HYPHENATING ION MOBILITY WITH MASS SPECTROMETRY TO INCREASE THE INFORMATION CONTENT OF TOP-DOWN ANALYSES

A Dissertation

by

NATHANAEL FREDERICK ZINNEL

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	David H. Russell
Committee Members,	Emile A. Schweikert
	Gyula Vigh
	Deborah Bell-Pedersen
Head of Department,	David H. Russell

May 2014

Major Subject: Chemistry

ABSTRACT

Mass spectrometry (MS) has been established as important analytical tool in the characterization of an array of analyte classes, including biological samples. However, without hyphenation with other techniques, the approach has limitations to the information that can be elucidated and the samples that can be analyzed. In an attempt to overcome these limitations, separation is performed prior to MS analysis to aid in alleviating sample complexity while dissociation is incorporated to increase the information content. Here, we employ ion mobility (IM), a gas-phase separation technique, to disperse product ions resulting from collision-induced dissociation (CID), denoted as MS-CID-IM-MS, for top-down analysis for a variety of applications, specifically, primary structure elucidation, disulfide bond identification, secondary structure characterization, and polymer characterization.

First, the fundamental attributes of this approach and the resulting information elucidated are investigated. Using this approach CID product ions are dispersed in two-dimensions, specifically size-to-charge (IM) and mass-to-charge (MS), and the resulting 2-D data display greatly facilitates the top-down information contents; (i) charge state specific trand lines, (ii) increased dynamic range, (iii) separation of overlapping ion signals. The increase in peak capacity allows for detection of low abundant fragment ions providing an increase in the primary sequence coverage and the confidence of ion assignments as demonstrated by melittin and ubiquitin.

Second, this general approach is applied to the top-down analysis for a variety of applications. MS-CID-IM-MS is used for the structural characterization of disulfide linked protein ions by monitoring the ATD of the ion pre- and post-collisional activation. Similarly, this approach can also be used to distinguish product ion type as well as, in some cases, specific secondary structural elements, *viz.* extended coils or helices providing rapid identification of the onset and termination of extended coil structure in peptides as demonstrated by insulin B-chain. Detect of low abundant ion signals associated with cross-ring cleavages allows this approach to be extended to determine regiochemistry of glucose derived polymers. As demonstrated, the MS-CID-IM-MS approach is highly versatile owing to the information content gained upon dispersion of ions in two-dimensions, providing an effective increase in experimental dynamic range as well as providing conformational information.

DEDICATION

For Josie, my ever faithful friend and for my sister Grace, for her love, support, and friendship.

ACKNOWLEDGEMENTS

I would like to thank my graduate research advisor, Professor David H. Russell for his guidance and support throughout the course of this research. He has allowed me the opportunity to pursue my interests and helped me develop as a scientist. Thank you to my entire research group, both past and present, for their help and guidance. I especially would like to thank Dr. Stephanie Cologna, Dr. Pei-Jing Pai, and Dr. Bill Russell for being outstanding mentors and the guidance provided.

Thank you to my parents, siblings, extended family, and friends for their love, support, and encouragement over the years. Their support helped me through the bad times and made the good times that much more sweet.

NOMENCLATURE

MS	mass spectrometry
MS-MS	tandem-mass spectrometry
PTM	post-translational modification
CID	collision-induced dissociation
ESI	electrospray ionization
ECD	electron capture dissociation
ETD	electron transfer dissociation
LC	liquid chromatography
FT-ICR	Fourier-transform ion cyclotron resonance
IEF	isoelectric focusing
IET	isoelectric trapping
MSWIFT	membrane separated wells for isoelectric focusing and trapping
IMS	ion mobility spectroscopy
IM	ion mobility
CCS	collision cross section
CE	capillary electrophoresis
MALDI	matrix assisted laser desorption/ionization
ToF	time of flight
SID	surface induced dissociation
3-D	three dimensional

2-D	two dimensional
MDS	molecular dynamic simulation
HISCU	human iron sulfur complex U
TWIG	traveling wave ion guide
TWIMS	traveling wave ion mobility spectrometry
TW	traveling wave
S/N	signal-to-noise
HPLC	high performance liquid chromatography
S-14	somatostatin 14
TCEP	*tris*(2-carboxyethyl)phosphine
MMTS	methyl methanethiosulfonate
DTT	dithiothreitol
NEM	*N*-ethylmaleimide
%SC	percentage sequence coverage
ATD	arrival time distribution
CD	circular dichroism
NMR	nuclear magnetic resonance
XRD	X-ray diffraction crystallography
He	helium
N_2	nitrogen
HDX	hydrogen-deuterium exchange
IoAc	iodoacetamide

kV	kilovolt
V	volt
Da	Dalton
kDa	kilodalton
m/z	mass-to-charge
MW	molecular weight
THF	tetrahydrofuran
HT	head-to-tail
TH	tail-to-head
HH	head-to-head
TT	tail-to-tail
mM	millimolar

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

CHAPTER I

INTRODUCTION

The last several decades have seen many advances in biological research, including the completion of the genomes for many organisms.[1] The goal of these genome projects was to define the genome of organisms in order to identify and study key genes in normal as well as disease pathways. An organism's genome may facilitate prediction of all possible proteins found in that organism; however, there are several key objections to biological studies solely relying on the monitoring of gene expression of mRNA. The levels of mRNA do not allow for the prediction of the level of protein expression. Moreover, the function of the protein is often controlled by post-translational modifications and the cells natural process of protein maturation and degradation alter the amount of active protein, independent of mRNA level.[2] As protein molecules are the functional molecules of the cell, most biological studies focus on the study of the proteome, the protein compliment of a cell or organism.

The proteome is diverse by comparison to the genome and presents a daunting task to researchers given the vast number of proteins present in a given organism and the dynamic range of protein concentration that is encountered. Early proteomic studies primarily focused on the identification of proteins present in an organism.[3] Current proteomic studies are not only limited to identification of proteins but also to localization, determination of post-translational modifications present, interactions, and quantitation. These studies are further complicated by the temporal changes associated

with all of these characteristics. Thus, proteome research requires a wide dynamic range of detection, high-throughput, sensitivity, and high confidence in protein identification and quantification. Mass spectrometry (MS) fulfills these requirements and has thus become an integral tool in proteomics. Currently, mass spectrometry–based proteomic studies are accomplished using one of two experimental methodologies: "bottom-up"[4] and "top-down."[5] As seen in **Figure 1**, the workflow for protein analysis from both approaches is complimentary.

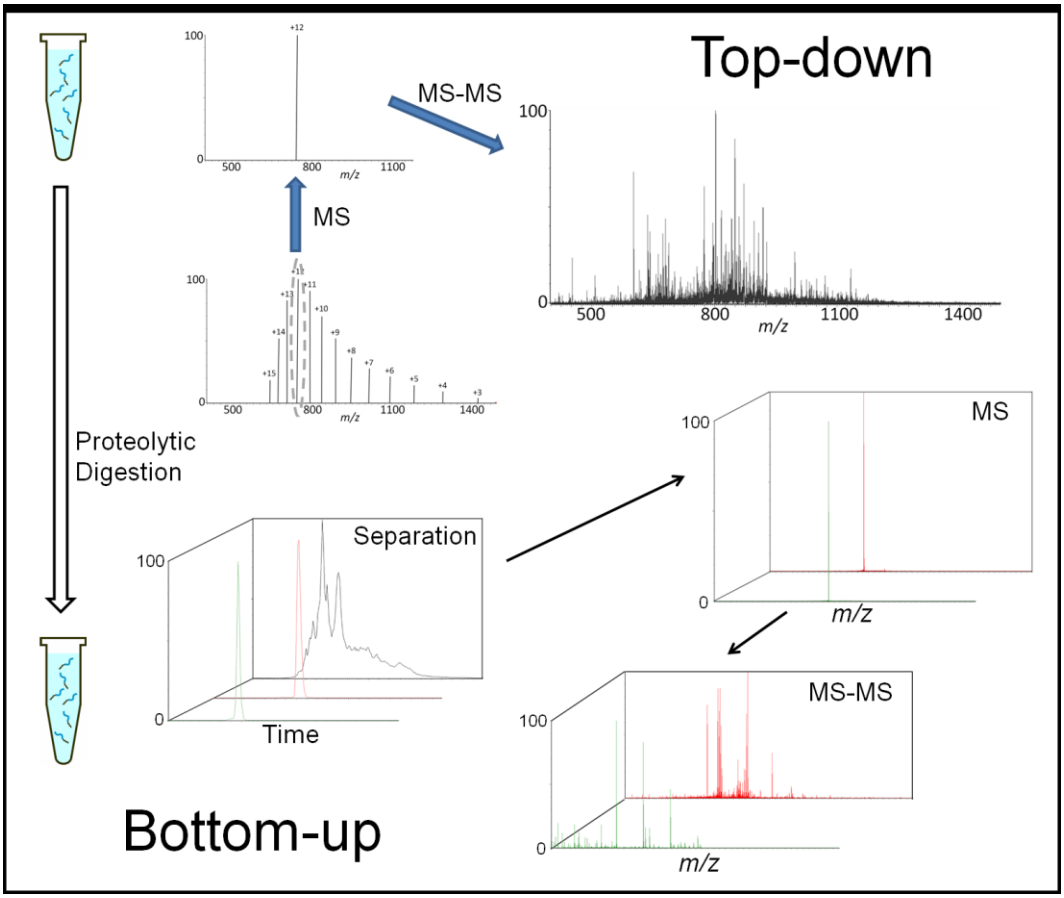


Figure 1. Comparison of top-down and bottom-up workflows.

Proteomics: Bottom-Up

Proteomic studies employing the bottom-up approach are the most common MS-based approach. In bottom-up proteomics, the sample is enzymatically digested with a protease resulting in a mixture of small peptides from the protein. The most commonly utilized proteolytic enzyme for bottom-up proteomics is trypsin. Trypsin is commonly used owing to: (i) its specificity to cleave C-terminal after the arginine and lysine residues which provides a basic amino acid residue charge carrying site, (ii) tryptic peptides commonly contain 6-12 amino acids and thus fall in a mass range which takes advantage of the mass spectrometers sensitivity and resolution, and (iii) *a priori* knowledge of the C-terminal residue assists in tandem MS (MS-MS) interpretation and database searching.

The proteolytically derived peptides are analyzed by MS and/or MS-MS to obtain sequence specific product ions. Using MS to obtain accurate mass measurement for protein identification is referred as "peptide mass mapping" as the *m/z* information is used to "map" the peptides that correlate to a specific protein. To increase confidence in these assignments MS-MS analysis is utilized to obtain peptide sequence information. The resulting mass spectra from MS and/or MS-MS are subsequently subjected to database search algorithms to identify the protein of origin[6]. This method allows for the straightforward sequence analysis of proteolytic peptides, but has inherent short comings such as; (i) extensive sample preparation steps and (ii) resulting proteolytic peptides generally fall into a narrow mass range resulting in greater complexity of the MS

spectrum. Furthermore, many low abundant peptide ion signals, some which may contain information such as identity and location of post-translational modifications (PTMs), may not be selected for MS/MS analysis owing to the complexity associated with the enzymatically digested solution. Additionally, collisionally-induced dissociation (CID) analysis often fails in preserving labile PTMs such as phosphorylation and glycosolation, cleaving off the side-chain substituents and destroying information on their backbone location. Though MS-based bottom-up proteomics are most common, obtaining complete sequence coverage of the protein is difficult to achieve which leads to information loss, especially for post-translational modification (PTM) studies.

Proteomics: Top-Down

The alternative to the bottom-up approach is the top-down method which relies on introduction and dissociation of intact proteins in a mass spectrometer. This approach typically relies on electrospray ionization (ESI) to create and subsequently introduce protein ions into the mass spectrometer. Once protein ions are introduced to the mass spectrometer, several methods exist for the dissociation of intact proteins the most common of which are: CID[7], electron capture dissociation (ECD), and electron transfer dissociation (ETD). As afore mentioned, CID often results in the loss of PTMs and is less often chosen as a dissociation technique compared to ECD and ETD, which are able to conserve PTMs[8,9]. Thus, ECD and ETD represent the main techniques employed for top-down analysis of proteins.

Top-down proteomic methods provide an appealing MS-based analysis technique owing to the need for minimal sample preparation as the top-down approach relies exclusively on fragment ions generated in the gas-phase from the intact protein. Additionally, this approach allows for intact mass measurement and elucidation of PTM identity and location.[10,11] Identification and characterization of a protein of interest can be made on a single ion abundance eliminating the large number of redundant protein identifications associated with bottom-up methods, *i.e.* all product ions observed correlate directly to the protein being analyzed and must be characteristic of the protein's sequence.

However, several challenges are associated with sequencing intact proteins by the top-down approach. The commonly applied top-down fragmentation techniques ECD and ETD suffer from low fragmentation efficiency, as conversion from parent to fragment ions is theoretically capped at 43% for doubly charged ions[12] and often results in fragmentation efficiencies of much less. Thus, to achieve substantial ion intensities, long acquisition times are required resulting in decreased throughput and hinders coupling to liquid chromatography (LC). As a result, high throughput top-down studies have resorted to CID for protein identification.[13,14] Furthermore, these techniques generally require a Fourier Transform-Ion Cyclotron Resonance (FT-ICR) mass spectrometer, an expensive instrument not readily available to many laboratories though recently ETD has become available on many commercial ion trap instruments. Another major problem associated with top-down proteomic analysis is that product ion spectra are typically composed of ions whose charges range from unity to that of the parent ion.

The determination of product ion charge state relies on the accurate mass measurement of isotopes of the product ion. The advantage of high magnetic field FT-ICR-MS is that the isotopic clusters of the various charge states can be separated, thereby facilitating identification of the charge state.[15-17] Thus, fragmentation of intact proteins in ion trap instruments typically results in fragments which lack significant resolution to assign higher order charge states. Interpretation of product ion spectra is further compounded owing to product ion charge state overlap.

Both bottom-up and top-down methods have proven useful for proteomic studies and have also been shown to be complimentary approaches for characterization, determination of amino acid sequence, and identification of post-translational modifications. It is important to note that both techniques have their strengths and limitations but the most daunting challenge shared by both techniques arises from the complexity and dynamic range that is attributed with a proteome sample as the concentration can span more than ten orders of magnitude.[18] Both methods approach this challenge by dispersion of sample in multiple dimensions prior to MS analysis. In the case of bottom-up proteomics, this can involve the electrophoretic separation of proteins in polyacrylamide gels in either one-dimension (molecular weight) or two-dimensions (isoelectric point and molecular weight).[19] The resulting band or spot is excised for digestion and subsequent mass spectrometric analysis.[20] More amenable to top-down analysis are solution phase sample pre-fractionation, though, it is equally applicable to bottom-up proteomic analysis. Several techniques have been used to decrease sample complexity including isoelectric focusing (IEF) which can be carried

out in capillary[21] or in multicompartment electrolyzers such as the OFFGEL™ Fractionator (Agilent)[22] as well as iosoelectric trapping (IET) in devices such as the membrane separated wells for isoelectric focusing and trapping (MSWIFT).[23,24] Typically, chromatographic methods are utilized in proteomic studies to alleviate sample complexity and often are combined with other fractionation techniques.[25] Chromatographic separations rely on interactions of analytes with a stationary phase. Dispersion of sample in another dimension helps to alleviate the complexity of sample prior to MS analysis but increases the analysis time (hours).

Ion Mobility

Similarly, ion mobility spectroscopy (IMS) serves to disperse ions in another dimension, much like solution phase pre-fractionation, and holds potential for use in proteomics. IMS is a gas-phase electrophoretic size separation technique, in which ions are separated based on their ability to traverse a chamber filled with inert neutrals under the influence of an electric field. Larger ions will undergo more collisions and have a larger collision cross-section (CCS) than compact ions with similar mass. This size-to-charge based separation occurs on the order of μs to ms. Thus, IMS analysis rapidly disperses ions in conformational space providing the size-to-charge information of the analytes allowing the technique to be applied to a wide range of analytical applications such as trace level detection of chemical weapons,[26] differentiation of molecular species in complex matrices,[27] and structural determination of hydrocarbons in crude oils.[28,29]

The last several decades have seen a marked increase in the interest of ion mobility, particularly the coupling of ion mobility (IM) with mass spectrometry (MS),

providing a means of mass identifying ions separated by IMS. The hyphenation of ion mobility and mass spectrometry (IM-MS) enhances the information content of the technique by providing data from two dimensions of analysis, specifically the analyte ion's size-to-charge ratio and mass-to-charge ratio. The orthogonality of IM-MS helps to alleviate sample complexity but to a lesser extent that other separation techniques coupled to MS such as LC and capillary electrophoresis (CE) owing to mass scaling near linearly with analyte size. Similarly, the low orthogonality of IM-MS will result in lower peak capacity than other separation techniques owing to analyte signals being clustered around size-mass correlation line rather than dispersed across the full area of analysis. Though the peak capacity of IM-MS is low comparatively, the speed of separation and the potential for information driven analysis is what drives IM-MS as an analytical tool.

Though IMS has been applied to a wide range of analytical applications, IM-MS has found a niche in biological analyses. IMS separations are based on a fundamental physical parameter of the analyte, the ion collision cross-section (CCS), allowing it to be utilized to study ion shape. A key component required in both mass and mobility spectrometers is the ability to create gas phase ions, thus, improvements in sample ionization techniques have promoted the study of biological molecules by IM-MS. Of paramount importance was the development of the soft ionization techniques of matrix assisted laser desorption ionization (MALDI)[30] and electrospray ionization (ESI).[31-33] Both MALDI and ESI have facilitated the introduction of large, nonvolatile molecules such as peptides and proteins into the mass spectrometer making them ideal ionization

methods for proteomic studies. Moreover, ESI not only produces multiple charge states for samples but also samples the analyte directly from solution resulting in ESI being more amenable to coupling with liquid chromatography (LC).

It is only in the last decade that a commercialized IM-MS platform was developed. Early studies employing IM-MS were performed on home-built instrumentation. These studies primarily employed IM as a structural probe of peptides as well to differentiate different biological molecule classes. Work by both Clemmer[34] and Gillig/Ruotolo[35,36] introduce the classical element of proteomics, confirmation of analyte sequence through tandem analysis (MS-MS) of the analyte. Clemmer *et al.* investigated the information content obtained from LC-IMS-CID-MS analysis of a mixture of proteins. The sample is first separated by LC prior to introduction to the instrument platform's drift tube. The LC-IMS separated ions are subjected to time-of-flight (ToF) to determine *m/z* values of protein parent ions. In addition, collisional activation could be carried out post-IM to allow dissociation product ions to be correlated with the specific charge states of the analyte ions in the form of "fragment lanes." Gillig/Ruotolo employed a similar instrument design incorporating surface induced dissociation (SID) post-IM separation to study peptides formed by MALDI. As in Clemmer's studies, the product ions are associated by total mobility drift time to the parent ion, *i.e.* "fragment lanes" (**Figure 2A**). Both studies underscore the increased peak capacity and gain in information content obtained from hyphenating IM and MS for proteomic research.

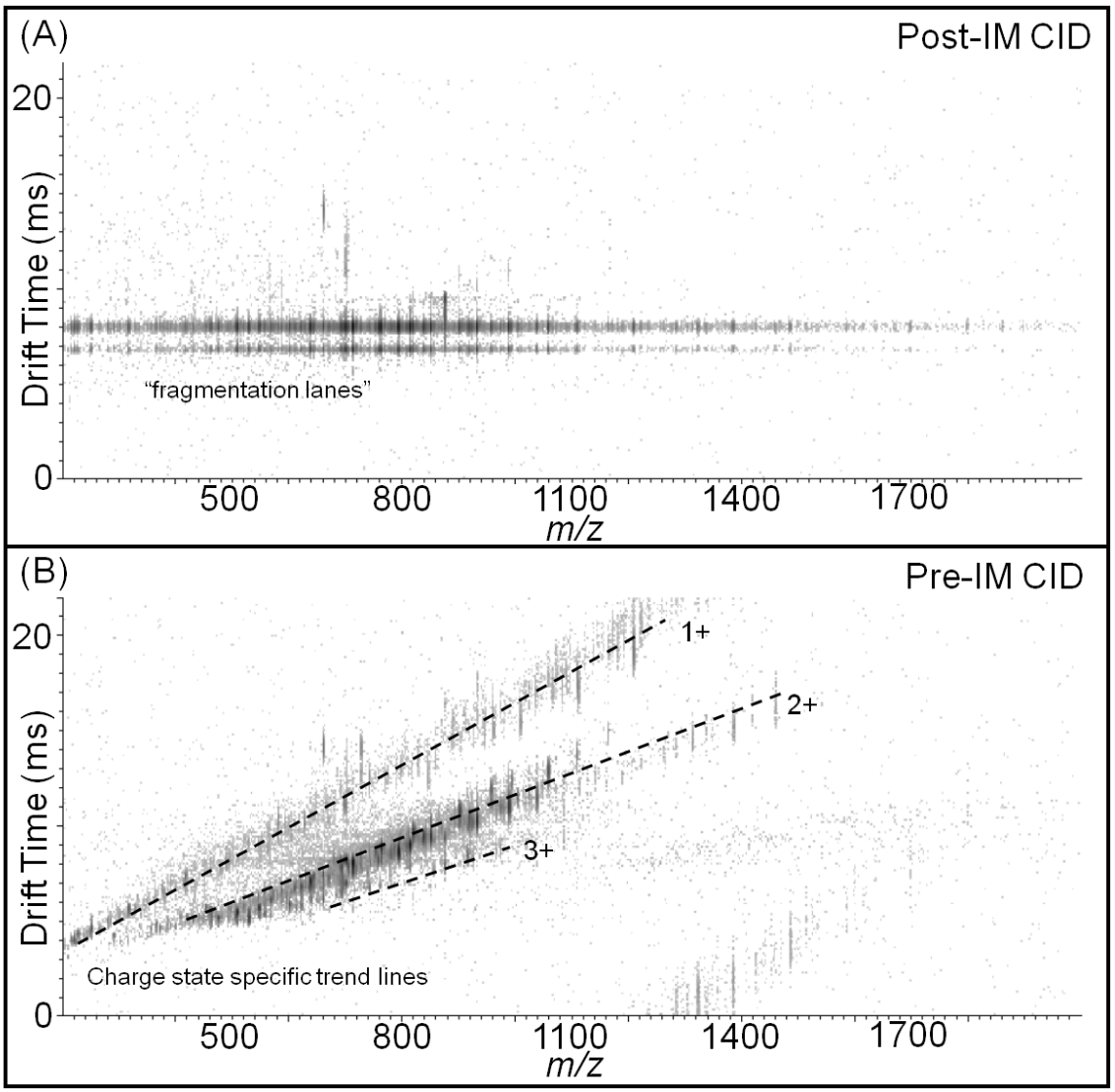


Figure 2. CID dissociation of the $[M + 5H]^{5+}$ ion of the B-chain of insulin (A) post-IM and (B) pre-IM separation. The same product ions are observed in both spectra, however, the dispersion in the mobility dimension are different. (A) Post-IM CID results in product ions clustered at the drift time of the parent ion while (B) pre-IM CID allows product ions to also be dispersed in conformational space.

