD.

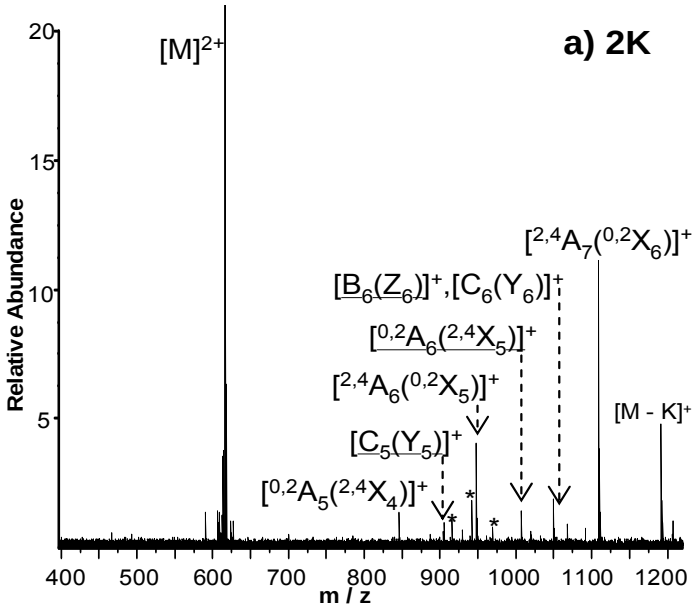


Figure 4.3. ECD FT-ICR (20 scans, 60 ms with a bias voltage of -1.0 V) tandem mass spectrum of 2K adducted maltoheptoase. * denotes electronic noise. Product ions are underlined if they were not observed following IRMPD.

maltoheptaose produces more extensive cross-ring cleavages (data not shown). However, glycosidic cleavage was still the favored fragmentation pathway. It is important to note that in a previous CAD examination of calcium-adducted maltoheptaose, several ^{0,2}A and ^{2,4}A-type ions were observed, some of which are not seen in our IRMPD spectrum (including $^{2,4}A_6$ and $^{2,4}A_5$).[27] Slight differences in fragmentation are due to the nature of the IRMPD process, which results in more extensive fragmentation compared to CAD. These larger ^{2,4}A-type ions are not observed following IRMPD because they are likely further fragmented into smaller product ions.

The ECD spectrum of several transition metal adducts of maltoheptaose are shown in Figure 4.4. Similar to previous examples, the most abundant product ions observed following ECD correspond to $^{2,4}A_7(^{0,2}X_6)$ and $^{2,4}A_6(^{0,2}X_5)$. The ECD spectrum of $[M + Mn]^{2+}$ is shown in Figure 4.4a, and demonstrates a mixture of glycosidic and cross-ring cleavage. The proton-stripped peak is also quite abundant, which is not the

case for any of the other ECD examples with maltoheptaose. The ECD spectrum of $[M + Co]^{2+}$ is shown in Figure 4.4b. Unlike previous ECD spectra, the occurrence of glycosidic cleavage is fairly abundant. However, all of the glycosidic cleavage product ions observed are also present following IRMPD. The ion labeled as $[M - 90]^{+}$ corresponds to loss of $C_3H_6O_3$, which could be a $^{0,3}A$ fragment in the reducing terminal saccharide. Figure 4.4c displays the ECD spectrum of $[M + Zn]^{2+}$. Zinc is not typically used as a cationizing agent for oligosaccharide analysis because it forms relatively weak complexes.[23] Consequently, several seconds of ion accumulation and several cell fill cycles were required for the zinc-adducted species, in order to produce an ion signal sufficient for ECD. These measures were necessary due to the lower fragmentation efficiency in ECD compared to traditional MS/MS techniques. Following IRMPD, only one cross-ring cleavage fragment is observed, ($[M - 60]^{+}$), and glycosidic cleavage fragments are much less abundant compared to the other examples discussed thus far. The ECD fragmentation pattern of the zinc adduct demonstrates that cross-ring fragmentation is the preferred fragmentation pathway. Several additional cross-ring fragments were observed, however, due to poor S/N ratio were not labeled. Although ECD of zinc-adducted maltoheptoase demonstrates that cross-ring fragmentation is extensive, selecting a metal ion which more effectively ionizes the oligosaccharide would shorten the ECD experimental sequence while still providing similar fragmentation information. During this examination, calcium was particularly effective in ionizing maltoheptaose. The ability of divalent metals to ionize neutral oligosaccharides has been documented by Harvey, whom also found that calcium was particularly effectual.[27]

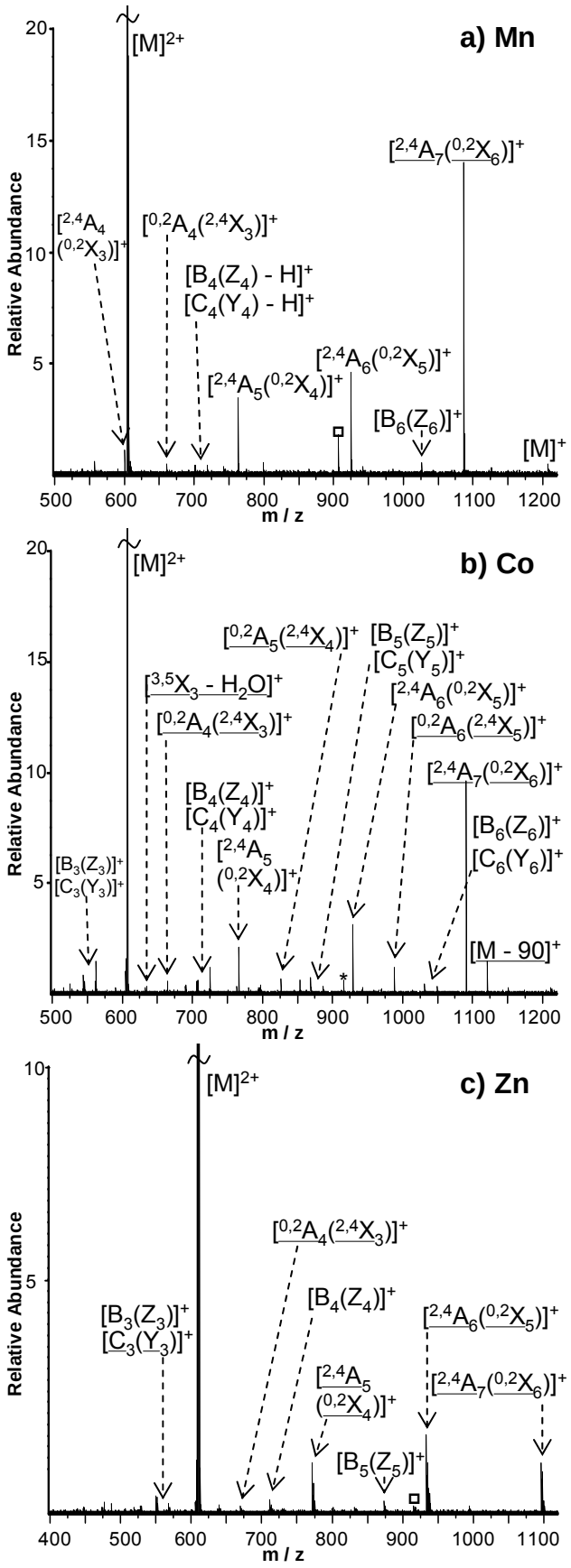


Figure 4.4. ECD FT-ICR tandem mass spectra of the a) Mn adduct of maltoheptaose (40 scans, 100 ms with a bias voltage of (- 0.2) V) b) Co adduct of maltoheptaose (40 scans, 125 ms with a bias voltage of (- 0.2) V), and c) Zn adduct of maltoheptaose (40 scans, 50 ms with a bias voltage of (- 0.1) V). * denotes electronic noise. □ indicates water loss from an adjacent peak. Product ions are underlined if they were not observed following IRMPD.

4.3.2 Linear Oligosaccharide: Para-lacto-*N*-hexaose

To further examine the ECD fragmentation patterns of metal-adducted oligosaccharides, a second linear oligosaccharide was examined. The IRMPD and ECD spectra of manganese-adducted para-lacto-*N*-hexaose (pLNH) are shown in Figures 4.5a and 4.5b, respectively. In the IRMPD spectrum of $[M + Mn]^{2+}$ in Figure 4.5a, the most abundant product ion corresponds to a protonated B_2 ion (protonated species are denoted by H^+). This ion along with a dehydroxy-*N*-acetyl glucosamine, labeled as $[GlcNAc - H_2O]^{*+}$, are observed in every IRMPD and ECD spectrum of pLNH. Cleavage to the reducing side of GlcNAc is a facile process, and energy imparted into the precursor ion during quadrupole and in-cell CHEF isolation was sufficient for formation of these ions. Following IRMPD, the two most abundant metal-adducted glycosidic cleavage fragments are B_2 and Y_4, which both correspond to cleavage to the reducing side of GlcNAc. In Figure 4.5a, three $^{2,4}A(^{0,2}X)$ and one $^{0,2}A(^{2,4}X)$-type cross-ring cleavage product are also observed. Following ECD, the most abundant product ions are $C_5(Y_5)$, $[C_5(Y_5) - 42]$, and $^{2,4}A_6(^{0,2}X_5)$. Loss of 42 Da corresponds to loss of a ketene molecule (CH_2CO) from *N*-acetylglucosamine, which was previously observed following ECD of protonated chitooligosaccharides.[37] The three most abundant product ions observed following ECD are not present in the IRMPD spectrum, thus demonstrating the complementary fragmentation patterns which can be obtained following ECD.

The ECD spectra of pLNH adducted with either Ca, Mg, Ba, Co, or Zn are shown in Figures 4.6 and 4.7. In all cases, several additional cross-ring cleavage product ions are produced following ECD. Those ions not observed in the corresponding IRMPD

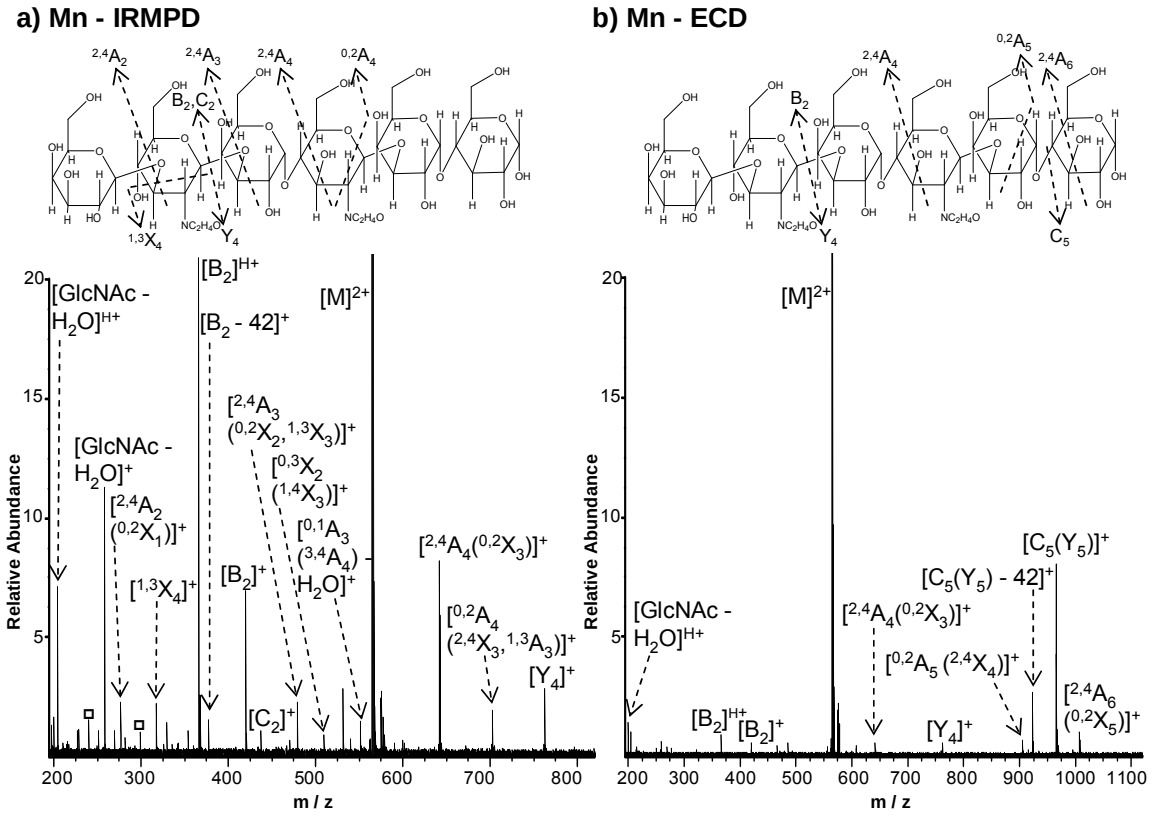


Figure 4.5. FT-ICR tandem mass spectra of Mn-adducted para-lacto-*N*-hexaose. a) IRMPD (40 scans, 150 ms irradiation with 7.5 W laser power) and b) ECD (80 scans, 50 ms with a bias voltage of (- 0.2) V). All product ions contain Mn, unless denoted by H^+, indicating protonation. ◘ indicates water loss from an adjacent peak. Specific cleavages are indicated in the oligosaccharide structure above each spectrum (product ions with neutral losses are not indicated in the upper figure). Due to the structure of the molecule, several product ions cannot be unambiguously assigned (indicated by parentheses in labels in each spectra). In the upper oligosaccharide structure, only one cleavage site is indicated if several are possible (to aid readability).

spectra are underlined in both Figures 4.6 and 4.7. Several neutral molecule losses are also observed in these spectra, including loss of 42 Da as previously discussed. In addition, loss of 57 Da is observed, which likely corresponds to loss of C_2H_3NO from *N*-acetyl glucosamine. The $[M + 2Na]^{2+}$ and $[M + 2K]^{2+}$ species of pLNH were not investigated, because the dominant fragmentation pathway following ECD is loss of a sodium or potassium ion, as demonstrated from the examples with maltoheptaose.

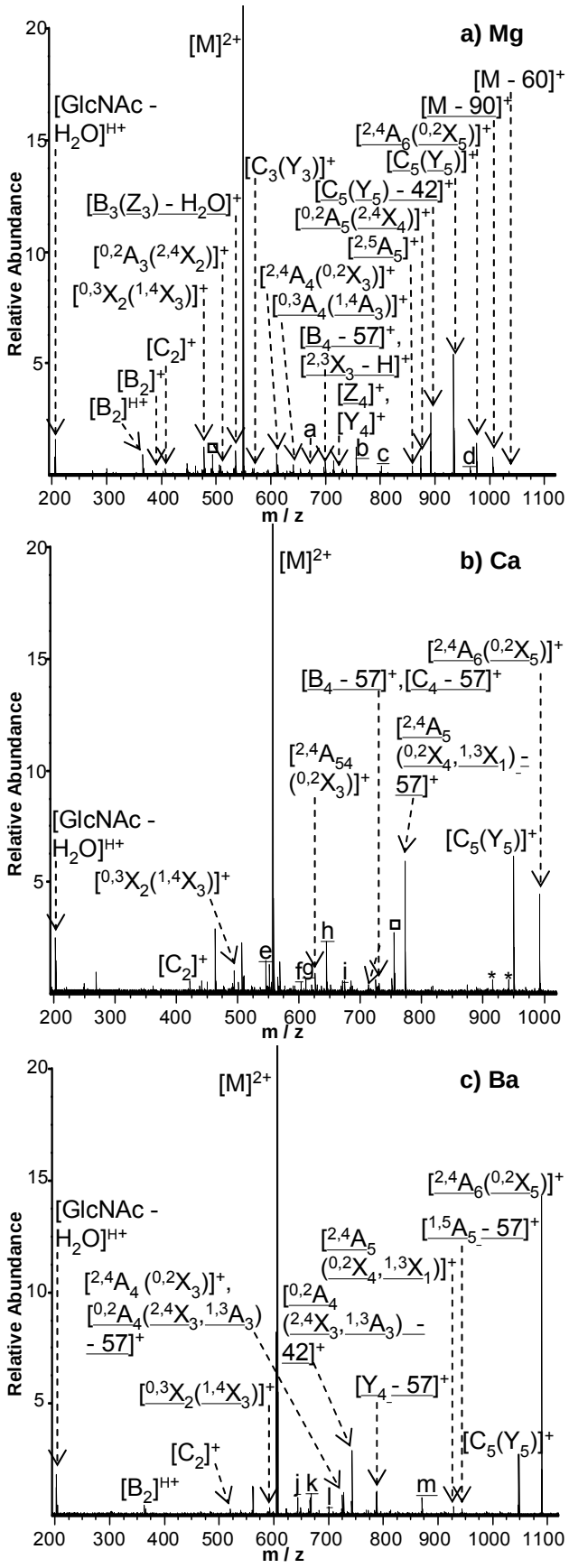


Figure 4.6. ECD FT-ICR tandem spectra of metal-adducted species of para-lacto-N-hexaose a) Mg adduct (80 scans, 150 ms with a bias voltage of -0.5 V). Product ion a = $[^{0,2}A_4(^{2,4}X_3,^{1,3}A_3)]^+$, b = $[^{2,4}A_5(^{0,2}X_4,^{1,3}X_1) - 57]^+$, c = $[^{2,5}A_5 - 57 - H]^+$, and d = $[M - 90 - 42]^+$. b) Ca adduct (60 scans, 75 ms with a bias voltage of -0.2 V). Product ion e = $[^{0,2}A_4(^{2,4}X_3,^{1,3}A_3) - 2(42) - 57]^+$, f = $[^{0,2}A_4(^{2,4}X_3,^{1,3}A_3) - 2(42)]^+$, g = $[^{2,4}A_5(^{0,2}X_4,^{1,3}X_1) - 57 - Hex]^+$, h = $[^{0,2}A_4(^{2,4}X_3,^{1,3}A_3) - 42]^+$, i = $[^{2,3}X_3 - 42]^+$. c) Ba adduct (60 scans, 125 ms with a bias voltage of -0.2 V). Product ion j = $[^{0,2}A_4(^{2,4}X_3,^{1,3}A_3) - 2*42 - 57]^+$, k = $[^{2,4}A_4(^{0,2}X_3) - 57]^+$, l = $[^{0,2}A_4(^{2,4}X_3,^{1,3}A_3) - 2*42]^+$, and m = $[^{2,4}A_5(^{0,2}X_4,^{1,3}X_1) - 57]^+$. * denotes electronic noise. □ indicates water loss from an adjacent peak. All product ions contain a metal, unless denoted by H^+, indicating protonation. Product ions are underlined if they were not observed following IRMPD.

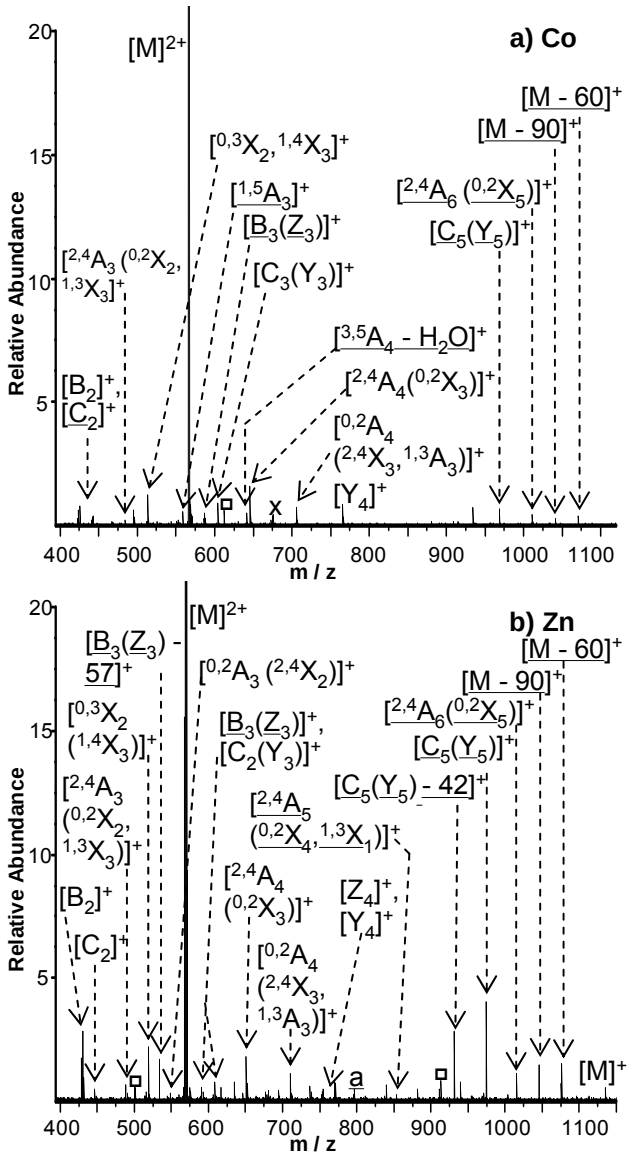


Figure 4.7. ECD FT-ICR tandem spectra of metal-adducted species of para-lacto-N-hexaose a) Co adduct (80 scans, 175 ms with a bias voltage of -0.75 V). b) Zn adduct (80 scans, 100 ms with a bias voltage -0.5 V). Product ion a= $[^{2,4}A_5(^{0,2}X_4, ^{1,3}X_1) - 57]^+$. * denotes electronic noise. □ indicates water loss from an adjacent peak. All product ions contain the metal unless denoted by H^+, indicating protonation. Product ions are underlined if they were not observed following IRMPD.

4.3.3 Branched N-linked Oligosaccharide

To determine whether ECD is applicable to non-linear oligosaccharides, a branched N-linked glycan was also examined. The structure of this glycan from human α_1-acid glycoprotein, with the general saccharide composition [Gal-GlcNAc]$_3$-Man$_3$-GlcNAc$_2$, is shown in Figure 4.8. The IRMPD and ECD spectra of the doubly protonated species are shown in Figures 4.8a and 4.8b. Unlike maltoheptaose and pLNH, a doubly protonated species was easily generated if the sample was electrosprayed under acidic

conditions. Following both IRMPD and ECD, several product ions are due to the combination of cleavages on both the reducing and non-reducing ends, such as the case with C_5/Z_4. Product ions corresponding to multiple cleavage sites are indicated by a front-slash. Several product ions in the ECD spectrum may be caused by vibrational excitation, since activated-ion ECD was necessary in order to fragment this oligosaccharide. Without ion activation, very little ECD fragmentation was observed.

Following IRMPD of the doubly protonated species, the most abundant product ions in the spectrum, B_2 and Y_4, correspond to cleavage adjacent to *N*-acetylglucosamine. One of the most diagnostic product ions observed in Figure 4.8a corresponds to "ion D", labeled as $[GlcNAcMan_3 - H_2O]^{+\cdot}$.[12] This ion is due to the loss of the 3-linked antennae and reducing end *N*-acetylglucosamines, and its presence allows for the determination of the saccharide composition of the 6-linked antenna. Following ECD of the protonated species, several additional product ions are observed; however most of these correspond to glycosidic cleavages. The product ion observed at *m/z* 1857.7 corresponds to loss of 149 Da from the neutral precursor ion, and could potentially be assigned as $[^{0,3}A_6 - H_2O]^+$. The product ion observed 162 mass units lower than this ion is due to loss of a non-reducing end saccharide (Y_5 cleavage). The product ion labeled as $[M]^+$ in Figure 4.8b corresponds to the proton-stripped species, not to be confused with the charge-reduced radical species.

Figure 4.9 shows the IRMPD and ECD spectra of the cobalt-adducted N-linked glycan. The ECD fragmentation patterns of several other metal adducts of this oligosaccharide were examined, however, they yielded similar results (data not shown).

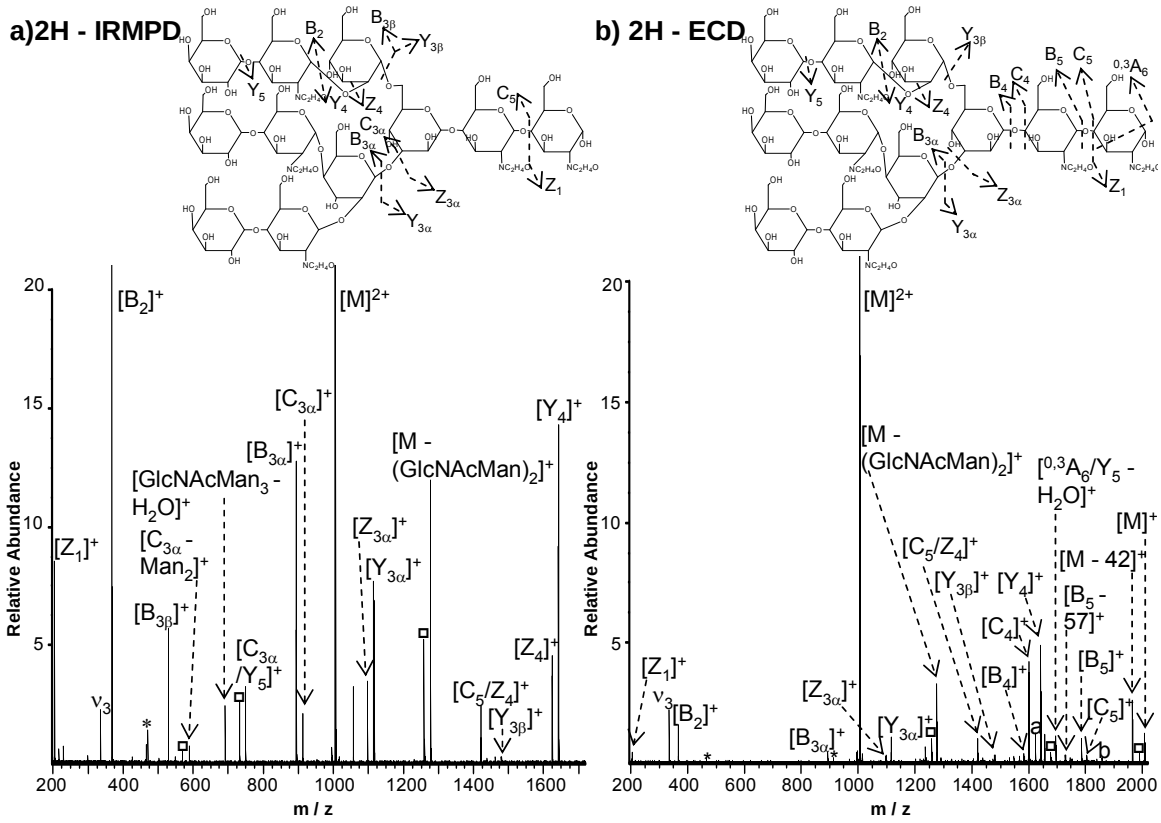


Figure 4.8. FT-ICR tandem mass spectra of a doubly protonated N-linked glycan. a) IRMPD (20 scans, 125 ms irradiation with 7.5 W laser power) and b) ECD (40 scans, 100 ms with a bias voltage of (- 0.2) V). ◘ indicates water loss from an adjacent peak. In the ECD spectrum, product ion a = $[Z_4]^+$, and b = $[^{0,3}A_6 - H_2O]^+$. Specific cleavages are indicated in the oligosaccharide structure above each spectrum (product ions with neutral losses are not shown in the upper figure). * indicates electronic noise. For several product ions, including Y_5, B_2, Y_4, and Z_4, cleavage may occur at several sites. In the upper oligosaccharide figure, only one site is indicated (to aid readability).

As with maltoheptaose and pLNH, ECD of the metal-adducted species always resulted in additional product ions compared to IRMPD and yielded complementary structural information. In Figure 4.9a, several diagnostic cross-ring cleavage product ions are observed. The cross-ring cleavage fragments that result from multiple bond scissions within the core branching mannose are especially useful for structural characterization. Several product ions correspond to a combination of $^{2,4}A$ cleavage within the reducing terminal GlcNAc, along with additional glycosidic cleavage (for example $^{2,4}A_6/Y_4$ and

$^{2,4}A_6/Z_4$). An earlier study by Leary and co-workers, who compared the fragmentation spectra of Cu^{2+}, Mn^{2+}, Co^{2+}, Zn^{2+}, and Ca^{2+}-adducted oligosaccharides, similarly demonstrated that a cobalt-coordinated oligosaccharide generated a unique fragmentation spectrum.

ECD of the cobalt-adducted species (Figure 4.9b) resulted in several additional product ions, yielding complementary information. Additional cross-ring cleavages include $^{0,2}A_5$, $^{1,5}X_4$, and $^{1,5}X_{3\alpha}$. Both high-energy CAD[11,31-33] and 157 nm photodissociation[38] of oligosaccharides generate ^{1,5}X-type product ions. Moreover, it should be noted that with maltoheptaose and pLNH, several ECD product ions could not be absolutely identified as A or X-type ions because of the symmetry of the molecule. However, the fragmentation patterns observed following ECD of an N-linked glycan indicate that previously observed ECD product ions correspond to ^{0,2}A and ^{2,4}A-type ions. With this knowledge, re-examination of the ECD fragmentation patterns of maltoheptaose and pLNH demonstrates that cross-ring cleavages preferably occur towards the reducing end of the oligosaccharide.

Two of the most abundant product ions in Figure 4.9b are the proton-stripped species, $[M]^+$, and $[2M]^{3+}$. The latter triply-charged species can be explained by dimerization. During the electrospray process, dimerization occurred, resulting in a non-covalent species containing two glycans and two cobalts. Inspection of the isotopic distribution of the precursor ion confirmed the presence of this dimer. ECD is able to maintain this non-covalent interaction, as demonstrated by the presence of $[2M]^{3+}$, which is due to gain of an electron and loss of a hydrogen, and $[2M - 42]^{3+}$. The $[2M - 42]^{3+}$ species is present following ECD of all metal adducted forms of the N-linked glycan.

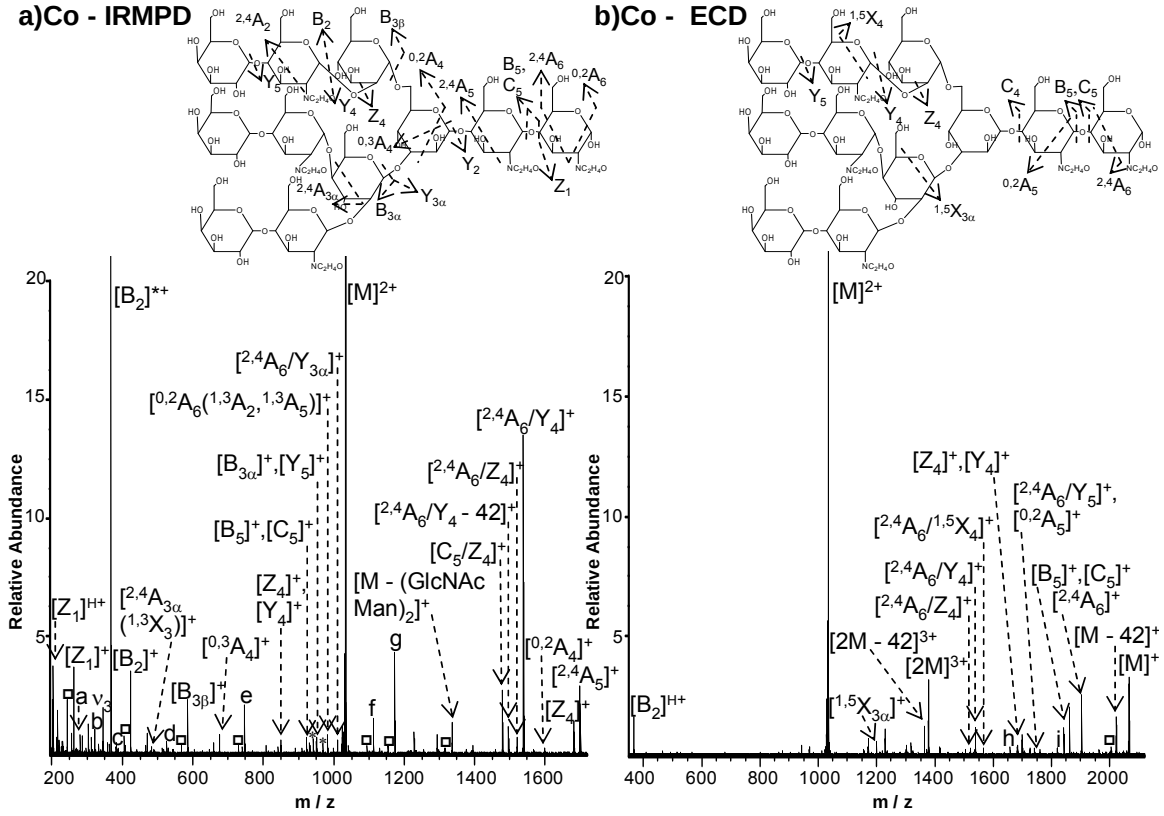


Figure 4.9. FT-ICR tandem mass spectra of a Co-adducted N-linked glycan. a) IRMPD (80 scans, 350 ms irradiation with 10 W laser power) and b) ai-ECD (80 scans, 60 ms irradiation with 7.5 W laser power, 75 ms with a bias voltage of (- 0.5) V). All product ions contain Co, unless denoted by H^+, indicating protonation. □ indicates water loss from an adjacent peak. In the IRMPD spectrum, product ion a = $[^{2,4}A_2]^+$, b = $[^{2,4}A_6/Y_2]^+$, c = $[B_2 - 42]^+$, d =$[B_{3\beta} - H_2O - 42]^+$, e = $[GlcNAcMan_3 - H_2O]^+$ ("ion D"), f = $[^{2,4}A_6/M - (GlcNAcMan)_2 - H_2O - 42]^+$, g = $[^{2,4}A_6/M - (GlcNAcMan)_2]^+$. In the ECD spectrum, product ion h = $[C_4]^+$ and i = $[C_5 - 42]^+$. Specific cleavages are indicated in the oligosaccharide structure above each spectrum (product ions with neutral losses are not shown in the upper figure). * indicates electronic noise. For several product ions, including Y_5, $^{2,4}A_2$, B_2, Y_4, Z_4, and $^{1,5}X_4$, cleavage may occur at several sites. In the upper oligosaccharide figure, only one site is indicated (to aid readability).

However in certain cases, such as with the Mg^{2+}-coordinated oligosaccharide, several other unique product ions are also observed (data not shown). These include $[M/^{2,4}A_6]^{3+}$,$[M/C_5]^{3+}$, and $[M/Y_4]^{3+}$. The ability of ECD to maintain non-covalent interactions is not a new finding, and it has been utilized for the characterization of both

peptides[72] and proteins.[73-75] However, these results are interesting because the application of ECD for the examination of non-covalent interactions of oligosaccharides has not been explored.

4.3.4 ECD Mechanism for Metal-Coordinated Oligosaccharides

Prior theoretical and experimental work aimed at establishing the nature of the ECD mechanism has focused on multiply-charged peptide ions. Most of that work only considered protonated species, with very few examinations devoted to ECD of metal-adducted peptides.[52-54,76] However, Heck and co-workers have recently proposed that in an oxytocin-transition metal complex, the metal ions serve as the main initial electron capture site during ECD. The captured electron is then transferred to other sites to form a hydrogen radical, which can drive subsequent fragmentation.[52] Previous work from our group examining metal-adducted Substance P indicates that hydrogen radicals are not necessarily involved in this process.[54] Those results imply that initial electron capture is at the metal and is followed by electron and proton transfer, which is rationalized by the "amide superbase" mechanism.[43] Furthermore, Chan and co-workers have also used metal ions as charge carriers for ECD of peptides.[53] They explained that due to the peptide existing predominantly in the zwitterionic form, electron capture leads to reduction of a mobile proton rather than the metal ion. In our investigation, a zwitterionic oligosaccharide is unlikely due to the absence of acidic sites, so the first two proposals may be more applicable here.

Metal coordination drastically changes the ECD fragmentation spectra of oligosaccharides compared to their protonated counterparts. The differences in ECD fragmentation of protonated versus metal-complexed oligosaccharides are likely due to

differences in metal coordination. While protonation is localized to the glycosidic oxygen, metals are able to undergo coordination with several oxygen atoms simultaneously.[15,22] If electron transfer is involved, multiple coordination could also explain why metal-adducted oligosaccharides fragment differently. According to the previously discussed mechanisms, the coordination of the metal to the oligosaccharide and the second ionization energy are expected to significantly affect ECD fragmentation. However, additional fundamental studies are required in order to understand the ECD mechanism in regards to oligosaccharide fragmentation.

4.4 Conclusions

We demonstrate that metal adduction allows ECD of oligosaccharides that do not contain basic sites. In addition, we show that complementary structural information can be obtained from ECD of such metal-adducted species as compared to IRMPD. Depending on the metal ion and oligosaccharide structure, in certain examples we have shown that cross-ring cleavage can be greatly enhanced (particularly for maltoheptaose). A comparison of the ECD spectra of linear oligosaccharides with varying metal charge carriers demonstrates subtle differences in fragmentation. These subtle differences are likely due to several factors, including coordination number and second ionization energy of each metal. The fragmentation pattern of a branched oligosaccharide was also examined to determine whether branching would significantly affect the ECD fragmentation pattern. The larger branched oligosaccharide was more difficult to fragment with ECD, presumably due to secondary structure, therefore a direct

comparison of ECD fragmentation efficiencies of linear versus branched oligosaccharides is difficult to make here.

Following ECD of oligosaccharides ionized with two sodium or potassium ions, the predominant fragmentation pathway was loss of sodium and potassium (with only a few low abundance glycosidic and cross-ring fragments). Of the divalent metal-adducted oligosaccharides examined, our results do not indicate that there is a clearly optimal metal which will maximize cross-ring fragmentation for both branched and linear oligosaccharides following ECD. ECD of all metal-adducted oligosaccharides consistently provided complementary fragmentation behavior as compared to IRMPD. However, for the characterization of oligosaccharides with the combination of ECD and metal adduction, it is advantageous to select a metal which ionizes the oligosaccharide most effectively. ECD fragmentation efficiency is lower compared to traditional MS/MS techniques, thus selecting a metal ion which maximizes precursor ion signal is an important factor to consider. In summary, when used in conjunction with IRMPD, ECD of metal-adducted oligosaccharides results in complementary fragmentation behavior, validating it as a useful tool for the structural characterization of oligosaccharides.

4.5 References

1. Varki, A. *Glycobiol.* **1993**, 97-130.
2. Parodi, A. J. *Annu. Rev. Biochem.* **2000**, *69*, 60-93.
3. Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *291*, 2370-2376.
4. Carr, S. A.; Reinhold, V. N.; Green, B. N.; Haas, J. R. *Biomed. Mass Spectrom.* **1985**, *12*, 288-295.
5. Müller, D. R.; Domon, B.; Blum, W.; Raschdorf, F.; Richter, W. J. *Biomed. Environ. Mass Spectrom.* **1988**, *15*, 441-446.
6. Domon, B.; Müller, D. R.; Richter, W. J. *Org. Mass Spectrom.* **1989**, *24*, 357-359.
7. Domon, B.; Müller, D. R.; Richter, W. J. *Biomed. Environ. Mass Spectrom.* **1990**, *19*, 390-392.
8. Laine, R. A.; Pamidimukkala, K. M.; French, A. D.; Hall, R. W.; Abbas, S. A.; Jain, R. K.; Matta, K. L. *J. Am. Chem. Soc.* **1988**, *110*, 6931-6939.
9. Domon, B.; Müller, D. R.; Richter, W. J. *Int. J. Mass Spectrom. Ion Processes* **1990**, *100*, 301-311.
10. Gillece-Castro, B. L.; Burlingame, A. L. *Methods Enzymol.* **1990**, *193*, 689-712.
11. Lemoine, J.; Strecker, G.; Leroy, Y.; Fournet, B.; Ricart, G. *Carbohydr. Res.* **1991**, *221*, 209-217.
12. Harvey, D. J. *Mass Spectrom. Rev.* **1999**, *18*, 349-451.
13. Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161-227.
14. Park, Y.; Lebrilla, C. B. *Mass Spectrom. Rev.* **2005**, *24*, 232-264.
15. Zhou, Z.; Ogden, S.; Leary, J. A. *J. Org. Chem.* **1990**, *55*, 5446-5448.
16. Orlando, R.; Bush, C. A.; Fenselau, C. *Biomed. Environ. Mass Spectrom.* **1990**, *19*, 747-754.
17. Hofmeister, G. E.; Zhou, Z.; Leary, J. A. *J. Am. Chem. Soc.* **1991**, *113*, 5964-5970.
18. Staempfli, A.; Zhou, Z.; Leary, J. A. *J. Org. Chem.* **1992**, *57*, 3590-3594.
19. Fura, A.; Leary, J. A. *Anal. Chem.* **1993**, *65*, 2805-2811.
20. Ngoka, L. C.; Gal, J.; Lebrilla, C. B. *Anal. Chem.* **1994**, *66*, 692-698.
21. Kohler, M.; Leary, J. A. *Anal. Chem.* **1995**, *67*, 3501-3508.
22. Cancilla, M. T.; Penn, S. G.; Carroll, J. A.; Lebrilla, C. B. *J. Am. Chem. Soc.* **1996**, *118*, 6736-6745.
23. Sible, E. M.; Brimmer, S. P.; Leary, J. A. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 32-42.
24. König, S.; Leary, J. A. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1125-1134.
25. Penn, S. G.; Cancilla, M. T.; Lebrilla, C. B. *Int. J. Mass Spectrom.* **2000**, *195/196*, 259-269.
26. Harvey, D. J. *J. Mass Spectrom.* **2000**, *35*, 1178-1190.
27. Harvey, D. J. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 926-937.
28. Wheeler, S. F.; Harvey, D. J. *Anal. Chem.* **2000**, *72*, 5027-5039.
29. Pfenninger, A.; Karas, M.; Finke, B.; Stahl, B. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1331-1340.
30. Chai, W.; Lawson, A. M.; Piskarev, V. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 670-679.

31. Harvey, D. J.; Naven, T. J. P.; Kuster, B.; Bateman, R. H.; Green, M. R.; Critchley, G. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1556-1561.
32. Harvey, D. J.; Bateman, R. H.; Green, B. N. *J. Mass Spectrom.* **1997**, *32*, 167-187.
33. Mechref, Y.; Novotny, M. V. *Anal. Chem.* **2003**, *75*, 4895-4903.
34. Shi, S. D.-H.; Hendrickson, C. L.; Marshall, A. G.; Seigel, M. M.; Kong, F.; Carter, G. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1285-1290.
35. Xie, Y. M.; Lebrilla, C. B. *Anal. Chem.* **2003**, *75*, 1590-1598.
36. Zhang, J. H.; Schubothe, K.; Li, B. S.; Russell, S.; Lebrilla, C. B. *Anal. Chem.* **2005**, *77*, 208-214.
37. Budnik, B. A.; Haselmann, K. F.; Elkin, Y. N.; Gorbach, V. I.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 5994-6001.
38. Devakumar, A.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2313-2320.
39. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
40. Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.
41. Zubarev, R. A.; Haselmann, K. F.; Budnik, B.; Kjeldsen, F.; Jensen, F. *Eur. Mass Spectrom.* **2002**, *8*, 337-349.
42. Syrstad, E. A.; Stephens, D. D.; Turecek, F. *J. Phys. Chem. A* **2003**, *107*, 115-126.
43. Syrstad, E. A.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 208-224.
44. Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201-222.
45. Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.
46. Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.
47. Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 2355-2361.
48. Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
49. Mormann, M.; Macek, B.; de Peredo, A. G.; Hofsteenge, J.; Peter-Katalinic, J. *Int. J. Mass Spectrom.* **2004**, *234*, 11-21.
50. Adamson, J. T.; Hakansson, K. *J. Proteome Res.* **2006**, *5*, 493-501.
51. Liu, H.; Hakansson, K. *Anal. Chem.* **2006**, *78*, 7570-7576.
52. Kleinnijenhuis, A. J.; Mihalca, R.; Heeren, R. M. A.; Heck, A. J. R. *Int. J. Mass Spectrom.* **2006**, *253*, 217-224.
53. Fung, Y. M.; Liu, H.; Chan, T. W. D. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 757-771.
54. Liu, H.; Håkansson, K. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1731-1741.
55. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2004**, *77*, 1876-1882.
56. Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.
57. Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26*, 514-518.

58. Heck, A. J. R.; Derrick, P. J. *Anal. Chem.* **1997**, *69*, 3603-3607.
59. de Koning, L. J.; Nibbering, N. M. M.; Orden, S. L.; Laukien, F. H. *Int. J. Mass Spectrom. Ion Processes* **1997**, *165/166*, 209-219.
60. Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.
61. Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.
62. Ledford, E. B., Jr.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **1984**, *56*, 2744-2748.
63. Lohmann, K. K.; von der Lieth, C. W. *Nucleic Acids Res.* **2004**, *32*, 261-266.
64. Domon, B.; Costello, C. E. *Glycoconj. J.* **1988**, *5*, 397-409.
65. Turecek, F. *J. Am. Chem. Soc.* **2003**, *125*, 5954-5963.
66. Turecek, F.; Syrstad, E. A. *J. Am. Chem. Soc.* **2003**, *125*, 3353-3369.
67. Leymarie, N.; Costello, C. E.; O'Connor, P. B. *J. Am. Chem. Soc.* **2003**, *125*, 8949-8958.
68. Breuker, K.; Oh, H. B.; Lin, C.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 14011-14016.
69. Cerda, B. A.; Horn, D. M.; Breuker, K.; Carpenter, B. K.; McLafferty, F. W. *Eur. Mass Spectrom.* **1999**, *5*, 335-338.
70. Cerda, B. A.; Breuker, K.; Horn, D. M.; McLafferty, F. W. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 565-570.
71. Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2006**, *18*, 113-120.
72. Haselmann, K. F.; Jorgensen, T. J. D.; Budnik, B. A.; Jensen, F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 2260-2265.
73. Oh, H.; Breuker, K.; Sze, S. K.; Ge, Y.; Carpenter, B. K.; McLafferty, F. W. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 15863-15868.
74. Breuker, K.; McLafferty, F. W. *Angew. Chem. Int. Ed.* **2003**, *42*, 4900-4904.
75. Xie, Y.; Zhang, J.; Yin, S.; Loo, J. A. *J. Am. Chem. Soc.* **2006**, *128*, 14432-14433.
76. Iavarone, A. T.; Paech, K.; Williams, E. R. *Anal. Chem.* **2004**, *76*, 2231-2238.

Chapter 5

Electron Detachment Dissociation of Neutral and Sialylated Oligosaccharides

5.1 Introduction

Glycosylation is a highly prevalent post-translational modification (PTM), whose role has been linked to a wide variety of biological activities,[1,2] ranging from protein folding[3] to immune system response.[4] Unlike most biomolecules, oligosaccharides often exist as several isomeric forms with diverse linkages, and may form linear or branched structures. Complete structural characterization of oligosaccharides requires the determination of constituent monosaccharides, their linkage, sequence, and branching patterns. Given their diversity and structural complexity, structural elucidation of oligosaccharides often relies upon a wide range of analytical methodologies, of which nuclear magnetic resonance spectroscopy and mass spectrometry are two vital techniques. Tandem mass spectrometry (MS/MS) is widely used for glycan structural characterization,[5-7] due to the advent of instruments that provide high quality spectra from even low abundance molecular species.

Tandem mass spectra of oligosaccharides consist mainly of glycosidic and cross-ring product ions (shown in Figure 1.5 in Chapter 1). Glycosidic cleavage occurs

between monosaccharide units and provides information regarding saccharide sequence and branching. Cross-ring cleavages can provide valuable information regarding saccharide linkage, particularly when occurring at branching residues. Several factors are known to affect oligosaccharide fragmentation and the degree of cross-ring fragmentation, such as the ionizing cation, the lifetime of the ion prior to detection, and the energy deposited into the ion.

Typically, neutral oligosaccharides are analyzed in positive ion mode, via their protonated forms or through metal ion adduction. In addition, chemical derivatization such as permethylation is widely used to increase sensitivity, reduce molecular ion lability, and produce structurally diagnostic product ions.[8-10] Low energy activation methods, such as collisional activated dissociation (CAD) and infrared multiphoton dissociation (IRMPD), applied to protonated oligosaccharides results in predominantly glycosidic cleavages. However, oligosaccharides ionized with alkali, alkaline earth, and transition metals often fragment to yield more cross-ring cleavages compared to their protonated counterparts.[11-14] Although fragmentation of neutral oligosaccharides in negative ion mode is not as frequently examined, it has been demonstrated that native, neutral oligosaccharide anions produce abundant C-type glycosidic cleavages and A-type cross-ring cleavages in CAD, as opposed to B- and Y-type glycosidic cleavages which are commonly observed for oligosaccharide cations (Domon and Costello nomenclature[15]).[16-19] Low energy CAD combined with negative ion mode electrospray ionization (ESI) has been used for examining neutral glycans from human urine and milk.[19-22] Less work has been done on negative ions from N-linked glycans, but it has

been shown that neutral, singly deprotonated *N*-glycans are highly susceptible to in-source fragmentation during electrospray ionization.[23]

Many oligosaccharides derived from glycoproteins and glycolipids contain sialic acids, monosaccharides which contain a carboxylic acid group at the C-1 position (shown in Figure 1.1 in Chapter 1). Oligosaccharides containing sialic acids typically yield more abundant signal in negative ion mode mass spectrometry. When sialylated oligosaccharides are analyzed in positive ion mode, sialic acids may be lost, resulting in an absence of information regarding sialic acid linkage.[5,6] Native sialylated oligosaccharides are particularly fragile in matrix-assisted laser desorption/ionization (MALDI) in which they often undergo in-source fragmentation, thereby suggesting electrospray ionization is more suitable for their analysis. The influence of sialylation on negative ion mode CAD for milk and *N*-glycans has been recently examined.[24] These results demonstrated that deprotonated sialylated oligosaccharides required more energy to fragment compared to either their deprotonated asialo or nitrate adducted counterparts. Zaia and co-workers explain that sialylated glycans are more stable in the gas phase because negative charge resides on the carboxyl group. However, for asialo oligosaccharides, a proton is extracted from a ring hydroxyl group during electrospray ionization, which imparts more energy than it would for a sialylated oligosaccharide.

Although most modern mass analyzers utilize low energy dissociation techniques such as low energy CAD, several alternative fragmentation techniques have been developed and applied to oligosaccharides including high-energy CAD,[25-27] IRMPD,[28,29] electron capture dissociation (ECD),[30,31] electron detachment dissociation (EDD),[32,33] and 157 nm photodissociation.[34] In particular, the use of ion-electron reactions for

biomolecular structural characterization is rapidly expanding. ECD has been used extensively for the characterization of peptide and protein cations,[35] due to its ability to promote extensive peptide backbone fragmentation without the loss of labile modifications such as glycosylation.[36-41] In ECD, polycationic molecules are irradiated with low-energy electrons (<1 eV), generating charge reduced radical species and product ions.[42,43] ECD has been applied to protonated chitooligosaccharides, which yielded primarily glycosidic cleavages corresponding to B and C-type ions,[30] and recently to permethylated oligosaccharides.[44] In Chapter 4, we demonstrated that ECD of oligosaccharides ionized with divalent metal ions can result in additional fragmentation not observed following vibrational excitation.[31]

Electron detachment dissociation (EDD) was introduced in 2001 as a fragmentation technique for polyanions.[45] In EDD, polyanions are irradiated with >10 eV electrons, resulting in electron detachment and subsequent product ions. EDD has been applied to peptides,[45-47] oligonucleotides,[48,49] gangliosides,[50] and recently glycosaminoglycans (GAGs).[32,33] GAGs are linear polysaccharides consisting of repeating disaccharide units and are frequently polysulfated. EDD of GAGs resulted in information-rich mass spectra with both cross-ring and glycosidic cleavages. In contrast, CAD and IRMPD of GAGs resulted in predominantly glycosidic cleavages.

EDD has, to our knowledge, not previously been applied to a wide variety of oligosaccharides, including neutral and sialylated glycans, as well as branched oligosaccharides. In the current manuscript, we examine the fragmentation patterns of neutral and sialylated oligosaccharides following EDD and compare these fragmentation patterns to those obtained from IRMPD and CAD. In addition, both linear and branched

oligosaccharides are examined to determine what effect branching has upon EDD fragmentation.

5.2 Experimental

5.2.1 Sample Preparation

Maltoheptaose (Sigma, St. Louis, MO), LS-tetrasaccharides (LSTa and LSTb), disialyllacto-*N*-tetraose (DSLNT), an asialo biantennary glycan (NA2) (V-labs, Covington, LA), and a monosialylated biantennary glycan (A1F) (Calbiochem, San Diego, CA) were prepared in a solution of 50 % methanol (Fisher, Fair Lawn, NJ) and 0.1 % ammonium hydroxide (Sigma) to a final concentration of 3 μM.

5.2.2 FT-ICR Mass Spectrometry

All experiments were performed with an actively shielded 7 T Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer with a quadrupole front-end (APEX-Q, Bruker Daltonics, Billerica, MA), as previously described.[51] Samples were infused via an Apollo II ion source at a flow rate of 60 μL/h with the assistance of N_2 nebulizing gas. Following ion accumulation in the first hexapole for 0.05 s, ions were mass selectively accumulated in the second hexapole for 2 - 6 s. Ions were then transferred through high-voltage ion optics and captured with dynamic trapping in an Infinity ICR cell.[52] The experimental sequence up to the ICR cell fill was looped 4 - 8 times to achieve maximum precursor ion signal for EDD. Precursor ions were further isolated in the ICR cell using correlated harmonic excitation fields (CHEF).[53] An indirectly heated hollow dispenser cathode was used to perform EDD.[54] IRMPD was performed with a vertically mounted 25-W, 10.6-μm, CO_2 laser (Synrad, Mukilteo, WA).

External CAD was performed in a hexapole following mass selective ion accumulation with argon as a collision gas. For EDD, the cathode heating current was kept constant at 2.0 A and the cathode voltage was pulsed during the EDD event to a bias voltage of (- 20) - (- 30) V for 1 - 2 s. A lens electrode located immediately in front of the cathode was kept 0.8 - 1 V higher than the cathode bias voltage.

5.2.3 Data Analysis

All mass spectra were acquired with XMASS software (Bruker Daltonics) with 256 data points and summed over 30 scans. Data processing was performed with MIDAS software.[55] Data were zero filled once, Hanning apodized, and exported to Microsoft Excel for internal frequency-to-mass calibration with a two-term calibration equation. The precursor ion and charge reduced radical species were used for calibration of EDD spectra. For CAD, IRMPD, and EDD spectra where the charge reduced species was not observed (or with only low abundance), an abundant glycosidic cleavage product ion and the precursor were used for calibration. Product ion spectra were interpreted with the aid of the web application GlycoFragment (www.dkfz.de/spec/projekte/fragments/).[56]

5.3 Results and Discussion

All product ions are labeled according to the Domon and Costello nomenclature.[15] Ions due to multiple cleavage sites are designated with a slash between sites of cleavage (e.g., $^{2,4}A_6/Y_5$). When several product ion assignments are possible, alternative assignments are indicated in parentheses (e.g., $Y_{4\alpha}(Y_{3\beta})$). For branched oligosaccharides, the letter α represents the largest branch.

5.3.1 MS/MS of Neutral Oligosaccharides

Two neutral oligosaccharides were examined to evaluate the EDD fragmentation patterns of oligosaccharides lacking acidic moieties. These oligosaccharides include maltoheptaose, a linear oligosaccharide containing seven α1→4 linked glucose molecules (Scheme 5.1), and an asialo, biantennary glycan (NA2) with a composition of Galβ4GlcNAcβ2Manα6(Galβ4GlcNAcβ2Manα3)Manβ4GlcNAcβ4-GlcNAc (Scheme 5.2). Due to the symmetric nature of maltoheptaose, a number of ions cannot be distinguished based on their m/z ratios (C and Y, and several A and X). However, native neutral oligosaccharide anions that are 1-4 or 1-6 linked have been shown to produce predominantly C-type and A-type ions during CAD.[16-20] The majority of ions observed following CAD (Figure 5.1a) and IRMPD (Figure 5.1b) of maltoheptaose are singly charged, with exceptions being denoted by an asterisk. The CAD and IRMPD spectra show very similar fragmentation patterns, although they are not identical. Several lower molecular weight glycosidic and cross-ring cleavage product ions were only observed following IRMPD. Although IRMPD and low-energy CAD are both low energy vibrational excitation techniques, Lebrilla and co-workers have shown that the fragmentation efficiency in IRMPD is greater than that in CAD for larger oligosaccharides.[57] An alternative explanation for the absence of low molecular weight species following CAD is the time-of-flight effect when transferring ions from the external hexapole to the ICR cell. Low *m/z* species will arrive at the cell sooner than higher *m/z* species, thereby displaying optimum capture at different experimental conditions. Cross-ring cleavages observed following both fragmentation techniques

corresponded either to ^{0,2}A-type or ^{2,4}A-type ions, consistent with α1-4 linked saccharides. CAD and IRMPD of singly deprotonated maltoheptaose were also

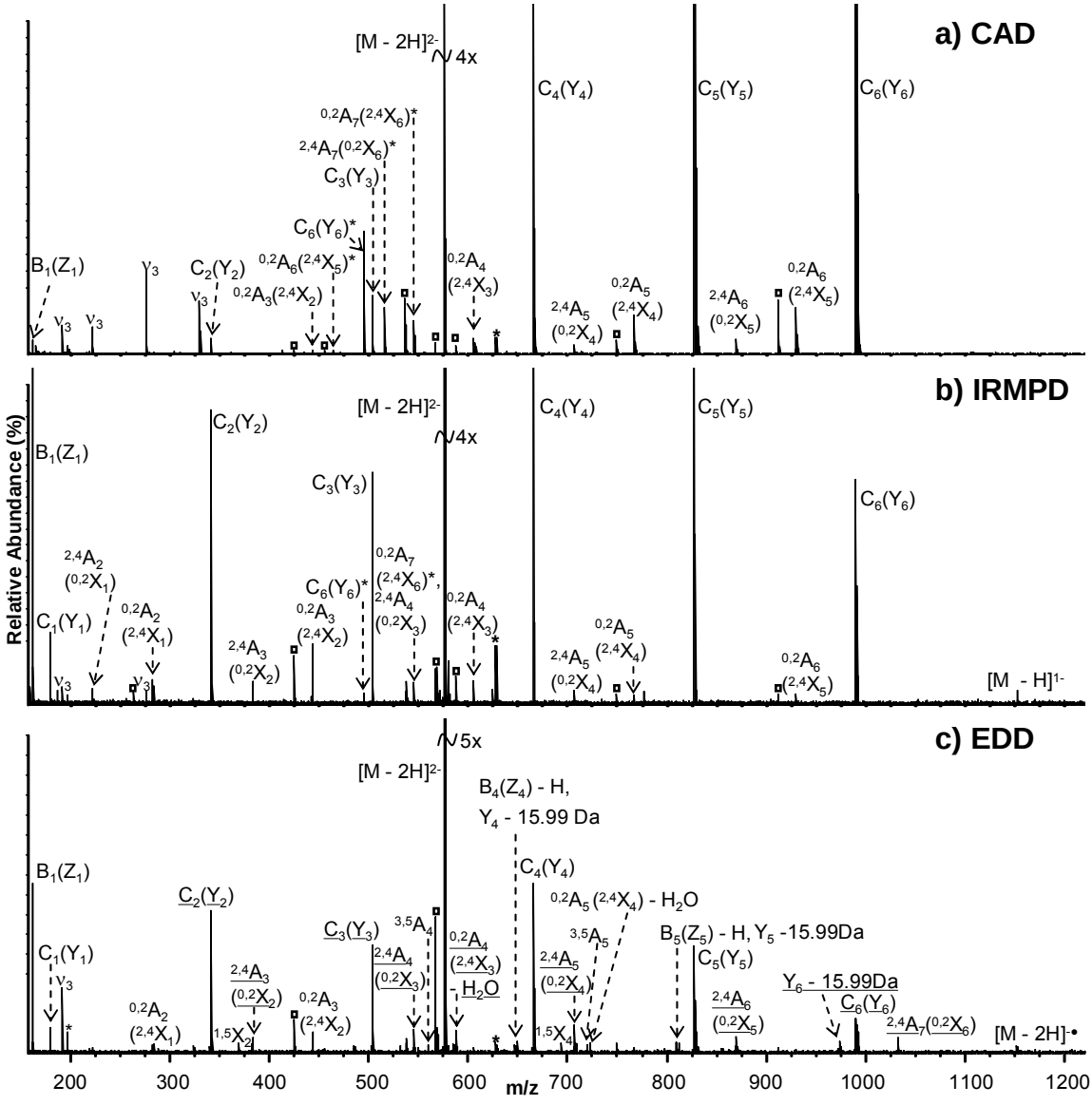
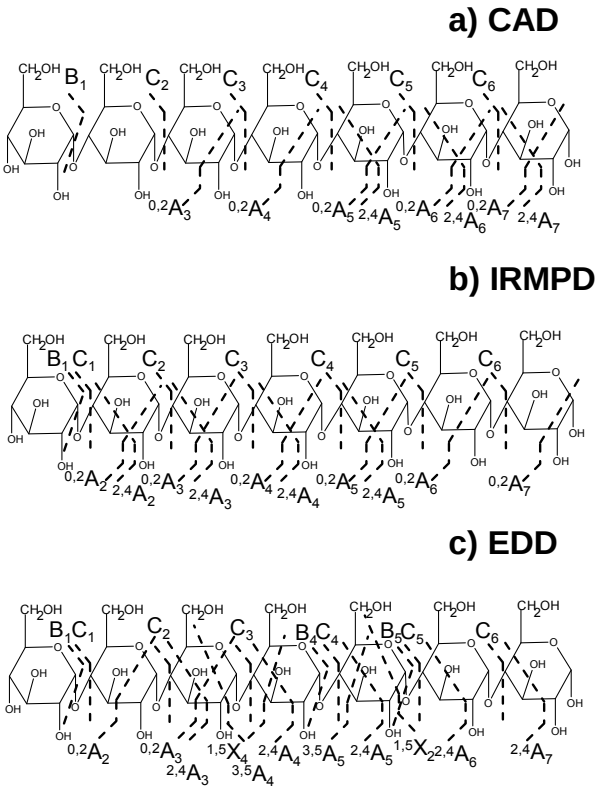


Figure 5.1. FT-ICR tandem mass spectra of doubly deprotonated maltoheptaose. a) CAD (30 scans, collision cell voltage 3 V) b) IRMPD (30 scans, 200 ms irradiation, 2.5 W) c) EDD (30 scans, 2 s, bias voltage of – 25 V). Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Asterisks above a peak indicate electronic noise. Squares indicate water loss from an adjacent product ion. Underlined product ions exhibit minor radical species (with a mass corresponding to hydrogen loss). Due to the symmetry of the molecule, several product ions cannot be unambiguously assigned (indicated by parentheses in the labels in each spectrum).

examined, and are shown in Figure 5.2. Product ions observed from the $[M - H]^-$ species were very similar to those from the $[M - 2H]^{2-}$ precursor ions. Although maltoheptaose is a neutral oligosaccharide, under the negative ion mode electrospray conditions described earlier, the relative abundance of the doubly deprotonated species was typically at least 50%. Thus EDD of this neutral oligosaccharide was not hindered by poor ion signal.