A similar experiment can be performed utilizing a commercially available mass spectrometer, the Synapt G2 HDMS (Waters Corp., Milford, MA, USA), in which mass selection and collisional activation are performed prior to ion mobility separation. Much

like post-IM dissociation, this technique allows all product ions to be associated with the parent ion as a single isotope cluster is mass selected for analysis. However, post-IM dissociation solely disperses product ions in terms of *m/z* as product ions all have the mobility drift time associated with the parent ion; limiting sequence coverage and detection of protein modification owing to the high complexity of the product ion spectrum from this approach. Dissociation prior to mobility separation allows the product ions to be separated by size-to-charge ratio and mass-to-charge ratio forming charge state specific trend lines as seen in **Figure 2B**. This increased dispersion of product ions provides reliable detection and assignment of low abundant CID product ions, rapid assignment of CID product ion charge states, providing greater sequence coverage and rapid identification of protein modifications.

Disulfide Analysis

The detection of post-translational modifications (PTMs) has become an increasingly important endeavor in proteomics. The identity of PTMs, localization, and temporal abundance are required to fully characterize a protein as well as gain understanding to protein and cellular function. For example, phosphorylation is used as a key signaling technique to propagate signals in a cell.[37,38] More recent studies have shown that the phosphorylation and subsequent dephosphorylation can influence the structure and protein-protein interactions.[39,40] Similarly, other PTMs such as cysteine oxidation and disulfide bonds can influence peptide/protein three-dimensional structure as well and thus the functionality of the molecule.[41-46] The formation of disulfide bonds is unique to the cysteine residues and the result of two cysteine residues undergoing

oxidation allowing a covalent bond to form between the two residues. These intra- and inter-disulfide bonds play a large part in stabilizing the overall structure of peptides/proteins[42,45,46] are often encountered in biological systems, found in toxins,[47] hormones,[48] and defensins.[49] Determining the presence of the disulfide bond as well as disulfide bond connectivity is complicated but an important task in the characterization of isolated and recombinantly expressed proteins.

MS has played an increasingly important role in the analysis of disulfide bonds, used to identify not only the disulfide bond presence but also their connectivity. Identifying cysteine residues that are involved in the formation with disulfide bonds with in a protein can be accomplished by a variety of MS-based methods: tandem mass spectrometry (MS-MS),[50,51] chemical reduction,[52,53] chemical oxidation,[54,55] and metal ion cleavage of disulfide bonds.[56,57] The most commonly employed MS-based approach for disulfide bond analysis involves the reduction of the disulfide bond and alkylation of the thiol providing a characteristic mass shift.[58-61] The reduced and non-reduced proteins are subjected to proteolytic digestion and the resulting peptides are then compared. Though development of advanced separation science and MS methodologies have been applied to identify disulfide bonds, this basic experimental strategy remains similar defining this as the classical approach to determining disulfide bond presence and connectivity.

In order to maintain the native disulfide bond network, several experimenatal variables must be careful controlled to prevent disulfide scrambling.[62-64] Several factors are known to promote disulfide bond rearrangement, *i.e.* the presence of free sulfhydryl

groups at neutral or alkaline pH. The disulfide bonds can undergo nucleophilic cleavage under alkaline conditions allowing them to react with a free thiol group to produce a scrambled/exchanged disulfide bond.[65] This phenomenon becomes an issue for the classical MS-based approach as the proteases commonly used in MS-based proteomics (chymotrypsin, trypsin, and endoproteinase AspN) are active at or near pH 8. This issue can be overcome through the use of alternate proteases such as pepsin that are active at acidic conditions (pH 2-3) or by performing the digestion under mildly acidic conditions (pH 6).[66]

The classical approach is hindered by the complexity of the experiment scheme. Additional reagents are required for scission of the disulfide bond as well as to alkylate the reduced cysteine residue, all of which occurs prior to enzymatic digestion creating a time consuming process with the potential for sample loss. Owing to these experimental factors and the disulfide bond's high electron affinity, disulfide bonds are being more often studied by top-down proteomic methods, specifically electron-mediated dissociation techniques (ECD and ETD).[15,67,68] In a top-down approach, intact protein ions enter into the mass spectrometer which are subsequently dissociated, typically either by electron capture dissociation (ECD)[16,69] or electron transfer dissociation (ETD)[70], into product ions, thus removing the need for enzymatic digestion. McLafferty[67] and McLuckey[71] demonstrated that ECD and ETD not only dissociates the backbone bonds but also results in dissociation of the disulfide bond. Low-energy CID provides an alternate dissociation technique for top-down analysis which has better efficiency than ECD/ETD. However, in disulfide bond containing peptides/proteins,

low-energy CID often generates little product ions from backbone cleavages between residues enclosed by a disulfide bond owing to the need to cleave additional bonds.[67,71-73] The limited sample preparation and fast experiment analysis allows top-down methods to circumvent the issues associated with bottom-up; *i.e*. analysis time, artificial modifications, and loss of information due to poor peptide recovery during sample processing.[74,75] Top-down methods also allow forgoing of reduction and thus provide the opportunity to characterize the connecting patterns of disulfide bonds.

The existence of intra-chain disulfide bonds cyclizes the peptide/protein backbone resulting in a more compact conformation. Therefore, cleaving the disulfide bond or the backbone to create a linear analog of the peptide or protein in the gas-phase is of importance as small differences in conformation of gas-phase ions can be distinguished using IM-MS. IM-MS can potentially differentiate between cyclic and linear peptides.[76-78] Previous studies by our group on PTM peptides demonstrate the potential of IM as a screening tool for PTMs.[79] Taking advantage of the phenomenon that low-energy CID produces a single cleavage and retains the disulfide bond intact, IM can be used to identify the presence of disulfide bonds in proteins. The changes in secondary/tertiary structure associated with the opening of cyclic disulfide peptides and protein loops provide a characteristic that can be monitored by IM-MS resulting in a rapid screening technique for the identification of peptides and proteins containing disulfide bonds.

Secondary Structure

Strong covalent bonds link amino acid residues into linear chains to comprise the proteins primary structure, however, the overall three-dimensional (3-D) structure of peptides and proteins is determined by weaker intramolecular interactions, both hydrophobic and hydrophilic interactions, as well as solute-solvent interactions. These weak interactions result in the native structure of a protein being comprised of multiple secondary structure elements and it is the organization of these elements in relation to one another that governs the 3-D structure of the protein.[80] In nature, the structures of proteins are often not the most stable structure so that proteins can be regulated or even dismantled in response to the life cycle of the cell or to an external stimulus.[81] Thus, the mechanism of protein folding has been the subject of much research.[82-84]

Work by Gianni *et al.* indicates that the stability of secondary structure influences the mechanism of protein folding.[85] Thus, understanding factors that affect secondary structure can allow insight into the factors that influence protein folding. The most common structural motif in proteins, and the most important secondary structural element in transmembrane proteins,[86] is the helix.[87-89] The α-helix is stabilized by backbone hydrogen bonds between the amide hydrogen and the carbonyl oxygen four residues away (termed i, i+4, respectively). Much of our understanding of helices and subsequently protein folding is based on experiments performed in aqueous environments. However, aqueous solution is not the only biologically important environment. Researchers have limited understanding of the hydrophobic and low dielectric environments that transmembrane proteins encounter in lipid bilayers.[90]

Research has shown that the stability of helical conformations depends on solvent but the specific role of the solvent is poorly understood.[91] Separately studying the intramolecular and solvent interactions would help to improve the role these interactions play at a fundamental level in the stabilization of structure.

Mass spectrometry (MS) provides a means to delineate intramolecular interactions and solvent interactions as MS can allow biological molecules to be transferred to the gas phase as anhydrous ions. Incorporation of IMS allows structural information to be obtained on these gas-phase anhydrous ions in the form of CCS measurements. These CCS measurements are often combined with molecular dynamic simulations (MDS) to assign a population of possible structures to IM-MS data. Gas-phase peptide ions predominantly form random coil conformations owing to charge salvation. The mobility-mass plot data obtain from IM-MS experiments allows random coil trend line to be defined. Positive and negative deviations from this predicted random coil trend line are then attributed to changes in secondary structure, *i.e.* helical conformations,[78,92] post-translational modifications, or peptide-small molecule interactions.[93]

Similar to the dispersion of different molecular types to form compound class specific trend lines, [27] CID product ions form charge state specific trend lines. Deviations from these charge state specific trend lines can be used to characterize not only product ion structure but also the structure of the parent ion. However, this structural characterization technique is dependent upon the gas-phase stability of the secondary structure; specifically, helices. Jarrold and co-workers used IM-MS in a

series of experiments to study the gas-phase stability of unsolvated helices from model peptides.[94] The helical peptide $[Ac\text{-}A_{15}K + H]^+$ and random coil peptide $[Ac\text{-}KA_{15} + H]^+$ were studied by Jarrold *et al.* as a function of temperature using a high-temperature (500 K) IM drift tube. As temperature is increased, neither the $[Ac\text{-}A_{15}K + H]^+$ helix nor the $[Ac\text{-}KA_{15} + H]^+$ random coil melt into a collapsed globular structure below the threshold for dissociation. Further studies by Jarrold *et al.* indicate that mobile protons as well as drift tube temperature (≥500 K) can influence the gas-phase stability of helices in designed peptides resulting in the "melting" of helices.[95,96] Jarrold's work indicates that the combined contributions of the non-covalent interactions stabilizing the helical conformation in the gas phase exceed the energy required to dissociate a covalent bond.

Polymer Characterization

Both natural and synthetic polymers play an essential and ubiquitous role in everyday life owing to their broad range of properties. Polymer materials, particularly synthetic, are used throughout the industrial world for a large variety of purposes owing to the wide range of properties that these materials can exhibit. More recently this demand for polymer materials has driven the development of synthetic polymers derived from renewable feedstocks resulting in structural and compositional complexity. As a result, polymer characterization is critical in understanding the relations of chemo-physical properties of these new synthetic polymers and their function properties.[97]

These new synthetic polymers can be analyzed by a variety of analytical techniques (*i.e.* separation techniques,[98,99] spectroscopic methods,[100,101] nuclear magnetic resonance (NMR) spectroscopy, X-ray,[102] microscopy,[103] and mass spectrometry

(MS)[104,105]) to provide information in regards to the chemical and physical properties. However, spectroscopic techniques as well as NMR provide information on the overall properties of the polymer system resulting in a combination of analytical techniques being required to fully characterize both individual polymers as well as the polymer system. To this end, MS has been gaining popularity and is quickly becoming an indispensable tool for polymer analysis. MS combined with soft ionization techniques provides the molecular mass and molecular mass distribution of polymers.[105] The information content can be further increased with the addition of MS-MS techniques which provides information on polymer composition[106] (including information about end groups)[107], structure[108], and can aid in understanding the fragmentation mechanism.[109]

One renewable feedstock for the production of synthetic polymers is that of glucose. Recently, the synthetic route of glucose based monomers to create polycarbonates with well-defined end groups has been demonstrated by Wooley *et al.*[110] Though similar to oligosaccharides, the polymers produced by this synthetic route are linear and do not branch. However, structural studies of oligosaccharides require not only composition and sequence analysis but also the assignment of branching and linkage.[111] Full characterization of a glucose derived polycarbonate polymer will require the same information. Thus, similar MS approaches to the analysis of oligosaccharides can be used to analyze these glucose derived polycarbonate polymers.

Studies have revealed that protonated oligosaccharides predominantly result in glycosidic cleavages providing sequence but not linkage information resulting in many oligosaccharide studies being performed using metal adducts to form ions. The

dissociation of oligosaccharide ions generally produces two main types of cleavage: glycosidic cleavages between neighboring residues and cross-ring cleavages involving the dissociation of the sugar ring. Glycosidic cleavages provide information related to composition and sequence while cross-ring cleavages provide information on the different linkage types. The most widely available dissociation technique employed is that of low-energy CID.[112,113] Low-energy CID analysis of oligosaccharides predominantly results in glycosidic cleavages but does produce some cross-ring cleavages, though in low abundance.[114] By comparison, high-energy CID produces a greater abundance of cross-ring cleavages providing more structural information.[115] More recently, the top-down dissociation techniques of ECD and ETD have been applied to the structural analysis of oligosaccharides.[116-118] These studies showed that cross-ring cleavages were the dominant fragmentation pathway, similar to results observed by high-energy CID. Though commonly applied to investigations in proteomics and glycoproteomics, few studies exist utilizing ETD for the structural analysis of oligosaccharides.

CHAPTER II

ION MOBILITY-MASS SPECTROMETRY (IM-MS) FOR TOP-DOWN PROTEOMICS: INCREASED DYNAMIC RANGE AFFORDS INCREASED SEQUENCE COVERAGE*

Introduction

Protein characterization, *i.e.,* identification, sequence analysis, and determination of modifications, is an essential part of mass spectrometry (MS) based proteomics; consequently, there has been a huge investment in the development of new methods and instrumentation that enhance the information content gained from MS analysis. Although a major contributor to the rapid advances in macromolecule MS can be traced to advances in ionization techniques, comparable gains have been achieved by increasing peak capacity afforded by combinations of separation, including pre-MS and MS-MS approaches, and high resolution MS. Much of this development has emphasized so-called "bottom-up MS" approaches, which involves the isolation and enzymatic digestion of the protein(s) to generate a set of peptides, which are analyzed by MS and/or MS-MS. The resulting mass spectral data are then searched against a protein database to identify the protein.[6,119,120] This method allows for the straightforward sequence analysis of proteolytic peptides, but there are drawbacks for this approach.

---

* Reprinted with permission from Zinnel, N. F.; Pai, P. J.; Russell, D. H. Ion Mobility-Mass Spectrometry (IM-MS) for Top-Down Proteomics: Increased Dynamic Range Affords Increased Sequence Coverage. *Anal Chem* **2012**, *84*, 3390. Copyright 2012 American Chemical Society.

Obtaining complete sequence coverage of the protein is difficult to achieve which leads to information loss, especially for post-translational modification (PTM) studies.

Although bottom-up proteomics has rapidly evolved as a high throughput, ultra-high sensitivity approach, the development of top-down approaches has moved rather quickly owing to the development of instruments designed specifically for such purposes as well as improved efficiencies of CID, electron-capture and electron-transfer dissociation (ECD and ETD) techniques. Top-down proteomic approaches are appealing owing to the reduction in sample preparation, the potential for high throughput and simplification of database searches.[10,72,119,121,122] Additionally, this approach allows for intact mass measurement and elucidation of PTM identity and location.[10,11] There are, however, significant challenges associated with sequencing intact proteins by using top-down MS, *i.e.,* the fragmentation efficiency of both ECD and ETD are low.[12] The low fragmentation efficiency can be compensated by using longer acquisition times, but the cost is decreased sample throughput and reduced compatibility with liquid chromatography (LC). CID, ECD and ETD fragment ion spectra are highly complex; typically composed of ions having charge states ranging from singly charged species to that of the parent ion. The advantage of high magnetic field FT-ICR-MS is that the isotopic clusters of the various charge states can be separated, thereby facilitating identification of the charge state.[15-17]

New top-down proteomic approaches have been proposed, which address issues of fragmentation efficiencies and ambiguities associated with protein sequence assignments. For example, McLuckey *et al.* have demonstrated that ion/ion proton-

transfer reactions on multiply charged product ion populations can convert the product ions to predominantly singly charged ions[123,124], which facilitates sequence information assignment of the product ions. Clemmer *et al.* developed an ion mobility approach that is compatible with on-line LC, *i.e.,* nano-LC-ESI-IM-CID-MS, for top-down analysis of protein mixtures[34]. For high throughput top-down studies such as Clemmer's, researchers have resorted to CID for protein identification.[13,14] IM separation was used for protein separation and CID was performed after IM separation. This approach allows CID fragment ions to be correlated with specific charge states of the analyte ions; however, the limited sequence coverage and the difficulty of protein modification analysis are not addressed by this approach due to the high complexity of the product ion spectrum. Longhi *et al.* have described a similar approach for probing the microheterogeneities for wild-type and K496C variant of recombinant X domain of the measles virus phosphoprotein (~6.5 kDa).[125] Here, a general top-down ion mobility approach is described. Ions formed by ESI are separated on the basis of *m/z*, the *m/z* selected ions are subjected to CID, the CID product ions are analyzed by ion mobility (size-to-charge), and the ions are then analyzed by a high resolution TOF mass spectrometer. The advantage of this approach is that it provides reliable detection and assignment of low abundance CID product ions, rapid assignment of CID product ion charge states, which provides greater sequence coverage and rapid identification of protein modifications.

Experimental

Sample Preparation

All solvents (HPLC grade), chemicals, and proteins were purchased from Sigma-Aldrich (St. Louis, MO, USA) and used as supplied unless otherwise stated. ZipTips™ with C18 resin were purchased from Millipore (Millipore, Bedford, MA, USA). Stock solutions of melittin and ubiquitin were prepared in water. For MS analysis, the concentrations of melittin and ubiquitin were 2 μM in 50% water:50% methanol solution containing 0.1% formic acid. Methyl esterification of carboxy- containing amino acids in ubiquitin was carried out as previously described[126]. The samples of human iron sulfur cluster U (HISCU) and the K101A mutant were prepared by the Barondeau group.[127]

Mass Spectrometry

All MS-CID-IM-MS experiments were performed using a SYNAPT G2 HDMS mass spectrometer equipped with a nano-electrospray ion source and MassLynx data processor (Waters Corp., Milford, MA, USA).[128] Instrument acquisition parameters used were as follows: an inlet capillary voltage of 1.85 kV, a sampling cone setting of 40 V, and a source temperature of 100°C. The argon pressure in the traveling wave ion guide trap (TWIG-trap) and the traveling wave ion guide transfer (TWIG-transfer) were 2.44×10^{-2} mbar and 2.61×10^{-2}, respectively. The wave height, the wave velocity, and the nitrogen pressure in the traveling wave (TW) IM drift cell were 32.0 V, 850 m/s, and 2.96 mbar, respectively. Samples were directly infused into the mass spectrometer at a rate of 0.5-0.8 μL/min. All IM-MS acquisitions were acquired for 2 mins.

All fragmentation was carried out by collision of ions with background gas (argon) and collision energy for CID of the proteins was optimized for each protein as well as each charge state, with trap collision voltage generally falling in a range of 23-50 V. CID can be potentially performed in three different regions of the instrument: the source, the TWIG-trap prior to the TW-IM drift cell, and the TWIG-transfer subsequent to the TW-IM drift cell. Precursor mass selection can only be performed in the quadrupole prior to the IM drift cell.