Scheme 5.1. Fragmentation pattern observed following CAD, IRMPD, and EDD of doubly deprotonated maltoheptaose.

The EDD spectrum of doubly deprotonated maltoheptaose (Figure 5.1c) contains only singly charged product ions, including the charged reduced species, $[M - 2H]^{-\bullet}$. Singly charged product ions can arise via two fragmentation mechanisms, either through direct decomposition of the activated precursor ion, or through electron detachment from the precursor and subsequent fragmentation.[32] A mixture of even- and odd-electron species are observed following EDD of maltoheptaose. Ions which exhibit minor hydrogen loss, resulting in odd-electron species, are underlined in Figure 5.1c. For these

product ions, hydrogen loss was a relatively minor fragmentation pathway. One exception is the ion at *m*/*z* 646.2, denoted $B_4(Z_4)$ - H, in which the radical species is more abundant than the even-electron species. Product ions corresponding to loss of hydrogen from B-type ions have also been observed following EDD of GAGs.[32] Odd-electron product ions can be formed following hydrogen loss from an even-electron species, or can be formed directly following fragmentation of the charge reduced radical species.

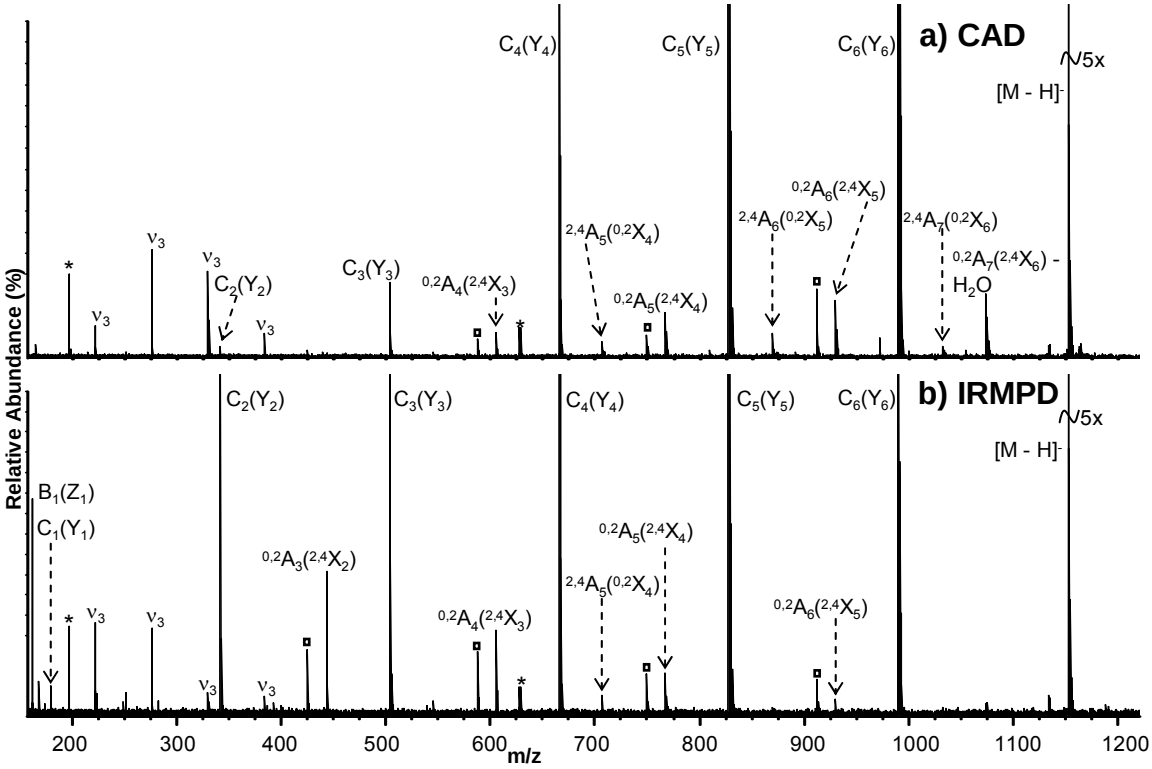


Figure 5.2. FT-ICR tandem mass spectra of singly deprotonated maltoheptaose. a) CAD (30 scans, collision cell voltage 14 V) b) IRMPD (30 scans, 150 ms irradiation 2.5 W). Squares indicate water loss from an adjacent product ion. Due to the symmetry of the molecule, several product ions cannot be unambiguously assigned (indicated by parentheses in the labels in each spectrum).

EDD of maltoheptaose results in a combination of glycosidic and cross-ring fragmentation, and exhibits several unique product ions not observed following either CAD or IRMPD. In addition to several $^{0,2}A(^{2,4}X)$ and $^{2,4}A(^{0,2}X)$-type cross-ring cleavages, $^{1,5}X$ and $^{3,5}A$-type ions are observed, including $^{1,5}X_2$, $^{1,5}X_4$, $^{3,5}A_4$, and $^{3,5}A_5$.

Both ^{1,5}X and ^{3,5}A-type cross-ring cleavages are commonly observed following positive ion mode high-energy CAD (heCAD) of oligosaccharides,[27,58] and were also noted following EDD of GAGs.[32] X-type cross-ring cleavages are typically not observed following low energy CAD or IRMPD.

Three unique product ions denoted Y_4 - 15.99 Da, Y_5 - 15.99 Da, and Y_6 - 15.99 Da were observed following EDD of maltoheptaose. These ions were not reported by Amster and co-workers following EDD of GAGs. Kováčik *et al.*[59] and Harvey *et. al.*[58] have reported Y - 16 Da ions following positive ion mode heCAD. In their experiments, only Y - 16 Da ions were observed, as opposed to C - 16 Da (a possible alternative assignment due to the symmetry of maltoheptaose). However, following EDD of other neutral and sialylated oligosaccharides (discussed below), only losses of 16 mass units are observed from Y-type ions. We propose that these ions correspond to Z + 2H ions. Kováčik *et al.* proposed that these ions were due to internal residue losses. However, for maltoheptaose, internal residue loss would not account for the molecular weight of these product ions. The same authors also discussed that these product ions could correspond to Z + 2H ions, but later disregarded this explanation. We believe there are several lines of evidence to support that these ions observed following EDD are Z + 2H species. Hydrogen addition and abstraction following ECD of peptides have been well documented,[60,61] and have also been reported for EDD of peptides.[45] Following EDD of GAGs, hydrogen abstraction from B- and C-type ions has also been noted, and observed as B - H, C - H, and C - 2H type ions.[32,33] Of particular interest is the product ion type C - 2H, which also has been reported by several investigators following positive ion mode heCAD[58,62-65] and, which is observed following EDD of sialylated oligosaccharides in our

examination. The complementary fragment for a C - 2H product ion would be a Z + 2H species. For both maltoheptaose and an asialo biantennary glycan (NA2) (Figure 5.4), Z + 2H (or Y - 16 Da) product ions are observed while their complementary C - 2H species are absent.

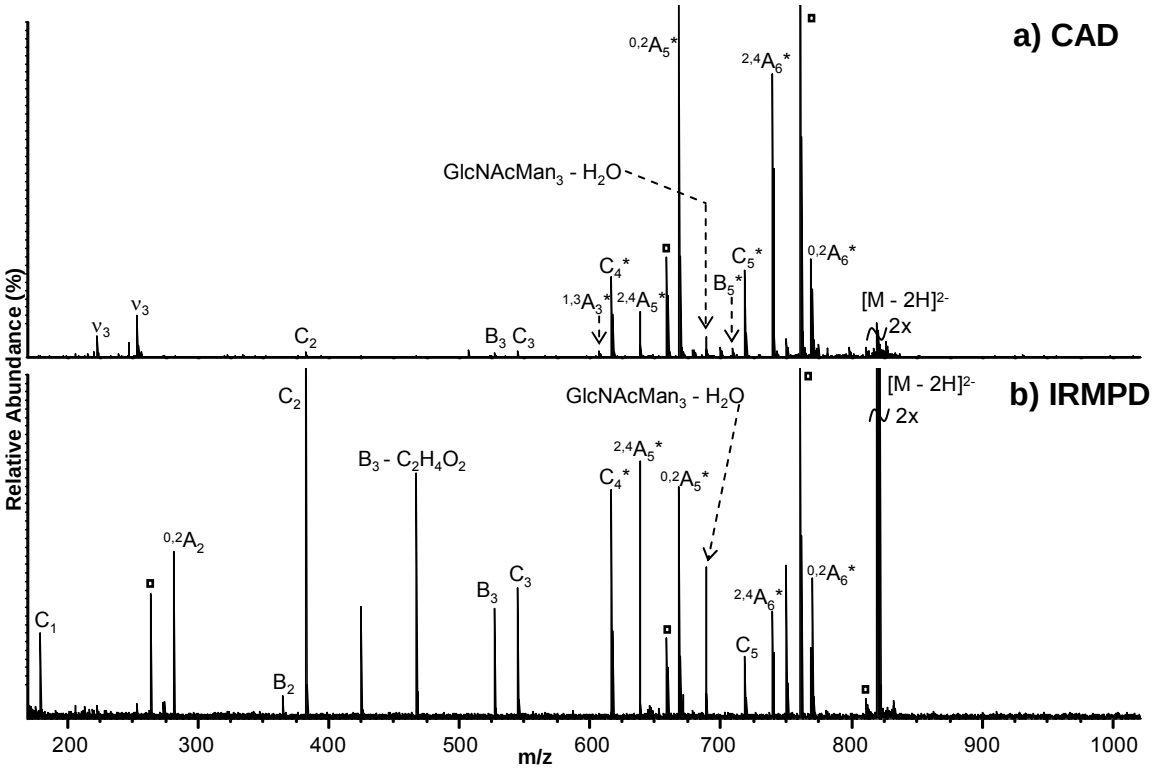


Figure 5.3. FT-ICR tandem mass spectra of doubly deprotonated NA2. a) CAD (30 scans, collision cell voltage 8 V) b) IRMPD (30 scans, 125 ms irradiation 5 W). Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Squares indicate water loss from an adjacent product ion.

The CAD and IRMPD spectra of a doubly deprotonated asialo, biantennary glycan (NA2) are shown in Figure 5.3. The corresponding singly deprotonated species was of relatively poor abundance following ESI, and was not examined. Using negative ion mode ESI conditions as described, the singly charged species had a relative abundance of less than 5%. Both CAD and IRMPD of the doubly deprotonated species resulted in a mixture of glycosidic and cross-ring cleavages, where glycosidic cleavages

included both C- and B-type ions (which could also correspond to water losses from C-type ions). Cross-ring fragments were almost exclusively ^{0,2}A and ^{2,4}A-type ions. The singly deprotonated glycan has also been examined in negative ion mode CAD following ESI by Harvey[66] and by Zaia and co-workers,[24] showing a similar fragmentation pattern to the doubly deprotonated species examined here.

Unlike maltoheptaose, a mixture of doubly and singly charged product ions was observed following EDD of NA2 (Figure 5.4 and Scheme 5.2). Doubly charged product ions, including C_4^*, C_5^*, $^{0,2}A_5^*$, $^{0,2}A_6^*$, and $^{2,4}A_6^*$ are likely due to direct decomposition of the activated precursor ion (an asterisk indicates doubly charged ions). The majority of product ions generated following EDD are even-electron species. Several product ions are due to multiple cleavages, many of which involve a combination of glycosidic and cross-ring fragmentation. In addition to ^{2,4}A and ^{0,2}A-type cross-ring cleavages, ^{1,5}X and ^{3,5}X-type product ions were observed. One of the most abundant product ions detected following EDD corresponds to $^{3,5}X_3$($^{3,5}X_5$), at *m*/*z* 1565.55. Also observed is the product ion at *m*/*z* 688.23 labeled GlcNAcMan$_3$ - H_2O, otherwise known as "ion D". This ion is due to loss of the 3-linked antenna and two reducing GlcNAc residues. Ion D allows for determination of the monosaccharide composition of the 6-linked antenna, and by subtraction also the 3-linked antenna. This ion was also observed following IRMPD and CAD. Two Y -16 Da type ions, or as we propose Z + 2H ions, are also observed following EDD (denoted as Y_4 - 15.99 Da and C_4/Y_5 - 15.99 Da). Similar to maltoheptaose, C - 2H ions were not detected following EDD.

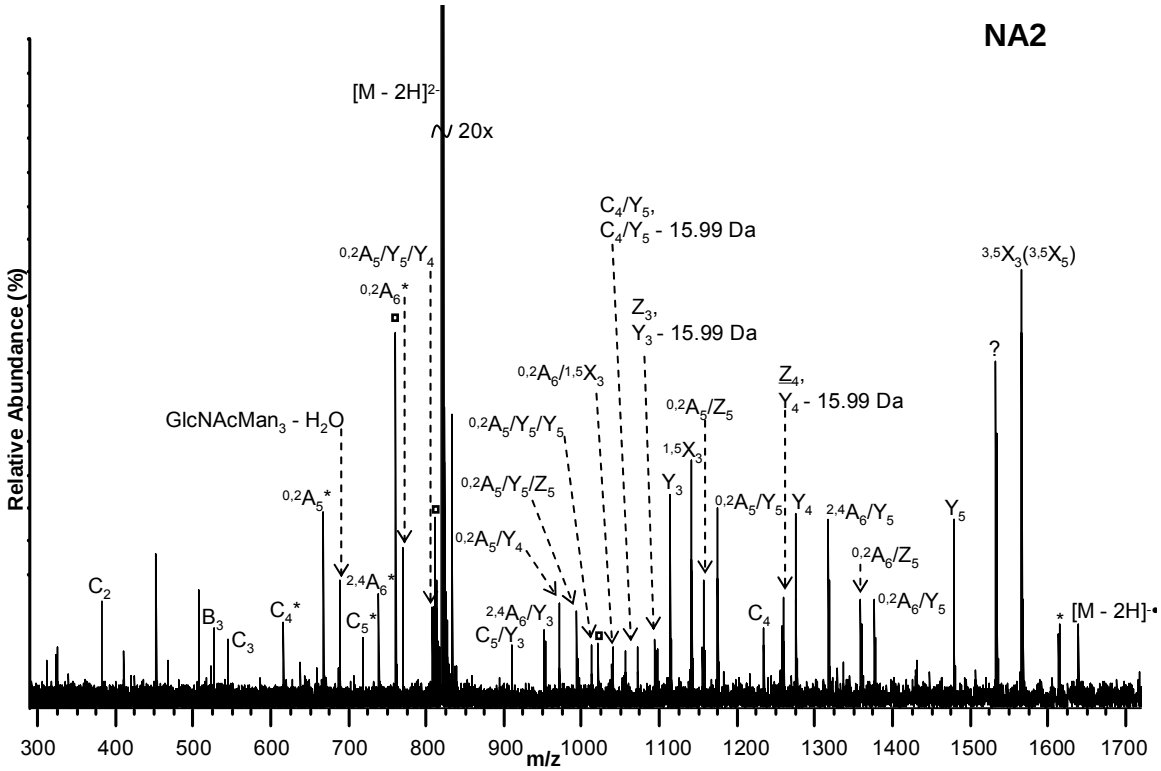
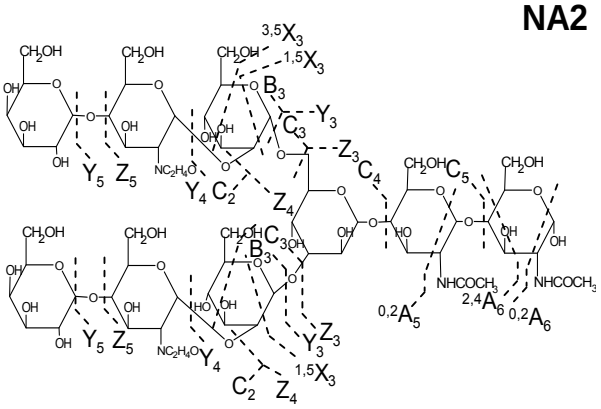


Figure 5.4. EDD FT-ICR (30 scans, 1 s, bias voltage of – 25 V) tandem mass spectrum of a doubly deprotonated asialo, biantennary glycan (NA2). Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Asterisks above a peak indicate electronic noise. Squares indicate water loss from an adjacent product ion. Underlined product ions exhibit minor radical species (with a mass corresponding to loss of hydrogen). Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels.



Scheme 5.2. Fragmentation pattern observed following EDD of a doubly deprotonated asialo, biantennary glycan (NA2).

5.3.2 MS/MS of Sialylated Oligosaccharides

The CAD, IRMPD, and EDD fragmentation patterns of four sialylated oligosaccharides were next examined. The first oligosaccharide examined, LSTa, is a linear oligosaccharide with the composition Neu5Acα3Galβ3GlcNAcβ3Galβ4Glc (Scheme 5.3a). A related branched oligosaccharide, LSTb, in which the sialic acid is instead α2→6 linked to GlcNAc (Scheme 5.3b) was also examined. The third oligosaccharide, a disialyllacto-*N*-tetraose oligosaccharide (DSLNT), is similar to LSTa but with a second sialic acid α2→6 linked to GlcNAc (Scheme 5.4). The last oligosaccharide examined was an N-linked monosialylated, biantennary glycan (A1F), with the composition Neu5Aca3/6Galβ4GlcNAcβ2Manα6(Galβ4GlcNAcβ2Manα3)-Manβ4GlcNAcβ4(Fucα6)-GlcNAc (Scheme 5.5).

CAD and IRMPD of both doubly and singly deprotonated LSTa were examined and are shown in Figures 5.5 and 5.6 respectively. Under standard negative ion mode electrospray ionization conditions, the average relative abundance of doubly deprotonated LSTa was 30%. Although the generation of this doubly deprotonated species is more difficult compared to maltoheptoase or the branched NA2 glycan, EDD of this species was facilitated by longer accumulation times and cell fills. CAD and IRMPD of both singly and doubly deprotonated LSTa resulted in ^{0,2}A and ^{2,4}A cross-ring cleavages in the reducing end glucose. Cross-ring cleavage in the reducing terminal residue and for internal 4-linked residues are commonly observed in negative ion mode low energy vibrational excitation fragmentation.[15,19,20] In addition, glycosidic cleavage (predominantly resulting in C-type ions) was observed between every residue. As has been shown for sialylated oligosaccharides by Zaia and co-workers, [24] singly

deprotonated LSTa (Figure 5.6) required more energy for fragmentation during CAD compared to the energies used for asialo oligosaccharides previously discussed (30 V for LSTa versus 8 – 14 V for maltoheptaose and NA2).

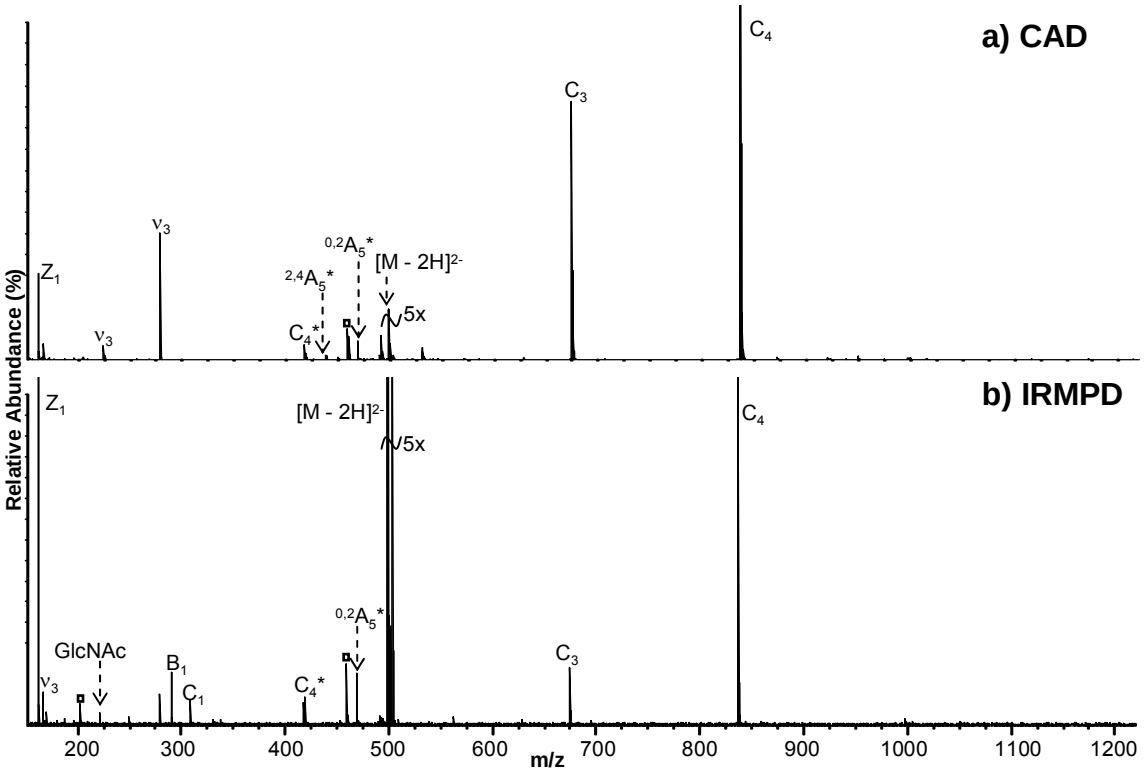


Figure 5.5. FT-ICR tandem mass spectra of doubly deprotonated LSTa. a) CAD (60 scans, collision cell voltage 2 V) b) IRMPD (30 scans, 100 ms irradiation 2.5 W). Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Squares indicate water loss from an adjacent product ion.

Only singly charged product ions were observed following EDD of LSTa (Figure 5.7a). Glycosidic cleavage was noted between every pair of saccharide residues and two cross-ring cleavage products, $^{2,4}X_3(^{1,3}A_2)$ and $^{3,5}A_5$ were also observed following EDD. These cross-ring cleavages were not seen following CAD or IRMPD of either the singly or doubly deprotonated species (Figures 5.5 and 5.6). Loss of CH_3O from the charge reduced species has been previously detected following EDD of GAGs. It is also

interesting to note that both Y - 16 Da species (Y_3 - 15.99 Da and Y_4 - 15.99 Da) and C - 2H species (C_3 - 2H, C_4 - 2H) are observed following EDD.

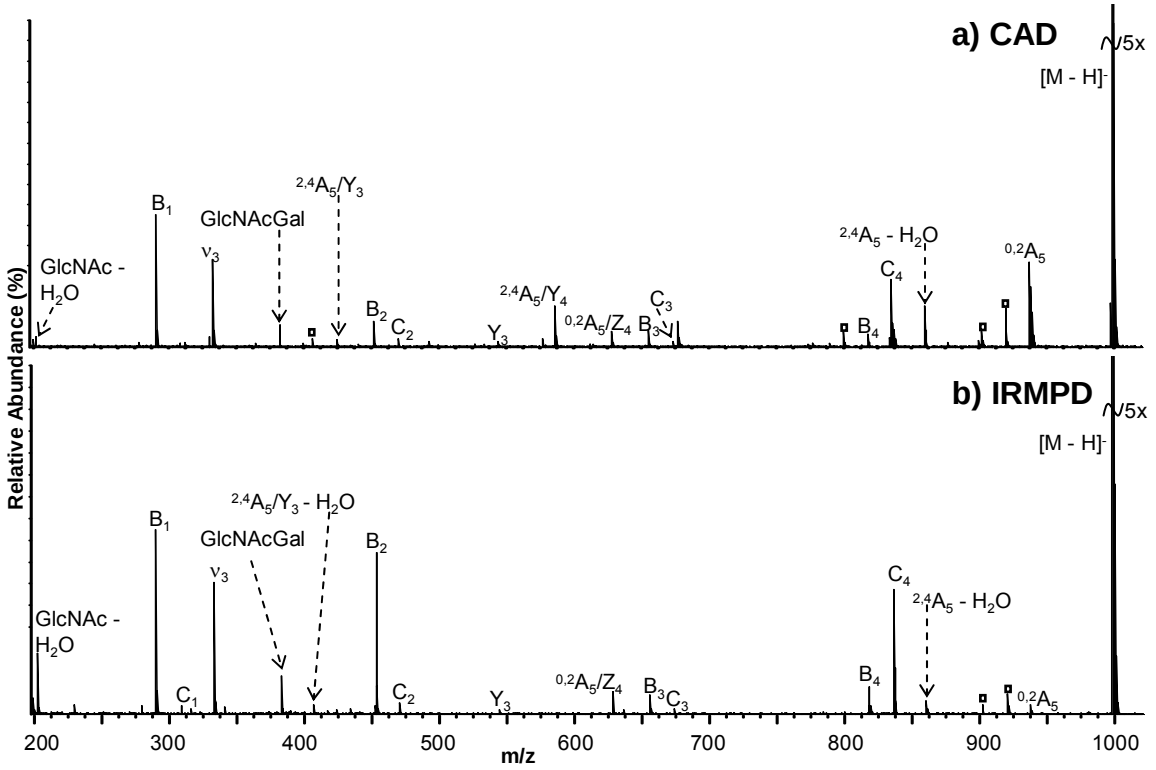


Figure 5.6. FT-ICR tandem mass spectra of singly deprotonated LSTa. a) CAD (30 scans, collision cell voltage 30 V) b) IRMPD (30 scans, 175 ms irradiation 5 W). Squares indicate water loss from an adjacent product ion. Ions due to multiple cleavage sites are designated with a slash between sites of cleavage.

As previously discussed, C - 2H species have been reported in several instances following heCAD in positive ion mode and recently following EDD of GAGs. During positive ion mode heCAD, the precise mechanism which generates these ions is somewhat unclear.[58] In EDD, Wolff *et. al.* have proposed that electron detachment at a carboxylate group is followed by hydrogen transfer to form either an oxy radical at C-2 or a radical at C-3, which can lead to the formation of C - 2H product ions by α-cleavage.[33] Following EDD of GAGs, the complementary Z + 2H ions (or Y - 16 Da species) were not observed. In our experiments, such species were observed following

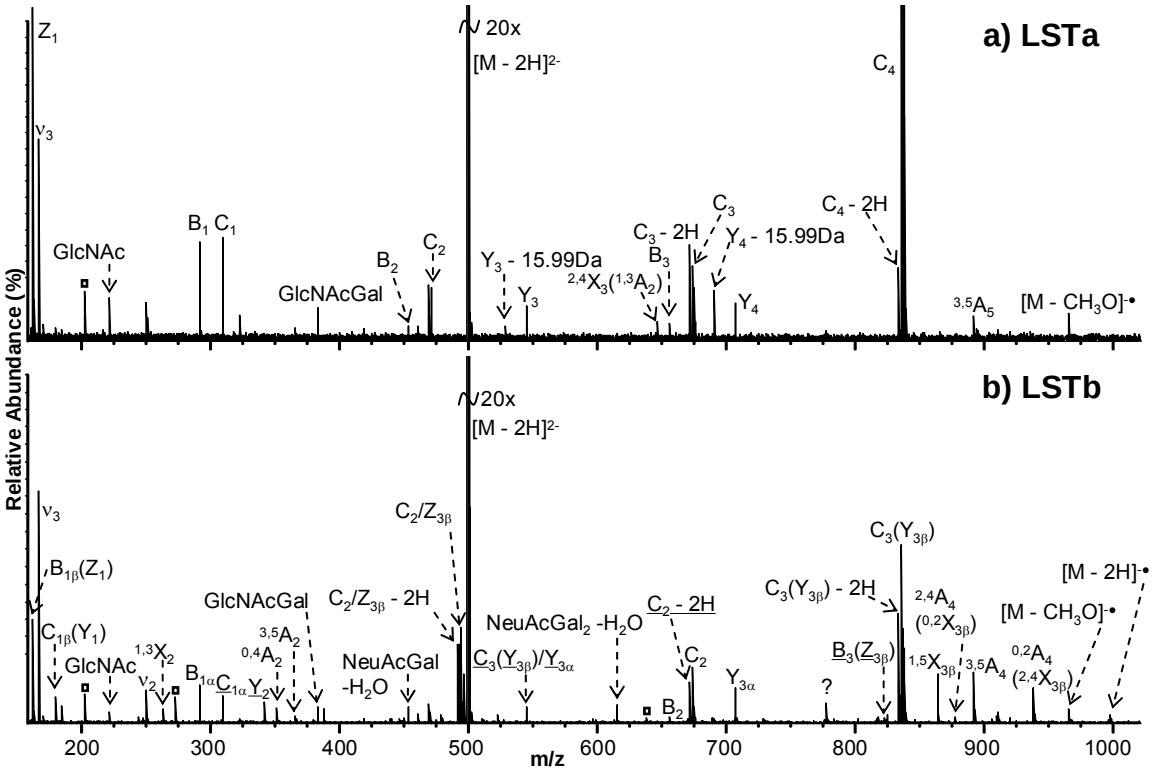
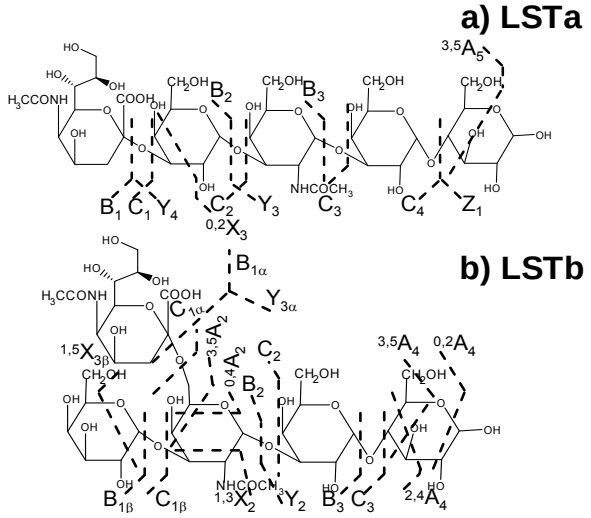


Figure 5.7. EDD FT-ICR tandem mass spectra of doubly deprotonated oligosaccharides a) LSTa (30 scans, 1 s, bias voltage of – 30 V) and b) LSTb (30 scans, 1 s, bias voltage of – 25 V). Squares indicate water loss from an adjacent product ion. Underlined product ions exhibit minor radical species (with a mass corresponding to loss of hydrogen). Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels. The branch containing sialic acid is referred to as the α-branch in product ion labels.