Data Analysis

All data processing was conducted using Waters MassLynx v4.1 and DriftScope v2.1. Product ion spectra were manually interpreted. To expedite manual interpretation and sequence assignment of mass spectra, an arbitrary cut-off threshold of 10% relative abundance was used in peak assignment for MS-CID-MS experiments as well as for extracted product ion spectra from MS-CID-IM-MS experiments. Internal fragment ions with a S/N >3 and an isotope cluster were assigned based on theoretical peak lists generated from ProteinProspector MS-Product (UCSF, San Francisco, CA, USA). Internal calibration based on *b*- and *y*-type ions mass as well as external calibration of the instrument were utilized to make accurate mass assignments. The percentage sequence coverage (%SC) is defined in Eq. (1):[129]

$$\%\mathrm{SC} = \frac{\text{observed amide bond clevages}}{\text{total amide bonds in the protein}} \times 100 \qquad (1)$$

Results and Discussion

Top-down proteomic approaches based on collision-induced dissociation (CID), electron-capture dissociation (ECD) and electron-transfer dissociation (ETD) yield a diverse population of product ions. Owing to the size and number of charges carried by large peptides and proteins product ion spectra (CID, ECD and ETD) are quite congested, which complicates data interpretation and detection of low abundant ion signals. Here, we report a general top-down ion mobility-mass spectrometry approach, denoted as MS-CID-IM-MS that disperses product ions in terms of both mass (*m/z*) and mobility (size-to-charge). This approach increases the experimental dynamic range and amino acid sequence coverage. For example, analyte ions that are mass-selected by MS-1 are dissociated by CID in the TWIG-trap (see **Figure 3A**) and the CID product ions are then analyzed by TW-ion mobility (separation based on size-to-charge) and *m/z* by the TOF mass analyzer. Previous reports combining CID with IM-MS were aimed at fragmentation of mobility-separated ions;[34] however, using this approach information regarding the size-to-charge of the fragment ions is lost. The size-to-charge information can be used in two ways: (i) to gain information concerning the conformation of the product ions[34,130,131] and (ii) to disperse the product ions along charge-state specific trend lines.[132] Here, we emphasize the use of the latter, *i.e.,* the mobility analysis of the CID product ions disperse the product ions on the basis of charge state, which simplifies the interpretation of the "top-down" product ion spectra.[16]

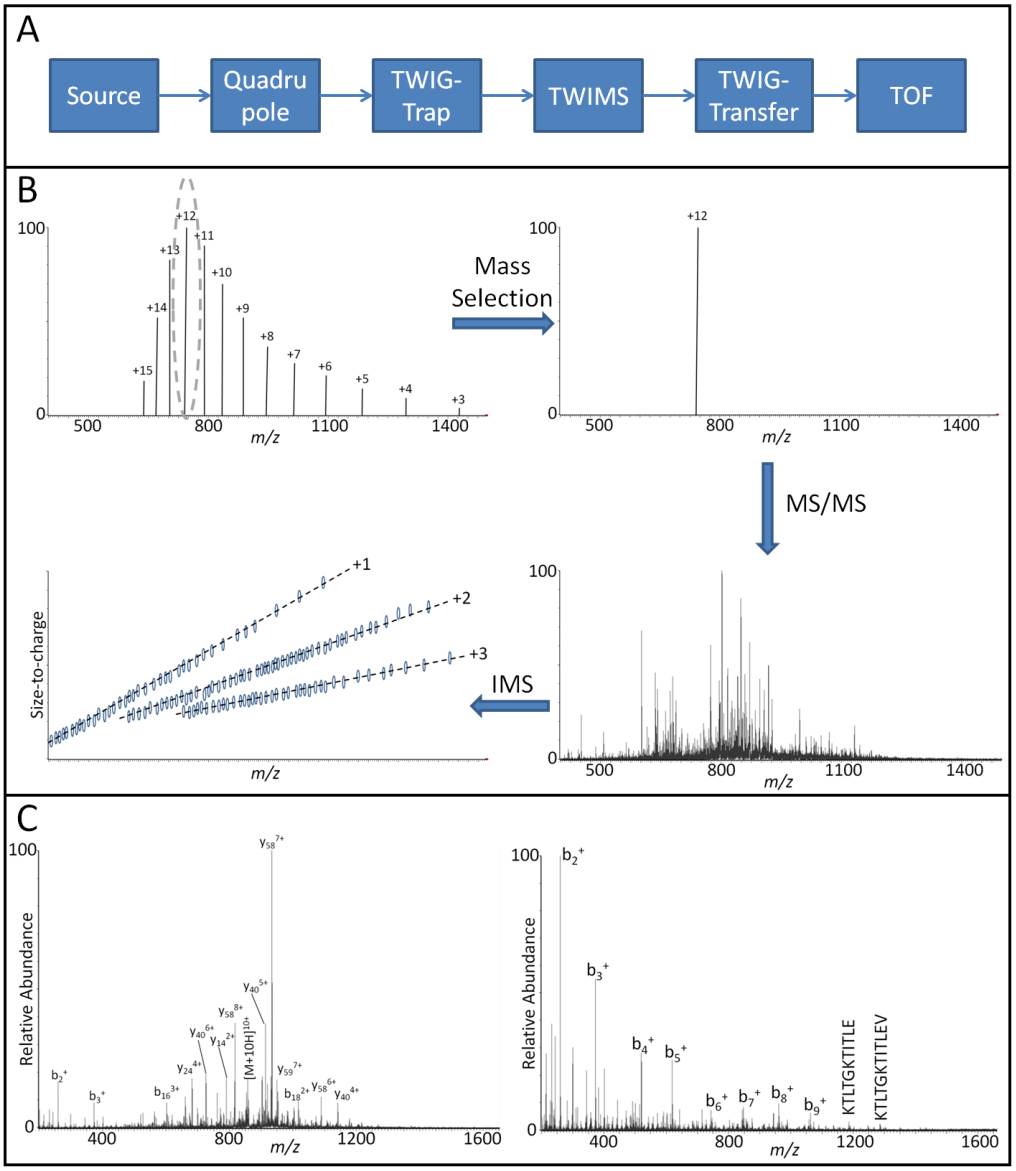


Figure 3. (A) Block schematic of the SYNAPT G2 HDMS mass spectrometer. CID can be performed in three different locations of the instrument: the source, the TWIG-trap, and the TWIG-transfer region. (B) MS-CID-IM-MS experimental workflow. The charge state envelope of a protein is detected by using normal *m/z* scans. A single charge state is isolated in the quadrupole and subjected to CID analysis in the TWIG-trap. The CID product ions are separated by ion mobility providing dispersion on the product ions in two dimensions (size-to-charge and mass-to-charge), which yield distinct charge-state trend lines.. (C) The left mass spectrum is a product ion spectrum obtained from MS-CID-MS of the $[M+10H]^{10+}$ ubiquitin ion and the right mass spectrum shows the extracted singly charged product ion spectrum. Note the reduction in spectral congestion and corresponding increase in experimental dynamic range.

The MS-CID-IM-MS workflow is illustrated in **Figure 3B**. The charge state envelope for the analyte protein is obtained by ESI-MS; a single charge state is mass-selected and collisionally activated to yield a broad m/z range of product ions, which are dispersed along charge-state specific trend lines in mobility-mass space, *i.e.,* a two-dimensional plot (2-D) correlating the product ions size-to-charge and *m/z*. Individual charge-state specific trend lines can be extracted to produce a charge-state specific mass spectrum, as illustrated by the spectrum shown in **Figure 3C**. It should be noted that the charge-state specific product ion spectra can be assembled by high-resolution high magnetic field FT-ICR-MS; however, product ions are dispersed along a single dimension, *i.e., m/z*, and the dynamic range for ion detection is limited.

CID-MS and CID-IM-MS Analysis of Model Peptide/Proteins

The utility of MS-CID-IM-MS for top-down protein identification was evaluated using several well-studied peptides and small proteins. For example, melittin (GIGAVLKVLTTGLPALISWIKRKRQQ-NH_2), a 2.8 kDa amphipathic polypeptide found in bee venom,[133,134] has been studied extensively using several dissociation techniques.[135-137] The ESI mass spectrum of melittin contains abundant signals corresponding to the $[M + 3H]^{3+}$ (*m/z* 949.3) and $[M + 4H]^{4+}$ (*m/z* 712.2) ions and low abundance signals for the $[M + 5H]^{5+}$ (*m/z* 570.0) ion. MS-CID-MS spectra for each of the three charge states demonstrate the proline effect[138-141] in which the most dominant product ions are formed by cleavage N-terminal to proline resulting in the formation of the y_{13} ion. Dissociation of the $[M + 3H]^{3+}$, $[M + 4H]^{4+}$, and $[M + 5H]^{5+}$ ions via CID in

the TWIG-trap results in an amino acid sequence coverage of 42.3%, 38.5%, and 7.7%, respectively.

The MS-CID-IM-MS significantly increases the amino acid sequence coverage of peptides. For example, the trend lines for singly, doubly, and triply charge state CID product ions (shown in the 2-D plots **Figure 4A**) of the $[M + 3H]^{3+}$ ion contain *y*-type ions ranging from y_3 to y_{24} and *b*-type ions from b_3 to b_{12}. These two ion series yields a sequence coverage of 84.6%, which represents 10 additional unique amide bond cleavages. In the case of the $[M + 4H]^{4+}$ and $[M + 5H]^{5+}$ ion the MS-CID-IM-MS data yield sequence coverage of 65.4% and 69.2%, respectively; *y*-series coverage from y_9 to y_{25} and *b*-series coverage from b_3 to b_{13} for the $[M + 4H]^{4+}$ ions and y_8 to y_{25}, a complete *b*-series coverage from b_3 to b_{13}, a b_{15}^{2+} ion, and a b_{16}^{2+} ion for the $[M + 5H]^{5+}$.

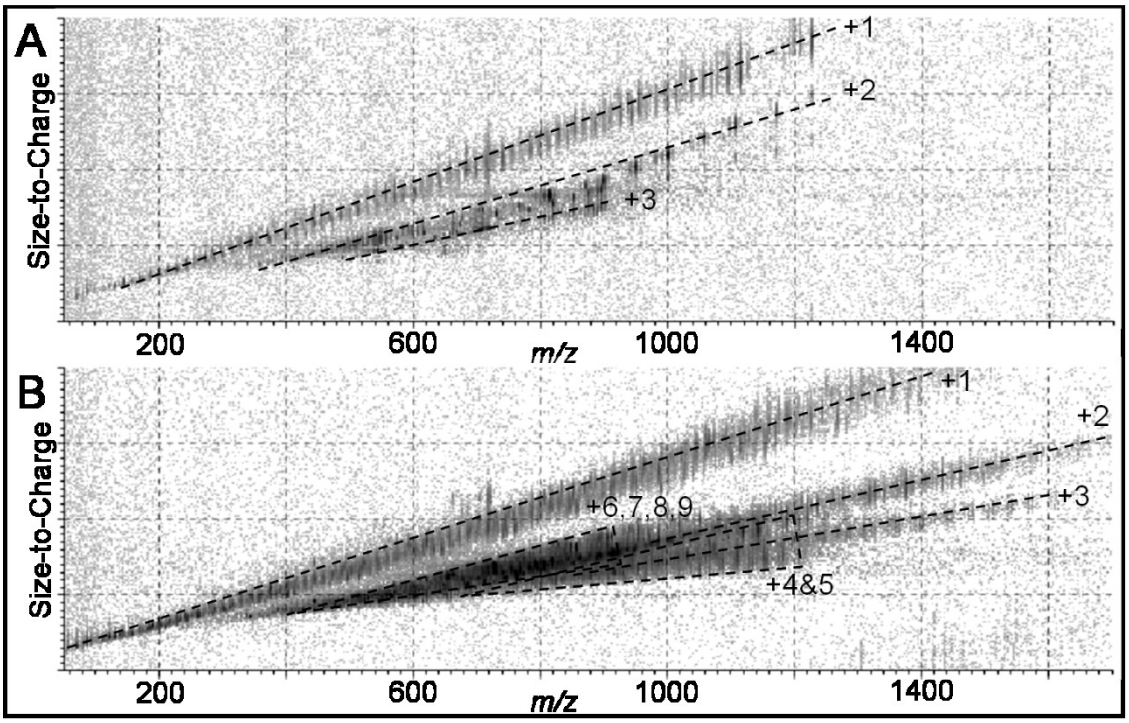


Figure 4. CID-IM-MS product ion spectra for (A) the $[M + 3H]^{3+}$ ion of melittin and (B) the $[M + 12H]^{12+}$ ion of ubiquitin. The $[M + 3H]^{3+}$ ion of melittin forms discrete trend lines according to charge state while the larger $[M + 12H]^{12+}$ ion of ubiquitin forms trend lines but spectral congestion remains.

Ubiquitin, a small 8.6 kDa protein containing 76 amino acids, has been well characterized using different top-down proteomic techniques (CID, ECD, and ETD),[137,140,142,143] and significant differences in the product ion spectra (fragment ions detected as well as their abundances) were noted.[123,144] Using 10% relative abundance as the threshold for ion detection, MS-CID-MS spectra contain abundant sequence informative fragment ions; however, the overall sequence coverage for specific charge states are quite low. On the other hand, if we sum all the fragment ions obtained from the $[M + 8H]^{8+}$ through $[M + 13H]^{13+}$ charge states, as suggested by McLuckey,[123] a

total of 30 out of 76 amide bond cleavages are detected, which corresponds to ~39% sequence coverage. For comparison MS-CID-IM-MS data for ubiquitin $[M + 12H]^{12+}$ ion (see **Figure 4B**) yield an overall sequence coverage of 76% (58 out of 76 amide bonds). It is important to note that the increased amino acid sequence coverage for MS-CID-IM-MS can be attributed to both reduction in chemical noise as well as an increased peak capacity for the IM-MS 2-D plot.[92,132] Ubiquitin also serves to illustrate a fundamental limitation of CID for top-down proteomics that has been previously noted as a limitation for bottom-up proteomics. That is, the most abundant product ions are formed by cleavage reactions that correspond to the lowest energy fragmentation channels. For example, the most abundant fragment ions (y_{58} and y_{40}) in the MS-CID-IM-MS spectra of ubiquitin correspond to cleavages at Pro_{19} and Pro_{37}. The y_{58} fragment ion is dominant in the MS-CID-IM-MS spectra of the higher charge states, *viz.* the $[M + 12H]^{12+}$ ion; however, for lower charge states, esp. $[M + 9H]^{9+}$ and $[M + 8H]^{8+}$, the y_{58} and y_{40} fragment ions are detected at approximately equal abundance., *i.e.,* y_{58}/y_{40} ~12:1 for $[M + 12H]^{12+}$ versus ~1.5:1 for $[M + 8H]^{8+}$ ions. This result underscores another important aspect of 2-D dispersion of the CID product ions. Specifically, MS-CID-IM-MS increases the dynamic range for ion detection as well as the confidence level for product ion assignment. The limited dynamic range arises owing to the energetically favored product ions, which dominate the MS-CID-MS spectra, and the chemical noise, which limits detection of low abundance, but sequence informative product ions. Such limitations are circumvented by viewing the data in 2-dimensions.

The data contained in **Figure 5A** illustrates an additional advantage for viewing the products of CID in 2-dimensions. Note that the isotope clusters for the b_{16}^{2+}-H_2O ion and the y_{39}^{5+} ion in the MS-CID-MS experiment overlap making it difficult to identify charge states and to make sequence assignments; however, the two product ions fall on different charge state trend lines in the 2-D plot, which allows for distinct isotope clusters as well as a high confidence-level sequence assignments **(Figure 5B and C)**.

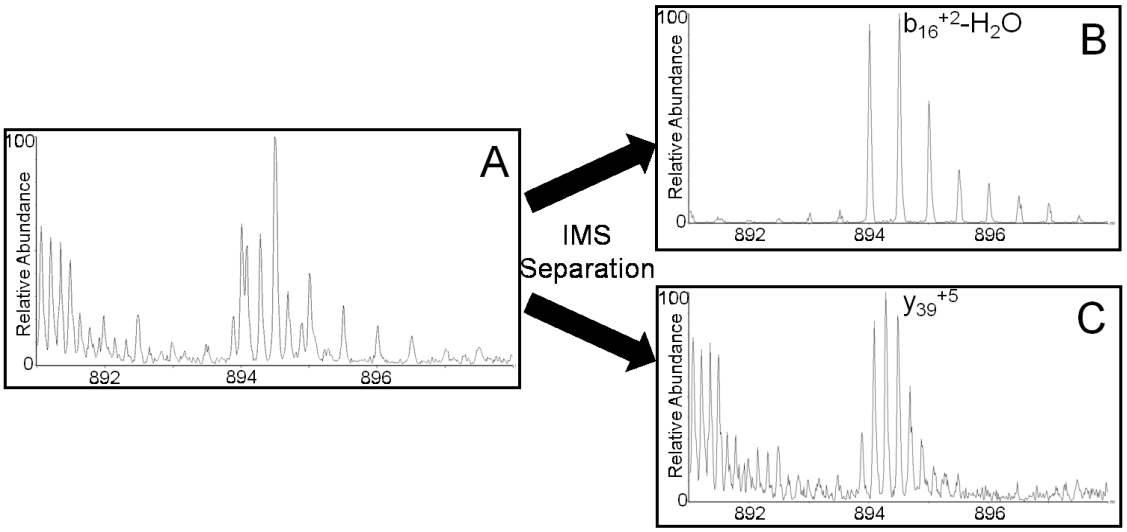
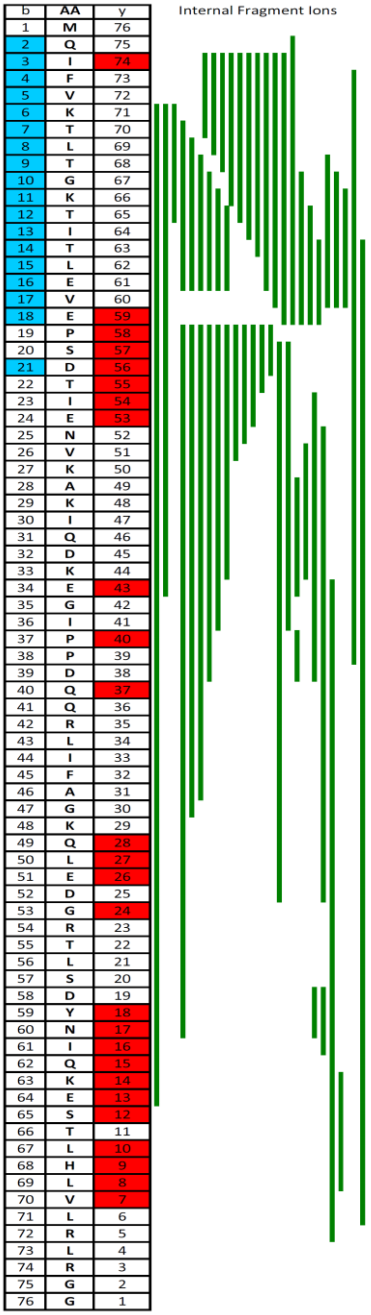


Figure 5. (A) A zoom-in product ion spectrum from the $[M + 12H]^{12+}$ ubiquitin ion without IM separation demonstrating the overlap in isotope cluster observed in non-IM separated product ions. (B and C) The extracted zoom-in product ion spectra from the $[M + 12H]^{12+}$ ubiquitin ion with IM separation. The b_{16}^{2+}-H_2O ion shown in (B) and the y_{39}^{5+} ion shown in (C) fall onto different trend lines upon IM separation thus allowing distinct isotope clusters for high certainty identification.

Both MS-CID-MS and MS-CID-IM-MS are subject to spectra congestion. In **Figure 5** we show that in some cases spectral congestion can be alleviated by the 2-D

data plots; however, there are examples where this may not be realized. In any event, the 2-D plots afford a means for a great depth of mining the CID spectra. This can be illustrated by assigning low abundance internal fragment ions, which are seldom employed for sequence analysis. For example, the MS-CID-IM-MS spectra for ubiquitin $[M + 12H]^{12+}$ ion contain 126 peaks with a *S/N* of ≥3 and an identifiable peptide isotope cluster; however, only 78 (62%) of these peaks could be assigned to *a*-, *b*-, or *y*-type fragment ions, and the remaining 38% of the product ions can be assigned as internal fragments (see **Figure 6**). The preference for cleavage N-terminal to Pro_{19} is illustrated by the frequency of fragment ions having Pro_{19} as the new N-terminus (see **Figure 6)**. Other trends that can readily be observed from the internal fragment ions include (i) preference for forming fragment ions C-terminal to Glu_{18}, which correspond to compliment fragment ions to y_{58}, (ii) preferences for cleavage C-terminal to Glu_{16} and (iii) preference for forming fragment ions N-terminal to Ile_3.[141,145] Similarly, the grouping of internal fragment ions demonstrates that ladder sequence information can be obtained from these internal fragment groups.



b	AA	y
1	M	76
2	Q	75
3	I	74
4	F	73
5	V	72
6	K	71
7	T	70
8	L	69
9	T	68
10	G	67
11	K	66
12	T	65
13	I	64
14	T	63
15	L	62
16	E	61
17	V	60
18	E	59
19	P	58
20	S	57
21	D	56
22	T	55
23	I	54
24	E	53
25	N	52
26	V	51
27	K	50
28	A	49
29	K	48
30	I	47
31	Q	46
32	D	45
33	K	44
34	E	43
35	G	42
36	I	41
37	P	40
38	P	39
39	D	38
40	Q	37
41	Q	36
42	R	35
43	L	34
44	I	33
45	F	32
46	A	31
47	G	30
48	K	29
49	Q	28
50	L	27
51	E	26
52	D	25
53	G	24
54	R	23
55	T	22
56	L	21
57	S	20
58	D	19
59	Y	18
60	N	17
61	I	16
62	Q	15
63	K	14
64	E	13
65	S	12
66	T	11
67	L	10
68	H	9
69	L	8
70	V	7
71	L	6
72	R	5
73	L	4
74	R	3
75	G	2
76	G	1

Figure 6. The chart summarizes the identified product ions from the ubiquitin $[M + 12H]^{12+}$ ion by MS-CID-IM-MS. The single letter amino acid code for ubiquitin is located in the middle column of boxes. To the left of the sequence is the *b*-type ions located in the spectra while to the right of the sequence is the *y*-type ions located in the product ion spectra. The vertical bars located to the right of the boxes correspond to internal fragment ions observed. The length corresponds to the sequence of amino acids comprised in the internal fragment ion.

CID-IM-MS Analysis for Protein Modifications

Top-down methodologies are often used to identify and locate protein modifications[11]. Using CID to induce fragmentation allows for a rapid screening technique for the presence and location of modifications. Although phosphate groups, esp. for serine phosphorylation, are not retained in CID spectra,[70,146] chemical modifications used in iTRAQ labeling[147] and/or alkylation, which is used for studies of folding/unfolding,[148,149] as well as methylation, acetylation, and oxidation are retained.[126] The utility of MS-CID-IM-MS for rapid screening of chemical modifications is illustrated by the following examples, methyl esterification of ubiquitin and single point mutation of human iron sulfur cluster U (HISCU). Ubiquitin contains 12 acidic sites, Glu_{16}, Glu_{18}, Asp_{21}, Glu_{24}, Asp_{32}, Glu_{34}, Asp_{39}, Glu_{51}, Asp_{52}, Asp_{58}, Glu_{64}, as well as the C-terminal carboxylic acid. From the *m/z* shifts of $[M + 11H]^{11+}$ ion of the intact protein it appears that 4 to 9 of the acidic groups are modified upon methyl esterification (see **Figure 7A**). Ubiquitin carrying six methyl ester groups was probed by MS-CID-IM-MS, and the sites of modification can be determined from the CID product ions. The sites for five methyl esterification reactions are assigned to Glu_{16}, Glu_{18}, Glu_{51}, Glu_{64}, and the C-terminal carboxylic acid. The sixth modification could not be precisely located but could be identified to reside between amino acids 19 through 36, of which there are 4 possible modification locations.

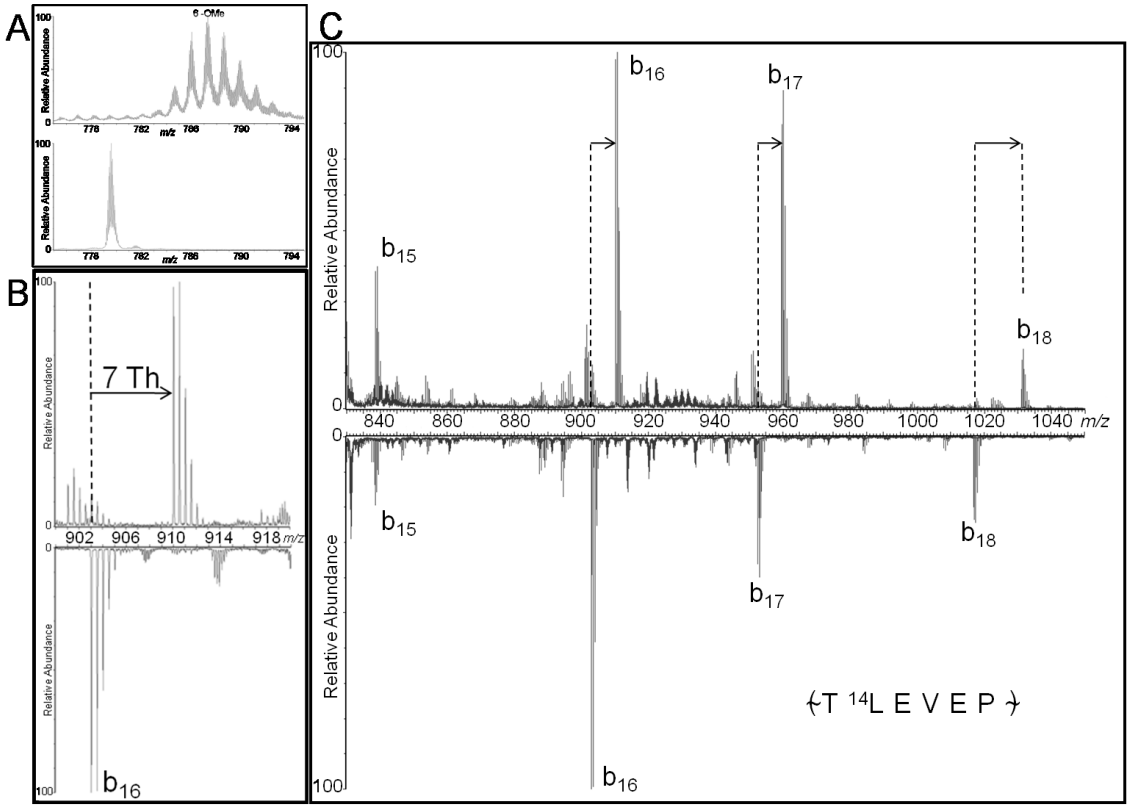


Figure 7. Spectra obtained for the top-down analysis of the methyl esterification of ubiquitin (top) and unmodified ubiquitin (bottom). (A) ESI-MS of the $[M + 11H]^{11+}$ of the modified and unmodified ubiquitin. (B and C) The extracted zoom-in product ion spectra from the doubly charged trend line from the modified (six methyl ester derivatives) and unmodified ubiquitin by MS-CID-IM-MS. Sites of modification show a shift in *m/z* corresponding to 14 Da as shown in (B) corresponding to the ion b_{16}^{2+}. When no modification is present at the site, there is no shift in *m/z* but the presense of modifications provide a shift in *m/z* as shown in (C).

Figure 7B and C demonstrates an example of the shift in *m/z* associated with both the methyl ester derivative and unmodified product ions of ubiquitin, which in this particular case are obtained from doubly charged trend lines. This example shows the quick identification of two modifications located on Glu_{16} and $Glu_{18.}$ Ion mobility allows for the separation of fragments ions into smaller data sets which assists in expediting the

interpretation process. The versatility of this instrument platform, however, could potentially make identification of this last modification site possible through incorporation of multiple stages of dissociation, either in-source CID and/or post-IM CID in TWIG-transfer region. This issue of modification site would traditionally be approached from a bottom-up method which is time consuming and if there is a variable amount of modifications provides location information of the entire ensemble. The identification of modification sites for a particular population from a variable amount of modifications is difficult to achieve by the bottom-up method. In contrast, MS-CID-IM-MS can yield the location of modification in less time as well as give location information on individual populations when variable amount of modifications exist.

MS-CID-IM-MS can also be used to locate single or multiple point mutations in the amino acid sequence of a protein.[125] Intact mass measurements of both the wild type and mutant can be used to confirm amino acid substitution; however, site-specific mutations generally require more extensive bottom-up proteomics. The bottom-up approach suffers from the same limitations as sequencing, specifically incomplete sequence coverage. Here, the methodology described above is used to analyze wild type human iron sulfur cluster U (HISCU), a 14.3 kDa protein, and the K101A mutant. The CID product ion spectra of the mutant variant shows a mass shift of 57 Da compared to that of the wild type. As noted above for ubiquitin, the proline effect limits the utility of top-down approaches owing to the dominance of specific fragment ions. For example, the CID-IM-MS spectra of HISCU $[M + 13H]^{13+}$ is dominated by the b_{32} and y_{101} product ions; therefore, in-source CID is used to generate the spectra shown in **Figure**

8A and B. The mass shift of the y_{36}^{4+} product ion (14.24 Th which is equal to 57 Da) is consistent with the substitution of lysine to alanine. In-source CID of the K101A variant of HISCU without mass selection was performed producing an abundant y_{36}^{4+} ion which was then mass selected and further dissociated by CID in the TWIG-trap prior to ion mobility separation. ESI-CID-MS-CID-IM-MS produces the 2-D plot observed in **Figure 8C**. The triply charged trend line is extracted to produce a product ion spectrum (shown in **Figure 8D**). The mass difference between y_{32} and y_{33} in the y_{36}^{4+} product ion spectra corresponds to alanine at position 101. The in-source CID analysis allows for MS^3 experiments to be performed further enhancing the top-down capabilities of this approach.

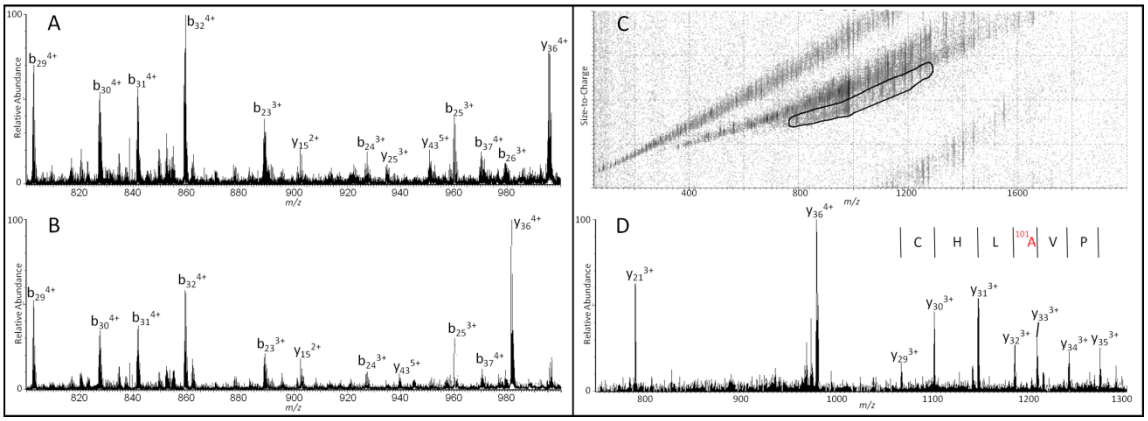


Figure 8. CID product ion spectrum obtained from the $[M+13H]^{13+}$ ion of wild-type (A) and the K101A variant (B) of human iron sulfur cluster U (HISCU). No mass shift is observed for the *b*-type fragment ions or the y_{15}^{2+} ion; however, a mass shift (57 Da which corresponds to the mass shift from lysine to alanine) is observed for the y_{36}^{4+} ion. In-source CID of the K101A variant of HISCU without mass selection was performed producing an abundant y_{36}^{4+} ion which was then mass selected and further dissociated by CID in the TWIG-trap prior to ion mobility separation. ESI-CID-MS-CID-IM-MS produces the 2-D plot observed in (C). The triply charged trend line is extracted to produce a product ion spectrum (D). The mass difference between y_{32} and y_{33} in the product ion spectra of the y_{36}^{4+} ion shows the position of amino acid substitution.

Conclusion

MS-CID-IM-MS has been demonstrated as a viable method for top-down proteomics. This approach alleviates the inherent limitations of top-down proteomics; limitations in dynamic range for fragment ion abundances owing to the number of fragmentation channels available to large ionic systems and spectral congestion. By incorporation of ion mobility to top-down proteomics, the complex mixture of protein CID product ions are dispersed in two-dimensions, specifically size-to-charge (IM) and mass-to-charge (MS), and the resulting 2-D data display greatly facilitates the top-down information contents; (i) charge state specific trend lines, (ii) increased dynamic range, and (iii) separation of overlapping ion signals.

Moreover, this approach also provides depth in biochemical information. The increase in peak capacity allows for the detection of low abundant fragment ions which increases the likelihood of identifying all the possible sequence variations and modifications in the protein of interest as seen with the methyl ester derivative of ubiquitin and the single point mutation of human iron sulfur cluster U. Detection of low abundant ions as well as separation of overlapping ion signals increases the primary sequence coverage and the confidence of sequence assignments. A hybrid top-down MS-CID-IM-MS approach has been demonstrated to obtain the depth of information in both analytical and biochemical contents for top-down proteomics research.

CHAPTER III

ION MOBILITY-TANDEM MASS SPECTROMETRY AS A STRUCTURAL PROBE FOR DISULFIDE CONTAINING BIOMOLECULES

Introduction

The presence of disulfide bonding is an important post-translational modification (PTM) that affects peptide/protein three-dimensional structure and thus their functionality.[42-46] Disulfide bonds are unique to the cysteine residues, which undergo oxidation and subsequent bond formation. These intra- and inter-disulfide bonds play a large part in stabilizing the overall structure of peptides/proteins.[42,45,46] Previous studies by Anfinsen *et al*. demonstrated that the absence of native disulfide bonds can result in some proteins not achieving the correct three-dimensional structure and therefore cannot perform their function.[42] Intra- and inter-chain disulfide bonds are often encountered in biological systems, found in toxins[47], hormones[48], and defensins[49]. Determining the presence of the disulfide bond as well as disulfide bond connectivity is complicated but an important task in the characterization of isolated and recombinantly expressed proteins.

Mass spectrometry (MS) has become a powerful tool in the analysis of disulfide bonds, both in identifying the presence as well as the connectivity. Locations of disulfide bonds within proteins can be determined by a variety of MS-based methods: tandem mass spectrometry (MS-MS),[50,51] chemical reduction,[52,53] chemical oxidation,[54,55] and metal ion cleavage of disulfide bonds.[56,57] The most commonly

employed strategy for disulfide bond analysis involves the reduction of the disulfide bond and alkylation of the thiol providing a characteristic mass shift.[58-61] The reduced and non-reduced proteins are subjected to proteolytic digestion and the resulting peptides are then compared. However, this modified bottom-up proteomics approach proves complicated owing to the requirement of addition of reagents to scission the disulfide bond and alkylate the cysteine residue as well as enzymatic digestion of the protein. Top-down proteomic mass spectrometry provides a complimentary technique to bottom-up proteomic methods for the analysis of disulfide bonds. For example, top-down proteomics introduces intact protein ions into the mass spectrometer which are subsequently dissociated, typically either by electron capture dissociation (ECD)[16,69] or electron transfer dissociation (ETD)[70], into product ions, thus removing the need for enzymatic digestion. The limited sample preparation and fast experiment analysis allows top-down methods to circumvent the issues associated with bottom-up; *i.e.* analysis time, artificial modifications, and loss of information due to poor peptide recovery during sample processing.[74,75] Top-down methods also allow forgoing of reduction and thus provide the opportunity to characterize the connecting patterns of disulfide bonds.

Top-down analysis of disulfide containing biomolecules provides an appealing method yet has seen limited application owing to the low fragmentation efficiency of both ECD and ETD.[12] McLafferty[67] and McLuckey[71] demonstrated that ECD and ETD not only dissociates the backbone bonds but also results in dissociation of the disulfide bond. Low-energy collision induced dissociation (CID) provides an alternate

dissociation technique for top-down analysis which has better efficiency than ECD/ETD. However, in disulfide bond containing peptides/proteins, low-energy CID often generates little product ions from backbone cleavages between residues enclosed by a disulfide bond owing to the need to cleave additional bonds.[67,71-73] This issue is further compounded as disulfide bonds rarely dissociate in the presence of mobile protons.[57] Thus, low-energy CID alone provides limited information in regards to the disulfide network and the primary sequence and must be integrated with a separation technique to increase peak capacity.

Integration of MS with ion mobility spectrometry (IMS) not only provides a rapid separation technique but also provides a powerful technique for studying peptide/protein secondary/tertiary structure. Ion mobility (IM) provides a rapid gas phase separation technique on the basis of ion-neutral collision cross sections, thus ion separation is dependent upon analyte ion conformational preferences. The existence of intra-chain disulfide bonds cyclizes the peptide/protein backbone resulting in a more compact conformation. Therefore, cleaving the disulfide bond or the backbone to create a linear analog of the peptide or protein in the gas-phase is of importance as small differences in conformation of gas-phase ions can be distinguished using IM-MS. IM-MS can potentially differentiate between cyclic and linear peptides.[76-78] Previous studies by our group on PTM peptides demonstrate the potential of IM as a screening tool for PTMs.[79]

Here, we describe a novel application of MS-CID-IM-MS[79,150] for structural characterization of disulfide linked protein ions.[78] This approach takes advantage of the

utility of tandem mass spectrometry (MS-MS) as a powerful technique for mapping peptide/protein primary structure and ion mobility spectrometry (IMS) as a powerful technique for separation of peptides/proteins on the basis of size, *viz.* secondary/tertiary structure. As previously demonstrated[79], integration of IMS and MS, whether it be IM-MS or MS-IMS, expands the scope of structural studies owing to the increase in peak capacity afforded by separation in two dimensions. Low-energy collisional activation predominantly promotes a single cleavage reaction as well as retention of any disulfide bonds. Here we take advantage of these two attributes of low-energy CID to identify the presence of disulfide bonds. IM is utilized to monitor structure of peptides and proteins ions pre- and post- low energy collisional activation thus allowing the monitoring of structural changes; specifically, the opening of cyclic disulfide bonded peptides and protein loops. These changes in secondary and tertiary structure potentially provide a rapid screening for disulfide bonds in peptides and proteins, as previously demonstrated for phospho-peptides.[79]

Experimental

Sample Preparation

All solvents (HPLC grade), chemicals, and proteins were purchased from Sigma-Aldrich (St. Louis, MO, USA) and used as supplied unless otherwise stated. ZipTips™ with C18 resin were purchased from Millipore (Millipore, Bedford, MA, USA). Stock solutions of somatostatin-14 (S-14), insulin, and lysozyme were prepared in water at various concentrations and diluted to a working concentration of 2μM in a 1:1 methanol:water mixture containing 0.1% formic acid. Glu-C digestion was performed

using standard protocols[151]. Briefly, 10 μL of 25 mM ammonium bicarbonate (pH 6.2) was added to 36 μg of protein followed by enzyme at a ratio of 20:1 protein:enzyme by mass. The digestion was carried out at 37°C for 2.5 hrs. The digestion was desalted and concentrated utilizing ZipTips™ prior to analysis. Reduction and alkylation of model peptides was performed as follows: 1 μL of 50 mM *tris*(2-carboxyethyl)phosphine (TCEP) was added to 50 μg of peptide and allowed to react for 30 mins at room temperature. 1 μL of 100nM methyl methanethiosulfonate (MMTS) was subsequently added and allowed to react for 10 minutes in the dark. Optimization of one reduced disulfide bond in insulin was obtained utilizing 6.5 μL of 3mM dithiothreitol (DTT) for 30 minutes at room temperature and subsequent addition 4.6 μL of 6mM *N*-ethylmaleimide (NEM). The reduced and alkylated peptide/protein was desalted and concentrated utilizing ZipTips™ prior to analysis.

Mass Spectrometry

All MS-CID-IM-MS experiments were performed using a SYNAPT G2 HDMS mass spectrometer equipped with a nano-electrospray ion source and MassLynx data processor (Waters Corp., Milford, MA, USA).[128] Instrument acquisition parameters used were as follows: an inlet capillary voltage of 1.85 kV, a sampling cone setting of 15 V for somatostatin-14 and 40 V for insulin and lysozyme, and a source temperature of 100°C. The argon pressure in the traveling wave ion guide trap (TWIG-trap) and the traveling wave ion guide transfer (TWIG-transfer) were 2.44×10^{-2} mbar and 2.61×10^{-2} mbar, respectively. The wave height, wave velocity, and nitrogen pressure in the traveling wave (TW) IM drift cell were 32.0 V, 850 m/s, and 2.96 mbar, respectively.

ESI source conditions were optimized to limit in-source collisional activation and subsequent ion heating. Samples were directly infused into the mass spectrometer at a rate of 0.5-0.8 μL/min. All IM-MS acquisitions were acquired for 2 mins.

All fragmentation was carried out by collision of ions with background gas (argon) and collision energy for CID of the proteins was optimized for each protein as well as each charge state, with trap collision voltage generally falling in a range of 23-50 V. CID can be potentially performed in three different regions of the instrument: the source, the TWIG-trap prior to the TW-IM drift cell, and the TWIG-transfer subsequent to the TW-IM drift cell. Precursor mass selection can only be performed in the quadrupole prior to the TWIG-trap and the TW-IM drift cell.

Data Analysis

All data processing was conducted using Waters MassLynx v4.1 and DriftScope v2.1. Product ion spectra were manually interpreted. To expedite manual interpretation and sequence assignment of mass spectra, an arbitrary cut-off threshold of 10% relative abundance was used in peak assignment for MS-CID-MS experiments as well as for extracted product ion spectra from MS-CID-IM-MS experiments. Internal fragment ions with a S/N >3 and an isotope cluster were assigned based on theoretical peak lists generated from ProteinProspector MS-Product (UCSF, San Francisco, CA, USA). The percentage sequence coverage (%SC) is defined in Eq. (1):[129]

$$\%\mathrm{SC} = \frac{\text{observed amide bond clevages}}{\text{total amide bonds in the protein}} \times 100 \qquad (1)$$

Results

Disulfide bonds are of particular importance in stabilizing and maintaining native protein structures, thus, elucidating the presence of disulfide bonds as well as the connectivity is of paramount significance. To this end, MS has become a powerful tool in the analysis of disulfide bonds. For example, Morris and Pucci developed a method in which the mass spectra of non-reduced and reduced enzymatic digests are compared.[152] Comparison of non-reduced versus reduced proteolytic digests of disulfide-containing proteins has become the classical approach to identifying the presence and connectivity of disulfide bonds. An alternative strategy to chemical reduction is chemical oxidation, which modifies each cysteine residue to the cysteic acid form. Both strategies are cumbersome to employ owing to the requirements of chemical reagents and multiple processing steps, making rapid disulfide bond presence elucidation difficult.

A novel approach to elucidating the presence and connectivity of disulfide bonds without chemical reagents is through the use of MS-CID-IM-MS. CID predominantly results in a single cleavage event along the peptide/protein backbone leaving the disulfide bond intact, opening loops and cyclic peptides. Using ion mobility spectroscopy to monitor the mobility profile of an ion pre- and post- low energy collisional activation allows for the identification of these secondary structure changes which are expressed as a shift in the mobility profile to a later drift time without a change in mass. Without the requirement of chemical reagents and sample handling steps, MS-CID-IM-MS holds potential as a rapid means for the identification for the presence of disulfide bonds.

MS-CID-IM-MS Analysis of Model Peptide/Proteins

The utility of MS-CID-IM-MS for disulfide bond identification was evaluated using several peptides and small proteins. Somatostatin-14 (AGCKNFFWKTFTSC), the 14 amino acid peptide/hormone involved in the inhibition of growth hormone as well as other secondary hormones[153,154] represents the smallest cyclic system studied. The disulfide bond between Cys_3 and Cys_{14} allows somatostatin-14 (S-14) to form a 12 amino acid cyclic structure. The ESI mass spectrum of S-14 contains abundant ion signals corresponding to the $[M + 2H]^{2+}$ (*m/z* 819.4) and $[M + 3H]^{3+}$ (*m/z* 546.9) ions. The mobility profiles associated with the $[M + 2H]^{2+}$ and the $[M + 3H]^{3+}$ ions are comprised of a single peak centered at 9.81 ms and 5.51 ms, respectively.

The MS-CID-IM-MS spectra for the $[M + 3H]^{3+}$ ion is dominated by two peaks, assigned to the $[M + 3H]^{3+}$-H_2O and $[M + 3H]^{3+}$-CO ions. Examination of the low abundance CID fragment ions are consistent with cleavage between the Phe_7 and Trp_8 residues. The ion mobility profile for the collisionally activated $[M + 3H]^{3+}$ ion shows two peaks; the most abundant centered at 5.73 ms and an additional lower intensity mobility peak centered at 6.28 ms. The occurrence of this later arrival time peak suggests an elongated structure.

Assignment of this secondary mobility peak as an elongated structure was based on comparison of the mobility profile of the reduced and alkylated peptide with that of the collisionally activated peptide. Reduction and alkylation of S-14 with MMTS results in the addition of two methylthiols corresponding to a 94 Da shift in mass of the peptide.