Scheme 5.3. Fragmentation patterns observed following EDD of doubly deprotonated oligosaccharides LSTA and LSTb.

EDD of maltoheptaose and NA2, while the C - 2H species were absent. This observation may be due to differences in the sites of deprotonation in neutral versus acidic oligosaccharides. For the neutral oligosaccharides examined, both sites of deprotonation should be hydroxyl groups. The allylic hydroxyl hydrogen[67] on C-1 and the C-3 hydroxyl hydrogen[68] have been shown to be acidic sites. LSTa, which contains one sialic acid, most likely has one deprotonation site at the carboxylate group while the other site should be a hydroxyl group.

The CAD and IRMPD spectra of doubly (Figure 5.8) and singly deprotonated (Figure 5.9) LSTb showed relatively similar fragmentation patterns. The relative abundance of doubly deprotonated LSTb was similar to LSTa, and EDD was facilitated by longer accumulation times and cell fills. CAD and IRMPD of the singly and doubly deprotonated species exhibited ^{0,2}A and ^{2,4}A-type cleavage in the reducing terminal glucose. The majority of observed product ions retained sialic acid and two of the most informative product ions observed were $C_2/Z_{3\beta}$ (D-type cleavage) and $^{0,4}A_2$. Residues which are 1-3 linked often produce C-type cleavage on the reducing side and Z-type cleavage on the non-reducing side in negative ion mode.[19] Both of these ions provide information regarding the linkage position of sialic acid.

EDD of LSTb (Figure 5.7b) resulted in more extensive fragmentation compared to CAD and IRMPD of the singly (Figure 5.9) and doubly deprotonated species (Figure 5.8). In addition to extensive glycosidic cleavage, several cross-ring cleavage product ions were observed following EDD. Along with ^{0,2}A and ^{2,4}A-type cleavage within the reducing end glucose (which were also observed from IRMPD and CAD), ^{1,3}X, ^{3,5}A, and ^{1,5}X-type cleavages were detected. Three cross-ring products were observed in the

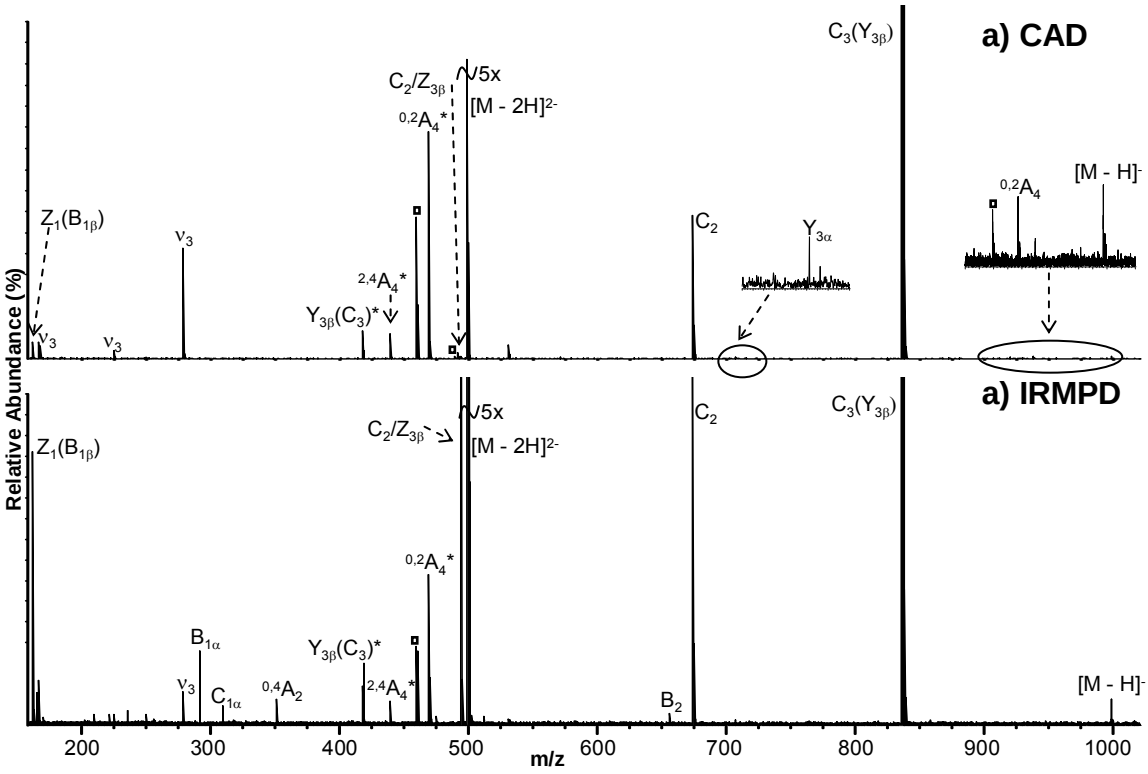


Figure 5.8. FT-ICR tandem mass spectra of doubly deprotonated LSTb. a) CAD (30 scans, collision cell voltage 3 V) b) IRMPD (30 scans, 100 ms irradiation 7.5 W). Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Squares indicate water loss from an adjacent product ion. Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels.

branching GlcNAc residue, including $^{1,3}X_2$, $^{3,5}A_2$, and $^{0,4}A_2$. The $^{0,4}A_2$ and $^{3,5}A_2$ product ions indicate the attachment of sialic acid to C-6 while the $^{1,3}X_2$ ion indicates a likely linkage position for the 3-linked galactose. Two hydrogen losses from several C-type ions (or hydrogen loss from a radical C-type ion) were also seen, including $C_2/Z_{3\beta}$ - 2H, C_2 - 2H, and $C_3(Y_3)$ - 2H. Two unique product ions, denoted NeuAcGal - H_2O and $NeuAcGal_2$ - H_2O are also interesting to note in the EDD spectrum. These ions can only be formed by saccharide rearrangement. Other carbohydrates have also been shown to undergo rearrangement reactions with losses of internal saccharide residues during CAD.[59,69-71]

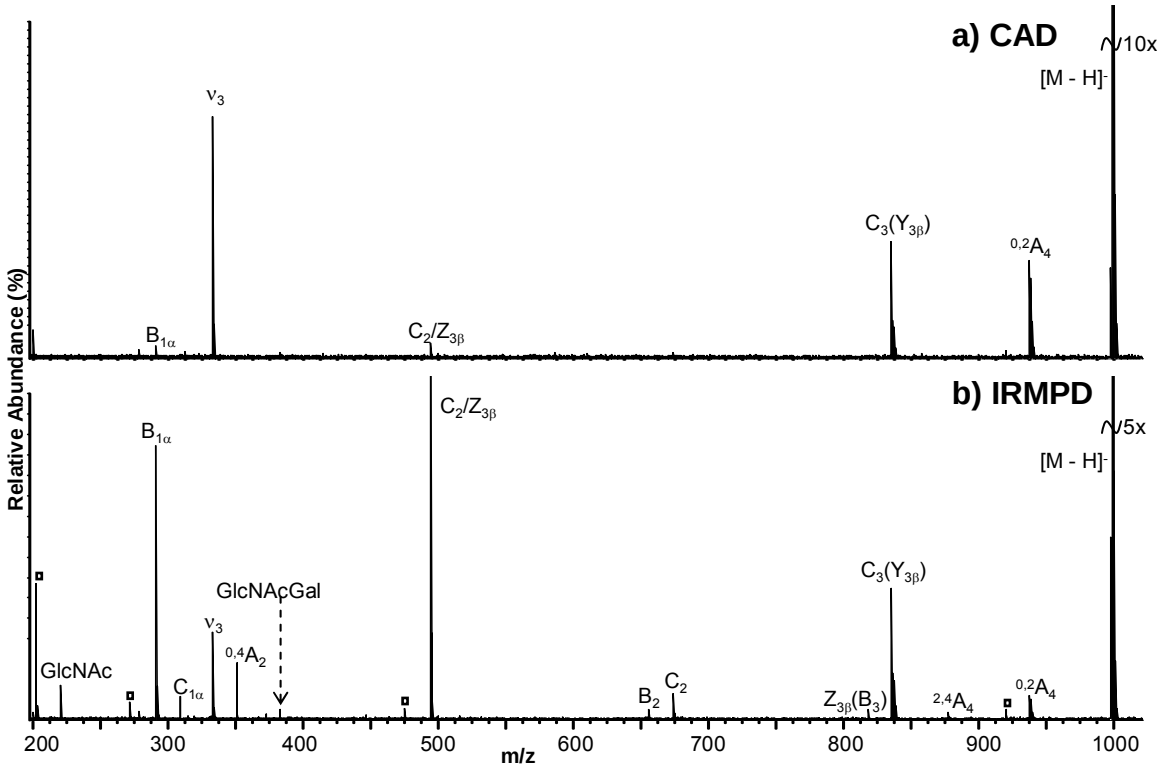


Figure 5.9. FT-ICR tandem mass spectra of singly deprotonated LSTb. a) CAD (30 scans, collision cell voltage 26 V) b) IRMPD (30 scans, 125 ms irradiation 10 W). Squares indicate water loss from an adjacent product ion. Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels.

CAD and IRMPD of a doubly (Figure 5.10) and singly deprotonated (Figure 5.11) disialylated oligosaccharide (DSLNT) were examined. As expected, the relative abundance of the doubly deprotonated form was always much larger than the singly deprotonated form. CAD and IRMPD of the singly deprotonated species resulted in less fragmentation than the doubly deprotonated species. In CAD and IRMPD of the doubly deprotonated species (Figure 5.10), both glycosidic and cross-ring cleavages were observed, similar to other oligosaccharides examined thus far. Both ^{0,2}A and ^{2,4}A-type cleavages were present in the reducing terminal saccharide. Similar to IRMPD and CAD, several product ions observed following EDD (Figure 5.12) are due to multiple cleavages. Neutral losses from the charge reduced species are also quite extensive,

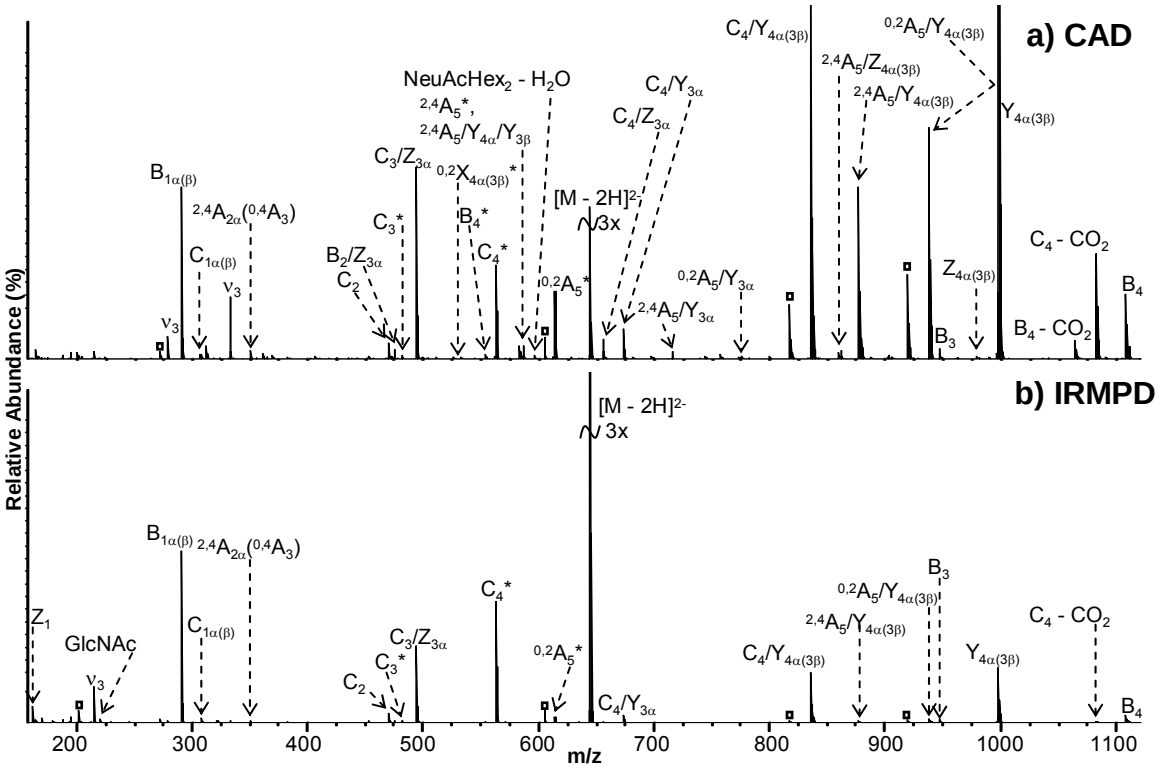


Figure 5.10. FT-ICR tandem mass spectra of doubly deprotonated DSLNT. a) CAD (30 scans, collision cell voltage 22 V) b) IRMPD (30 scans, 75 ms irradiation 10 W). Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Squares indicate water loss from an adjacent product ion. Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels.

with losses of CO_2 and CH_3O being the most predominant. In addition to a ^{0,2}A-type cleavage in the reducing terminal saccharide, several additional cross-ring cleavages are observed following EDD which were not present in CAD or IRMPD spectra. These correspond to ^{3,5}A, ^{1,5}A, ^{0,2}X, ^{4,5}X, and ^{1,5}X type product ions. One of the most abundant product ions observed following EDD is at *m*/*z* 977.34 ($Y_{3\alpha}(Y_{3\beta})$), due to loss of sialic acid. Another product ion is observed that is 2 Da lighter, and would appear to correspond to a loss of two hydrogens from this Y-type ion (or loss of one hydrogen from the radical Y-type species). Y - 2H product ions have also been observed following positive mode heCAD and are likely formed via a mechanism similar to C - 2H ions.[63]

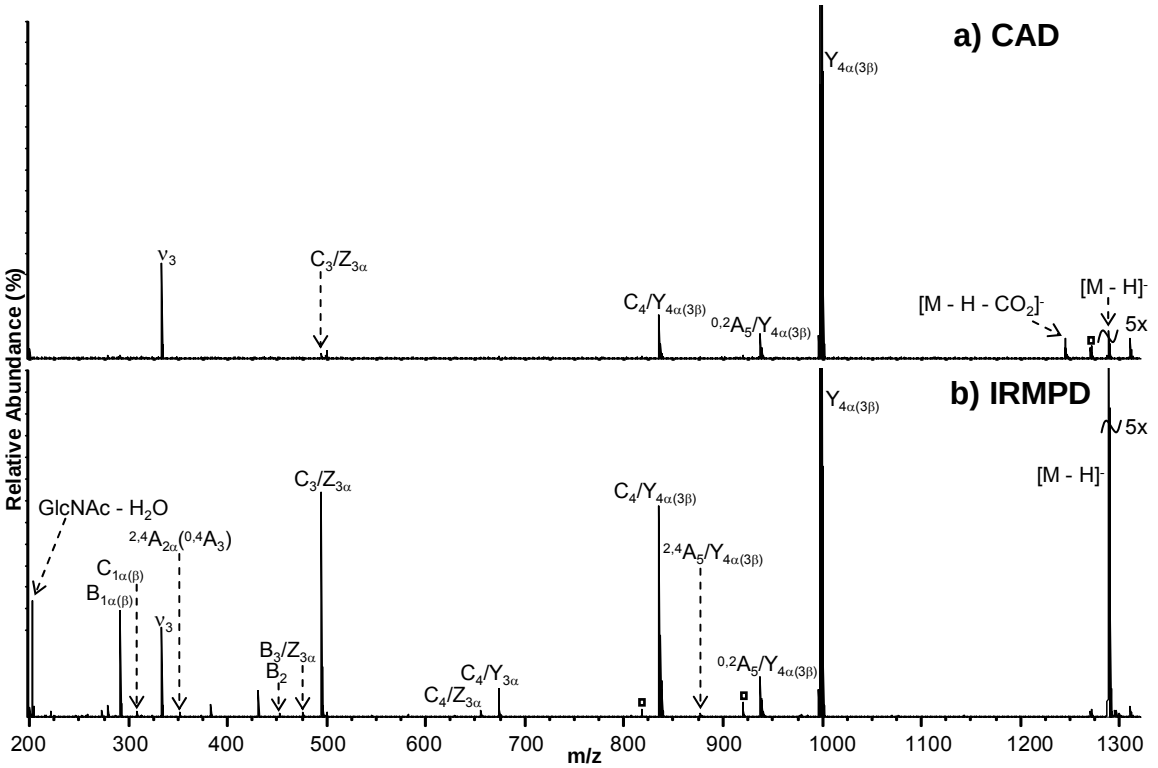


Figure 5.11. FT-ICR tandem mass spectra of singly deprotonated DSLNT. a) CAD (30 scans, collision cell voltage 34 V) b) IRMPD (30 scans, 75 ms irradiation 10 W). Squares indicate water loss from an adjacent product ion. Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels.

The product ion at *m/z* 833.27 has been assigned as a Y - 2H type product ion ($Y_{3\alpha}$ - 2H), but given that EDD of DSLNT resulted in significant internal cleavages, this ion could also be assigned C_4/Y_4 - 2H. Similar to other EDD spectra, C - 2H, Y - 16 Da, and B - H product ions are also seen following EDD of DSLNT. Many of these product ion types have been observed in heCAD, but only C - 2H and B - H ions have been noted following EDD of GAGs. However, the electron energies used in our examination were higher (20 - 30 eV) compared to those used in EDD of GAGs (19 eV), which may result in energy deposition more similar to that in heCAD.

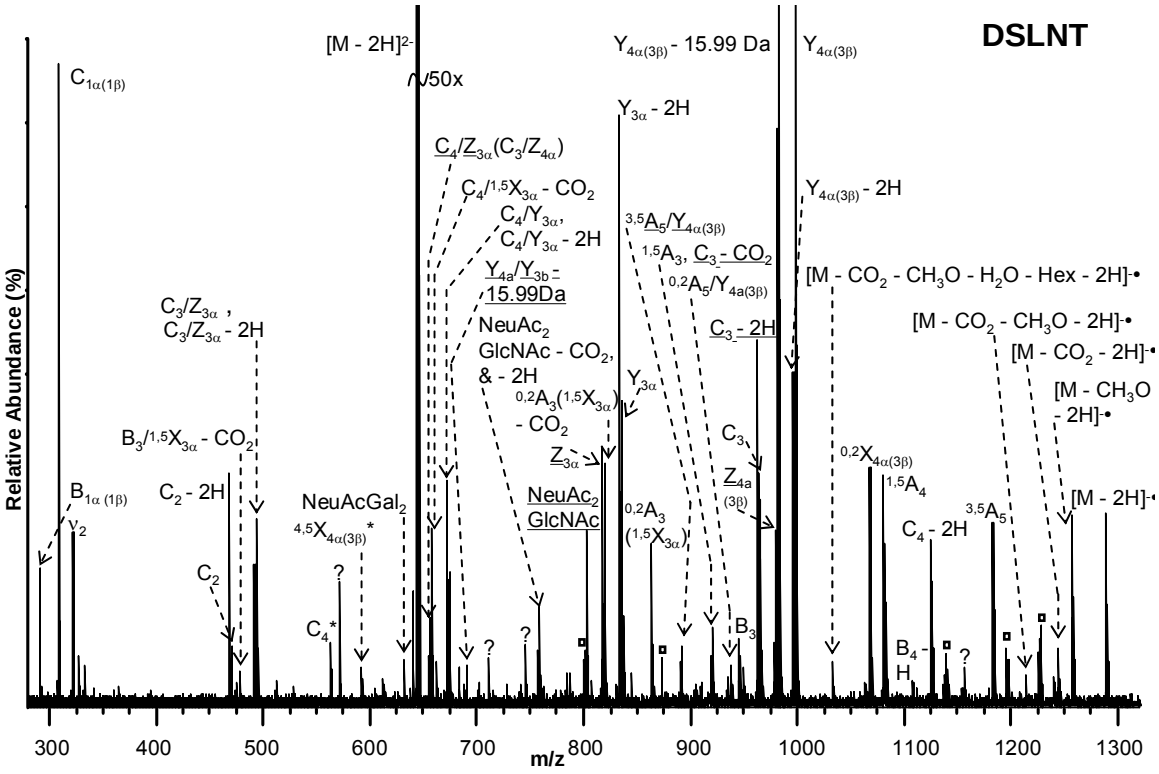
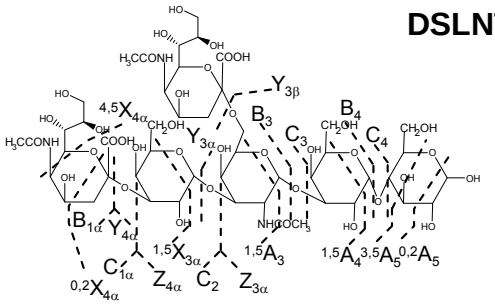


Figure 5.12. EDD FT-ICR tandem mass spectrum of doubly deprotonated DSLNT (30 scans, 2 s, bias voltage of – 20 V). Squares indicate water loss from an adjacent product ion. Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Underlined product ions exhibit minor radical species (with a mass corresponding to loss of hydrogen). Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels. The α-branch refers to the heavier branch, containing sialic acid and galactose.



Scheme 5.4. Fragmentation pattern observed following EDD of doubly deprotonated DSLNT.

The singly deprotonated sialylated N-linked glycan (A1F) was of relatively low abundance following negative mode electrospray ionization (<5%), and was therefore not examined. Using negative ion mode ESI conditions as described, the singly charged species had a relative abundance of less than 1%. CAD and IRMPD of the doubly deprotonated species (shown in Figure 5.13) resulted in several glycosidic cleavages, which corresponded mainly to C- and B-type ions (which may also be due to water loss from C-type ions). ^{0,2}A and ^{2,4}A-type cleavages were observed in both reducing terminal GlcNAc residues. These ions are useful for determining the presence of a core fucose residue. Fragmentation following EDD (Figure 5.14 and Scheme 5.5) was more extensive than either CAD or IRMPD, and resulted in several glycosidic and cross-ring product ions. The most abundant product ions correspond to C-type cleavages. In many instances the C - 2H species were more abundant than typical C-type product ions (for example C_6 - 2H ad C_5 - 2H). This behavior is similar to what was observed in EDD of DSLNT. Similar to other EDD spectra of neutral and sialylated oligosaccharides, EDD of A1F resulted in several internal cleavages. Several additional cross-ring cleavages were observed following EDD compared to CAD and IRMPD, including those corresponding to ^{1,5}X, ^{3,5}X, ^{4,5}X, ^{1,5}A, ^{0,2}A, and ^{2,4}A product ions. Two cross-ring fragments within the branching mannose residue were observed, including $^{1,5}A_5$ and $^{4,5}X_2$. The latter ion provides information regarding the linkage position of the α-branch. Similar to other EDD spectra, Y - 16 Da product ions are also observed. Also detected was an ion due to loss of a fucose residue, $Y_{1\gamma}$, which was not observed following CAD or IRMPD. In negative ion mode, it has been observed that fucose tends not to be lost following CAD,[23] a trend that is not observed in EDD. EDD of this branched glycan

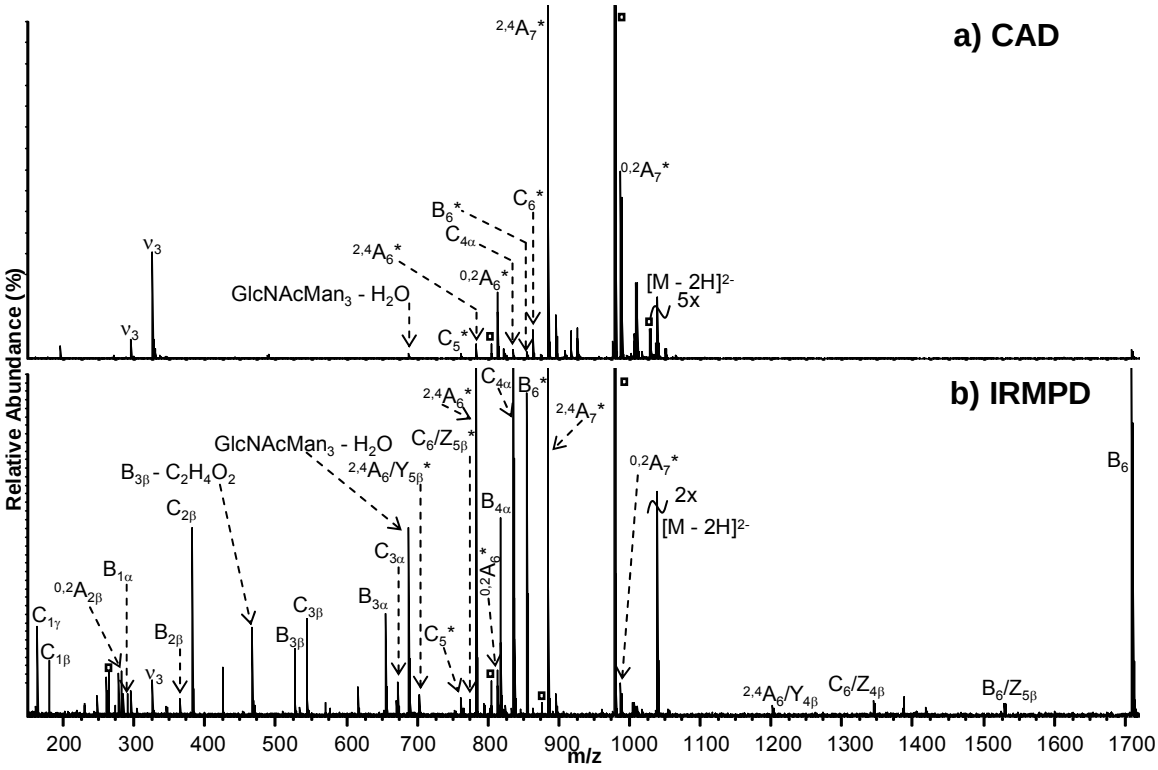


Figure 5.13. FT-ICR tandem mass spectra of doubly deprotonated A1F. a) CAD (30 scans, collision cell voltage 9 V) b) IRMPD (30 scans, 85 ms irradiation 10 W). Doubly deprotonated product ions are indicated with an asterisk next to a product ion assignment. Squares indicate water loss from an adjacent product ion. Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels.

demonstrates that EDD fragmentation efficiency does not seem to be affected by oligosaccharide branching. In Chapter 4, we showed that ECD of larger, branched oligosaccharides requires ion activation (presumably due to secondary structure) to fragment efficiently.[31] However, electron energies used for EDD (20 – 30 eV in the current experiments) are much higher than those typically used for ECD (< 1eV). These higher electron energies likely contribute to disruption of any intramolecular non-covalent interactions.

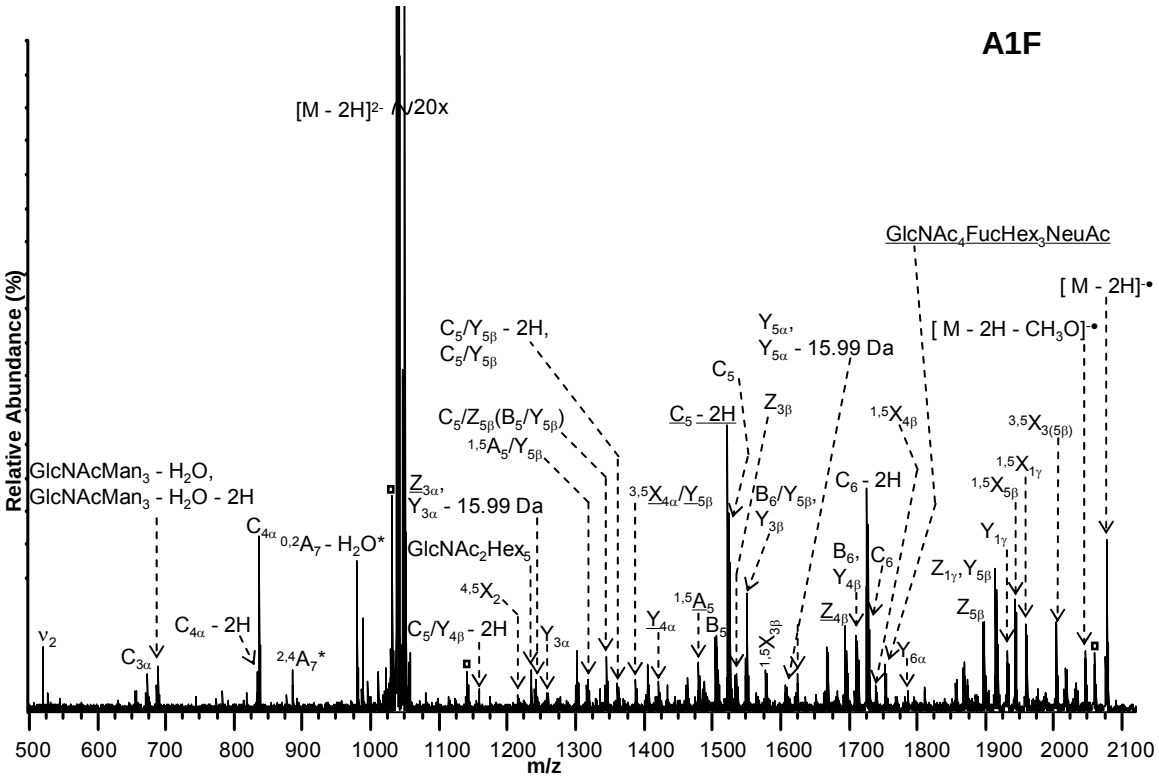
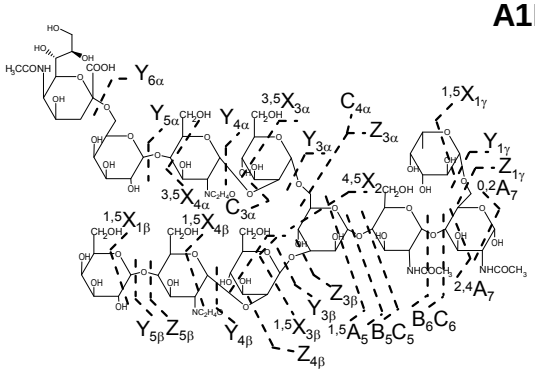


Figure 5.14. EDD FT-ICR tandem mass spectrum of doubly deprotonated A1F (30 scans, 2 s, bias voltage of – 30 V). Squares indicate water loss from an adjacent product ion. Underlined product ions exhibit minor radical species (with a mass corresponding to loss of hydrogen). Ions due to multiple cleavage sites are designated with a slash between sites of cleavage. Product ions which cannot be unambiguously assigned are indicated by parentheses in the labels.

A1F



Scheme 5.5. Fragmentation pattern observed following EDD of a doubly deprotonated A1F glycan. This glycan sample contains both 2→3 and 2→6 linked sialic acid, but only the 2→6 linkage is shown. The α-branch contains the sialic acid residue while the γ-branch contains fucose.