Importantly, reduction and alkylation cleaves the disulfide bond to create an elongated conformation. As seen in **Figure 9B**, the mobility profile for the $[M + 3H]^{3+}$ ion of the reduced and alkylated S-14 contains two peaks. These two peaks are shifted in the time domain owing to the increase in molecular weight from alkylation.[155]

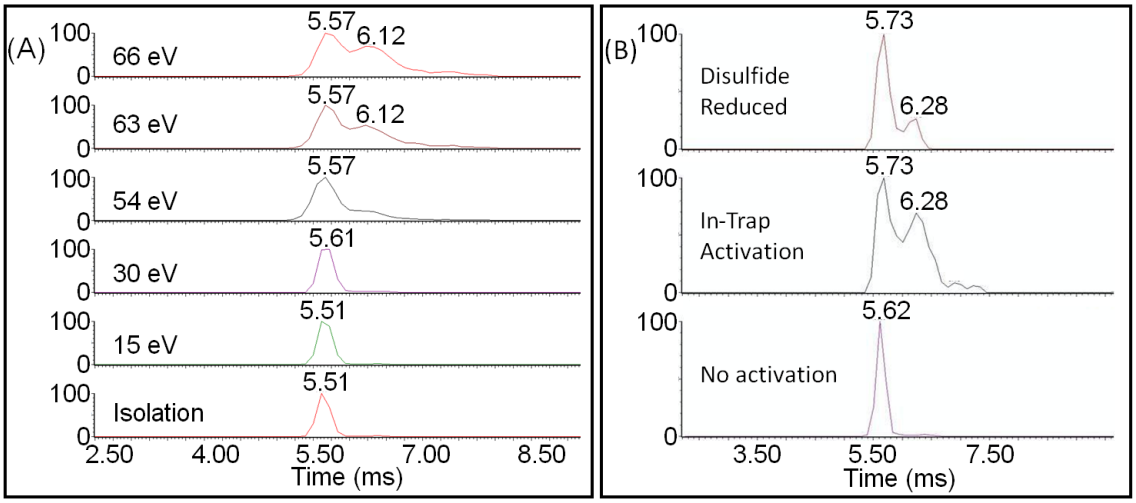


Figure 9. (a) Drift time distributions of the $[M+3H]^{+3}$ ion of somatostatin-14 at varying collsional energies. (b) Comparison of drift time distributions obtained for the reduced (top), collsionally activate (middle), and non-activated (bottom) forms of the $[M+3H]^{+3}$ ion of somatostatin-14. Owing to the increase in mass, the time scale of the disulfide reduced has been adjusted for comparison purposes.

Bovine insulin consists of two chains (A and B) that are connected by two inter-chain disulfide bonds (A-Cys_6 to B-Cys_6 and A-Cys_{19} to B-Cys_{19}) and an intra-disulfide bond contained in the A-chain (Cys_5 to Cys_{10}) to form a cyclic system. The ESI mass spectrum of insulin contains abundant signal from the $[M +5 H]^{5+}$ ion and low abundance signals for the $[M + 4H]^{4+}$ and $[M + 6H]^{6+}$ ions. Obtaining primary structure

information from the intact insulin proves challenging owing to the rigid cyclic backbone formed by the two inter-disulfide bonds as the cyclic structure contains approximately one-half of the amino acids of the protein. McLafferty and co-workers noted that CID analysis of the $[M +5 H]^{5+}$ ion demonstrated only exocyclic structural information; specifically, 8 *b*- and 6 *y*-type product ions which all correspond to exocyclic cleavages in the two C-termini; formation of none of the product ions required cleavage of the inter-chain ring.[67]

The MS-CID-IM-MS spectra for the $[M + 5H]^{5+}$ ion forms several charge-state specific trend lines.[132] As seen in **Figure 10A**, the dissociation of the $[M + 5H]^{5+}$ ion of insulin shows an unexpected truncated trend line located between the doubly and triply charged trend line. Ions located on this truncated trend line contain the same *m/z* as product ions found in the 4+ charge trend line. Extraction of the mobility profile for the product ion corresponding to the loss of the b_2 ion from the B-chain, the ABy_{28}^{4+} ion, yields a mobility profile composed of two peaks centered at 12.90 ms and 16.1 ms. This result is similar to that observed for S-14 and prompted investigation of the mobility profile of the $[M + 5H]^{5+}$ ion of insulin. Monitoring the mobility profile of the $[M + 5H]^{5+}$ ion of insulin **(Figure 11)** allows for the investigation into secondary and tertiary structural changes that are occurring during collisional activation. Isolation of the $[M + 5H]^{5+}$ ion produces a mobility profile which contains a single peak centered at 11.25 ms. As seen in **Figure 11**, as the collision energy is increased across the range of 175 to 225 eV, a second peak centered at 12.57 ms corresponding to a change of 11.7% in arrival time. The development of this secondary peak is consistent with an elongated structure.

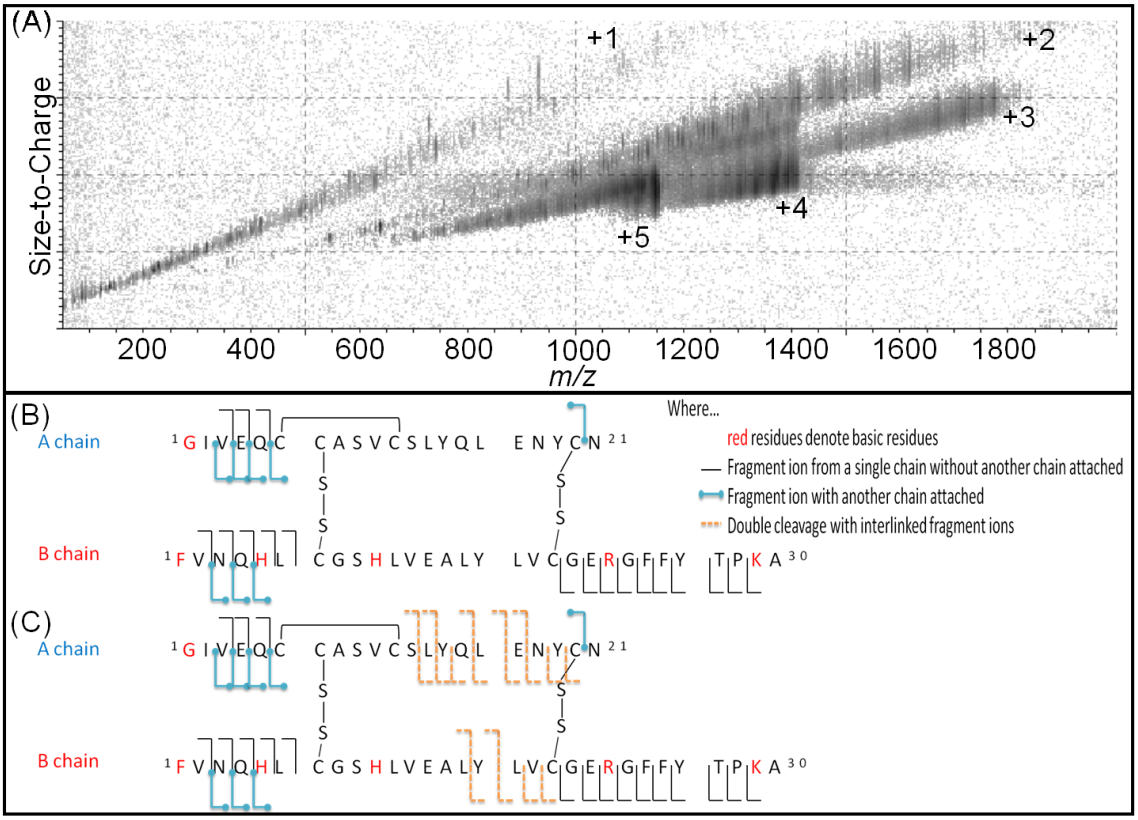


Figure 10. (a) Two-dimensional plot of the product ions with respective charge state formed from the dissociation of the $[M+5H]^{+5}$ ion of insulin. The associated sequence coverage obtained from CID (b) and CID-IMS (c).

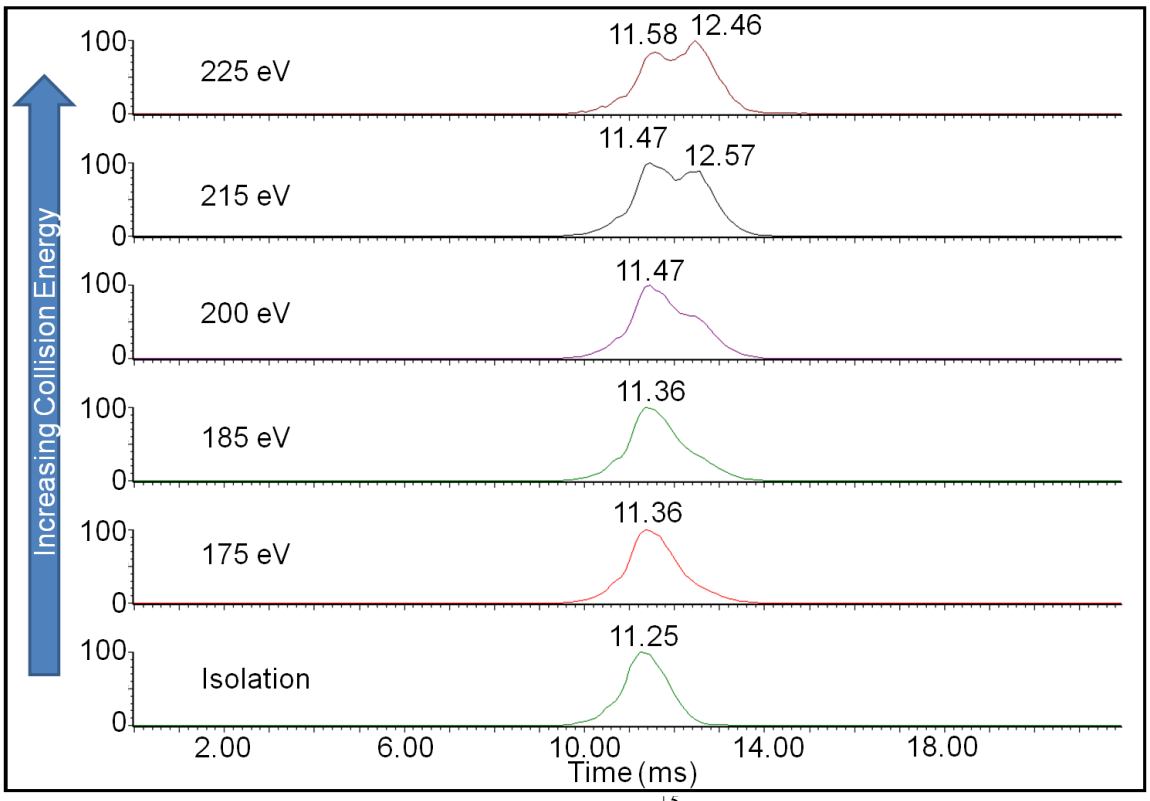


Figure 11. Drift time distributions of the $[M+5H]^{+5}$ ion of insulin at varying collsional energies.

As described previously, insulin was reduced and alkylated under mild conditions, in which, reduction of a single inter-disulfide bond results in a ring opening event forming an elongated structure, similar to the proposed structure, herein. Reduction and alkylation under mild conditions, however, produces a mixture of products including one and two disulfide reduced insulin, intact insulin, and free A- and B-chains. Therefore, confirmation of the location of the reduced disulfide bond was achieved utilizing CID analysis of the +5 charge state of the ion containing one and two reduced disulfide bonds (*m/z* 1165 and 1185, respectively). Analysis of product ions from the two reduced disulfide bond parent ion indicates reduction of the intra-disulfide

bond of the A-chain and reduction of the inter-disulfide bond located towards the N-terminus. The analysis of the product ions formed from the CID analysis of the parent ion containing one reduced disulfide bond indicates reduction of the inter-disulfide bond located towards the N-terminus. Mobility profiles associated with the intact, collisionally activated, and one disulfide reduced of the $[M + 5H]^{5+}$ ion of insulin are displayed in **Figure 12**.

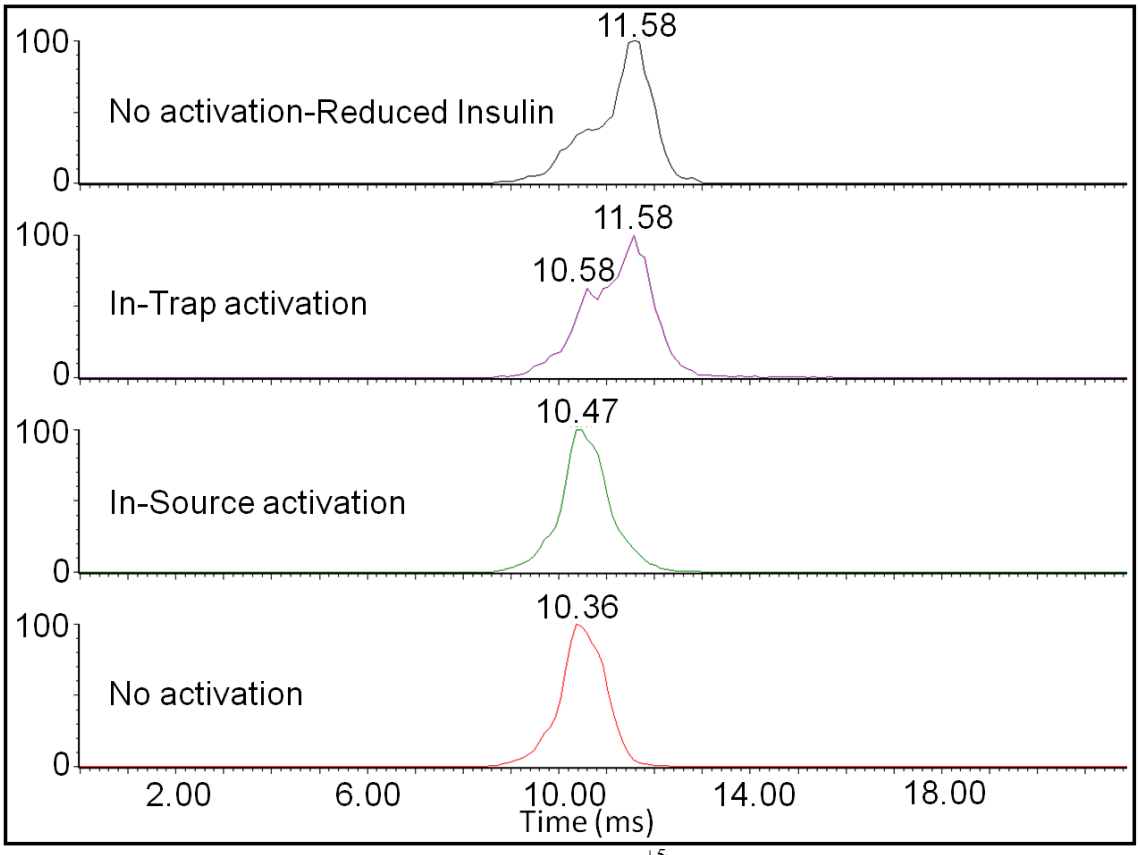


Figure 12. Drift time distributions for the $[M+5H]^{+5}$ ion of insulin with no activation, in-trap activation, and the reduced form of insulin. The secondary mobility peak at 14.99 ms corresponds to a background peak of the fully reduced and alkylated B-chain which occurs at the same *m/z*.

Observation by ion mobility of ring scission arising from collisional activation has applications to larger proteins as well. For example, lysozyme, a 14.3 kDa protein that contains four covalent disulfide bonds, which stabilize and restrict both the solution and gas-phase conformations by creating several cyclic structural elements within the molecule. The ESI spectrum of intact lysozyme contains the +7 to +12 charge states with the +10 being the most abundant.[156,157] **Figure 13** shows ion mobility profiles for the hydrolyzed, collisionally activated, and non-activated lysozyme +10 charge state. The non-activated +10 charge state of lysozyme produces a single peak centered at 12.02 ms; however, a second peak develops at a later arrival time upon collisional activation. Assignment of this secondary mobility peak as an elongated structure was not based on reduction and alkylation owing to the likelihood of a mixture of reduced products and the associated spectral complexity. Rather, assignment of the secondary mobility peak was based on a comparison with a Glu-C digestion. Lysozyme contains only two glutamic acid residues (Glu_7 and Glu_{35}) thus limiting the potential sites of hydrolysis and the resulting spectral complexity. Interrogation of the ESI mass spectrum shows an increase in mass of lysozyme of 18 Da indicating a single site of hydrolysis. Analysis by CID indicates that the site of hydrolysis is Glu_{35}. The mobility profile for the hydrolyzed lysozyme is shown in **Figure 13**. Like collisionally activated lysozyme, the mobility profile of the hydrolyzed lysozyme contains two distinct peaks. However, the separation between the peaks of the hydrolyzed lysozyme is not as great as the peak separation observed in the mobility profile of the collisionally activated.

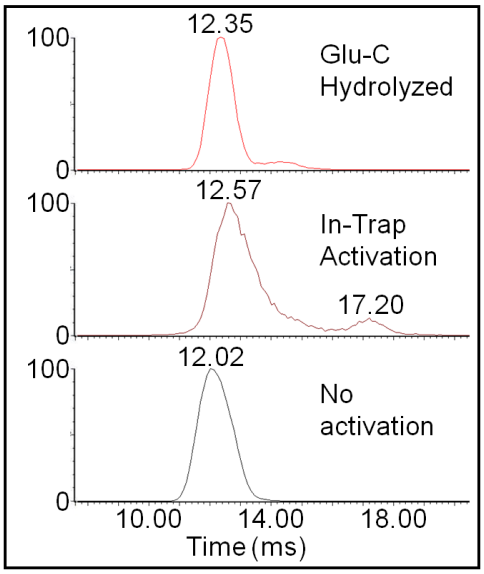


Figure 13. Comparison of drift time distributions obtained for the reduced (top), collsionally activate (middle), and non-activated (bottom) forms of the $[M+10H]^{+10}$ ion of lysozyme.

Discussion

Incorporation of ion mobility with mass spectrometry increases the information content of the analysis by increasing peak capacity, reducing spectral complexity and facilitating assignment of low abundant ion signals. No change in mass is observed during the collisional activation of S-14, insulin, and lysozyme indicating that the change in the mobility is a result of a backbone cleavage which does not produce a product ion. A single cleavage results in the opening of the ring structure forming an elongated structure with a later arrival time distribution (ATD). A similar structure to that of the ring opening by collisional cleavage can be achieved by utilizing reduction and alkylation and therefore can be used to support the current model for the changes in secondary and tertiary structure. This is most easily seen in the simplest system studied,

the $[M + 3H]^{3+}$ ion of S-14. Collisional activation of the $[M + 3H]^{3+}$ ion results in cleavage of the back bone away from the disulfide bond, between the Phe_7 and Trp_8 residues, producing an open structure that does not have a change in *m/z*. Reduction and alkylation of the disulfide bond alters the cyclic structure of the peptide resulting in an elongated peptide structure, similar to the one that occurs during collisional activation of the unmodified form. As seen in **Figure 9**, the mobility profile produced from collisional activation and reduction are similar. The most discernible difference is the relative abundance associated with the elongated forms.

Insulin contains three disulfide bonds which provide a constrained 3-D structure making significant rearrangement to form a secondary IM peak difficult. Specifically, the C-terminus of the B-chain contains the longest exocyclic amino acid sequence (11 amino acids) which potentially could undergo structural rearrangement. Observation of the ABy_{28}^{4+} ion demonstrates both a cleavage as well as a structural rearrangement which corresponds to a ~25% increase in mobility drift time. Recent reports by Desaire *et al.*[158] denotes the prevalence of disulfide bonded peptides to undergo secondary cleavages, suggesting that this later arrival time peak represents an elongated structure which is created by a secondary cleavage that occurs within the cyclic structure. A similar trend is observed by monitoring the intact $[M + 5H]^{5+}$ ion as a function of collisional activation energy. The isotope cluster and *m/z* remain unchanged in the collisionally activated $[M + 5H]^{5+}$ ion, indicating that the structure rearrangement that is occurring to produce a secondary mobility peak is the result of a cleavage on either the A- or B-chain between the inter-disulfide bonds thus producing an open ring structure. A

suitable candidate structure to mimic the proposed ring opening was obtained by reduction and alkylation of a single disulfide bond. As seen in **Figure 12**, the mobility profiles produced from the collisionally activated and single reduced disulfide bond are similar with a lower intensity front shoulder peak and a second peak centered at 11.58 ms. This observation would suggest that the second peak observed in the ion mobility profile of the collisionally activated $[M + 5H]^{5+}$ ion is the result of a cleavage which causes the ring structure to open.

The appearance of a second peak in the mobility profile is also observed in the collisionally activated $[M + 10H]^{10+}$ ion of lysozyme. This result is similar to observations made by Clemmer *et al.* in which an increase in injection voltage resulted in the observation of two resolved mobility peaks for lysozyme.[157] In this study, the later arrival time peak has ~48% increase in ion mobility drift time compared to that of the inactivated $[M + 10H]^{10+}$ ion of lysozyme. The appearance of this secondary peak indicates a significant structural rearrangement is occurring to form a more open conformer owing to the structural constraints imposed by the four disulfide bonds. It should be noted that the most abundant mobility peak in the collisionally activated lysozyme shows a shift in drift time compared to that of the inactivated lysozyme ion. This shift of 0.55 ms is likely the result of small rearrangements caused by ion heating. However, S-14 and insulin indicate that the significant change in drift time associated with the lower intensity second peak is the result of a cleavage along the back-bone allowing for a more elongated structure without a change in *m/z*.

Lysozyme presents a challenge in mimicking the observations from collisional activation of unmodified lysozyme owing to the significant number of protein loops and possible sites of cleavage. However, backbone cleavage and subsequent formation of a model elongated structure can be achieved in a controlled fashion through a Glu-C digestion of the intact lysozyme. There are two possible sites of hydrolysis: Glu_7 and Glu_{35}, both of which result in the formation of a more open conformer. Analysis of the CID product ions of the hydrolyzed lysozyme indicates that the site of hydrolysis is Glu_{35}. Specifically, a *b*-type ion series corresponding to an N-terminus at residue Ser_{36} is observed ranging from b_2 through b_{17}. As seen in **Figure 13**, the mobility profile for the intact and hydrolyzed $[M + 10H]^{10+}$ ion produce a mobility profile with two peaks similar to the mobility profile observed for the collisionally activated lysozyme. Though similar in profile, the most notable difference observed is the separation between the resolved mobility peaks in the profiles. The observed differences in peak separation between the hydrolyzed and collisionally activated lysozyme is likely a result of differences in location of ring opening. Glu-C hydrolysis cleaves the back-bone at the Glu_{35} while collisionally activation can potentially cleave anywhere along the protein backbone. Furthermore, collisional activation imparts internal energy to the protein ion allowing the ion to sample a greater number of conformations; the mobility peaks are broader than the mobility peaks observed in the non-activated and hydrolyzed lysozyme. Increasing molecular size and number of disulfide bonds does not affect the formation of a secondary peak in the mobility profile upon collisional activation making this technique of screening for disulfide bonds widely applicable.