5.4 Conclusions

Our results demonstrate that EDD of neutral and sialylated oligosaccharides provides structural information that is complementary to that obtained following IRMPD and CAD. One potential issue with this approach is that, for neutral and singly sialylated oligosaccharides, doubly charged anions (which are required for EDD) are sometimes more difficult to obtain compared to singly charged anions. However, for both the neutral glycan (NA2) and the singly sialylated glycan (A1F) examined here, the doubly deprotonated species was more easily produced in negative ion mode ESI compared to the singly deprotonated species. These results indicate that EDD is a potentially valuable tool for the characterization of biologically relevant N-linked glycans, even if they are not highly acidic. Cross-ring cleavages observed following IRMPD and CAD of doubly and singly deprotonated oligosaccharides were typically either ^{0,2}A or ^{2,4}A type cleavages. These product ion types can also be observed following IRMPD and low energy CAD in positive ion mode MS/MS, however cross-ring fragments are generally more prevalent for metal adducted oligosaccharides compared to protonated species. Following EDD, several additional types of cross-ring cleavages were observed, such as ^{3,5}A, ^{1,5}A, ^{1,5}X, and ^{3,5}X type ions. These product ion types are usually not produced from either positive or negative ion mode IRMPD or low energy CAD. In most cases, cross-ring fragmentation in EDD was more extensive than in CAD or IRMPD. A comparison of the fragmentation patterns of branched and linear oligosaccharides demonstrates that branching does not seem to affect EDD fragmentation. Throughout these experiments, several unique product ions were observed. These include C - 2H, Y - 16Da, and Y - 2H

product ions. These ion types have all been observed following heCAD, but only C - 2H ions have been noted following EDD of GAGs.

5.5 References

1. Varki, A. *Glycobiol.* **1993**, 97-130.
2. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
3. Parodi, A. J. *Annu. Rev. Biochem.* **2000**, *69*, 60-93.
4. Rudd, P. M.; Elliott, T.; Cresswell, P.; Wilson, I. A.; Dwek, R. A. *Science* **2001**, *291*, 2370-2376.
5. Harvey, D. J. *Mass Spectrom. Rev.* **1999**, *18*, 349-451.
6. Zaia, J. *Mass Spectrom. Rev.* **2004**, *23*, 161-227.
7. Park, Y.; Lebrilla, C. B. *Mass Spectrom. Rev.* **2005**, *24*, 232-264.
8. Reinhold, V. N.; Reinhold, B. B.; Costello, C. E. *Anal. Chem.* **1995**, *67*, 1772-1784.
9. Weiskopf, A. S.; Vouros, P.; Harvey, D. J. *Rapid Commun. Mass Spectrom.* **1997**, *11*, 1493-1504.
10. Viseux, N.; de Hoffmann, E.; Domon, B. *Anal. Chem.* **1997**, *69*, 3139-3198.
11. Zhou, Z.; Ogden, S.; Leary, J. A. *J. Org. Chem.* **1990**, *55*, 5446-5448.
12. Orlando, R.; Bush, C. A.; Fenselau, C. *Biomed. Environ. Mass Spectrom.* **1990**, *19*, 747-754.
13. Fura, A.; Leary, J. A. *Anal. Chem.* **1993**, *65*, 2805-2811.
14. Cancilla, M. T.; Penn, S. G.; Carroll, J. A.; Lebrilla, C. B. *J. Am. Chem. Soc.* **1996**, *118*, 6736-6745.
15. Domon, B.; Costello, C. E. *Glycoconj. J.* **1988**, *5*, 397-409.
16. Dallinga, J.; Heerma, W. *Biol. Mass Spectrom.* **1991**, *20*, 215-231.
17. Li, D. T.; Her, G. R. *Anal. Biochem.* **1993**, *211*, 250-257.
18. Carroll, J. A.; Ngoka, L. C.; Beggs, C. G.; Lebrilla, C. B. *Anal. Chem.* **1993**, *65*, 1582-1587.
19. Chai, W.; Piskarev, V.; Lawson, A. M. *Anal. Chem.* **2001**, *73*, 651-657.
20. Pfenninger, A.; Karas, M.; Finke, B.; Stahl, B. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1331-1340.
21. Pfenninger, A.; Karas, M.; Finke, B.; Stahl, B. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1341-1348.
22. Chai, W.; Lawson, A. M.; Piskarev, V. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 670-679.
23. Harvey, D. J. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 622-630.
24. Seymour, J. L.; Costello, C. E.; Zaia, J. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 844-854.
25. Lemoine, J.; Strecker, G.; Leroy, Y.; Fournet, B.; Ricart, G. *Carbohydr. Res.* **1991**, *221*, 209-217.
26. Harvey, D. J.; Naven, T. J. P.; Kuster, B.; Bateman, R. H.; Green, M. R.; Critchley, G. *Rapid Commun. Mass Spectrom.* **1995**, *9*, 1556-1561.
27. Mechref, Y.; Novotny, M. V. *Anal. Chem.* **2003**, *75*, 4895-4903.
28. Shi, S. D.-H.; Hendrickson, C. L.; Marshall, A. G.; Seigel, M. M.; Kong, F.; Carter, G. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 1285-1290.
29. Xie, Y. M.; Lebrilla, C. B. *Anal. Chem.* **2003**, *75*, 1590-1598.
30. Budnik, B. A.; Haselmann, K. F.; Elkin, Y. N.; Gorbach, V. I.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 5994-6001.

31. Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 2901-2910.
32. Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. *J. Am. Soc. Mass Spectrom.* **2006**, *18*, 234-244.
33. Wolff, J. J.; Chi, L.; Linhardt, R. J.; Amster, I. J. *Anal. Chem.* **2007**, *79*, 2015-2022.
34. Devakumar, A.; Thompson, M. S.; Reilly, J. P. *Rapid Commun. Mass Spectrom.* **2005**, *19*, 2313-2320.
35. Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrom. Rev.* **2005**, *24*, 201-222.
36. Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.
37. Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.
38. Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 2355-2361.
39. Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
40. Mormann, M.; Macek, B.; de Peredo, A. G.; Hofsteenge, J.; Peter-Katalinic, J. *Int. J. Mass Spectrom.* **2004**, *234*, 11-21.
41. Adamson, J. T.; Hakansson, K. *J. Proteome Res.* **2006**, *5*, 493-501.
42. Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.* **1998**, *120*, 3265-3266.
43. Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *J. Am. Chem. Soc.* **1999**, *121*, 2857-2862.
44. Zhao, C.; Xie, B.; Chan, S. Y.; Costello, C. E.; O'Connor, P. B. *J. Am. Soc. Mass Spectrom.* **2007**, *19*, 138-150.
45. Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
46. Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803-1812.
47. Kalli, A.; Hakansson, K. *Int. J. Mass Spectrom.* **2007**, *263*, 71-81.
48. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2005**, *77*, 1876-1882.
49. Mo, J.; Hakansson, K. *Anal. Bioanal. Chem.* **2006**, *386*, 675-681.
50. McFarland, M. A.; Marshall, A. G.; Hendrickson, C. L.; Nilsson, C. L.; Fredman, P.; Mansson, J. E. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 752-762.
51. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2004**, *77*, 1876-1882.
52. Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26*, 514-518.
53. Heck, A. J. R.; Derrick, P. J. *Anal. Chem.* **1997**, *69*, 3603-3607.
54. Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.
55. Senko, M. W.; Canterbury, J. D.; Guan, S.; Marshall, A. G. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1839-1844.
56. Lohmann, K. K.; von der Lieth, C. W. *Nucleic Acids Res.* **2004**, *32*, 261-266.

57. Zhang, J. H.; Schubothe, K.; Li, B. S.; Russell, S.; Lebrilla, C. B. *Anal. Chem.* **2005**, *77*, 208-214.
58. Harvey, D. J.; Bateman, R. H.; Green, B. N. *J. Mass Spectrom.* **1997**, *32*, 167-187.
59. Kovacik, V.; Hirsch, J.; Kovac, P.; Heerma, W.; Thomas-Oates, J.; Haverkamp, J. *J. Mass Spectrom.* **1995**, *30*, 949-958.
60. O'Connor, P. B.; Lin, C.; Cournoyer, J. J.; Pittman, J. L.; Belyayev, M.; Budnik, B. A. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 576-585.
61. Savitski, M. M.; Kjeldsen, F.; Nielsen, M. L.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2006**, *18*, 113-120.
62. Küster, B.; Naven, T. J. P.; Harvey, D. J. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1645-1651.
63. Stephens, E.; Maslen, S. L.; Green, L. G.; Williams, D. H. *Anal. Chem.* **2004**, *76*, 2343-2354.
64. Lewandrowski, U.; Resemann, A.; Sickmann, A. *Anal. Chem.* **2005**, *77*, 3274-3283.
65. Yu, S.; Wu, S.; Khoo, K. *Glyconj. J.* **2006**, *23*, 355-369.
66. Harvey, D. J. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 647-659.
67. Hardy, M. R.; Townsend, R. R. *Methods Enzymol.* **1994**, *230*, 208-225.
68. Neuberger, A.; Wilson, B. M. *Carbohydr. Res.* **1971**, *17*, 89-95.
69. Ernst, B.; Müller, R. T.; Richter, W. J. *Int. J. Mass Spectrom. Ion Processes* **1997**, *160*, 283-290.
70. Brüll, L. P.; Heerma, W.; Thomas-Oates, J.; Haverkamp, J.; Kovácik, V.; Kovác, P. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 43-49.
71. Harvey, D. J.; Mattu, T. S.; Wormald, M. R.; Royle, L.; Dwek, R. A.; Rudd, P. M. *Anal. Chem.* **2002**, *74*, 734-740.

Chapter 6

Strategies for Characterizing Pancreatic Cancer Associated O-Linked Glycans

6.1 Introduction

Glycosylation is one of the most prevalent post-translational modifications (PTMs) in eukaryotes, and glycoproteins have been found to play a wide variety of roles in biological systems.[1,2] Protein glycan expression is highly affected by cellular conditions, and it has often been shown that glycan structures change with the onset of cancer and inflammation. Such alterations may include increased glycan branching and elevated levels of sialic acid.[3-6] In cancer, glycosylation alterations may affect growth, differentiation, transformation, adhesion, metastasis, and immune surveillance of the tumor.[7] Currently, many glycoproteins act as clinical biomarkers and therapeutic agents in several forms of cancer. For example, in ovarian cancer the CA 125 assay (a mucin-type glycoprotein) is used as a blood test for ovarian cancer.[8] Unfortunately, in many cases, no reliable diagnostic biomarkers exist for early disease detection.

Pancreatic cancer is one of the most frequent causes of cancer death in the United States and Europe.[9] This disease has a poor prognosis, with few treatment options available for patients. No reliable biomarkers have been identified that allow for early

detection. Currently, the serum glycoprotein marker CA 19-9 is the most commonly used biomarker for pancreatic cancer. However, this biomarker lacks sensitivity and specificity, and is not recommended for diagnostic purposes.[10] In addition, CA 19-9 is not useful for distinguishing between pancreatic cancer or a less severe case of chronic pancreatitis.

It has been known for several years that oligosaccharides on cancer cells differ from those on normal cells.[11-15] As disease markers, oligosaccharides are an attractive option because there are fewer potential oligosaccharide biomarkers compared to the total number of potential protein and peptide biomarkers. There are several reports which have examined the glycan profiles of serum from pancreatic cancer patients. In 2006, Okuyama and co-workers identified fucosylated haptoglobin as a novel marker for pancreatic cancer.[16] In addition, Lubman and co-workers have developed a screening approach for comprehensive analysis of *N*-glycans and glycosylation sites in serum proteins. They reported forty-four N-linked glycans which were unique to pancreatic cancer serum.[17]

On the surface of cancer cells, mucin glycoproteins (a family of glycoproteins, often highly O-glycosylated) are typically over-expressed and also aberrantly glycosylated. These glycoproteins are often shed into intracellular space. Recently, Lebrilla and co-workers have examined O-linked oligosaccharides in conditioned media from ovarian cancer cells and compared them to oligosaccharides in serum from ovarian cancer patients.[8] Conditioned media (the supernatant) of various ovarian cancer cell lines were extracted and examined without the cells. Due to the presence of shed glycoproteins in the growth media, many of the same oligosaccharides were observed in

both media and serum samples. These O-linked oligosaccharides have the potential to be used as markers for ovarian cancer. Although no protein information is obtained through this type of analysis, their approach is simpler than monitoring the presence of intact glycoproteins.

Here, we present a strategy for the characterization of O-linked glycans shed from pancreatic cancer cells into growth media. Given the success of the recent studies done by Lebrilla and co-workers on ovarian cancer and breast cancer,[18] it may be possible to find diagnostic glycans for pancreatic cancer in conditioned media. Although sera *N*-glycan profiles of pancreatic cancer[5,17,19] have been recently characterized, these studies focused on overall glycan composition (not specific structural information). In addition, there has been very little work done on the analysis of O-linked glycans associated with this disease. Here, we apply an approach similar to that used by Lebrilla and co-workers towards pancreatic cancer. One major difference, however, is that in their experiments a matrix-assisted laser desorption/ionization (MALDI) source was used for the generation of gas-phase glycan ions. In the work presented here, an electrospray ionization (ESI) source is utilized. ESI has a lower tolerance to salt contamination compared to MALDI, and glycan ionization efficiency is not as high. However, native sialylated oligosaccharides and sulfated oligosaccharides are particularly fragile in MALDI and may undergo in-source fragmentation, whereas ESI tends to be a gentler ionization source. Another advantage of ESI is its ability to generate multiply charged ions, allowing for fragmentation techniques such as electron capture dissociation (ECD) and electron detachment dissociation (EDD) to be utilized. To examine the O-linked oligosaccharides present in conditioned media, several clean-up strategies were explored

with a model glycoprotein, bovine fetuin. Discussions of these strategies and the oligosaccharides observed from the conditioned media of two cell lines, BxPC-3 (from pancreatic adenocarcinoma) and HDPE (from normal pancreatic ductal epithelial cells), are presented.

6.2 Experimental

6.2.1 Bovine Fetuin Oligosaccharides Sample Preparation

Between 1-10 nmoles of bovine fetuin (Sigma, St Louis, MI) was treated for approximately 15 hours at 42°C with 500 μL of alkaline borohydride solution (a 1:1 mixture of 1M $NaBH_4$ and 0.1 M NaOH).[8] After the reaction, acetic acid was slowly added to each sample on ice until a pH of 6-7 was achieved. These samples were then treated via three different strategies; a) samples were desalted with graphitized carbon cartridges (Alltech, Deerfield, IL); b) Pepclean C-18 Spin Columns (Pierce, Rockford, IL) were used and samples were further desalted with graphitized carbon cartridges; or c) phenol and chloroform extraction was performed and samples were further desalted with graphitized carbon cartridges.

Graphitized carbon solid-phase extraction (SPE) cartridges were activated with 5 mL of 0.1% formic acid in 80% acetonitrile/water (v/v) and washed with 5 mL of water. The solution of released oligosaccharides was loaded onto the cartridge and washed with 10 mL of water. Oligosaccharides were eluted with 1.5 mL of 0.5% formic acid in 40% acetonitrile/water (v/v) and concentrated in vacuo. Borate was removed with repeated addition and evaporation (50 μL each) of 1% acetic acid in methanol.[20] Prior to mass

spectrometric analysis, 200 μL of negative ion mode electrospray solution was added to each sample (50% methanol with 0.1% ammonium hydroxide).

Pepclean C-18 Spin Columns (Pierce, Rockford, IL) were equilibrated with and washed according to the manufacturers recommendations. Sample was loaded and the flow-through was collected. The flow-through was dried in vacuo, reconstituted in 400 μl of water, and further desalted with a graphitized carbon cartridge as previously described.

For phenol/chloroform extraction, 500 μl of phenol/chloroform/isoamyl alcohol (25:24:1) was added to samples of released sugars. The samples were vortexed and centrifuged, and the aqueous layer was retained. This procedure was repeated until precipitate was no longer observed (2-3 times). An equal volume of chloroform was added to the aqueous layer, followed by vortexing and centrifugation. The aqueous layer was removed and further desalted with a graphitized carbon cartridge as described previously.

6.2.2 Conditioned Media Oligosaccharides Sample Preparation

Conditioned media from BxPC-3 and HDPE cell lines was provided by the laboratory of Dr. Diane Simeone (Department of Surgery, University of Michigan).[21] The BxPC-3 cell line is from a pancreatic adenocarcinoma, and was cultured in RPMI-1640 media supplemented with 10% fetal bovine serum. The HDPE cell line is from normal pancreatic ductal epithelial cells and is cultured in Dulbecco's Modified Eagle's Medium (DMEM) media without fetal bovine serum. RPMI-1640 media supplemented with fetal bovine serum and DMEM media were used as control samples. Conditioned media and control samples were sterile filtered (0.2 μM syringe filters) and frozen at

-80°C. Samples were thawed and 12 mL of each sample was concentrated to approximately 500 μL with Vivaspin 20 (5,000 MW cutoff) concentrators (VivaScience, Edgewood, NY). Concentrated solutions of $NaBH_4$ and NaOH were added to each sample to obtain a final concentration of 0.5 M NaBH4 and 0.05 M NaOH. Beta-elimination proceeded for approximately 15 hours at 42°C in a water bath.

Following oligosaccharide release, samples were placed on ice and adjusted to pH 6-7 with acetic acid and then desalted using a combination of C-18 SPE and graphitized carbon SPE. Sep-Pack C18 cartridges (Waters, Milford, MA) were activated with 5 mL 50% acetonitrile/water (v/v) and equilibrated with 5 mL of 0.1% formic acid in 5% acetonitrile/water (v/v). Samples were loaded onto the cartridge and the flow-through was retained. The cartridge was washed with an additional 4 mL of water and this flow-through was also retained and pooled. The flow-through was further desalted with a graphitized carbon cartridge. The procedure is as previously described, except oligosaccharides were stepwise eluted with 0.5 mL 10% acetonitrile in water, 0.5 mL 20% acetonitrile in water, and 0.5 mL 40% acetonitrile/0.05% formic acid in water. Samples were dried in vacuo, and borate was removed as previously described. Samples were reconstituted in 100 μl of water. For direct infusion mass spectrometric analysis, 10 μl aliquots were diluted 4-12 fold in negative ion mode electrospray solution. For liquid chromatography-mass spectrometry, a 50 μl aliquot was dried in vacuo and reconstituted in 5 μl of 95% acetonitrile/water (v/v) and injected onto the column.

6.2.3 Liquid Chromatography-Mass Spectrometry

Oligosaccharides were separated on a Tosoh TSK-GEL Amide-80 column (5 μm; 100 Å; 1 mm x 25 cm; Montgomeryville, PA) with an Agilent 1100 HPLC. Solvent A

was acetonitrile with 2.5 mM $NH_4C_2H_3O_2$ and solvent B was water with 2.5 mM $NH_4C_2H_3O_2$. The following gradient was used: t = 0 min, 0% solvent B; t = 10 min, 30% solvent B; t = 90 min, 35% solvent B, t = 95 min, 95% solvent B; t = 115 min, 95% solvent B. The flow rate was 30 μL/min (with the use of a pre-column flow splitter). The HPLC was directly coupled to the mass spectrometer via an Apollo II ion source.

6.2.4 FT-ICR Mass Spectrometry

All experiments were performed with an actively shielded 7 T FT-ICR mass spectrometer with a quadrupole front-end (APEX-Q, Bruker Daltonics), as previously described.[22] Bovine fetuin oligosaccharide samples were electrosprayed with an Apollo II ion source at 60 μL/hour, while oligosaccharides from media samples were nanosprayed at 15 μl/hr with 30 μm i.d. fused-silica PicoTips (New Objective, Woburn, MA). External CAD was performed in a hexapole following mass selective ion accumulation.

Mass spectra were acquired with XMASS software (Bruker Daltonics) with 256k data points and summed over 10 scans (direct infusion experiments). LC-MS data were acquired with Bruker Hystar software. LC-MS and MS data were analyzed with Bruker Data Analysis software. Potential oligosaccharide structures in conditioned media samples were determined using Glycomod (http://www.expasy.ch/tools/glycomod).[23]

6.3 Results and Discussion

6.3.1 Protocol Optimization with Bovine Fetuin

Bovine fetuin was used as a model protein to optimize β-elimination and desalting of O-linked oligosaccharides. O-linked glycans account for 20% of fetuin-

oligosaccharides, while N-linked glycans comprise the remaining 80%.[24] Bovine fetuin has three O-linked glycans, including a trisaccharide NeuAc(α-2-3)Gal(β1-3)GalNAc, a tetrasaccharide NeuAc(α2-3)Gal(β1-3)[NeuAc(α2-6)]GalNAc, and a hexasaccharide NeuAc(α2-3)Gal(β1-4)GlcNAc(β1-6)[NeuAc(α 2-3)Gal(β1-3)].[25] The N-linked glycans of bovine fetuin consist of a wide variety of di- and tri-branched sialylated structures. Although it is often stated that harsher alkali conditions (1M NaOH) are necessary during β-elimination to observe the release of N-linked glycans, in these experiments a wide variety of N-linked glycans and N-linked glycopeptides were observed from bovine fetuin with only 0.05 M NaOH present during β-elimination. Similar instances of N-glycan release from β-elimination have been previously reported.[20,26,27]

Several strategies were explored for the purification and desalting of oligosaccharides released from bovine fetuin. As an initial strategy, samples were desalted with graphitized carbon SPE cartridges. Graphitized carbon is commonly used for the chromatographic separation of oligosaccharides, but is also useful for the isolation and clean-up of mixtures of oligosaccharides. It has been shown that graphitized carbon SPE can separate oligosaccharides from contaminants such as salts, detergents, proteins, as well as reagents commonly used during β-elimination (such as sodium borohydride).[28] In addition, neutral oligosaccharides can be fractionated from acidic oligosaccharides (such as sialylated species) by adjusting elution conditions. The mass spectrum of a bovine fetuin sample desalted with graphitized carbon SPE following β-elimination is shown in Figure 6.1a. One O-linked glycan was observed, a tetrasaccharide at *m/z* 675. However, the majority of ions shown in Figure 6.1a are singly charged species, and are likely due to remaining salts and β-elimination reagents (NaOH and NaBH4) as well as

bovine fetuin peptides. The recent work done by Lebrilla and co-workers utilizing a similar approach was done with a MALDI ionization source. As previously discussed, MALDI is more tolerant to salt contamination than ESI, thus this was not a major obstacle in their work.

To remove peptides and protein remaining following β-elimination, two strategies were explored. First, C18 spin-columns were utilized for peptide/protein removal prior to desalting with graphitized carbon SPE cartridges. A similar strategy has been utilized for sample cleanup of human serum samples.[29] C18 resin shows excellent retention of hydrophobic species such as proteins and peptides, but has less affinity for hydrophilic species such as oligosaccharides. Figure 6.1b shows the spectrum of bovine fetuin oligosaccharides observed following C18 spin-column treatment and graphitized carbon SPE. All three O-linked oligosaccharides are observed as either singly or doubly deprotonated species at *m/z* 482 (doubly deprotonated tetrasaccharide), 665 (doubly deprotonated hexasaccharide), 675 (singly deprotonated trisaccharide), and at 966 (singly deprotonated tetrasaccharide). In addition, several abundant N-linked glycans and N-linked glycopeptides are observed. Similar results have been previously reported.[20] It is also interesting to note that in Figure 6.1a, 10 nmoles of bovine fetuin were subjected to β-elimination and only resulted in the observation of one O-linked glycan, whereas in Figures 1b and 1c only 1 nmole of protein was utilized. These results demonstrate the necessity of protein and peptide removal.

As an alternative strategy for protein and peptide removal, a phenol/chloroform extraction was utilized prior to graphitized carbon SPE. This technique is commonly used for purifying DNA contaminated by proteins, however there are very few reports of

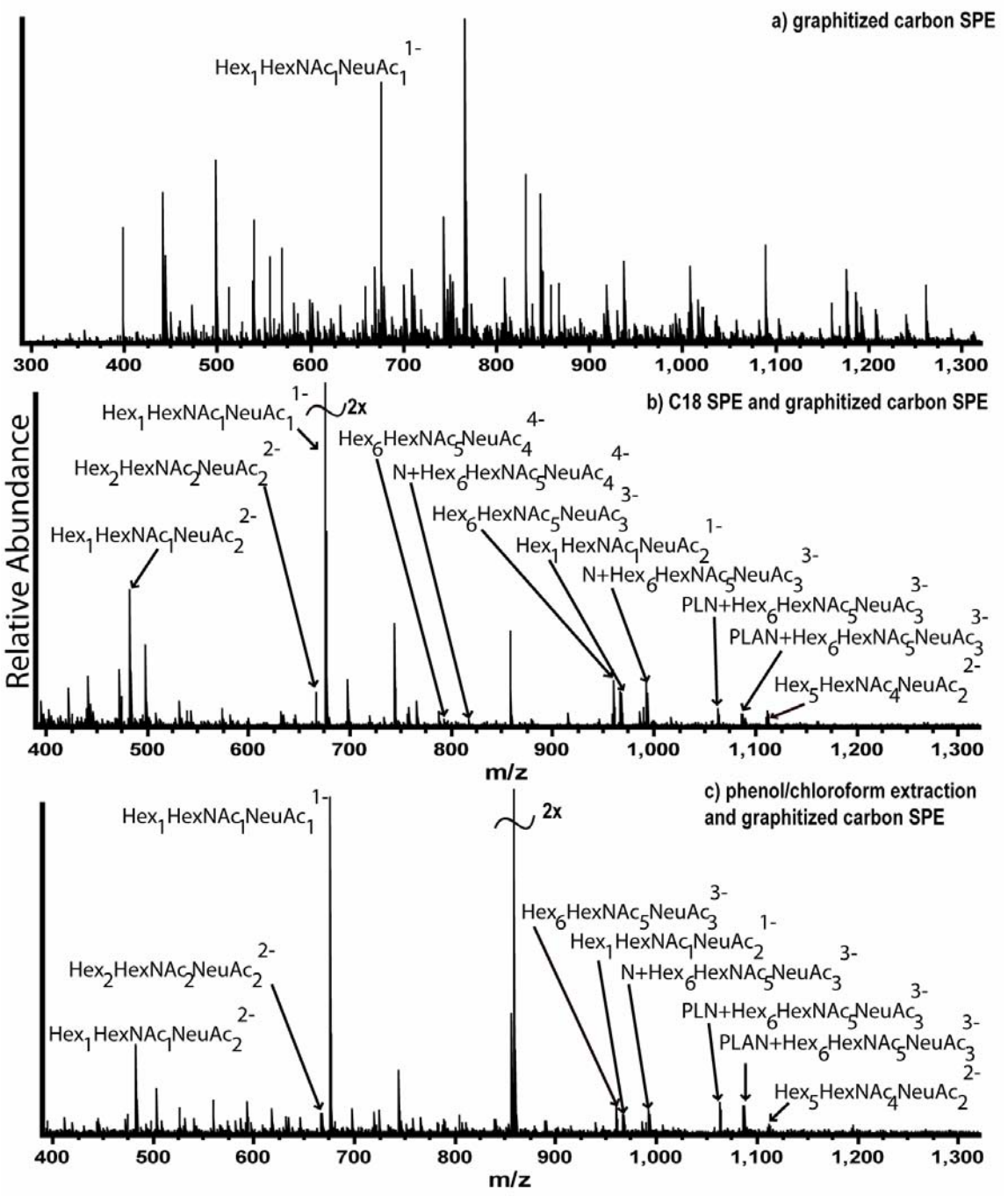


Figure 6.1. FT-ICR mass spectra (10 scans) of bovine fetuin oligosaccharides released via β-elimination. a) Bovine fetuin samples desalted with graphitized carbon SPE. b) Bovine fetuin samples following peptide/protein removal with C18 SPE and desalting with graphitized carbon SPE. c) Bovine fetuin samples following peptide/protein removal with phenol/chloroform extraction and desalting with graphitized carbon SPE.

this strategy for the clean-up of oligosaccharides released from glycoproteins.[30] Figure 6.1c shows the glycan profile of bovine fetuin following phenol/chloroform extraction and desalting with graphitized carbon SPE. All three O-linked glycans are observed, as well as several N-linked glycans and N-linked glycopeptides. The most abundant ion in Figure 6.1c is at *m/z* 857.1, and could not be identified. MS/MS of this ion (via collision activated dissociation (CAD)) did not indicate this species was an oligosaccharide. This ion was often one of the more abundant peaks observed following phenol/chloroform extraction and graphitized carbon SPE of bovine fetuin samples. This species was not observed in Figure 6.1b (although an abundant ion of similar mass at *m/z* 857.4 was present). These results indicate that the combination of C18 SPE and graphitized carbon SPE provided the best clean-up strategy for oligosaccharides released from bovine fetuin. A similar strategy was utilized for sample cleanup of conditioned media oligosaccharides.

6.3.2 Direct Infusion of Conditioned Media Oligosaccharides

Conditioned media from two pancreatic cell lines, BxPC-3 and HDPE, were examined. HDPE is from normal pancreatic ductal epithelial cells whereas the source of the BxPC-3 cell line is pancreatic adenocarcinoma. One of the reported cellular products of the latter cell line is mucins. Mucins are a family of glycoproteins that help maintain the integrity of, lubricate, and protect epithelial surfaces. Alterations in mucin glycosylation have been observed in several forms of cancer.[12,13,31] In pancreatic cancer, several reports have indicated that many mucins demonstrate over-expression and aberrant glycosylation. [32-37]

The protocol optimized with bovine fetuin was applied to oligosaccharides in conditioned media from the BxPC-3 and HDPE cell lines. Following β-elimination,

samples were treated with C18 SPE cartridges and further desalted with graphitized carbon SPE. Samples were stepwise eluted from the graphitized carbon SPE cartridge with 10% acetonitrile, 20% acetonitrile, and 40% acetonitrile. Lebrilla and co-workers utilized a similar protocol, and observed that small anionic oligosaccharides were observed in fractions eluted with a lower acetonitrile concentration, while larger oligosaccharides were observed in the 40% acetonitrile fraction.[8]

Figure 6.2 shows the mass spectra of two graphitized carbon SPE fractions (10% and 20% acetonitrile) from the BxPC-3 cell line. The fraction eluted with 40% acetonitrile (data not shown) appeared similar to the fraction eluted with 20% acetonitrile. Spectra were deconvoluted, and ion masses were submitted to the online software tool GlycoMod, in order to predict potential oligosaccharide structures.[23] Potential oligosaccharide species are indicated with black circles in Figure 6.2. Potential oligosaccharides with structures that have been reported in the literature (and have been added to the GlycoSuiteDB database) are indicated with asterisks in Figure 6.2. All potential oligosaccharides in Figure 6.2a (top) correspond to O-linked oligosaccharides, except for the ion at *m/z* 1519.567, which may correspond to an N-linked structure. In Figure 6.2b (bottom), many of the potential oligosaccharides observed would correspond to N-linked glycans. These results are consistent with the behavior of bovine fetuin, which demonstrated that the β-elimination procedure was able to release both O- and N-linked oligosaccharides.

To verify that these ions are oligosaccharides, tandem mass spectrometry was necessary. In the 10% acetonitrile fraction, ions at *m/z* 675.248, 966.342, 1351.501, 1533.597, and 1644.588 were fragmented with CAD and, in the 20% fraction, ions at *m/z*

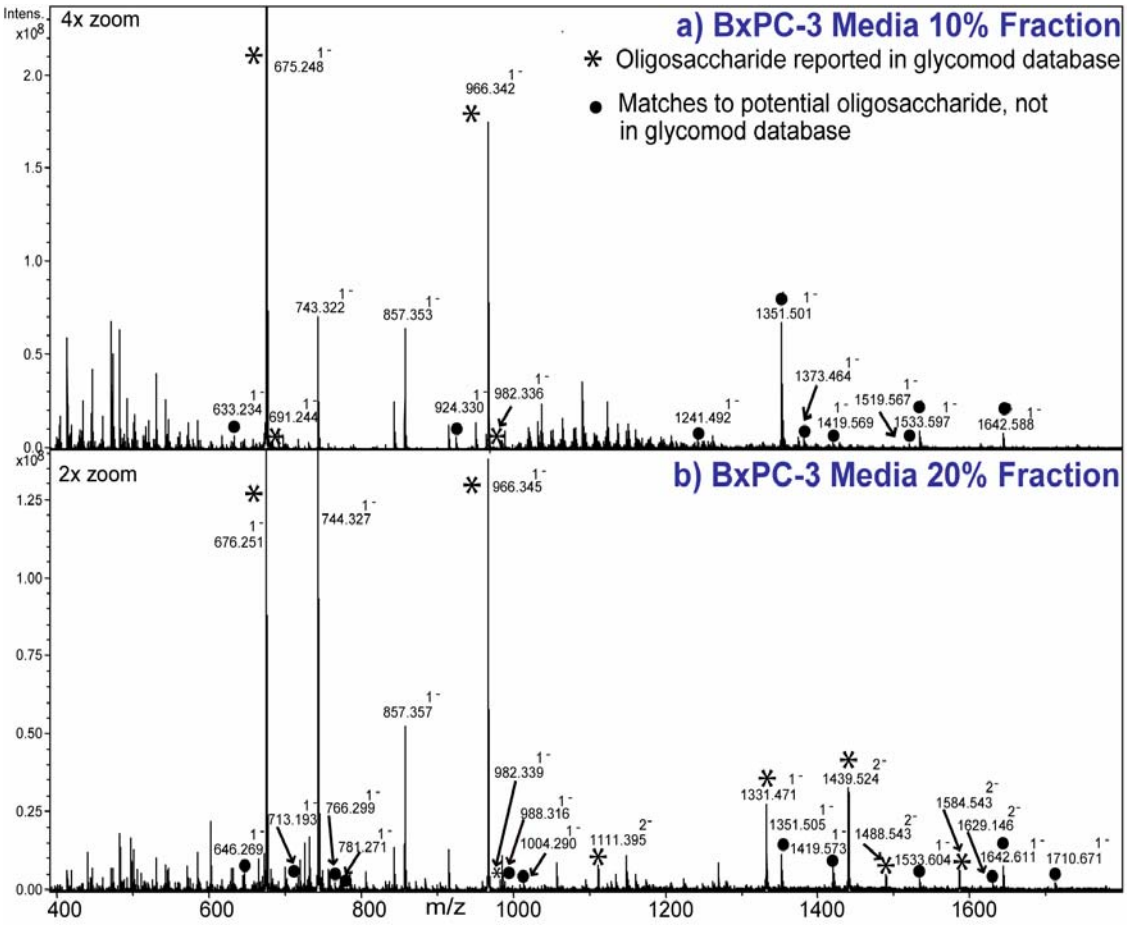


Figure 6.2. Negative ion mode FT-ICR mass spectra (10 scans) of BxPC-3 conditioned media oligosaccharides released via β-elimination. a) Fraction eluted from graphitized carbon SPE with 10% acetonitrile/water b) Fraction eluted from graphitized carbon SPE with 20% acetonitrile/water. Neutral masses were submitted to GlycoMod, to identify potential oligosaccharides. Black dots indicate potential oligosaccharides, while asterisks indicate potential oligosaccharides with structures that have been reported in the GlycoSuiteDB database.

1111.395, 1331.472, 1419.573, 1439.019, 1488.543, 1584.543, and 1639.146 were fragmented. CAD spectra of all ions exhibited characteristic oligosaccharide fragmentation patterns (glycosidic cleavages were almost exclusively observed). Several of the species observed in Figure 6.2 were assigned as potential oligosaccharides, but with structures that had not been reported in the GlycoSuiteDB database. CAD fragmentation patterns indicated that some of these species may actually be due to

dimerization in the gas-phase, as was previously reported in Chapter 4 for metal-adduced oligosaccharides (although in this instance, there is no evidence to indicate metal adduction). As an example, the CAD spectra of the most abundant ions in Figure 6.2a (at *m/z* 675.248, 966.342, and 1351.501) are shown in Figure 6.3. The former two ions may be assigned as a sialylated trisaccharide, $(Hex)_1(HexNAc)_1(NeuAc)_1$ and a disialylated tetrasaccharide, $(Hex)_1(HexNAc)_1(NeuAc)_2$, respectively. Following CAD, loss of sialic acid is the predominant fragmentation pathway for both of these ions (Figure 6.3a and 6.3b). The ion at *m/z* 1351.501 has two potential oligosaccharide structures, either $(Hex)_4$ $(HexNAc)_2$ $(Deoxyhexose)_1$ $(Pent)_1$, or $(Hex)_2$ $(Deoxyhexose)_6$ $(Pent)_1$. However, CAD of this species disproves both of these assignments, and indicates this oligosaccharide is a sialylated species. Considering the possibility of dimerization in the gas-phase, a possible assignment for this species is a dimer consisting of two $(Hex)_1$ $(HexNAc)_1$ $(NeuAc)_1$ oligosaccharides. The ion at *m/z* 1351.501 was fragmented with a much lower energy compared to the other oligosaccharide species shown in Figure 6.3 (collision voltage of 10 V versus 18-20 V). In addition, one of the major product ions observed in Figure 6.3c is a trisaccharide at *m/z* 675.259. This sialylated trisaccharide is the most abundant species observed in the 10% acetonitrile fraction (at *m/z* 675.248 in Figure 6.2a). Its relatively high concentration could explain the presence of a gas-phase dimer.

As a control sample, RPMI-1640 media supplemented with 10% fetal bovine serum was subjected to the same treatment as the BxPC-3 conditioned media samples and examined. All potential oligosaccharides observed in the BxPC-3 conditioned media sample were also present in the control sample. These results were surprising, because Lebrilla and co-workers also used conditioned media from cell lines that were

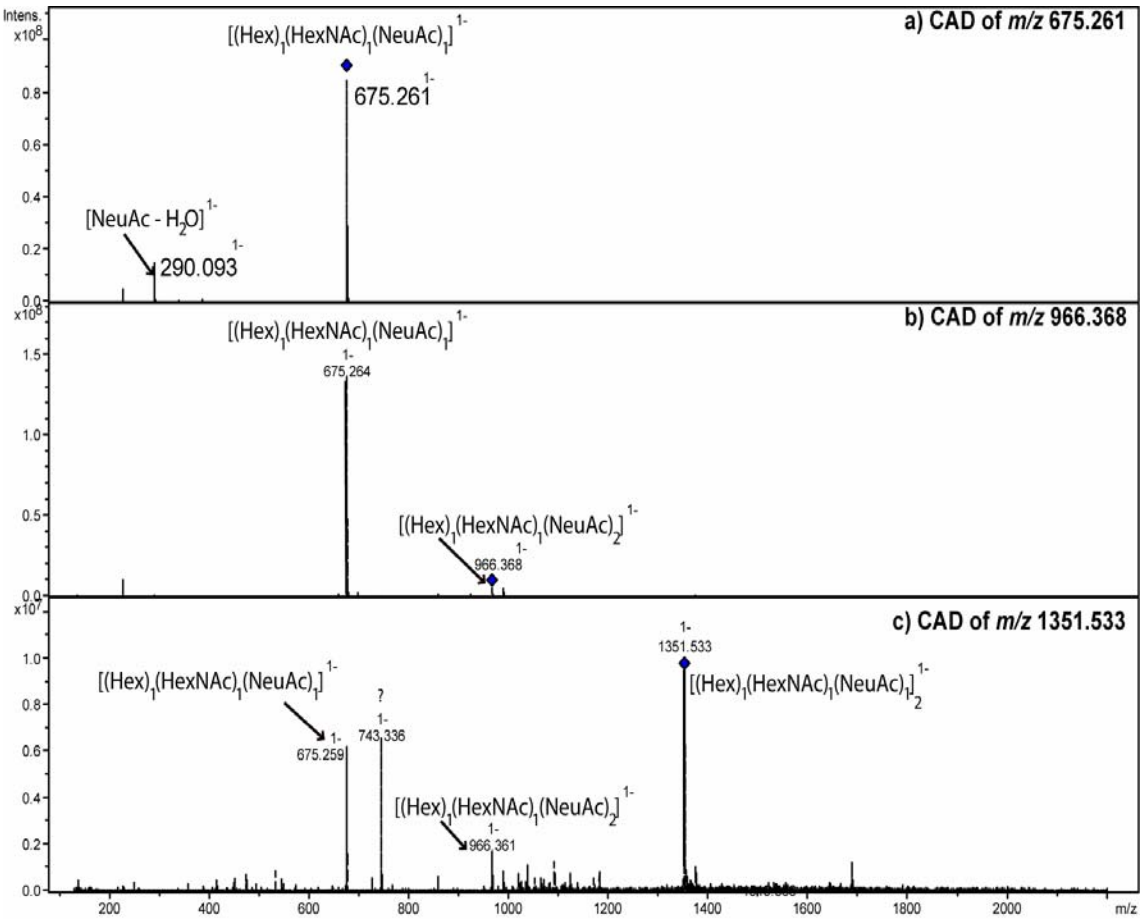


Figure 6.3. CAD FT-ICR tandem mass spectra (10 scans) of several oligosaccharides observed in the 10% elution fraction from BxPC-3 conditioned media. a) CAD spectrum (collision voltage 18 V) of an oligosaccharide with the potential structure $(Hex)_1(HexNAc)_1(NeuAc)_1$ b) CAD spectrum (collision voltage 20 V) of an oligosaccharide with the potential structure $(Hex)_1(HexNAc)_1(NeuAc)_2$ c) CAD spectrum (collision voltage 10 V) of a potential oligosaccharide dimer consisting of two $(Hex)_1(HexNAc)_1(NeuAc)_1$ trisaccharides. A black diamond indicates the parent ion in each CAD spectrum.

supplemented with 10% fetal bovine serum. They reported very few oligosaccharides originating from fetal bovine serum, whereas our results indicate that a plethora of O-linked and N-linked oligosaccharides are present. To determine whether there are any low-abundance oligosaccharides unique to conditioned media from the BxPC-3 cell line, liquid chromatographic separation was also utilized (see section 6.3.3). Due to ionization

suppression by fetal bovine serum oligosaccharides, it may be difficult to observe pancreatic cancer associated oligosaccharides in a direct infusion experiment.

HDPE conditioned media was also examined. Several potential oligosaccharides were observed, and several of these had structures which are reported in the GlycoSuiteDB database (data not shown). However, CAD did not indicate any of these species are oligosaccharides. As a control sample, DMEM media was also subjected to the same treatment as conditioned media. Similar to the BxPC-3 conditioned media, the majority of potential oligosaccharides observed in the HDPE conditioned media samples were also present in the control DMEM media. DMEM media is not supplemented with fetal bovine serum, or any other glycoprotein sources. Unlike the BxPC-3 cell line, the HDPE cell line is from normal pancreatic ductal epithelial cells, which should not exhibit overexpression of mucins. It may therefore be more difficult to observe oligosaccharides in HDPE conditioned media compared to pancreatic cancer tumor cell lines.

6.3.3 Liquid-Chromatography Mass Spectrometry of Oligosaccharides from BxPC-3 Conditioned Media

Online-LC-ESI-MS characterization of underivatized oligosaccharides is often performed using graphitized carbon chromatography.[20,38-41] However, Wuhrer and co-workers have demonstrated that normal-phase liquid chromatography of underivatized oligosaccharides is also a useful separation tool for complex oligosaccharide mixtures.[42] Their results demonstrated that normal-phase chromatography is a powerful alternative to graphitized carbon chromatography for LC-MS of oligosaccharides.

Oligosaccharides released from conditioned media from the BxPC-3 cell line were separated with normal-phase liquid chromatography, which was directly coupled to

an FT-ICR mass spectrometer. Figure 6.4 shows the total ion chromatogram of conditioned media oligosaccharides from the BxPC-3 cell line. The majority of oligosaccharides elute between 70-100 minutes. Retention of oligosaccharides is caused by polar interactions, and typically increases as glycan chain length increases. Thus, smaller O-linked oligosaccharides elute earlier while larger N-linked glycans elute later. A list of potential oligosaccharides observed throughout the LC-MS run is shown in Table 6.1 at the end of this Chapter. Only potential oligosaccharides that have been reported in the GlycoSuiteDB database are indicated in this Table, which lists a mixture of both O-linked and N-linked glycans, as well as some glycopeptides that may have a single asparagine attached to the glycan portion (which was also seen with bovine fetuin samples).

As a control sample, oligosaccharides released from RMPI-1640 media supplemented with fetal bovine serum were also separated with normal phase LC/FT-ICR MS. All potential oligosaccharides shown in Table 6.1 from conditioned media samples were also observed following chromatographic separation of the control sample. These results indicate that with both a direct infusion experiment, and with an online liquid chromatographic separation, it is extremely difficult to identify oligosaccharides specifically associated with the BxPC-3 cell line due to the presence of fetal bovine serum in the media.

6.4 Conclusions

An optimal strategy for examining O-linked oligosaccharides from conditioned media from pancreatic cell lines has been developed. This strategy is based upon the treatment of released oligosaccharides with reversed-phase SPE and desalting with

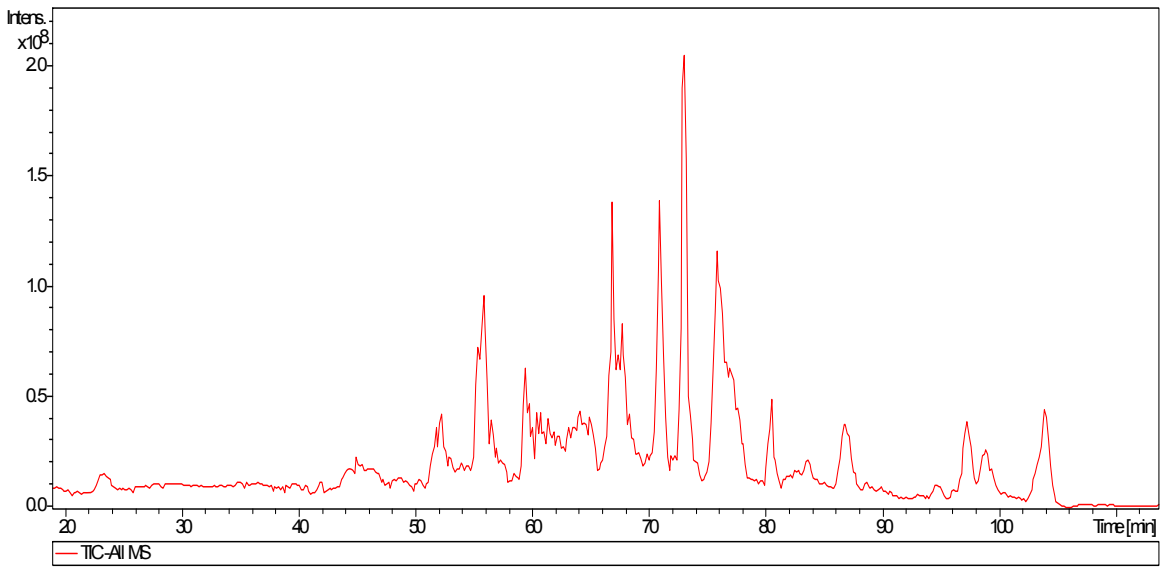


Figure 6.4. Normal-phase LC-ESI FT-ICR MS total ion chromatogram of oligosaccharides released from BxPC-3 conditioned media. Table 6.1 includes a list of potential oligosaccharides observed throughout the chromatogram.

graphitized carbon SPE. This approach allows for the examination of oligosaccharides from complex samples with an ESI mass spectrometer. Oligosaccharides from conditioned media of two cell lines, HDPE and BxPC-3 were examined. Currently, no oligosaccharides have been found in conditioned media from the HDPE cell line. This negative result may be due to a lack of mucin overexpression in this cell line. In conditioned media from the BxPC-3 cell line, normal-phase LC-ESI-FT-ICR MS revealed several O-linked and N-linked oligosaccharides. However, all of these species were also found in a control sample consisting of RPMI-1640 supplemented with fetal bovine serum. These results were surprising, because Lebrilla and co-workers reported very few oligosaccharides originating from fetal bovine serum. However, our results clearly demonstrate that fetal bovine serum oligosaccharides are a major interference in these samples. One possible solution is to avoid supplementing growth media with fetal bovine serum. A wide variety of serum-free media are commercially available.

Although it has already been demonstrated that human pancreatic cancer cell lines can be adapted for growth in serum-free media,[43,44] it is unclear what effect a change in media will have upon glycosylation expression.

6.5 References

1. Varki, A. *Glycobiol.* **1993**, 97-130.
2. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
3. Demetriou, M.; Nabi, I. R.; Coppolino, M.; Dedhar, S.; Dennis, J. W. *J. Cell Biol.* **1995**, *130*, 383-392.
4. Orntoft, T. F.; Vestergaard, E. M. *Electrophoresis* **1999**, *20*, 362-371.
5. Patwa, T. H.; Zhao, J.; Anderson, M. A.; Simeone, D. M.; Lubman, D. M. *Anal. Chem.* **2006**, *78*, 6411-6421.
6. Kim, Y. J.; Varki, A. *Glycoconj. J.* **1997**, *14*, 569-576.
7. Choudhury, A.; Moniaux, N.; Ulrich, A. B.; Schmied, B. M.; Standop, J.; Pour, P. M.; Gendler, S. J.; Hollingsworth, M. A.; Aubert, J. P.; Batra, S. K. *Br. J. Cancer* **2004**, *90*, 657-664.
8. An, H. J.; Miyamoto, S.; Lancaster, K. S.; Kirmiz, C.; Li, B.; Lam, K. S.; Leiserowitz, G. S.; Lebrilla, C. B. *J. Proteome Res.* **2006**, *5*, 1626-1635.
9. Niederhuber, J. E.; Brennan, M. F.; Menck, H. R. *Cancer* **1995**, *76*, 1671-1677.
10. Grote, T.; Logsdon, C. D. *Curr. Opin. Gastroenterol.* **2007**, *23*, 508-514.
11. Dall'Olio, F.; Chiricoclo, M.; Mariani, E.; Facchini, A. *Eur. J. Biochem.* **2001**, *268*, 5876-5884.
12. Hollingsworth, M. A.; Swanson, B. J. *Nat. Rev. Cancer* **2004**, *4*, 45-60.
13. Brockhausen, I. *EMBO rep.* **2006**, *7*, 599-604.
14. Yamori, T.; Kimura, H.; Stewart, K.; Ota, D. M.; Cleary, K. R.; Irimura, T. *Cancer Res.* **1987**, *47*, 2741-2747.
15. Gorelik, E.; Galili, U.; Raz, A. *Cancer Metast. Rev.* **2001**, *20*, 245-277.
16. Okuyama, N.; Ide, Y.; Nakano, M.; Nakagawa, T.; Yamanak, K.; Moriwai, K.; Murata, K.; Ohigashi, H.; Yokoyama, S.; Eguchi, H.; Ishikawa, O.; Ito, T.; Kato, M.; Kasahara, A.; Kawano, S.; Gu, J.; Taniguchi, N.; Miyoshi, E. *Int. J. Cancer* **2006**, *118*, 2803-2808.
17. Zhao, J.; Qiu, W.; Simeone, D. M.; Lubman, D. M. *J. Proteome Res.* **2007**, *6*, 1126-1138.
18. Kirmiz, C.; Li, B.; An, H. J.; Clowers, B. H.; Chew, H. K.; Lam, K. S.; Ferrige, A.; Alecio, R.; Borowsky, A. D.; Sulalmon, S.; Lebrilla, C. B.; Miyamoto, S. *Mol. Cell Proteomics* **2007**, *6*, 43-55.
19. Zhao, J.; Patwa, T. H.; Qiu, W.; Shedden, K.; Hinderer, R.; Misek, D. E.; Anderson, M. A.; Simeone, D. M.; Lubman, D. M. *J. Proteome Res.* **2007**, *6*, 1864-1874.
20. Schulz, B. L.; Packer, N. H.; Karlsson, N. G. *Anal. Chem.* **2002**, *74*, 6088-6097.
21. Arumugam, T.; Simeone, D. M.; Golen, K. V.; Logsdon, C. D. *Clin. Cancer Res.* **2005**, *11*, 5356-5364.
22. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2004**, *77*, 1876-1882.
23. Cooper, C. A.; Gasteiger, E.; Packer, N. H. *Proteomics* **2001**, *1*, 340-349.
24. Green, E. D.; Adelt, G.; Baenziger, J. U.; Wilson, S.; Vanhalbeek, H. *J. Biol. Chem.* **1988**, *263*, 18253-18268.
25. Edge, A. S. B.; Spiro, R. G. *J. Biol. Chem.* **1987**, *262*, 16135-16141.

26. Hounsell, E. F.; Pickering, J. J.; Stoll, M. S.; Lawson, A. M.; Feizi, T. *Biochem. Soc. Trans.* **1984**, 607-610.
27. Likhosherstov, L. M.; Novikova, O. S.; Derevitskaya, V. A.; Kochetkov, N. K. *Carbohydr. Res.* **1990**, *199*, 67-76.
28. Packer, N. H.; Lawson, M. A.; Jardine, D. R.; Redmond, J. W. *Glycoconj. J.* **1998**, *15*, 737-747.
29. Kyselova, Z.; Mechref, Y.; Al Bataineh, M. M.; Dobrolecki, L. E.; Hickey, R. J.; Vinson, J.; Sweeney, C. J.; Novotny, M. V. *J. Proteome Res.* **2007**, *6*, 1822-1832.
30. Tokugawa, K.; Oguri, S.; Takeuchi, M. *Glycoconj. J.* **1996**, *13*, 53-56.
31. Fuster, M. M.; Esko, J. D. *Nat. Rev. Cancer* **2005**, *5*, 526-542.
32. Yonezawa, S.; Byrd, J. C.; Dahiya, R.; Ho, J. J. L.; Gum, J. R.; Griffiths, B.; Swallow, D. M.; Kim, Y. S. *Biochem. J.* **1991**, *276*, 599-605.
33. Osako, M.; Yonezawa, S.; Siddiki, B.; Huang, J.; Ho, J. J. L.; Kim, Y. S. *Cancer* **1993**, *71*, 2191-2199.
34. Burdick, M. D.; Harris, A.; Reid, C. J.; Iwamura, T.; Hollingsworth, M. A. *J. Biol. Chem.* **1997**, *272*, 24198-24202.
35. Andrianifahanana, M.; Moniaux, N.; Schmied, B. M.; Ringel, J.; Friess, H.; Hollingsworth, M. A.; Buchler, M. W.; Aubert, J. P.; Batra, S. K. *Clin. Cancer Res.* **2001**, *7*, 4033-4040.
36. Hinoda, Y.; Ikematsu, Y.; Horinochi, M.; Sato, S.; Yamamoto, K.; Nakano, T.; Fukui, M.; Suehiro, Y.; Hamanaka, Y.; Nishikawa, Y.; Kida, H.; Waki, S.; Oka, M.; Imai, K.; Yonezawa, S. *J. Gastroenterol.* **2003**, *38*, 1162-1166.
37. Moniaux, N.; Andrianifahanana, M.; Brand, R. E.; Batra, S. K. *Br. J. Cancer* **2004**, *91*, 1633-1638.
38. Kawasaki, N.; Ohta, M.; Hyuga, S.; Hashimoto, O.; Hayakawa, T. J. *Anal. Biochem.* **1999**, *269*, 297-303.
39. Itoh, S.; Kawasaki, N.; Ohta, M.; Hyuga, M.; Huyuga, S.; Hayakawa, T. J. *J. Chrom. A* **2002**, *968*, 89-100.
40. Thomsson, K. A.; Karlsson, N. C.; Hansson, G. C. *J. Chrom. A* **1999**, *854*, 131-139.
41. Barroso, B.; Dijkstra, R.; Geerts, M.; Lagerwerf, F.; van Veelen, P.; de Ru, A. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1320-1329.
42. Wuhrer, M.; Koeleman, C. A. M.; Deelder, A. M.; Hokke, C. H. *Anal. Chem.* **2004**, *76*, 833-838.
43. Smith, J. P.; Fantaskey, A. P.; Zagon, I. S. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **1995**, *268*, 135-141.
44. Smith, J. P.; Shish, A.; Wu, Y.; McLaughlin, P. J.; Zagon, I. S. *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **1996**, *270*, 1078-1084.

Table 6.1 Potential oligosaccharides observed in a normal-phase LC-ESI-MS chromatogram of conditioned media from BxPC-3. Only structures which have been reported in the GlycoSuiteDB database are indicated. The approximate elution time, neutral mass, and relative abundance are indicated. As a control sample, RPMI-1640 media supplemented with 10% fetal bovine serum was subjected to the same treatment as conditioned media. All oliogsaccharides observed in conditioned media samples (and shown in this table) were also present in control samples.

Elution time (range)	Mass	Relative Abundance	Present in Control?	Potential Structural Assignment
70-75 min				
	676.255	100.0	Y	(Hex)1 (HexNAc)1 (NeuAc)1* or (HexNAc)1 (Deoxyhexose)1 (NeuGc)1*
75-80 min				
	676.262	1.6	Y	(Hex)1 (HexNAc)1 (NeuAc)1* or (HexNAc)1 (Deoxyhexose)1 (NeuGc)1*
	692.256	0.7	Y	(Hex)1 (HexNAc)1 (NeuGc)1*
	838.312	0.2	Y	(Hex)2 (HexNAc)1 (NeuAc)1* or (Hex)1 (HexNAc)1 (Deoxyhexose)1 (NeuGc)1*
	967.350	31.9	Y	(Hex)1 (HexNAc)1 (NeuAc)2* or (Hex)1 (HexNAc)3 (HexA)1* or (HexNAc)1 (Deoxyhexose)1 (NeuAc)1 (NeuGc)1*
	983.348	0.9	Y	(Hex)1 (HexNAc)1 (NeuAc)1 (NeuGc)1* or (HexNAc)1 (Deoxyhexose)1 (NeuGc)2*
	999.352	0.6	Y	(Hex)1 (HexNAc)1 (NeuGc)2* or (Hex)1 (HexNAc)1 (Deoxyhexose)3 (HexA)1*
	1332.478	1.4	Y	(Hex)2 (HexNAc)2 (NeuAc)2*
80 -85 min				
	1041.386	2.0	Y	(Hex)2 (HexNAc)2 (NeuAc)1* or (Hex)1 (HexNAc)2 (Deoxyhexose)1 (NeuGc)1*
	1332.488	79.6	Y	(Hex)2 (HexNAc)2 (NeuAc)2*
	1364.485	3.7	Y	(Hex)4 (HexNAc)2 (Deoxyhexose)2*
	1623.588	2.2	Y	(Hex)1 (HexNAc)2 (Deoxyhexose)1 + (Man)3(GlcNAc)2*
	1908.613	0.5	Y	(Hex)1 (HexNAc)3 (Deoxyhexose)1 (Phos)1 + (Man)3(GlcNAc)2*
85-90 min				
	880.346	1.8	Y	(Hex)1 (HexNAc)2 (Deoxyhexose)2*
	1057.383	1.0	Y	(Hex)2 (HexNAc)2 (NeuGc)1*
	1083.435	0.3	Y	(Hex)1 (HexNAc)3 (Deoxyhexose)2*
	1203.439	1.0	Y	(Hex)3 (HexNAc)2 (NeuAc)1* or (Hex)2 (HexNAc)2 (Deoxyhexose)1 (NeuGc)1*
90-95 min				
	750.293	4.7	Y	(Hex)2 (HexNAc)2*
	1083.431	2.0	Y	(Hex)1 (HexNAc)3 (Deoxyhexose)2*
	2266.807	2.0	Y	(Hex)1 (HexNAc)3 (Deoxyhexose)2 (NeuAc)1 + (Man)3(GlcNAc)2*
	2401.827	1.6	Y	(Hex)1 (HexNAc)4 (Deoxyhexose)1 (NeuAc)1 (Sulph)1 + (Man)3(GlcNAc)2*

	2515.907	7.7	Y	(Hex)2 (HexNAc)2 (NeuAc)3 + (Man)3(GlcNAc)2*, (Hex)2 (HexNAc)2 (Deoxyhexose)2 (NeuAc)2 + (Man)3(GlcNAc)2*
	2881.041	1.7	Y	$(Hex)_3$ $(HexNAc)_3$ $(NeuAc)_3$ + $(Man)_3$(GlcNAc)2*
	2923.042	7.9	Y	N+ (Hex)5 (HexNAc)4 (NeuAc)1 + (Man)3(GlcNAc)2*
	3172.142	71.1	Y	(Hex)3 (HexNAc)3 (NeuAc)4 + (Man)3(GlcNAc)2* or N+ (Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuGc)3 + (Man)3(GlcNAc)2*
	3214.132	2.4	Y	N + (Hex)5 (HexNAc)4 (NeuAc)2 + (Man)3(GlcNAc)2*
	3463.224	5.8	Y	(Hex)3 (HexNAc)3 (NeuAc)5 + (Man)3(GlcNAc)2*
95-100 min				
	668.245	1.3	Y	(Hex)4*
	2224.804	29.2	Y	(Hex)2 (HexNAc)2 (NeuAc)2 + (Man)3(GlcNAc)2* or (Hex)2 (HexNAc)2 (Deoxyhexose)2 (NeuAc)1 + (Man)3(GlcNAc)2*
	2240.799	1.0	Y	(Hex)2 (HexNAc)2 (NeuAc)1 (NeuGc)1 + (Man)3(GlcNAc)2* or (Hex)3 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2*
	2524.927	0.8	Y	N+ (Hex)2 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2*
	2881.034	100.0	Y	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2*
	2897.039	3.0	Y	(Hex)4 (HexNAc)3 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2*
	2976.985	0.3	Y	(Hex)4 (HexNAc)3 (Deoxyhexose)1 (NeuAc)2 (Sulph)1 + (Man)3(GlcNAc)2*
	3036.096	0.7	Y	N+ (Hex)3 (HexNAc)4 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)*
	3172.141	11.5	Y	(Hex)3 (HexNAc)3 (NeuAc)4 + (Man)3(GlcNAc)2* or N+ (Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuGc)3 + (Man)3(GlcNAc)2*
	3181.140	1.2	Y	N+ (Hex)3 (HexNAc)4 (NeuAc)3 + (Man)3(GlcNAc)2*
	3327.195	1.1	Y	N + (Hex)3 (HexNAc)4 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2*
	3457.168	1.8	Y	(Hex)5 (HexNAc)6 (Deoxyhexose)1 (NeuAc)1 (Sulph)1 + (Man)3(GlcNAc)2*
	3537.250	1.0	Y	(Hex)4 (HexNAc)4 (NeuAc)4 + (Man)3(GlcNAc)2*
100-105 min				
	912.347	4.2	Y	(Hex)3 (HexNAc)2*
	2223.817	38.8	Y	(Hex)2 (HexNAc)2 (NeuAc)2 + (Man)3(GlcNAc)2*
	2240.809	8.5	Y	(Hex)2 (HexNAc)2 (NeuAc)1 (NeuGc)1 + (Man)3(GlcNAc)2* or (Hex)3 (HexNAc)2 (Deoxyhexose)1 (NeuAc)1 + (Man)3(GlcNAc)2*
	2322.861	14.6	Y	N+ (Hex)2 (HexNAc)2 (NeuAc)2 + (Man)3(GlcNAc)2* or N+ (Hex)2 (HexNAc)2 (Deoxyhexose)2 (NeuAc)1 + (Man)3(GlcNAc)2*
	2514.924	2.0	Y	(Hex)2 (HexNAc)2 (NeuAc)3 + (Man)3(GlcNAc)2*
	2588.929	5.2	Y	(Hex)3 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2*
	2687.953	1.6	Y	N+ (Hex)3 (HexNAc)3 (NeuAc)2 + (Man)3(GlcNAc)2*
	2734.978	3.0	Y	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2*
	2838.023	24.6	Y	(Hex)3 (HexNAc)3 (Deoxyhexose)1 (NeuAc)2 (NeuGc)1 + (Man)3(GlcNAc)1*or (Hex)4 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)1*

	2880.056	100.0	Y	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2*
	2896.045	20.2	Y	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)1*
	2979.077	41.9	Y	N+ (Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2*
	3036.107	5.1	Y	N+ (Hex)3 (HexNAc)4 (Deoxyhexose)1 (NeuAc)2 + (Man)3(GlcNAc)2*
	3171.167	67.0	Y	(Hex)3 (HexNAc)3 (NeuAc)4 + (Man)3(GlcNAc)2*
	3181.143	3.8	Y	N+ (Hex)3 (HexNAc)4 (NeuAc)3 + (Man)3(GlcNAc)2*
	3270.197	21.4	Y	N+ (Hex)3 (HexNAc)3 (NeuAc)4 + (Man)3(GlcNAc)2*
	3391.292	2.9	Y	(Hex)4 (HexNAc)4 (Deoxyhexose)1 (NeuAc)3 + (Man)3(GlcNAc)2*
	3536.224	2.0	Y	(Hex)4 (HexNAc)4 (NeuAc)4 + (Man)3(GlcNAc)2*
105-115 min				
	2880.026	2.1	Y	(Hex)3 (HexNAc)3 (NeuAc)3 + (Man)3(GlcNAc)2*

Chapter 7

Conclusions and Prospects for Future Work

7.1 Challenges of Glycosylation Characterization and Goals of this Dissertation

Due to the vast complexity of glycans, their structural characterization and profiling remain as difficult tasks. Several disease states (such as cancer and inflammation) exhibit aberrant glycosylation, a fact which has fueled research pertaining to glycan profiling and structural analysis. Complete glycosylation characterization demands knowledge of saccharide linkage, branching, sequence, glycosylation location, heterogeneity, and occupancy. Several analytical techniques are available for glycosylation characterization and were described in Chapter 1. Amongst this wide variety of techniques, mass spectrometry is one of the most versatile, due to its sensitivity and low sample consumption. In particular, tandem mass spectrometry is a powerful tool for the structural characterization of both proteins and oligosaccharides.