Primary Structure Characterization

The most abundant charge state of insulin (+5) was subjected to CID-IM-MS with the dominant product ions corresponding to *y*-type fragment ions from the B chain, the complimentary *b*-type ions with an intact A chain attached, and un-dissociated parent ion, all of which represent exocyclic cleavages which is in agreement with previous reports.[67,159] IM separation increases peak capacity by dispersion of product ions in two dimensions **(Figure 10A)**. A noticeable advantage to this technique is the formation of trend lines of product ions based on their charge, from that of the parent ion to unity. Separation in this manner allows for analysis of individual charge trend lines thus allowing for identification of cleavage at 20 out of a possible 23 exocyclic amide bonds (87%).

As seen in **Figure 10A**, there is a doubly charged trend line which contains a significant population of ions. The product ions of this doubly charged trend line are of interest as results from McLafferty[67] and from McLuckey[159] indicate a lack of doubly charged fragment ions in the mass spectra. Initially, no correlation could be established between these doubly charged ions and possible *b*-, *c*-, *y*-, and *z*-type ions based on amino acid sequence. The ABy_{28}^{+4} ion mobility result indicates that ion formation could be the result of cleavage of more than one bond as recently reported by Desaire *et al.*[158] Considering this result, the doubly charged product ions from *m/z* 850 to 1800 were assigned as product ions resulting from two backbone cleavages (note: *b*- *b*- type or *y*- *y*- type). These product ions are formed endocyclic and thus provide additional sequence coverage. In total, 32 out of 49 possible amide backbone cleavages (65%) are observed.

Mapping the Disulfide Network

Single and double bond cleavages can provide valuable insight into the disulfide bond network of proteins. The presence of multiple disulfide bonds complicates peptide/protein analysis, thus disulfide bonds are often reduced and alkylated under general proteomics protocols which results in the loss of disulfide connectivity, information that cannot be inferred. Insulin contains six cysteine residues, all of which are involved in disulfide bonds. Exocyclic cleavages of insulin provide primary structure information for both the A- and B-chain from the N- and C-terminus to the inter-disulfide bonds forming the cyclic structure of insulin. Therefore, single dissociations provide no information in regards to disulfide connectivity. However, double cleavages of the ring structure provide connectivity information for the C-terminal disulfide bond (A-Cys_{19} to B-Cys_{19}). The CID analysis of the $[M + 5H]^{5+}$ ion of insulin produces a significant number of doubly charged fragment ions that correspond to double cleavage ions. Based on molecular weight, these ions are *y*-type ions from the A- and B-chains connected via an inter-disulfide bond. The close proximity of cysteine residues on the A-chain and the intra-disulfide bond makes explicit determination of the remaining four cysteines difficult. From the MS-MS, the presence of an inter-disulfide can be verified. As no dissociation is observed from residue 4 though 10 of the A-chain, it could be inferred that a disulfide exists between residue 5 and residue 10 of the A-chain which serves to protect that region from dissociation. This would indicate that Cys_6 from the A-chain is connected to Cys_6 of the B-chain. Though

all disulfide assignments cannot explicitly be determined, the MS-MS can be used to infer connectivity in insulin.

Conclusion

Disulfide bonds act to stabilize as well as constrain peptides and proteins limiting the degrees of freedom of the molecule and results in the formation of cyclic peptides or loops in proteins. CID analysis of molecules containing disulfide bonds with cyclic elements result in the retention of the disulfide bond and cleavage along the back-bone, thus opening up the ring structure. Ion mobility studies of somatostatin-14, insulin, and lysozyme ions demonstrate structural changes upon CID-IM-MS analysis. Ions from all three molecules demonstrate the formation of a peak at a later ATD which is indicative of an elongated structure. S-14 and insulin are highly restricted by the presence of the disulfide bonds, thus significant structural changes are limited to the dissociation of bonds. Therefore, formation of elongated structures is the result of sufficient collisions to induce cleavage along the backbone. From these results we propose that a qualitative identification of peptides and proteins which contain disulfide bonds can be made by the appearance of this later arrival time peak during CID-IM-MS analysis. Therefore, changes in secondary and tertiary structure potentially may provide a screening technique for identification of the presence of disulfide bonds in proteins.

CHAPTER IV

SIZE-TO-CHARGE DISPERSION OF COLLISION-INDUCED DISSOCIATION (CID) PRODUCT IONS ENHANCES STRUCTURAL INFORMATION AND PRODUCT ION IDENTIFICATION

Introduction

The three-dimensional (3-D) structure of peptides and proteins is determined by intramolecular interactions, both hydrophobic and hydrophilic interactions, as well as solute-solvent interactions, and these interactions play key roles in determining biological activity.[43,44,46,160] The 3-D structure of a protein comprises multiple secondary structure elements as well as the organization of these elements in relation to one another. The most common structural motif in proteins, and the most important secondary structural element in transmembrane proteins,[86] is the helix,[87-89] which can be characterized by using circular dichroism (CD),[161,162] X-ray diffraction crystallography (XRD),[163] and nuclear magnetic resonance (NMR);[164] however, mass spectrometry (MS)-based approaches provide similar structural information and require much less analysis time than do atomic-level resolution techniques such as XRD and NMR. Although MS is most frequently used for determination of molecular weight,[165] amino acid sequence,[166] and post-translational modification, which have served as major driving forces for the rapid developments of tandem MS approaches,[70,143] there have been increasing numbers of studies utilizing MS as a means for probing higher order structure, specifically secondary, tertiary, and quaternary.[167-169] These studies have

employed rather unique approaches for selection of specific charge states and for inducing fragmentation[170] as well as specialized techniques for probing ion conformations, *viz.* HDX-MS and chemical labeling.[171-173] Ion mobility-MS (IM-MS)[174] is also frequently used for studies of gas phase ion structure because the size of the ions can be correlated to three-dimensional (3-D) shape (conformation) of the ions, including distinguishing structural preferences (cyclic vs. linear structures) of *b*-type peptide fragment ions.[175,176]

IM spectrometry, a gas-phase electrophoretic separation technique, separates ions on the basis of ion-neutral collision cross sections (CCS) with an inert gas (typically He or N_2). The CCS information and accurate mass-to-charge (*m/z*) measurements are then correlated to candidate structures derived using molecular dynamic simulations (MDS) in order to gain insight into the gas-phase structure.[177,178] IM-MS disperses ions in two-dimensions, size-to-charge ratio and *m/z*, thus plots of IM arrival-time distributions (ATD) or CCS can be used to separate ions on the basis of compound class.[27] For example, different types of molecules (peptides, lipids, nucleic acids, etc.) possess different packing efficiencies (atomic densities), thus the relationship between IM ATD (or CCS) and molecular weight are characteristic of the compound type. In addition, mobility-mass plots can be used to identify conformations, *i.e.*, extended coil (or helix) that deviate from the trend line for random coil conformations.[78,92] In previous studies these deviations in mobility-mass trend lines have been used to identify specific secondary structure elements, post-translational modifications, or peptide-small molecule interactions.[93]

We previously reported on the utility of combining collision-induced dissociation (CID) with IM-MS to disperse CID product ions along charge-state specific trend lines.[132,150,179] In this study we take advantage of deviations from product ion charge-state trend lines formed by CID to determine the presence and location of specific secondary structural elements, *i.e.* extended coils (or helices). The advantage of this approach is that product ion identity can be correlated to gas-phase ion structure, which provides rapid identification of the onset and termination of extended coil conformations of the peptide ions.

Experimental

Sample Preparation

All solvents (HPLC grade), chemicals, and proteins were purchased from Sigma-Aldrich (St. Louis, MO, USA) and used as received unless otherwise stated. The initial investigations of B-chain insulin were carried out by reduction of 25μg bovine insulin with *tris*(2-carboxyethyl)phosphine (TCEP) and alkylation with methyl methanethiosulfonate (MMTS) following iTraq protocol.[180] Bovine insulin consists of two chains (A and B) that are connected by two inter-chain disulfide bonds (A-Cys_6 to B-Cys_6 and A-Cys_{19} to B-Cys_{19}). Chemical reduction of insulin cleaves the inter-chain disulfide bonds thereby creating free A- and B-chain insulin. Reduction and alkylation of insulin was performed as follows: 1 μL of 50 mM *tris*(2-carboxyethyl)phosphine (TCEP) was added to 50 μg of peptide and allowed to react for 30 mins at room temperature. The cysteine modification effect studies were performed using dithiothreitol (DTT) as the reducing agent and MMTS, iodoacetamide (IoAc), and *N*-

ethylmaleimde (NEM) as the alkylating agents. The reduced and alkylated sample was diluted with a 1:1 water: methanol solution containing 0.1% formic acid to a working concentration of 2 μM. Somatostatin was reduced with TCEP and alkylated with MMTS following the iTraq protocol.[180] The peptide HPIK(Acetyl)HQGLPQEVLNENLLR was prepared by enzymatic digestion of alpha-S1-casein. Specifically, 10μg of alpha-S1-casein was dissolved in 10μL of water and trypsin was added 50:1 protein:enzyme and allowed to incubate overnight in a 37°C water bath.

Mass Spectrometry

All MS-CID-IM-MS experiments were performed using a SYNAPT G2 HDMS mass spectrometer equipped with a nano-electrospray ion source and MassLynx data processor system (Waters Corp., Milford, MA, USA).[128] Instrument acquisition parameters used were as follows: an inlet capillary voltage of 1.85 kV, a sampling cone setting of 15, and a source temperature of 100°C. The argon pressure in the traveling wave ion guide trap (TWIG-trap) and the traveling wave ion guide transfer (TWIG-transfer) were 2.44 x 10^{-2} mbar and 2.61 x 10^{-2} mbar, respectively. The wave height, wave velocity, and nitrogen pressure in the traveling wave (TW) IM drift cell were 25.0 V, 650 m/s, and 2.96 mbar, respectively. ESI source conditions were optimized to limit in-source collisional activation and subsequent ion heating. Samples were directly infused into the mass spectrometer at a rate of 0.5-0.8 μL/min. All IM-MS acquisitions were acquired for 2 mins.

Dissociation was performed in the TWIG-trap prior to the TW-IM drift cell. All dissociation was carried out by collision of ions with background gas (argon). Activation energy for the dissociation of the B-chain of insulin was optimized for each charge state, with trap collision voltage generally falling in a range of 25-60 V. Data processing was conducted using Waters MassLynx v4.1 and DriftScope v2.1. Product ion spectra were manually interpreted. The ion neutral collision cross section (CCS) values were calculated using the method described previously by Ruotolo et al.[181,182] Ubiquitin and melittin were used as the CCS calibration standards.

Results and Discussion

We have previously described a general top-down approach that is based on ion mobility-MS, specifically MS-CID-IM-MS.[150,183] For example, ions formed by ESI are separated on the basis of *m/z*, the *m/z* selected ions are then subjected to CID, and the CID product ions are analyzed by ion mobility (size-to-charge); this is followed by *m/z* analysis by high resolution TOF-MS. Thus MS-CID-IM-MS is used to correlate CID product ions of *m/z* selected ions on the basis of both *m/z* and size-to-charge. The advantage of this approach is that it provides reliable detection and assignment of low abundance CID product ions and rapid assignment of CID product ion charge states, which provides greater sequence coverage and increases the experimental dynamic range.[150,183] The size-to-charge information obtained from this approach can be used in three ways: (i) to facilitate top-down peptide/protein sequencing from the increased dynamic range and sequence coverage, (ii) to gain information concerning the conformation of the product ions,[34,130,131] and (iii) to disperse the product ions along

charge-state specific trend lines.[132] Here, we emphasize the use of this approach for the identification of product ion conformations, *i.e.*, additional conformational information can be obtained regarding product ions that do not fall along charge state specific trend lines that can be correlated to random coil conformations.

ESI-IM-MS and ESI-MS-CID-IM-MS Analysis of the B-chain of Insulin

The utility of MS-CID-IM-MS for identification and localization of secondary structure was evaluated using the B-chain of insulin (MW 3397.7 Da; amino acid sequence FVNQHLCG**SH**^{10} **LVEA**LYLVCG^{20} ERGFFYTPKA^{30}). Previous X-ray crystallography studies show that the B-chain of insulin contains a partial helix in the region Ser_9 to Ala_{14} (shown in bold font) and the remaining portions of the sequence are best described as random coil.[184]

The ESI mass spectrum of the B-chain of insulin (generated following reduction with DTT) contains low abundant signals corresponding to the $[M + 3H]^{3+}$ (*m/z* 1134.3) and abundant signals corresponding to the $[M + 4H]^{4+}$ (*m/z* 850.4) and the $[M + 5H]^{5+}$ (*m/z* 680.5) ions. The arrival time distribution (ATD) for the $[M + 3H]^{3+}$ and $[M + 4H]^{4+}$ ions each contain a single peak whereas the ATD of the $[M + 5H]^{5+}$ ion contains two baseline resolved peaks **(Figure 14A)**. The two-component ATD of the 5^+ ion is consistent with the presence of two distinct conformations, *i.e.*, the collision cross section for the first peak agrees well with the expected value for a random coil conformation and the later peak, which deviates from the expected random coil value by 7.2%, is assigned as an "extended coil" conformation. The extended coil conformation suggests the possibility that the partial helical conformation of insulin B-chain noted in

X-ray diffraction studies may be retained in the gas phase ion and quite possibly the CID product ions. It is important to note that the IM profile for the $[M + 5H]^{5+}$ precursor ion that survives collisional activation is unchanged from that of the initially formed ion. That is, collisional activation does not promote structural rearrangement of the gas phase ion.

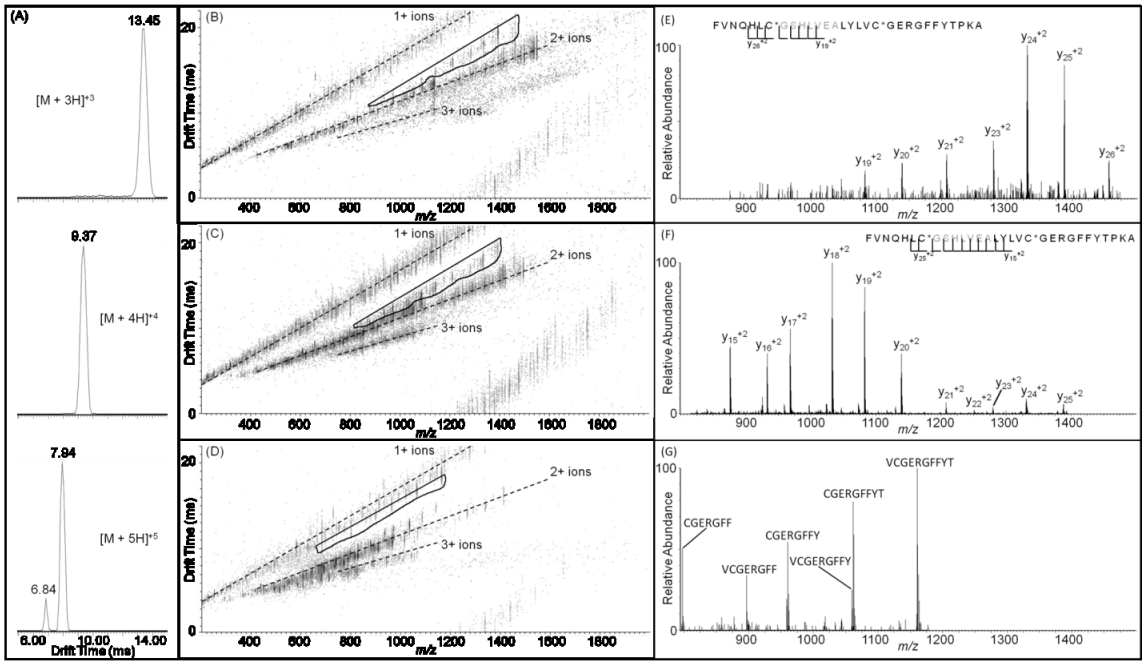


Figure 14. (A) ATD and the mobility-mass plots for the $[M + 3H]^{3+}$, $[M + 4H]^{4+}$, and $[M + 5H]^{5+}$ ions (B-D, respectively) of B-chain of insulin reduced with DTT. Deviations from the doubly charged random-coil trend line can clearly be observed in the IM analysis of the product ions, regardless of precursor ion charge state. (E) and (F) show the extracted MS of the positive deviations from the doubly charged random coil trend line circled in (B) and (C), respectively. Specifically, *y*-type product ions fall above the trend line for random coil conformers. The onset for deviation from the doubly charged trend line occurs as the *y*-ion series contains amino acids of the helix region (noted in gray in (E) and (F)). Thus, it appears that these *y*-type product ions retain a significant portion of helical conformation. Conversely, *b*-type product ions encompassing the amino acids of the helix region show no deviation from the doubly charged trend line. (G) shows the negative deviating product ions from the single charged random coil trend line circled in (D). These negative deviating product ions are identified as internal fragment ions and indicate a more compact conformer than the random-coil conformer.

MS-CID-IM-MS was employed to probe the conformation of the insulin B-chain $[M + 3H]^{3+}$, $[M + 4H]^{4+}$, and $[M + 5H]^{5+}$ ions as well as the conformations of the CID product ions. Each precursor ion produces abundant *b*- and *y*-type product ions. The mobility-mass plots for the $[M + 3H]^{3+}$, $[M + 4H]^{4+}$, and $[M + 5H]^{5+}$ ions are shown in **Figure 14B, C, and D,** respectively**.** Note that all three mobility-mass plots contain charge-state specific trend lines corresponding to singly, doubly, and triply charged product ions. Although dissociation of each precursor ion produces similar *m/z* product ion spectra, the specific product ions and their abundance are charge state dependent.[123] For example, the product ions from the $[M + 3H]^{3+}$ ion (**Figure 14E)** contains the y_{26}^{2+} ion that is not observed in the product ion spectra of the $[M + 4H]^{4+}$ ion **(Figure 14F)**. Similarly, the y_{15}^{2+} through y_{18}^{2+} ions are observed as product ion for $[M + 4H]^{4+}$ ions **(Figure 14F)** but not for the $[M + 3H]^{3+}$ ions **(Figure 14E)**. These differences in observed product ions can be rationalized on the basis of the site of protonation. That is, the charges are located on the N-terminus, Arg_{22}, and Lys_{29} for the $[M + 3H]^{3+}$ ion whereas the $[M + 4H]^{4+}$ ion has an additional charge on one of the histidine residues (His_5 or His_{10}), thus decreasing the likelihood of forming large doubly charged product ions. Interestingly, the mobility-mass plots contain ion series that deviate from the expected ATD for the doubly charged, random coil ion conformers. We interpret these results as evidence that product ions falling along the random coil trend line do not retain extended conformations, conversely ions that fall above the random coil trend line retain the extended conformations, suggesting that helical character is retained by the product ions and may be used to identifying the helical region.

Product Ion Mobility Deviations and Implications for Conformer Assignment

As noted in a previous paper,[150] 2-D mobility-mass plots reduce spectral congestion and chemical noise, thereby increasing the signal-to-noise of the product ion spectrum. In addition, the 2-D plots yield correlated data for specific charge states as well as conformer families. For example, **Figure 14E, F, and G** contain the extracted mass spectrum of the areas circled in **Figure 14B, C, and D,** respectively. **Figure 14B and C** contain mass spectra for ions that deviate (in terms of ATD) from the expected ATD, *i.e.*, the ATD fall above the expected trend line. **Figure 14E and F** contain the resulting mass spectrum and ion identities from the extracted areas in **Figure 14B and C**, respectively. In both cases, the ions that deviate from the expected random coil ATD for doubly charged random coil conformers are *y*-type ions, with combined sequence coverage from y_{15} through y_{26}. It is interesting to note that the y_{17}^{2+} through y_{22}^{2+} ions correlate to the helical region of the B-chain of insulin, which suggests that the helical character of the peptide is conserved following collisional activation and subsequent dissociation. Although rearrangement of the collisionally activated ion prior to dissociation cannot be excluded, it is important to note that the ATD of the collisionally activated $[M + 5H]^{5+}$ ion remains unchanged (**Figure 15**). That is, rearrangement to form an extended coil conformer would yield an increase in the abundance of later arrival time signals. Interestingly, the doubly charged *b*-type product ions encompassing the amino acids of the helical region show no deviation from the doubly charged trend line. This result is consistent with the *b*-type product ions having a random coil conformation.

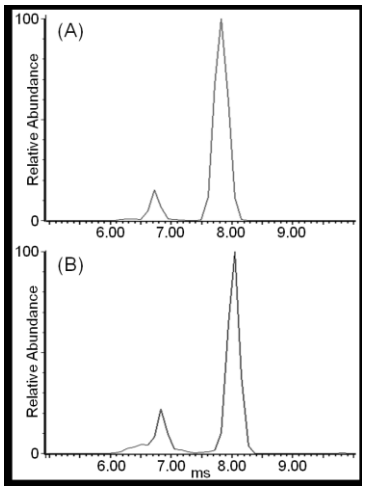


Figure 15. ATD of the $[M + 5H]^{5+}$ ion upon (A) MS-IM-MS and (B) MS-CID-IM-MS. The ATD of the parent ion remains unchanged indicating that conformational rearrangement as a result of collisional activation is not a contributing factor to the observed extended coil conformation of the doubly charged y-type product ions.

Figure 14G contains the extracted mass spectrum for the area extracted in **Figure 14D**, which includes product ions having a negative deviation. Product ions that fall below the random coil trend line (negative deviation) have conformations that are more compact than a random coil.[79] In this case the product ions are neither *b*- or *y*-type ions, which suggests that they are internal fragment ions. Interestingly, all the singly charged internal fragment ions contain arginine, which is located near the middle of the fragment ion sequence (**see Figure 14G)**. It appears that the charge is localized on the arginine side chain and the backbone of the remaining sequence solvates the charge producing a collapsed globule.