Many approaches for glycosylation characterization have been utilized, although typically glycosylation is examined at the glycopeptide level or following enzymatic or chemical removal of the glycan. The advantage of analyzing glycosylation at the glycopeptide level is that glycosylation location and heterogeneity can be determined. However, for specific structural information, tandem mass spectrometry of glycopeptides

is necessary. With conventional collisional activated dissociation (CAD) of glycopeptides, preferential cleavage of glycosidic bonds is often observed. This tendency makes the localization of glycosylation sites within the peptide difficult to determine. However, within the past decade several groups have utilized an alternative fragmentation technique, electron capture dissociation (ECD), for the characterization of glycopeptides. ECD, unlike CAD, tends to preferentially cleave the peptide backbone, allowing for localization of glycosylation sites.

In addition to glycopeptide analysis, another common approach for glycosylation characterization is the enzymatic or chemical removal of glycans from the glycoprotein, and subsequent analysis of the glycan mixture. The principal disadvantage of this technique is that information regarding glycosylation location and heterogeneity is lost. Following glycan release, there are several mass spectrometric approaches available. In positive ion mode, the protonated or metal adducted form of the oligosaccharide can be examined. Oligosaccharide anions can also be examined with mass spectrometry, although this approach is typically utilized to a lesser degree. Upon CAD of oligosaccharides (in both positive and negative ion mode) glycosidic cleavages are often the most abundant product ions observed. Although these ions provide information regarding the monosaccharide composition of oligosaccharides, cross-ring cleavages provide linkage information. Strategies such as metal-adduction with CAD have been utilized to increase cross-ring fragmentation.

The main goal behind the research presented in this dissertation was to explore the utility of ion-electron and ion-photon based fragmentation for glycosylation characterization. Several topics are explored, including lectin *de novo* sequencing,

glycopeptide and oligosaccharide structural characterization. To prove the validity of ion-electron reactions for glycosylation characterization, extensive comparisons were made between vibrational excitation fragmentation patterns (CAD and infrared multiphoton dissociation (IRMPD)) and ion-electron based fragmentation reactions (ECD and electron detachment dissociation (EDD)).

7.2 Summary of Results

Lectins are an extremely useful tool for glycoprotein isolation and purification. Several lectins are currently commercially available, however new ones are constantly being isolated. In Chapter 2, we illustrated how the combination of vibrational excitation (here IRMPD) and electron based fragmentation (ECD) allows for *de novo* sequencing of a lectin. The overlap between infrared multiphoton and electron capture dissociation data, indicated by so called "golden pairs", is a powerful tool for *de novo* sequencing of proteins originating from organisms with unsequenced genomes.

The characterization of high-mannose type glycopeptides using both vibrational excitation (here IRMPD) and electron-based fragmentation (ECD) was also examined. Previously, IRMPD was understood to selectivity cleave glycosidic bonds in the gas phase. Our results demonstrate that IRMPD of high-mannose type glycopeptides exhibits a unique fragmentation behavior, and often results in a mixture of glycan and peptide backbone cleavage. In contrast, ECD of high-mannose type glycopeptides results in extensive peptide backbone cleavage. The fragmentation spectrum of the largest glycopeptide examined with ECD to date was illustrated in Chapter 3.

Glycopeptide anion fragmentation behavior was also examined (Appendix A). Our results indicate that negative ion mode vibrational excitation results in a highly complicated mass spectrum, due to a mixture of glycan and peptide backbone cleavage. In addition, electron-based fragmentation in negative ion mode with EDD demonstrated that glycan cleavages are predominant. These results illustrate that glycopeptide characterization in positive ion mode results in more useful fragmentation information.

Both Chapter 4 and Chapter 5 focus on oligosaccharide structural characterization with a combination of vibrational excitation and electron based fragmentation. Although it is well established that CAD of metal-adducted oligosaccharides often exhibits more cross-ring fragments compared to protonated species, very little work had focused on ECD of oligosaccharides. We show that complementary structural information can be obtained from ECD of metal-adducted species as compared to IRMPD. Depending on the metal ion and oligosaccharide structure, in certain examples cross-ring cleavage can be greatly enhanced.

Many oligosaccharides are neutral or contain acidic saccharides, such as sialic acid. In these instances, negative ion mode mass spectrometry would be more suitable than positive ion mode characterization. Electron detachment dissociation (EDD) has recently been shown by Amster and co-workers to constitute a valuable analytical approach for the structural characterization of glycosaminoglycans (linear sulfated oligosaccharides).[1] We have explored the application of EDD to neutral and sialylated oligosaccharides. Our results demonstrate that EDD of both neutral and sialylated oligosaccharides provides structural information that is complementary to that obtained

from vibrational excitation. In most cases, cross-ring fragmentation obtained via EDD is more extensive than that obtained from vibrational excitation.

Finally, Chapter 6 focuses on strategies for the examination of O-linked glycans found in conditioned media from pancreatic cancer cell lines. Lebrilla and co-workers have demonstrated for ovarian and breast cancer that cancer associated O-linked glycans are present in both conditioned media and serum.[2,3] To examine the O-linked oligosaccharides present in conditioned media, several clean-up strategies were explored with a model glycoprotein, bovine fetuin. We have developed an optimal strategy for examining O-linked oligosaccharides from conditioned media from pancreatic cell lines. However, our results demonstrate that fetal bovine serum oligosaccharides are a major interference in these samples.

7.3 Prospects for Future Work

Our results clearly demonstrate the utility of both metal-adduction in combination with ECD for oligosaccharide structural characterization, as well as EDD for the analysis of neutral and sialylated oligosaccharides. This work was performed on model oligosaccharides. Having proven the utility of these techniques, the next logical step would be the implementation of these strategies for the structural analysis of oligosaccharides from biological samples.

Chapter 6 demonstrates that fetal bovine serum oligosaccharides present in growth media hinders the analysis of pancreatic cancer associated *O*-glycans in conditioned media. One possible solution is to use serum-free media. However, another approach to characterize cell-surface glycosylation would be to analyze glycans present

in the cell membrane fractions of pancreatic cancer cells. Kakehi and co-workers recently described an approach for glycan profiling of cell surface glycosylation from a variety of cancerous cell lines.[4] In the future, optimization and application of a similar protocol can be applied to pancreatic cancer cells. Although there are already several reports on the glycan profiles associated with pancreatic cancer from serum and pancreatic cancer cells (as described in Chapter 6), the majority of these reports focused on *N*-glycosylation. As previously discussed, the over-expression of mucins in pancreatic cancer indicates the relevance of *O*-glycans. However, very few studies have examined *O*-glycans in this disease. In addition, the majority of work performed on pancreatic cancer glycan profiles focused more on overall glycan composition, and not specific structural information. Given that pancreatic cancer exhibits a unique glycosylation profile, a potential future goal would be to design a vaccine based on these cancer-associated glycans in order to illicit an immune system response. However, in order to do so, complete structural characterization of these glycans is necessary.

The combination of vibrational excitation and ion-electron based fragmentation should provide extensive structural characterization for pancreatic cancer associated oligosaccharides. However, one main concern with utilizing ion-electron based fragmentation techniques for biomolecular structural characterization is that both ECD and EDD are not easily combined with on-line liquid chromatographic separation due to relatively long ECD and EDD fragmentation events. For the analysis of *N*- and *O*-glycans from pancreatic cancer cells, liquid chromatography is necessary due to the complexity of these samples. Chapter 6 demonstrates that normal-phase chromatography of underivatized oligosaccharides would be an appropriate choice for separation.

However, in order to utilize ECD and EDD for structural characterization, off-line liquid chromatography would be required. For metal-adducted ECD, metal salts also need to be added prior to mass spectrometric analysis. In the future, optimization of the off-line separation of *N*- and *O*-glycans from pancreatic cancer cells is needed.

Another variable which needs consideration is whether positive ion mode ESI-MS or negative ion mode ESI-MS will be utilized for glycan characterization. ECD is only feasible for glycan cations, while EDD is only applicable to glycan anions. As previously discussed, biologically relevant glycans are often neutral or contain acidic saccharides. This acidity would suggest that negative ion mode characterization would be more appropriate for glycans containing sialic acid or sulfated monosaccharides. Although EDD of neutral and sialylated oligosaccharides provides very extensive cross-ring fragmentation, due to the lower fragmentation efficiency associated with EDD, this technique may be difficult to implement with biological samples if sample quantities are minute. The work presented in this dissertation on ECD of oligosaccharides focused on neutral species, but recent findings in our group have also illustrated how ECD of metal-adducted acidic *N*-glycans provides complementary structural information to vibrational excitation. In these instances, metal-adducted ECD in combination with vibrational excitation fragmentation would be an attractive alternative to negative ion mode characterization.

Overwhelming evidence indicates the functional role of glycosylation in various forms of cancer. In order to fully comprehend how changes in glycosylation are related to cancer progression, detailed structural characterization of *N*- and *O*-glycans is essential. The work presented in this dissertation demonstrates how ion-electron

reactions, ECD and EDD, can provide additional structural information for glycans. With the detailed structural characterization of pancreatic cancer *N*- and *O*-glycans, several potential future goals can be envisioned. Glycoconjugates are already used as targets in many therapeutic applications such as cancer chemotherapy, diabetes therapy, antibiotics, antivirals, and anti-inflammatories.[5-10] A detailed characterization of pancreatic cancer associated glycans may permit the identification of cancer vaccine candidates. Given the link between aberrant glycosylation and malignancy, it may also be feasible to design therapeutics which prevent or perturb the formation of specific cancer associated glycans. This strategy is somewhat complex, because in addition to knowing specific glycan structures, these glycans must also demonstrate a specific role in cancer progression. Currently, the precise link between specific glycan structures and disease states, such as cancer, are often poorly understood. Finally, as previously discussed in Chapter 6, no reliable biomarkers exist for early detection of pancreatic cancer. A complete glycan profile of *N*- and *O*-glycans from pancreatic cancer cells may indicate which glycans can be used as diagnostic markers in this fatal disease.

7.4 References

1. Wolff, J. J.; Amster, I. J.; Chi, L.; Linhardt, R. J. *J. Am. Soc. Mass Spectrom.* **2006**, *18*, 234-244.
2. An, H. J.; Miyamoto, S.; Lancaster, K. S.; Kirmiz, C.; Li, B.; Lam, K. S.; Leiserowitz, G. S.; Lebrilla, C. B. *J. Proteome Res.* **2006**, *5*, 1626-1635.
3. Kirmiz, C.; Li, B.; An, H. J.; Clowers, B. H.; Chew, H. K.; Lam, K. S.; Ferrige, A.; Alecio, R.; Borowsky, A. D.; Sulalmon, S.; Lebrilla, C. B.; Miyamoto, S. *Mol. Cell Proteomics* **2007**, *6*, 43-55.
4. Naka, R.; Kamoda, S.; Ishizuka, A.; Kinoshita, M.; Kakehi, K. *J. Proteome Res.* **2006**, *5*, 88-97.
5. McAuliffe, J. C.; Hindsgaul, O. *Chem. Ind.* **1997**, *5*, 170-174.
6. Yarema, K. J.; Bertozzi, C. R. *Curr. Opin. Chem. Biol.* **1998**, *2*, 49-61.
7. Koeller, K. M.; Wong, C.-H. *Nat. Biotechnol.* **2000**, *18*, 835-841.
8. Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357-2364.
9. Dove, A. *Nat. Biotechnol.* **2001**, *19*, 913-917.
10. Dube, D. H.; Bertozzi, C. R. *Nat. Rev. Drug Discov.* **2005**, *4*, 477-488.

Appendix A

Infrared Multiphoton Dissociation and Electron Detachment Dissociation of Glycopeptide Anions

A.1 Introduction

Chapter 3 explores the fragmentation behavior of glycopeptide cations with infrared multiphoton dissociation (IRMPD) and electron capture dissociation (ECD). ECD of high-mannose type glycopeptide cations results in extensive backbone peptide cleavage and no cleavage within the glycan portion. This trend has also been observed with other varieties of N-linked glycopeptides and O-linked glycopeptides.[1-5] Chapter 3 also demonstrates that IRMPD of high-mannose type glycopeptides may result in a mixture of both peptide and glycosidic cleavages. However, IRMPD of complex type[1,2,6] and hybrid type[7] glycopeptides typically results in exclusively glycosidic cleavages. The investigations presented thus far on glycopeptide fragmentation have pertained to glycopeptide cations; however, the fragmentation of glycopeptide anions with IRMPD has yet to be examined. In addition, while ECD is only applicable towards multiply charged positive ions, electron detachment dissociation (EDD) is only possible with polyanionic species. EDD has been shown to cleave the peptide backbone while retaining post-translational modifications such as tyrosine phosphorylation and

sulfation.[8-10] Since many peptides contain acidic residues, EDD would be a useful fragmentation technique. In addition, many glycopeptides contain acidic saccharides, such as sialic acid and sulfated residues. For these reasons, the fragmentation behavior of glycopeptide anions with IRMPD and EDD is explored here. A xylose type and high-mannose type glycopeptide from model glycoproteins, a lectin from *Erythrina cristagalli* and ribonuclease B, are examined.

A.2 Experimental

Approximately 1.25 nmols of ribonuclease B and lectin from *Erythrina cristagalli* (Sigma, St. Louis, MO) were used for enzymatic digestion. Ribonuclease B was digested (after reduction and alkylation) with GluC (Roche Applied Science, Indianapolis, IN) for either 6 hours (1:20 GluC:protein w/w) or for 13 hours (1:100) at 25°C. The lectin was digested with trypsin (Princeton Separations, Adelphia, NJ) at 37°C for 15 hours (1:30). Digests were desalted with C18 ZipTips (Millipore, Billerica, MA) and diluted to 10^{-5} M in 1:1 water:isopropanol (Fisher, Fair Lawn, NJ) with 20 mM tripropylamine (Acros Organics, Morris Plains, NJ).

Samples were electrosprayed (60 μL/h) into a 7.0 Tesla Fourier transform ion cyclotron resonance mass spectrometer (APEX-Q, Bruker Daltonics, Billerica, MA). Glycopeptide precursor ions were isolated using the quadrupole and externally accumulated, followed by injection and dynamic trapping in a cylindrical ICR cell. EDD (17-18 eV electrons) was performed with an indirectly heated hollow dispenser cathode (Heat Wave, Watsonville, CA) with irradiation times ranging from 0.75 - 1.0 seconds.

IRMPD was performed with a 25 W CO_2 laser (Synrad, Mukilteo, WA) at 30% power with 90 ms irradiation.

A.3 Results and Discussion

A.3.1 IRMPD and EDD of a Xylose Type Glycopeptide

Following a trypsin digestion, a xylose type glycopeptide from the lectin *Erythrina cristagalli* was examined with both IRMPD and EDD. This lectin was chosen as a model glycoprotein because the N-linked oligosaccharides[11] and amino acid sequence[12] of this protein have been previously characterized. In addition, the IRMPD and ECD fragmentation behavior of this glycopeptide were examined in Chapter 3. The IRMPD mass spectrum of a doubly deprotonated xylose type glycopeptide is shown in Figure A.1 along with its structure. The most abundant product ion observed following IRMPD is the loss of the entire glycan (labeled as $[M - glycan]^-$). The majority of other product ions correspond to peptide backbone cleavage, including several *b*- and *y*-type ions. In some cases, the glycan portion remains attached to these ions (indicated with the label [*b*/*y* + g]). However, in many cases the glycan portion has been lost (indicated with the label [*b*/*y* - g]). Neutral molecule losses, such as loss of H_2O and NH_3, as well as internal fragmentation product ions (indicated by [i] and [i_N]) are also observed. Following IRMPD, several product ions due to cleavage within the glycan portion of the glycopeptide are also noted. The intact glycan is observed at *m/z* 1186. One of the most abundant product ions, at *m/z* 677, corresponds to a $^{2,4}A_3$ cross-ring cleavage in an *N*-acetylglucosamine residue (cleavage is indicated in Figure A.1). A product ion at *m/z* 545 corresponds to loss of xylose from this product ion. In addition, C-type cleavage is

observed at *m/z* 635. Unlike IRMPD of the cationic form (as shown in Chapter 3), IRMPD of this species results in a highly complicated mass spectrum.

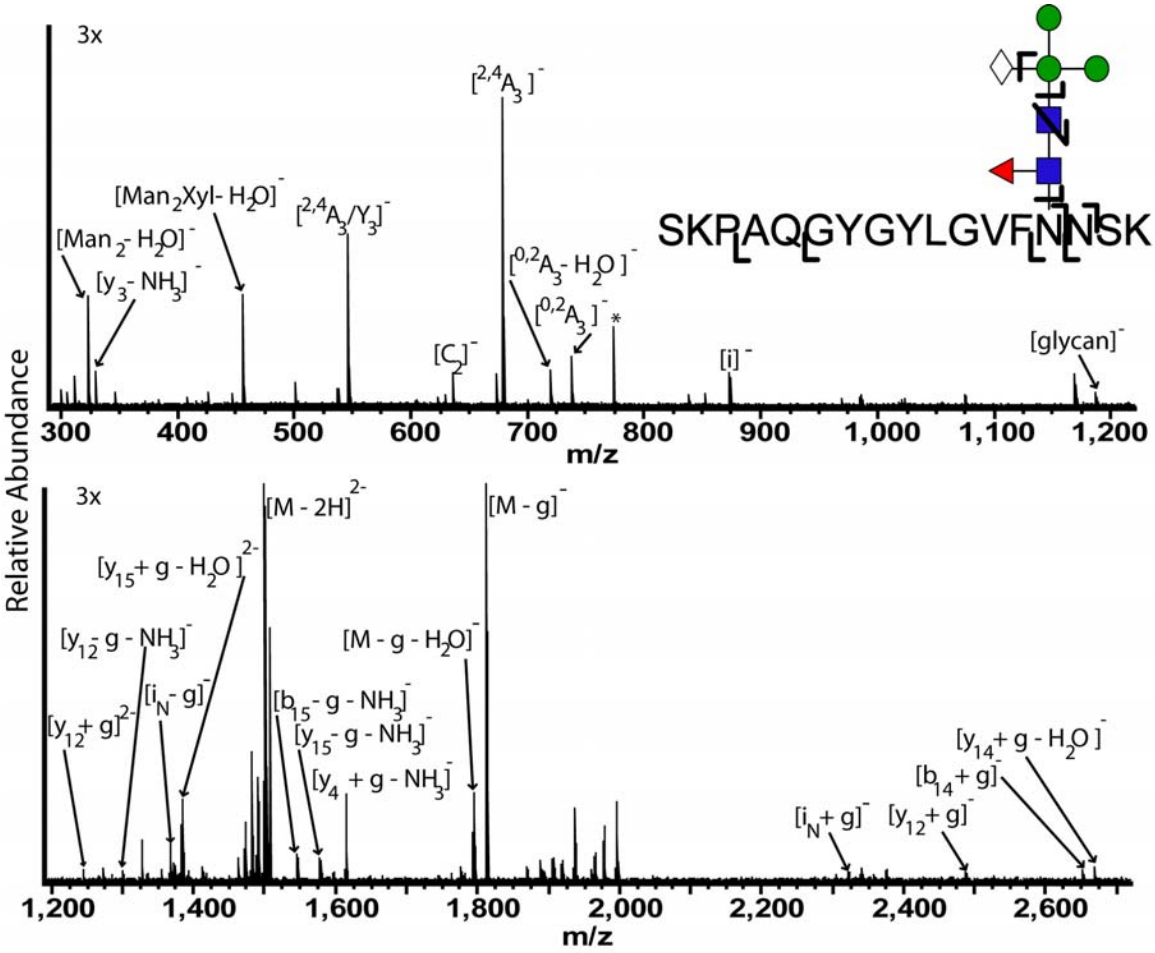


Figure A.1. IRMPD FT-ICR tandem mass spectrum (100 scans, 90 ms with 7.5 W) of a doubly deprotonated xylose type lectin glycopeptide at *m/z* 1499. The top figure shows *m/z* 300-1,200 while the bottom shows 1,200 – 2,700. The most abundant product ion corresponds to loss of the entire glycan (indicated as $[M - g]^-$). g is the abbreviation for glycan. [i] indicates an internal fragment product ion without the asparagine to which the glycan is originally attached, whereas $[i_N]$ indicates an ion that contains this asparagine. (■=GlcNAc, ●=Man, ◄=Fuc, ◊=Xyl).

EDD of the same xylose type glycopeptide examined in Figure A.1 is shown in Figure A.2. Following EDD, several product ions are observed in the upper *m/z* region. These product ions correspond to the charge reduced non-fragmented species (at *m/z* 2997), as well as several ions due to the loss of mannose, fucose, xylose, and small

molecules from the charge reduced species. Unlike ECD of glycopeptides, EDD of this glycopeptide anion resulted in exclusively glycan cleavage. These EDD results contradict those which have been observed for sulfated and phosphorylated peptides, where the post-translational modification was retained following fragmentation. Although the EDD mechanism remains highly debated, cleavages are postulated to occur at or near sites of negative charge or charge solvation for peptides.[9] The xylose type glycopeptide examined in Figure A.2 does not contain acidic residues such as aspartic or glutamic acid. Exclusive glycan cleavage following EDD may be due to deprotonation being localized within the glycan portion of the molecule.

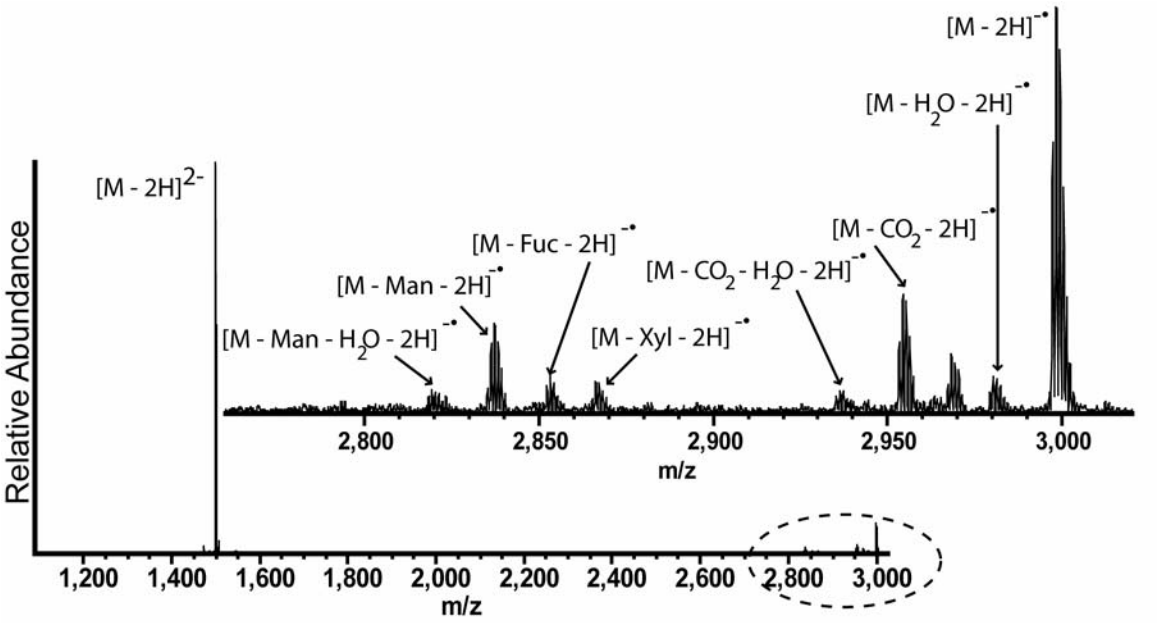


Figure A.2. EDD FT-ICR tandem mass spectrum (64 scans, 750 ms with a bias voltage of -18 V) of a doubly deprotonated xylose type lectin glycopeptide at *m/z* 1499. The most abundant product ion corresponds to the charge reduced non-fragmented radical species, indicated by $[M - 2H]^{-\bullet}$. Other product ions include loss of mannose, fucose, xylose, and small molecules from the charge reduced species. No peptide backbone cleavage is observed. (■=GlcNAc, ●=Man).

A.3.2 EDD of a High-Mannose Type Glycopeptide

The EDD fragmentation behavior of another glycopeptide, a high-mannose type glycopeptide from ribonuclease B, was also examined. ECD of high-mannose type glycopeptides (examined in Chapter 3) results in extensive peptide backbone cleavages. EDD was performed on a GluC digest glycopeptide from ribonuclease B, with a molecular weight of 5919 Da. The EDD spectrum of this glycopeptide, as well as the glycopeptide structure is shown in Figure A.3. Following EDD, several product ions are observed in the upper *m/z* region. The most abundant product ion, at *m/z* 2,959, corresponds to the charge reduced non-fragmented radical species. Several product ions due to glycan cleavages, as well as small molecule losses, are also observed. Glycan cleavages include the loss of mannose and an X-type sugar cross-ring cleavage. Several small molecule losses are also denoted in Figure A.3, including the loss of 31 Da (which could correspond to the loss of CH_3O), loss of 45 Da (which could correspond to the loss of C_2H_5O), and loss of C_2H_4ONS (presumably from carboxyamidomethylation of cysteine). These results are similar to those for EDD of a xylose type glycopeptide. However, unlike the previously examined xylose-type glycopeptide, this high-mannose type glycopeptide contains several acidic residues (2 aspartic acids and 1 glutamic acid). This peptide acidity implies deprotonation sites should be along the peptide backbone and not within the glycan.

A.4 Conclusions

Here, it is shown that IRMPD of a xylose type glycopeptide anion results in a highly complicated fragmentation pattern, with a mixture of both glycan and peptide

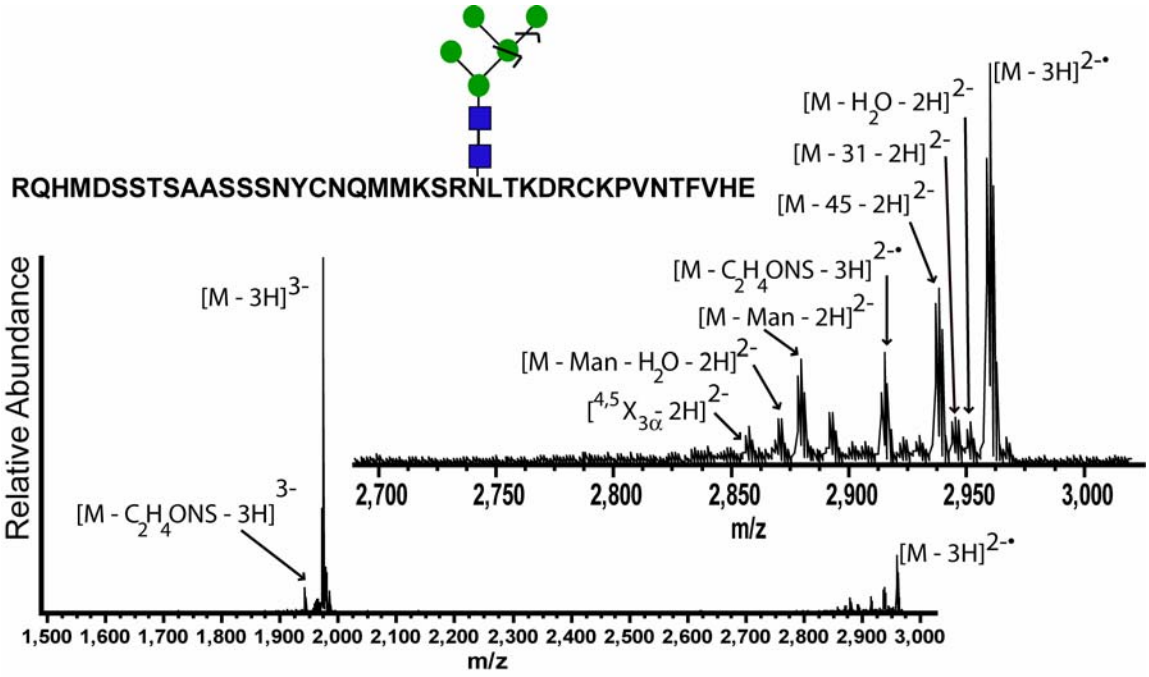


Figure A.3. EDD FT-ICR tandem mass spectrum (100 scans, 1 s with 18 eV electrons) of a triply deprotonated high-mannose type ribonuclease B glycopeptide at *m/z* 1973. The most abundant product ion corresponds to the charge reduced non-fragmented radical species, indicated by $[M - 3H]^{2-\bullet}$. Other product ions include loss of mannose, as well as several neutral molecule losses. These products include the loss of 31 Da (which could correspond to the loss of CH_3O), loss of 45 Da (which could correspond to the loss of C_2H_5O), and loss of $C2H_4ONS$ (presumably from carboxyamidomethylation of cysteine).

backbone cleavage. In addition, several internal fragments and neutral molecule losses are observed. In general, negative ion mode tandem mass spectrometry of peptides and proteins is rarely used for obtaining structural information,[13] one reason being that CAD (and vibrational excitation techniques in general) of deprotonated peptides often result in relatively complex fragmentation patterns. Figure A.1 demonstrates that this same trend is also observed following glycopeptide fragmentation.