To investigate whether the observed deviations in ATD of product ions are a result of specific conformational preference due to cysteine modification, the cysteine residues were modified using several reagents. **Figure 16** contains the mobility-mass

plots of the MS-CID-IM-MS of the $[M + 4H]^{4+}$ ion of the B-chain of insulin with modified cysteines: (**A**) DTT, (**B**) MMTS, (**C**) IoAc, and (**D**) NEM. Although the relative abundances of CID product ions for the modified B-chain insulin differ, the deviations in the mobility-mass plot trend lines for the doubly charged random coil conformers are clearly present. For example, doubly charged *y*-type ions are observed to have positive deviations. This result is consistent with retention of helical character in the observed product ions and is consistent with the conformational preference of product ions not being related to cysteine modification.

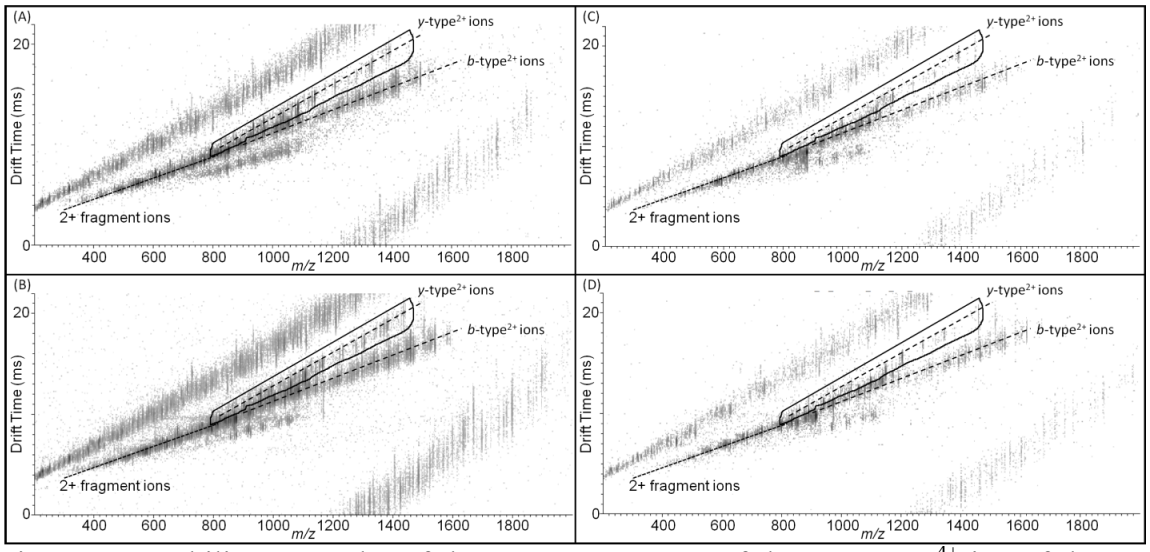


Figure 16. Mobility-mass plot of the MS-CID-IM-MS of the $[M + 4H]^{4+}$ ion of the B-chain of insulin with several alkylating reagents: (A) DTT, (B) MMTS, (C) IoAc, and (D) NEM. Regardless of cysteine modification, the outliers from the doubly charged random coil conformer trendline are present. Thus, these outliers are a result of secondary structure and not a result of the cysteine modification.

Comparisons with Other Peptide Ions

The apparent stability of the helical region of B-chain insulin following CID is observed for other helix containing peptide ions. The native conformation of melittin (GIGAVLKVLT^{10} TGLPALISWI^{20} KRKRQQ-NH_2), a 2.8 kDa amphipathic polypeptide found in bee venom, contains two helices separated by Pro_{14}.[133,134] Abundant CID product ions are observed corresponding to cleavage of the amide bond at Pro_{14} owing to the proline effect,[138-141] thus the CID spectrum of melittin contains only low *m/z* *b*-type product ions and an abundant series of *y*-type product ions (**Figure 17A**). The MS-CID-IM-MS data reveal that these *y*-type product ions fall along a trend line having positive deviations from that for the doubly charged random coil ions (**Figure 17B**), which is consistent with retention of the helical region located near the N-terminus. The ATD of these ions contains two distinct peaks **(Figure 17C)**, the most abundant ion signal corresponds to a later arrival time peak that is consistent with a conformational preference for a partially helical or elongated conformation. Observation of an elongated conformation by MS-CID-IM-MS analysis suggests that the helical character of the peptide that is observed in the solution phase can be preserved upon transfer of the molecule to the gas phase.

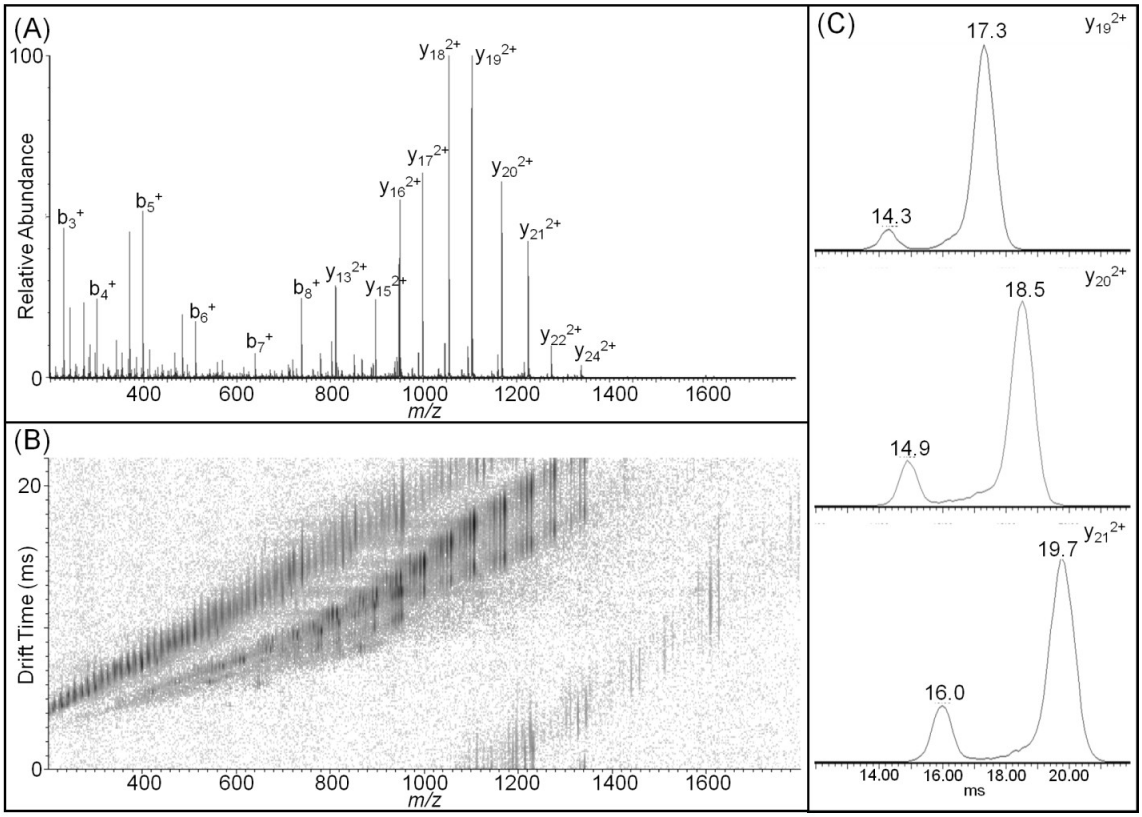


Figure 17. (A) The mass spectrum of MS-CID-IM-MS of the $[M + 3H]^{3+}$ ion of melittin. The most abundant product ions observed are the doubly charged *y*-type ion series from y_{13} to y_{24} owing to the proline effect, specifically, the facile cleavage of the amide bond at Pro_{14}. (B) Mobility-mass plot of the MS-CID-IM-MS of the $[M + 3H]^{3+}$ ion of melittin. As observed with the B-chain of insulin, positive deviations are produced from the doubly charged random coil trend line upon MS-CID-IM-MS analysis. (C) The ATD of these ions contain two distinct peaks, the most abundant occurring at a later arrival time and consistent with a conformational preference for a partially helical or elongated conformation.

The retention of helical character upon dissociation is consistent with studies reported by Jarrold *et al.* on the stability of unsolvated helices.[94] Jarrold *et al.* studied the helical peptide $[Ac\text{-}A_{15}K + H]^+$ and random coil peptide $[Ac\text{-}KA_{15} + H]^+$ as a function of temperature using a high-temperature (500 K) IM drift tube. As temperature is increased, neither the $[Ac\text{-}A_{15}K + H]^+$ helix nor the $[Ac\text{-}KA_{15} + H]^+$ random coil melt

into a collapsed globular structure below the threshold for dissociation. Moreover, the $[Ac\text{-}KA_{15} + H]^+$ peptide ion follows the dominant dissociation pathway to produce *b*-type product ions, whereas the $[Ac\text{-}A_{15}K + H]^+$ peptide ion dissociates to yield the charged lysine product ions. It is important to note that further studies by Jarrold *et al.* indicate that mobile protons as well as drift tube temperature (≥500 K) can influence the gas-phase stability of helices in designed peptides.[95,96] The results from the MS-CID-IM-MS study of B-chain insulin closely agree with Jarrold's findings. That is, the CID spectrum contains abundant product ions that have helical content. Jarrold's observation of a random coil peptide producing *b*-type ions whereas the helical peptide produces *y*-type ions is similar to our results for the doubly charged product ions, specifically, product ions with helical character were identified as *y*-type and those that were random coil were identified as *b*-type (**Figure 18)**. Both Jarrold's work and this current study indicate that the combined contributions of the non-covalent interactions stabilizing the helical conformation in the gas phase exceed the energy required to dissociate a covalent bond. Thus, helices observed in the solution phase and transferred to the gas phase can be identified through MS-CID-IM-MS analysis.

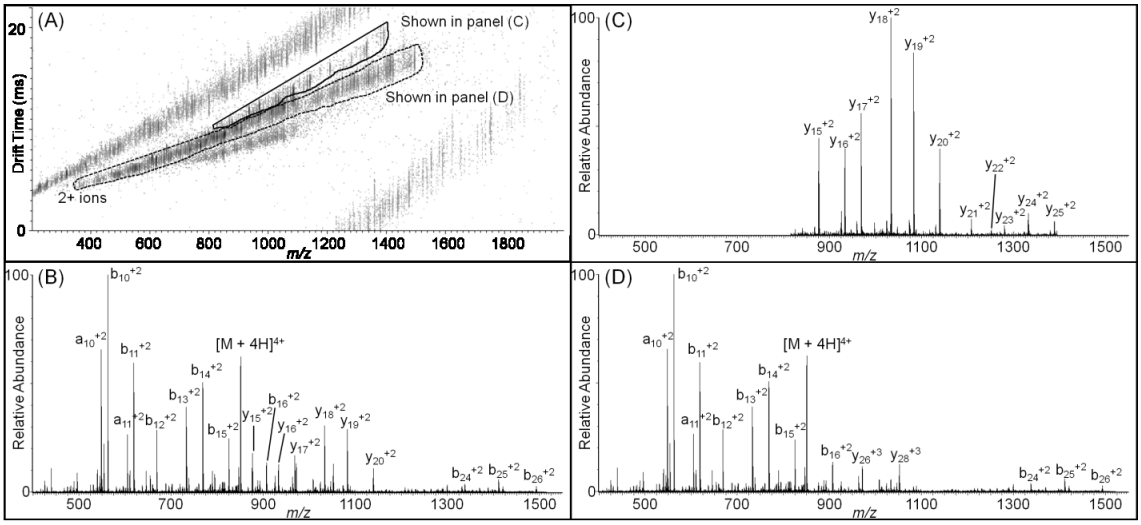


Figure 18. (A) Mobility-mass plot of the MS-CID-IM-MS of the $[M + 4H]^{4+}$ ion of the B-chain of insulin. (B) Composite production spectra showing all doubly charged productions. (C) The extracted MS of the positive deviations from the doubly charged random coil trend line circled (solid line) in (A). These product ions are identified as doubly charged *y*-type ions. (D) The extracted MS of the doubly charged random coil trend line circled (dashed line) in (A). These ions are identified as doubly charged *b*-type ions with minor overlap of triply charged *y*-type ions.

To further evaluate this hypothesis, *i.e.*, that the solution phase helical character is retained by the gas phase ions, as opposed to the helical character being induced by loss of solvent or collisional heating of the gas phase ion, peptides known to exist as random coil or have disordered structure were also examined by MS-CID-IM-MS. For example, somatostatin, the 14 amino acid peptide/hormone (AGCKNFFWKT[10] FTSC) involved in the inhibition of growth hormone as well as other secondary hormones[153,154] was also examined. The $[M + 3H]^{3+}$ ion of somatostatin (generated following reduction with TCEP and alkylation with MMTS) yielded abundant singly and doubly charged product ions, all of which fall along the random coil trend line (see **Figure 19A**). Proteolytic digestion using trypsin of alpha-S1-casein, a 24kDa protein found in milk

that is known to have a disordered structure,[185] resulted in peptides with random coil or disordered structures. **Figure 19C** contains the 2-D mobility-mass plot from the MS-CID-IM-MS analysis of the tryptic peptide HPIK(Acetyl)HQGLPQ[10] EVLNENLLR (2.3 kDa). The singly, doubly, and triply charged CID product ions fall along random coil trend lines indicating that extended coil or helical character are not induced by collisional activation and that the random coil conformations of the peptides are conserved.

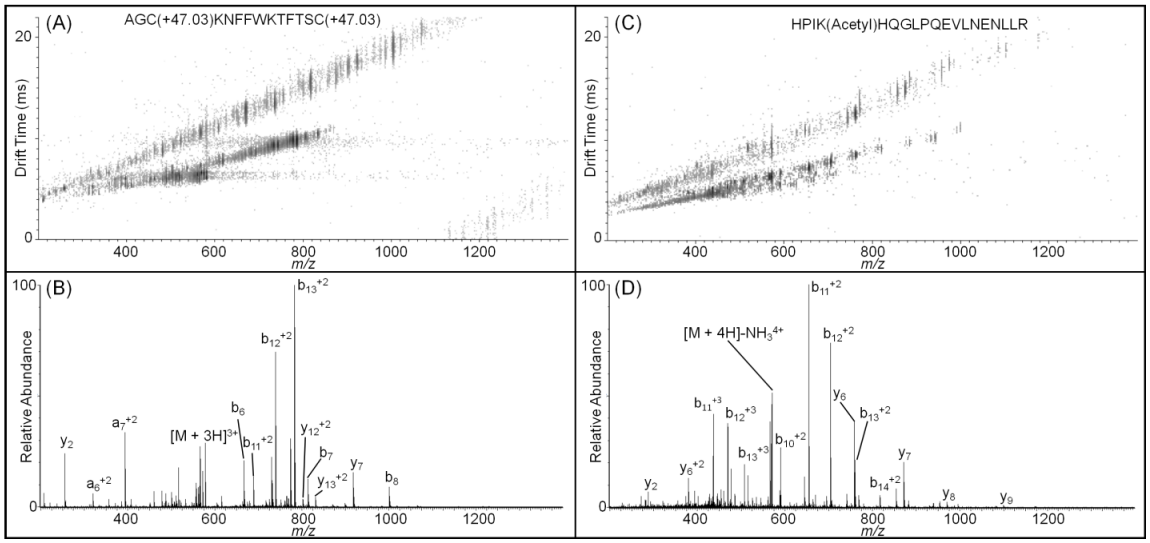


Figure 19. Mobility-mass plots of the MS-CID-IM-MS of (A) the $[M + 3H]^{3+}$ ion of somatostatin (AGC(+47.03)KNFFWKTFTSC(+47.03)) and (C) the $[M + 4H]^{4+}$ ion of the peptide HPIK(Acetyl)HQGLPQEVLNENLLR, both of which have random coil structures. No deviations are observed from the charge state specific trend lines indicating that the product ion conformations have not changed significantly. (B) and (D) show the correspond product ion spectra for each peptide, respectively.

Significance of Collision Cross Section-Molecular Weight Plots

To more fully evaluate the conformational preference of product ions formed by CID of the $[M + 4H]^{4+}$ and $[M + 5H]^{5+}$ ions of B-chain insulin the product ion collision cross sections (CCSs) were calculated from their TWIMS ATD.[181,182] The CCSs are grouped based on product ion charge state in **Figure 20A-C**. **Figure 20A** contains plots of singly charged product ion CCSs as a function of molecular weight (MW). The singly charged product ions fall along a single trend line, regardless of product ion identity, *i.e. b*- or *y*-type ions, indicating that the conformations of the singly charged ions are similar. **Figure 20B** contains similar plots for the doubly charged product ions. Note that two distinct CID product ion trend lines are observed and that the trend lines have very different slopes. In particular, the *y*-type product ions have CCS values that are greater than those for both *a*- and *b*-type product ions, a result that is consistent with a helical conformation for the *y*-type ions. As in the mobility-mass plot, the onset of divergence corresponds to a region of the sequence (-S^{9}HLVEA^{14}-) where the product ions (y^{19}-y^{26}) contain the amino acids that form the α-helix, which facilitates rapid identification and localization of the helical region. The CCS vs. MW plots for the triply charged product ions (**Figure 20C**) also yield two trend lines with different slopes, and the CCS values that fall above the expected random coil trend line are consistent with a helical conformation. Interestingly, the doubly charged CID product ions trend line of helical product ions contained a single ion type, specifically *y*-type product ions, whereas the triply charged helical product ion trend line also contained *a*- and *b*-type ions. Thus, the triply charged product ions can facilitate the identification of the onset of

the helix from both the N- and C- terminus and are complementary to conformational information gained from the doubly charged ion CCS plot. The smallest identified triply charged product ions from the N-terminus correspond to the a/b_{10} ion, which contains two amino acids from the helix region. Deviations from higher charge state random coil trend lines can be sufficient enough to cause overlap with the random coil trend line of a lower charge state. Specifically, the charge state trend lines for the triply charged helical product ions overlap with the doubly charged random coil trend line; however, the overlap in the charge state trend lines is not observed when the ion mobility data is plotted as CCS **(Figure 20C)**. The combination of the doubly and triply charged product ions can be used as an indication of helical conformation from either the N- or C-terminus of the peptide allowing for the identification of the onset and termination of the helical region.

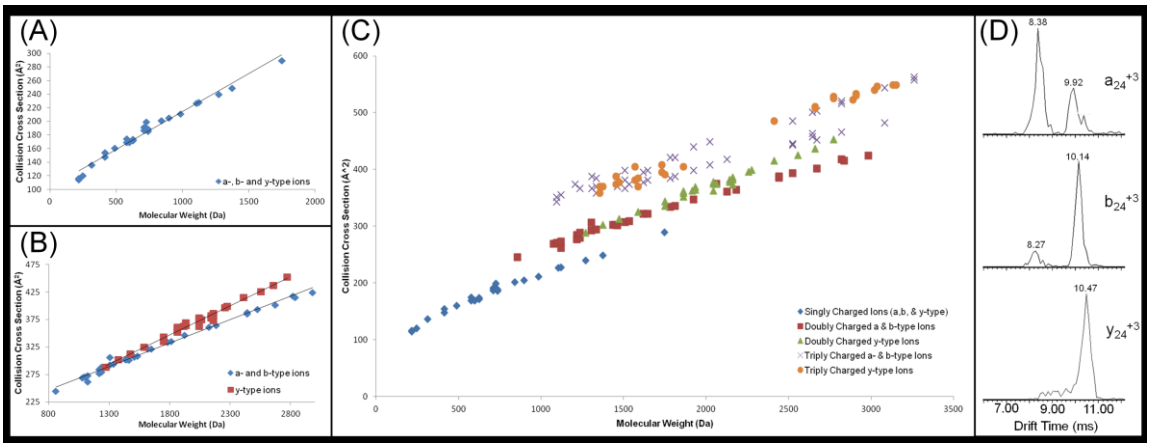


Figure 20. (A-C) Plot of collision cross section (CCS) of product ions from the $[M + 4H]^{4+}$ and $[M + 5H]^{5+}$ ions of B-chain insulin (reduced with DTT) as a function of molecular weight. (A) Singly charged product ions form one trendline, regardless of product ion identity, *i.e. b*- or *y*-type ions. (B) Doubly charged product ions form two distinct trendlines. Product ions from the c-termini (*y*-type) have CCS values greater than those product ions with charge on the n-termini (*a*- and *b*-type) indicating that *y*-type ions are possibly partially helical. (C) CCS-molecular weight plots increase dispersion of product ions when compared with the mobility-mass plot preventing overlap of different charge state conformers. Similar to the doubly charged, the triply charged product ions show two trendlines. *b*- and *y*-type ions encompassing the helix region show two distinct trendlines, *i.e.* product ions with *m/z* greater than ~1700. (D) ATD of triply charged product ions with similar mass. *y*-type ions have one distinct conformer preference whereas *a*- and *b*-type ions have two distinct conformer preferences.

The product ion abundances are not accounted for in the CCS-molecular weight plots. **Figure 20D** contains the ATD of triply charged product ions with similar mass, specifically, the a_{24}, b_{24}, and the y_{24} ions (2737.3, 2765.3, and 2752.3 Da, respectively). The *y*-type ions have a single distinct conformer preference whereas *a*- and *b*-type ions have two distinct conformer preferences. Interestingly, the abundance of the compact conformer is higher for the *a*-type ions, whereas the abundance of the elongated conformer is higher in the ATD for the *b*-type ion. These results are consistent with the previously suggested effects of the macro-dipole on the helical preferences. That is, for

helical conformations the positive charge of the *b*-type product ion is located on the C-terminal acetyl ions and aligned with the macrodipole of the peptide backbone. Independent of the mechanisms by which *a*-type fragment ions are formed, *viz.* by direct bond cleavage of the C_α-keto carbon bond or by initial cleavage of the amide bond followed by loss of CO, the charge on the *a*-type fragment ion is likely delocalized, which would destabilize the helical conformation and promote formation of a random coil conformer.

Conclusion

Incorporation of ion mobility with CID allows product ions to be dispersed in two-dimensions, specifically size-to-charge (IM) and mass-to-charge (MS). This approach allows CID product ions to be separated on the basis of charge state specific trend lines and greatly facilitates the identification of product ions, *viz. b-/y*-type, and internal fragment ions. Moreover, the product ions are separated in conformational space providing additional information about the product ions (random coil, helical/extended coil, and collapsed globule) and in some cases result in distinct trend lines for *b*- and *y*-type ions.