Figures A.2 and A.3 show that EDD of glycopeptides results in glycan cleavage and small molecule losses. Unlike EDD of phosphorylated and sulfated peptides, the post-translational modification is cleaved during the fragmentation process. For the xylose type glycopeptide, it was hypothesized that since there are no acidic residues

located within the glycopeptide, deprotonation within the glycan portion of the molecule may explain why glycan cleavage is observed following EDD. However, for the high-mannose type glycopeptide, several acidic residues are located within the peptide. It is unclear how the location of deprotonation sites affects EDD fragmentation behavior, because similar trends are observed with both glycopeptides.

A.5 References

1. Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.
2. Hakansson, K.; Chalmers, M. J.; Quinn, J. P.; McFarland, M. A.; Hendrickson, C. L.; Marshall, A. G. *Anal. Chem.* **2003**, *75*, 3256-3262.
3. Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal. Chem.* **2003**, *75*, 2355-2361.
4. Mirgorodskaya, E.; Roepstorff, P.; Zubarev, R. A. *Anal. Chem.* **1999**, *71*, 4431-4436.
5. Renfrow, M. B.; Cooper, H. J.; Tomana, M.; Kulhavy, R.; Hiki, Y.; Toma, K.; Emmett, M. R.; Mestecky, J.; Marshall, A. G.; Novak, J. *J. Biol. Chem.* **2005**, *280*, 19136-19145.
6. Hakansson, K.; Emmett, M. R.; Marshall, A. G.; Davidsson, P.; Nilsson, C. L. *J. Proteome Res.* **2003**, *2*, 581-588.
7. Adamson, J. T.; Hakansson, K. *Proc. 231st American Chemical Society Meeting & Exposition*, Atlanta, GA, March 26-30 2005.
8. Budnik, B. A.; Haselmann, K. F.; Zubarev, R. A. *Chem. Phys. Lett.* **2001**, *342*, 299-302.
9. Kjeldsen, F.; Silivra, O. A.; Ivonin, I. A.; Haselmann, K. F.; Gorshkov, M.; Zubarev, R. A. *Chem. Eur. J.* **2005**, *11*, 1803-1812.
10. Kweon, H. K.; Hakansson, K. *J. Proteome Res.* **2008**, *7*, 749-755.
11. Ashford, D.; Dwek, R. A.; Welply, J. K.; Amataykul, S.; Homans, S. W.; Lis, H.; Taylor, G. K.; Sharon, N.; Rademacher, T. W. *Eur. J. Biochem.* **1987**, *166*, 311-320.
12. Svensson, C.; Teneberg, S.; Nilsson, C. L.; Kjellberg, A.; Schwarz, F. P.; Sharon, N.; Krengel, U. *J. Mol. Biol.* **2002**, *321*, 69-83.
13. Bowie, J. H.; Brinkworth, C. S.; Dua, S. *Mass Spectrom. Rev.* **2002**, *21*, 87-107.

Appendix B

Collision Activated Dissociation and Electron Detachment Dissociation of a Chloride Adducted Oligosaccharide

B.1 Introduction

Chapter 5 explores the fragmentation behavior of oligosaccharides in negative ion mode mass spectrometry, with collision activated dissociation (CAD), infrared multiphoton dissociation (IRMPD), and electron detachment dissociation (EDD). Our results demonstrate that EDD of both neutral and sialylated oligosaccharides provides structural information that is complementary to that obtained from both CAD and IRMPD. In all cases, EDD resulted in additional cross-ring cleavages. In most cases, cross-ring fragmentation obtained via EDD is more extensive than that obtained from IRMPD or CAD.

Chapter 5 examines the fragmentation behavior of singly and doubly-deprotonated oligosaccharide species. As an alternative to deprotonated species, Cole and co-workers have introduced a promising negative ion method for the analysis of neutral and acidic oligosaccharides utilizing anion attachment.[1-3] It has been extensively demonstrated that certain metal ions can aid the ionization of oligosaccharides via metal

adduction, and often provide more informative fragmentation mass spectra (as discussed in Chapter 4). However, there are very few reports on the effects upon oligosaccharide fragmentation with anion attachment in negative ion mode mass spectrometry. In 2005, Cole and co-workers examined the CAD fragmentation of singly deprotonated disaccharides along with their fluoride, acetate, and chloride adducted counterparts.[3] Although CAD fragmentation patterns for all species were similar, chloride adducts provided greater precursor ion signal than the singly deprotonated species. This signal enhancement indicates their potential for improving the ionization of neutral oligosaccharides. In addition, the fragmentation behavior of chloride adducted stereoisomers was examined. The ratio of chloride ions to non-chloride CAD product ions indicated their potential for differentiating the anomeric configuration of the glycosidic oxygen.

Here, the fragmentation behavior of a chloride adducted linear oligosaccharide, maltoheptoase, is examined. The CAD and EDD fragmentation patterns of the $[M + Cl - H]^{2-}$ and $[M + 2Cl]^{2-}$ species are explored. CAD and EDD fragmentation patterns of chloride-adducted species are compared to the corresponding $[M - 2H]^{2-}$ species, as shown in Chapter 5. Unlike the previous study performed by Cole and co-workers, a larger oligosaccharide (7 monosaccharide residues) is examined here. Furthermore, this work constitutes the first examination of an anion adducted oligosaccharide with EDD.

B.2 Experimental

B.2.1 Sample Preparation

Maltoheptaose (Sigma, St. Louis, MO) was prepared in a solution of 50%

methanol (Fisher, Fair Lawn, NJ) and 20 μM ammonium chloride (Sigma) to a final concentration of 5 μM.

B.2.2 FT-ICR Mass Spectrometry

All experiments were performed with an actively shielded 7 T Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer with a quadrupole front-end (APEX-Q, Bruker Daltonics, Billerica, MA), as previously described.[4] Samples were infused via an Apollo II ion source at a flow rate of 60 μL/h with the assistance of N_2 nebulizing gas. Ions were mass selectively accumulated with a quadrupole, transferred through high-voltage ion optics and captured with dynamic trapping in an Infinity ICR cell.[5] External CAD was performed in a hexapole following mass selective ion accumulation with argon as a collision gas. An indirectly heated hollow dispenser cathode was used to perform EDD.[6] For EDD, the cathode heating current was kept constant at 2.0 A and the cathode voltage was pulsed during the EDD event to a bias voltage of (- 20) - (- 30) V for 3 - 5 s. A lens electrode located immediately in front of the cathode was kept 0.8 V higher than the cathode bias voltage.

B.3 Results and Discussion

B.3.1 CAD and EDD of $[M + Cl - H]^{2-}$

The CAD fragmentation of singly deprotonated, chloride adducted maltoheptaose, $[M + Cl - H]^{2-}$, is shown in Figure B.1. Following CAD, several glycosidic and cross-ring cleavages are observed. Due to the symmetric nature of maltoheptaose, a number of ions cannot be distinguished based on their m/z ratios (C and Y, and several A and X ions). However, native neutral oligosaccharide anions that are 1-4 or 1-6 linked have

been shown to produce predominantly C-type and A-type ions during CAD.[7-11] Following CAD of $[M + Cl - H]^{2-}$, several glycosidic and cross-ring cleavages (^{0,2}A- and ^{2,4}A-type) are observed. Product ions consist of both singly charged and doubly charged species. Singly charged species are either deprotonated (indicated by H in the labels) or contain a chloride ion (indicated by Cl in the labels).

Upon CAD, there are several potential decomposition pathways. Cole and co-workers have previously examined the fragmentation behavior of a chloride adducted disaccharide. They proposed that upon CAD there are two predominant fragmentation pathways 1) loss of chloride or 2) loss of HCl and subsequent fragmentation. In the example presented here, because the oligosaccharide is both deprotonated and chloride adducted, the fragmentation behavior is slightly more complex. The presence of doubly charged product ions, due to a chloride adduct and deprotonation site (indicated by * in Figure B.1), indicate that direct decomposition of the oligosaccharide is one observed fragmentation pathway. Singly charged chloride adducted product ions also confirm retention of a chloride ion, perhaps due to direct decomposition of the precursor. Singly charged, deprotonated, product ions may arise via two pathways. These product ions may be simply due to direct decomposition of the precursor ion. Alternatively, these product ions may be due to loss of HCl and subsequent fragmentation. The presence of an ion at *m/z* 575, which is assigned as $[M - 2H]^{2-}$, confirms that some loss of HCl does occur upon CAD. The ion at *m/z* 1151, assigned as $[M - H]^{-}$, indicates that loss of chloride is also observed. Overall, the CAD fragmentation pattern of $[M + Cl - H]^{2-}$ is similar to that of $[M -2H]^{2-}$ (shown in Chapter 5). Several glycosidic and cross-ring

cleavages are observed, all of which are also observed from CAD of doubly deprotonated maltoheptaose.

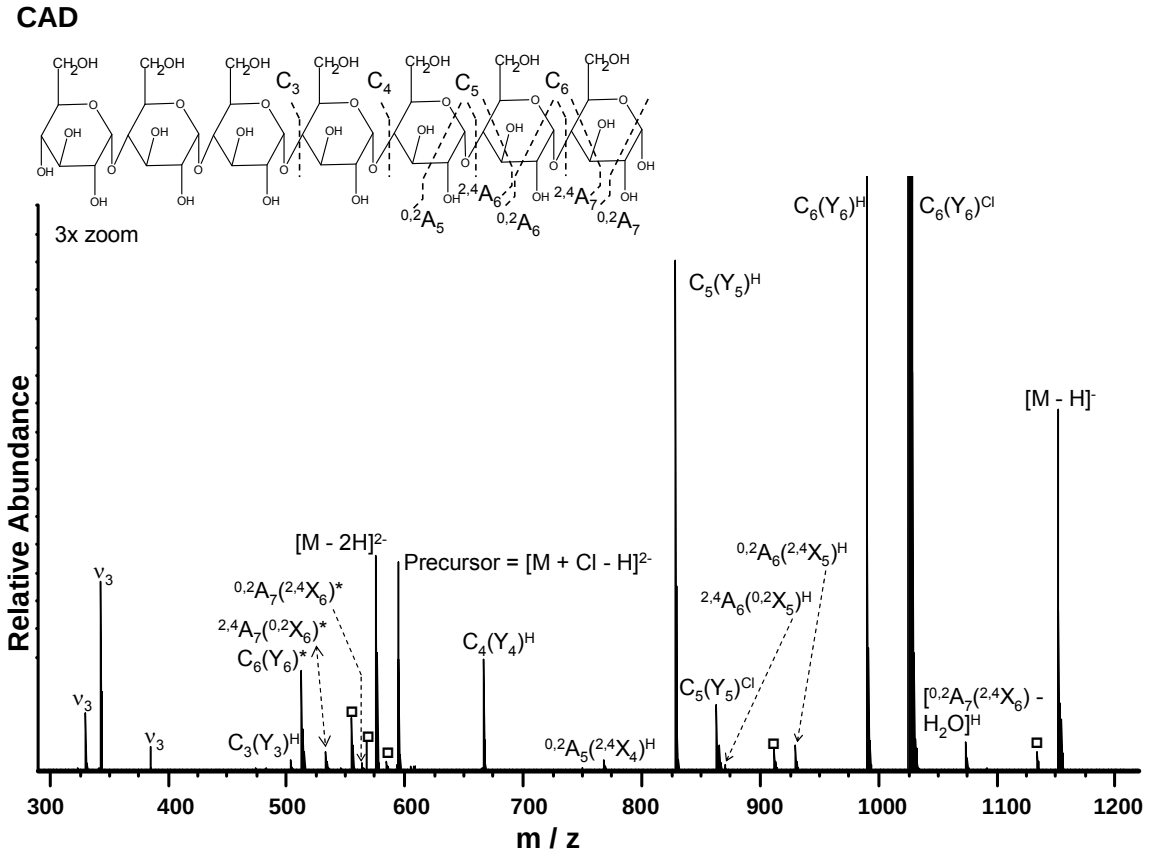


Figure B.1. CAD FT-ICR (30 scans, collision cell voltage 6V) tandem mass spectrum of deprotonated, chloride adducted maltoheptaose. Doubly charged product ions, due to deprotonation and chloride adduction, are indicated with an asterisk. Singly charged product ions due to deprotonation are indicated with "H" in the labels, while those containing chloride are indicated by "Cl". Squares indicate water loss from an adjacent product ion. ν_3 indicates a harmonic.

EDD of the $[M + Cl - H]^{2-}$ species of maltoheptaose is shown in Figure B.2. Following EDD, only singly charged product ions are observed. These ions are a mixture of both deprotonated and chloride adducted product ions, similar to the CAD spectrum in Figure B.1. A $[M - 2H]^{-\bullet}$ (at *m/z* 1150) ion is observed following EDD, indicating the loss of HCl and an electron (or due to loss of a chloride and a hydrogen). The ion [M +

$Cl - H - CH_3O]^{-\bullet}$ at *m/z* 1155 is also observed. Loss of CH_3O from the charge reduced species was also noted following EDD of sialylated oligosaccharides in Chapter 5.

In Figure B.2, product ions consist of a mixture of both glycosidic and cross-ring fragments. Glycosidic cleavages are more extensive following EDD than following CAD (shown previously in Figure B.1). Both C-type (or Y) type cleavages and B-type (or Z) type cleavages are observed. Several additional cross-ring cleavages are seen following EDD, which are not observed following CAD of this same species. In particular, two X-type cross-ring cleavages are observed ($^{1,5}X_5$ and $^{1,5}X_4$). EDD of doubly deprotonated maltoheptaose also results in extensive cross-ring fragmentation, with both A-type and

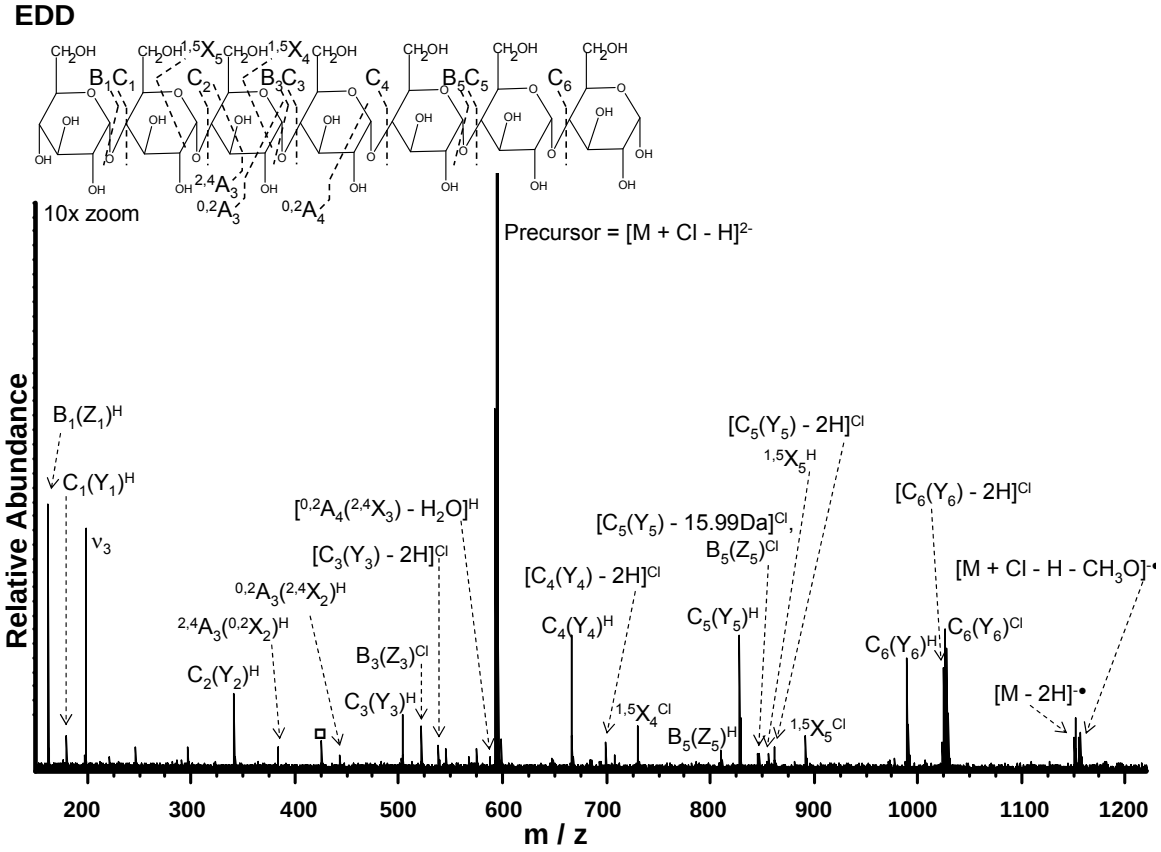


Figure B.2. EDD FT-ICR (30 scans, 3 sec with a bias voltage - 20 eV) tandem mass spectrum of deprotonated, chloride adducted maltoheptaose. Singly charged product ions due to deprotonation are indicated with "H" in the labels, while those containing chloride are indicated by "Cl". Squares indicate water loss from an adjacent product ion. ν_3 indicates a harmonic.

X-type cross-ring cleavages (see Chapter 5). However, following EDD of the doubly deprotonated species, ^{0,2}A and ^{2,4}A-type cleavages are typically more abundant than other types of cross-ring fragmentation. Here, X-type product ions observed following EDD of $[M + Cl - H]^{2-}$ are more abundant than the A-type product ions. These results indicate that the formation of X-type product ions is a preferred fragmentation pathway. Also observed in Figure B.2 are product ions due to hydrogen losses and loss of 15.99 Da from C-type ions. These same types of product ions were reported following EDD of doubly deprotonated maltoheptaose and discussed in Chapter 5.

B.3.2 CAD and EDD of $[M + 2Cl]^{2-}$

The CAD fragmentation behavior of the $[M + Cl]^{2-}$ species of maltoheptaose was examined and is shown in Figure B.3. Following CAD, both singly and doubly charged product ions are observed. Doubly charged product ions due to both chloride adduction and deprotonation are formed via loss of HCl from the precursor (indicated by * in Figure B.3). Doubly charged product ions from the loss of two protons are the result of the loss of two HCl molecules. No product ions are observed which contain two chloride ions. Several singly charged product ions are also observed, including $[M + Cl]^-$ at *m/z* 1187 and $[M - H]^-$ at *m/z* 1151. The former ion is due to loss of a chloride from the precursor, while the latter is due to loss of chloride and HCl. Other singly charged product ions observed following CAD predominantly consist of singly deprotonated species (indicated by H in the label). However, two chloride adducted C-type ions are also detected (indicated by Cl in the label). Overall, the fragmentation behavior of the $[M + 2Cl]^{2-}$ maltoheptaose species is very similar to the doubly deprotonated species (as shown in

Chapter 5), and results in a mixture of glycosidic (C-type) and cross-ring cleavages (^{0,2}A and ^{2,4}A-type ions only).

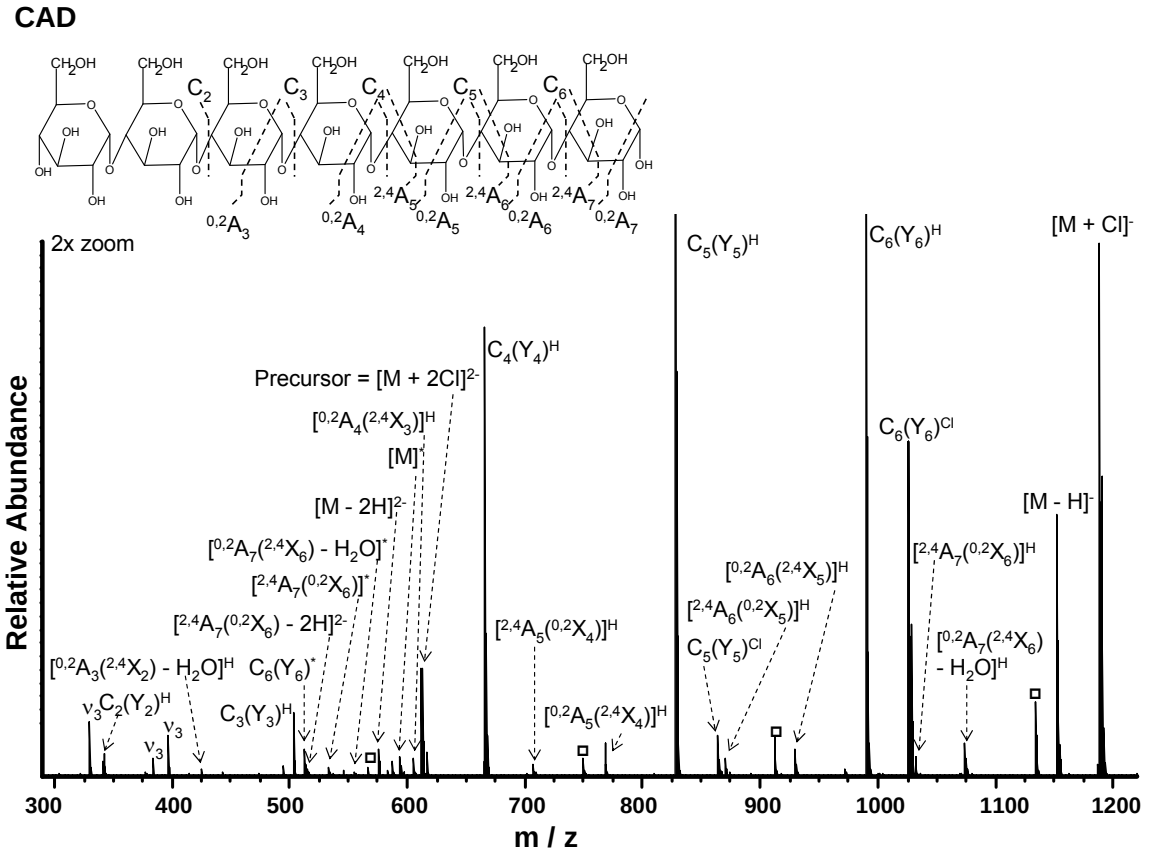


Figure B.3. CAD FT-ICR (30 scans, collision cell voltage 8V) tandem mass spectrum of maltoheptaose adducted with two chloride ions. Doubly charged product ions due to deprotonation and chloride adduction, are indicated with an asterisk. Singly charged product ions due to deprotonation are indicated with "H" in the labels, while those containing chloride are indicated by "Cl". Squares indicate water loss from an adjacent product ion. ν_3 indicates a harmonic.

The EDD fragmentation behavior of the $[M + Cl]^{2-}$ species of maltoheptaose was examined and is shown in Figure B.4. Almost all product ions observed following EDD are singly charged, and are either singly deprotonated or contain a chloride ion. However, two doubly charged product ions are observed, $[M - 2H]^{2-}$ and $[M]^{*}$ (the latter has both a deprotonation site and a chloride ion adduct). Doubly charged product ions likely arise from vibrational excitation or electronic excitation, and not from the EDD

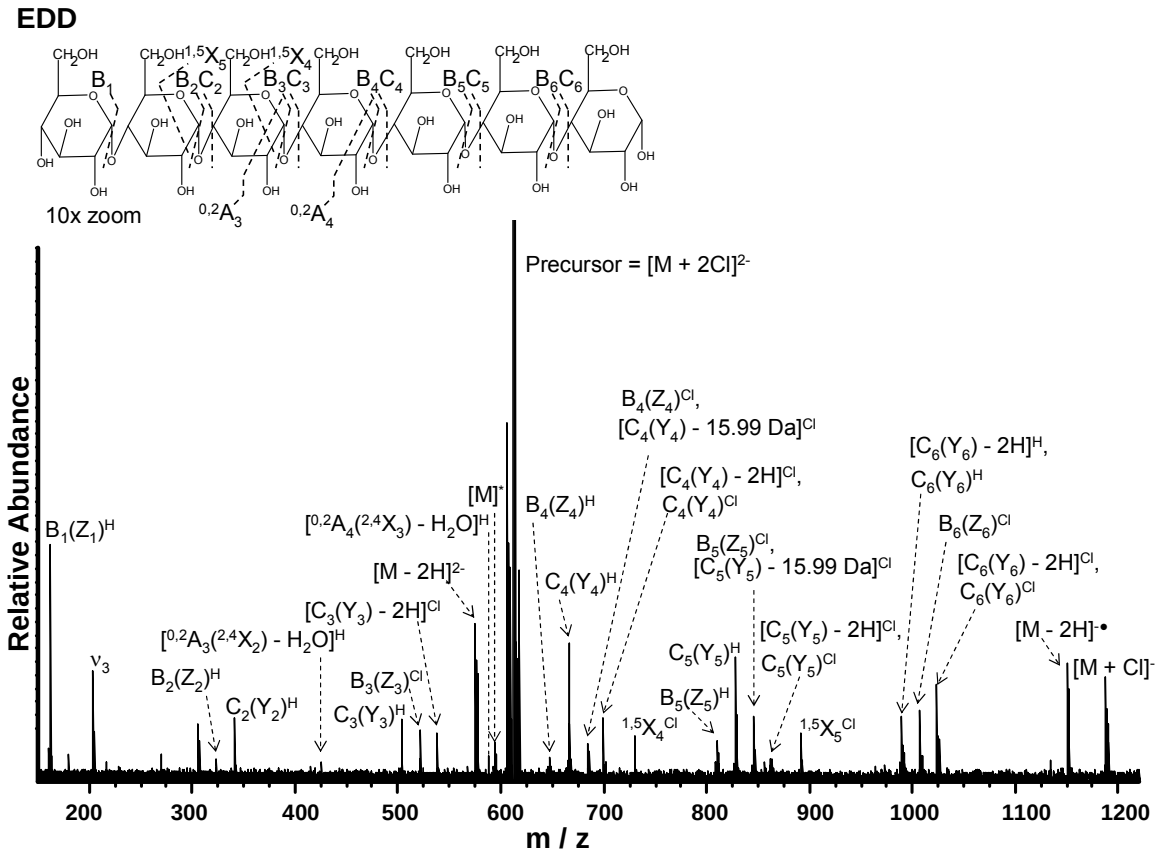


Figure B.4. EDD FT-ICR (30 scans, 5 s with a bias voltage of - 30 eV) tandem mass spectrum of maltoheptaose adducted with two chloride ions. Doubly charged product ions due to deprotonation and chloride adduction are indicated with an asterisk. Singly charged product ions due to deprotonation are indicated with "H" in the labels, while those containing chloride are indicated by "Cl". Squares indicate water loss from an adjacent product ion. ν_3 indicates a harmonic.

process. These same ions are also observed following CAD, implying they are due to vibrational excitation. Both glycosidic and cross-ring cleavages are observed following EDD of this oligosaccharide. Glycosidic cleavage is more extensive compared to CAD of this same species (shown in Figure B.3). A mixture of A- type and X-type product ions are also observed. Similar to EDD of $[M + Cl - H]^{2-}$, the X-type product ions ($^{1,5}X_4$ and $^{1,5}X_5$) are more abundant than A-type product ions. This higher abundance further indicates that the formation of X-type ions is a preferred fragmentation pathway for chloride adducted species. Similar to EDD of doubly deprotonated maltoheptaose and

the $[M + Cl - H]^{2-}$ species, hydrogen losses and losses of 15.99 Da are observed for several C-type ions.

B.4 Conclusions

The fragmentation behavior of the $[M + Cl - H]^{2-}$ and $[M + 2Cl]^{2-}$ forms of maltoheptaose were examined. The CAD fragmentation patterns of these ions are similar to the corresponding doubly deprotonated species. EDD of chloride adducted maltoheptaose results in a mixture of glycosidic and cross-ring fragmentation. For both species examined, cross-ring fragmentation is less extensive compared to EDD of the doubly deprotonated species. However, our results indicate that X-type ions are preferentially formed over A-type ions following EDD of chloride adducted maltoheptaose. This behavior contrasts with data shown in Chapter 5, where ^{0,2}A and ^{2,4}A-type ions were the most abundant cross-ring cleavages observed. To determine if this cleavage preference is generally true for chloride adducted oligosaccharides, further investigation is necessary.

B.5 References

1. Zhu, J.; Cole, R. B. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 1193-1204.
2. Cai, Y.; Jiang, Y.; Cole, R. B. *Anal. Chem.* **2003**, *75*, 1638-1644.
3. Jiang, Y.; Cole, R. B. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 60-70.
4. Yang, J.; Mo, J.; Adamson, J. T.; Hakansson, K. *Anal. Chem.* **2004**, *77*, 1876-1882.
5. Caravatti, P.; Allemann, M. *Org. Mass Spectrom.* **1991**, *26*, 514-518.
6. Tsybin, Y. O.; Witt, M.; Baykut, G.; Kjeldsen, F.; Hakansson, P. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 1759-1768.
7. Dallinga, J.; Heerma, W. *Biol. Mass Spectrom.* **1991**, *20*, 215-231.
8. Li, D. T.; Her, G. R. *Anal. Biochem.* **1993**, *211*, 250-257.
9. Carroll, J. A.; Ngoka, L. C.; Beggs, C. G.; Lebrilla, C. B. *Anal. Chem.* **1993**, *65*, 1582-1587.
10. Chai, W.; Piskarev, V.; Lawson, A. M. *Anal. Chem.* **2001**, *73*, 651-657.
11. Pfenninger, A.; Karas, M.; Finke, B.; Stahl, B. *J. Am. Soc. Mass Spectrom.* **2002**, *13*, 1331-1340.