ESI-MS-CID-IM-MS analysis affords rapid and accurate determination of CID product ion CCS. A distinct advantage of including the CCS data in mobility-mass mode is the increased dispersion of product ions compared with the other mobility and/or mass analysis modes. The CCS data illustrates that helical character is retained upon dissociation,[94,130] which suggests that some elements of solution phase conformations are retained by the gas phase ions. This claim is consistent with previous

reports by Jarrold *et al.* on the stability of unsolvated helices.[94-96] Both Jarrold's work and the current study indicate that the combination of contributions of the non-covalent interactions that stabilize the helical conformation of the gas phase ion exceeds the energy required to dissociate a covalent bond.

CHAPTER V

REGIOREGULARITY OF GLUCOSE DERIVED POLYCARBONATES PROBED BY MASS SPECTROMETRY

Introduction

The drive to develop synthetic polymers derived from renewable feedstocks has resulted in the increase in structural and compositional complexity of synthetic polymers, thus making detailed characterization of these newly developed polymeric materials a challenge. Though challenging, polymer analysis remains a key step in understanding the relations of chemo-physical properties of a polymeric material and its function.[97] Recently, the synthetic route of glucose based monomers to create polycarbonates with well-defined end groups has been demonstrated.[110] A suitable analytical protocol was developed to characterize the regiochemistry of this polymeric system.

A variety of analytical techniques exist to fully characterize new synthetic polymers including: separation techniques,[98,99] spectroscopic methods,[100,101] NMR, X-ray,[102] microscopy,[103] and mass spectrometry (MS).[104,105] Each technique has advantages and limitations over others and detailed characterization of the polymer often involves a combination of techniques. Among these techniques, MS has become an indispensable tool for polymer analysis providing molecular mass and molecular mass distribution.[105] Tandem mass spectrometry (MS-MS), including low- and high-energy collision-induced dissociation (CID), provides information on polymer composition as

well as structure.[106,107] CID analysis of saccharides generally produce two main types of cleavage: glycosidic cleavages between neighboring residues and cross-ring cleavages involving the dissociation of the sugar ring. Glycosidic cleavages provide information related to composition and sequence while cross-ring cleavages provide information on the different linkage types.

Low-energy CID analysis of protonated oligosaccharides predominantly results in glycosidic cleavages but does produce some cross-ring cleavages, though in low abundance.[114] By comparison, high-energy CID produces a greater abundance of cross-ring cleavages providing more structural information.[115] Complementary to CID, electron capture dissociation (ECD) and electron transfer dissociation (ETD) have been applied recently to the structural analysis of oligosaccharides.[116-118] These studies showed that cross-ring cleavages were the dominant fragmentation pathway, similar to results observed by high-energy CID. Though commonly applied to investigations in proteomics and glycoproteomics, few studies exist utilizing ETD for the structural analysis of oligosaccharides.

In this study, the polycarbonate polymer derived from glucose monomers is studied by ETD, low- and high-energy CID to determine the regiochemistry. Analysis of the mono- and doubly sodiated polycarbonate polymer by ETD, low- and high-energy CID shows only the presence of the head-to-tail orientation of the subunits. Upon analysis of the triply sodiated polymer by ETD, the subunits are shown to contain both a head-to-tail and tail-to-head orientation. ETD analysis provides reliable structural information of polymers comprised of saccharide subunits.

Experimental

Sample Preparation

Glucose derived polycarbonate samples were synthesized and obtained from the Wooley Group.[110] Sample stock solutions were prepared to a concentration of 10 μg/μL utilizing -20°C tetrahydro furan (THF). Immediately prior to ESI-MS analysis, sample solutions were diluted with 50/50 methanol/water (v/v) to a working concentration of 2μM. For ETD analysis, sample solutions were diluted as previously to a working concentration of 15 μM.

Mass Spectrometry

All ESI-MS experiments were performed using a SYNAPT G2 HDMS mass spectrometer equipped with a nano-electrospray ion source and MassLynx data processor (Waters Corp., Milford, MA, USA). Instrument acquisition parameters used were as follows: an inlet capillary voltage of +1.85 kV, a sampling cone setting of 15 V, a source temperature of 100°C. The ETD reagent utilized was 1,3-dicyanobenzene. ETD parameters were as follows: 25 ml/min helium make-up gas, 32 V discharge current and wave heights of 0.38 V and 0.43 V for the doubly and triply sodiated analytes, respectively. The triply sodiated 15-mer and 16-mer were mass selected for ETD analysis in the quadrupole and transferred to the trap for ETD anlysis. ETD acquisitions were acquired for 30 minutes for the doubly sodiated and 20 minutes for the triply sodiated analyte to obtain abundant product ion signal. All data processing was conducted using Waters MassLynx v4.1. All low-energy CID fragmentation was carried out by collision of ions with background gas (argon) and collision energy for CID of the

polymer was optimized for each polymer, with trap collision voltage generally falling in a range of 75-85 V.

All MALDI-MS experiments were performed using a model 4800 Proteomics Analyzer MALDI-TOF/TOF (Applied Biosystems, Framingham, MA). A 2 μL aliquot of the polymer stock was mixed 1:1 with the MALDI matrix 2,5-dihydroybenzoic acid (DHB) and spotted on a stainless steel MALDI plate. Samples were analyzed in reflectron-mode. The instrument was operated for detection of only positively charged ions.

Product Ion Assignment

Ion assignments were made by determining all potential cross-ring cleavages for both the head-to-tail (HT) and tail-to-head (TH) orientation of the polymer subunit and formulating a table associated with each orientation (**Figure 21A and B**). Each orientation produces unique cross-ring cleavage fragment ions which can be used as diagnostic ions to determine regiochemistry for a particular subunit.

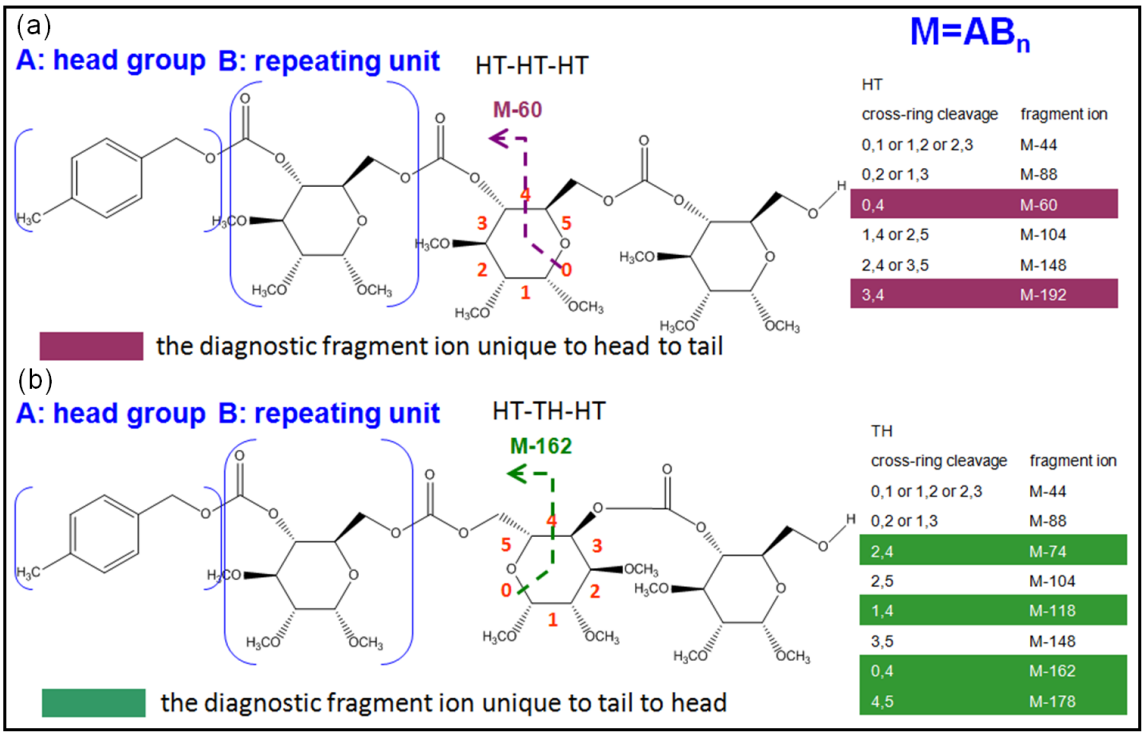


Figure 21. Schematic of all the possible fragment ions with cross-ring cleavages for a head-to-tail (HT) subunit orientation. M-60 and M-192 are the diagnostic fragment ions for HT orientation. (b) Schematic of all the possible fragment ions with cross-ring cleavages for a tail-to-head (TH) subunit orientation. M-74, M-118, M-162 and M-178 are the diagnostic fragment ions for TH orientation.

Results and Discussion

Mass spectrometry is an important tool in the analysis of polymers allowing mass determination as well as mass distribution. Tandem mass spectrometry provides compositional information as well as structural information. However, MS-MS analysis of protonated oligosaccharides via CID predominantly produces glycosidic cleavages providing sequence information but not structural information (regiochemistry). Metal-adducted oligosaccharides have been shown to produce more cross-ring cleavages in low

energy CID.[116,186] Here, we report a general approach for the analysis of polymers derived from saccharide subunits.

ETD Analysis

Two distinct ion series are observed in the MS analysis spectra correlating to the doubly and triply sodiated species of the polymer. For the doubly sodiated series, the most abundant ion signal was the 13-mer (*m/z* 1697.1) while the 15-mer (*m/z* 1304.5) was the most abundant ion signal of the triply sodiated species. In order to determine the regiochemistry of this glucose derived polymer, ETD analysis of the most abundant ion from each series was performed.

Examination of the ETD product ion spectra for the doubly sodiated 13-mer (**Figure 22A**) and triply sodiated 15-mer **(Figure 22B)** indicate several distinct ion series. Though both ions produce charge reduced species and predominantly singly charged product ions, the product ion series observed in the doubly sodiated analysis is different from those observed in the ETD analysis of the triply sodiated species. As seen in **Figure 22A**, four distinct ion series can be observed corresponding to whole subunit loss (248 Da), loss of 104 Da, loss of 148 Da, and loss of 205 Da. Loss of 104 and 148 Da corresponds to cross ring cleavages of the glucose subunits while the loss of 205 Da corresponds to cleavage at the alpha position of the ether linker resulting in the loss of the glucose ring structure. These four series indicate that ETD analysis results in significant dissociation along the polymer as well as cross ring cleavages but does not provide information to determine regiochemistry. The ETD analysis of the triply sodiated 15-mer indicates the presence of five distinct ion series. As with the doubly

sodiated species, whole subunit loss is observed as well as the loss of 104 Da. Unique to the dissociation of the triply sodiated species are the ion series corresponding to loss of 60 Da, loss of 162 Da and loss of 180 Da. The loss of 60 Da corresponds to the cross-ring cleavage 0,4 of the head-to-tail (HT) subunit orientation while the loss of 162 Da and 180 Da corresponds to the cross-ring cleavage 0,4 of the tail-to-head (TH) subunit orientation and loss of water, respectively.

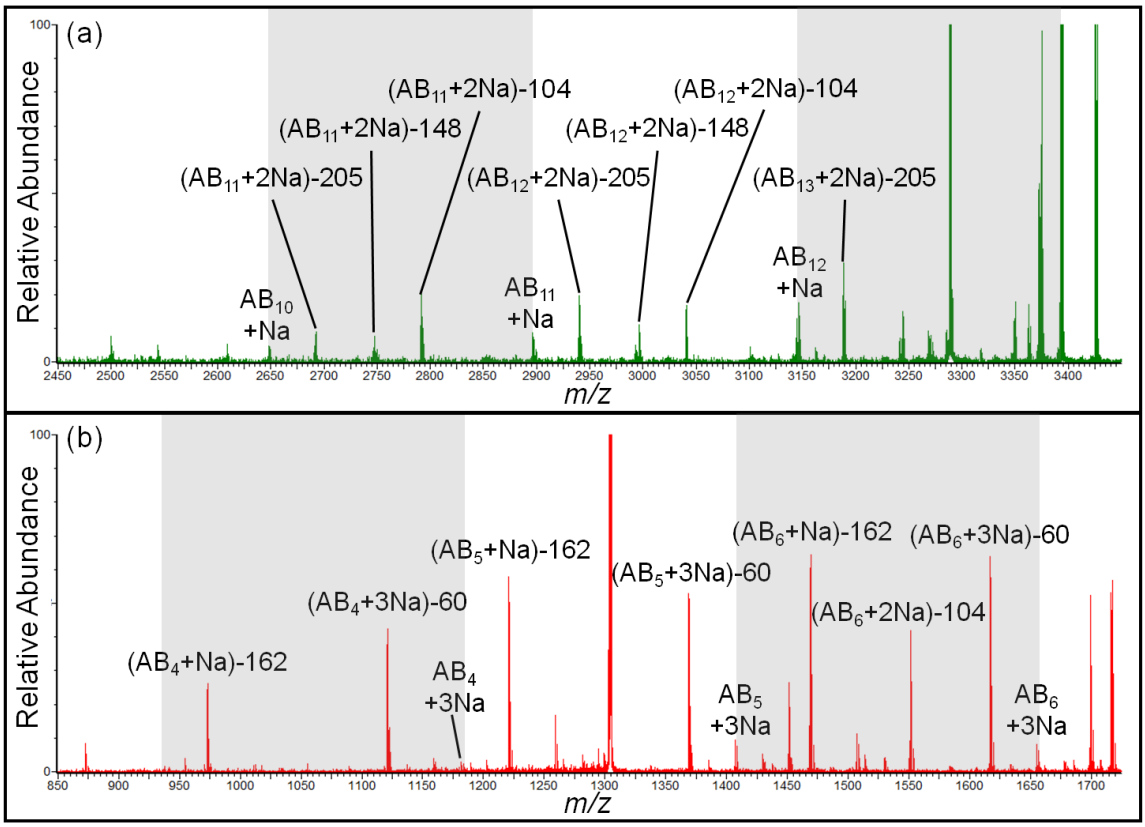


Figure 22. ETD of (A) doubly sodiated 13-mer and (B) triply sodiated 15-mer. Only ETD of the triply sodiated polymer produces the diagnostic ions associated with both the HT and TH subunit orientations.

These distinct ion series found in the doubly and triply sodiated ETD product ion spectra provide information on the regiochemistry which occurs during the polymerization reaction. Ion assignments were made by determining all potential cross-ring cleavages for both the head-to-tail (HT) and tail-to-head (TH) orientation of the polymer subunit and formulating a table associated with each orientation (**Figure 21A and B**). Each orientation produces unique cross-ring cleavage fragment ions which can be used as diagnostic ions to determine regiochemistry for a particular subunit and are highlighted in **Figure 21**. Dissociation of the doubly sodiated species produces several ion series that correspond to either subunit orientation and are therefore not informative in determining regiochemistry. Unlike the doubly sodiated species, the dissociation of the triply sodiated species produces several ion series which can be used to determine regiochemistry, specifically the series corresponding to loss of 60 and 162 Da. The loss of 60 and 162 Da ion series correspond to the 0,4 cross-ring cleavages of the HT and TH subunit orientation, respectively. As seen in **Figure 22B**, the diagnostic ions for both the HT and TH subunit orientation were observed for each subunit observed which indicates that both orientations are present in each subunit. The tandem MS analysis results indicate that the polymerization produces a randomization of orientation making all three regiochemistries (HH, HT, and TT) possible.

Low-energy CID Analysis

Low-energy CID was also used to aid in the determination of regiochemistry. The doubly sodiated 11-mer and 13-mer were both isolated and dissociated. As seen in **Figure 23**, the product ions produced upon collisional activation of the two polymers are

identical with small differences in abundance and the presence of low abundant ion signals. Similar to the ETD product ion spectrum, several ion series can be observed corresponding to the cross-ring cleavages of the glucose subunits. The most abundant ion series is that of the loss of 148 Da which is the 3,5 cross-ring cleavage. As was observed in the ETD product ion spectrum of the doubly sodiated 13-mer, the ion series generated upon collisional activation are not informative in determining the regiochemistry of the polymer as the majority of the ion series correspond to cross-ring cleavages observed in both subunit orientations. The ion series corresponding to loss of 192 Da provides information of the regiochemistry as this cleavage is a diagnostic ion of the HT subunit orientation confirming the presence of this orientation. Interestingly, the loss of 180 Da ion series does not correspond to any cross-ring cleavage associated with either subunit orientation. A possible explanation for this series is that it corresponds to the loss of 162 Da as well as loss of water of the polymer. This explanation is supported by the abundance of water loss peaks observed in **Figure 23A and B**. However, the identity of the ion series cannot be confirmed as this as all other ions identified as loss of water also are accompanied by the ion series without water loss. The 180 Da does not contain such an isotope cluster and therefore can only be proposed as a possible indication of the TH orientation. Lower-energy CID of the doubly sodiated polymers help in the determination of regiochemistry by supporting the presence of the HT subunit orientation but cannot be used to definitively determine regiochemistry.

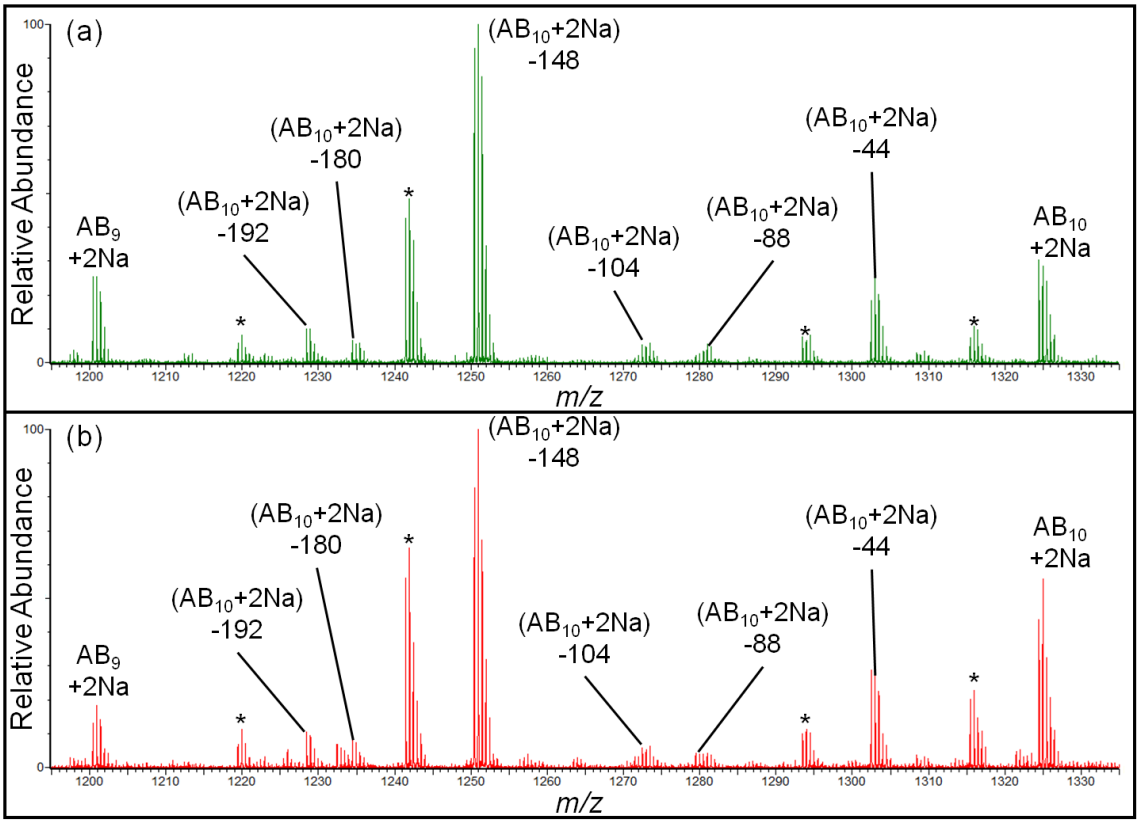


Figure 23. CID of (A) doubly sodiated 11-mer and (B) doubly sodiated 13-mer. Note that regardless of polymer length, sequence informative fragment ions are the same. Isotope clusters marked with a "*" denote loss of water.

High-energy CID

High-energy has been shown to produce cross-ring cleavages of poly-saccharides and was therefore used to help determine the regiochemistry of the glucose derived polymer. MALDI analysis of the polymer produces mono sodiated ions ranging in polydispersion from the 9-mer to the 18-mer in low abundance. As was observed during the ETD and low-energy CID analysis, several ion series are observed. Whole subunit loss, loss of 104 Da, loss of 148 Da, loss of 166 Da, and loss of 205 Da are the observed ion series. Loss of 104, 148, and 205 Da are cross-ring cleavages that are observed in

both orientations while the loss of 166 Da corresponds to the 148 cross-ring cleavage as well as a loss of water. As these series are observed in both subunit orientations, the high-energy CID analysis does not allow us to determine the regiochemistry of the polymer.

Polymer Conformation

Discrepancy with ions associated with different dissociation techniques is expected owning to the different dissociation channels accessed. However, the dissociation products from different sequential charge states should not vary substantially. The difference in observed product ions from the ETD analysis of the doubly and triply sodiated polymer could be a result of each species having a different conformation as electron mediated dissociation techniques have been shown to be influenced by conformation.[187] **Figure 24** shows the charge state normalized ATD for the doubly sodiate (top) and triply sodiated (bottom) 15-mer. Charge normalization of the peaks allows peaks observed in different charge states to qualitatively be compared in order to determine if conformations are the same. Both ATD contain an abundant peak and a low abundant peak. For the doubly sodiated species, the most peak occurs at the later arrival time indicating that the doubly sodiated species has an extended conformation. Conversely, the abundant peak of the triply sodiated sepecies occurs at a shorter arrival time, consistent with the triply sodiated species have a compact or globular conformation. The differences observed product ions formed during ETD analysis could be a result of the difference in conformation associated with each species.

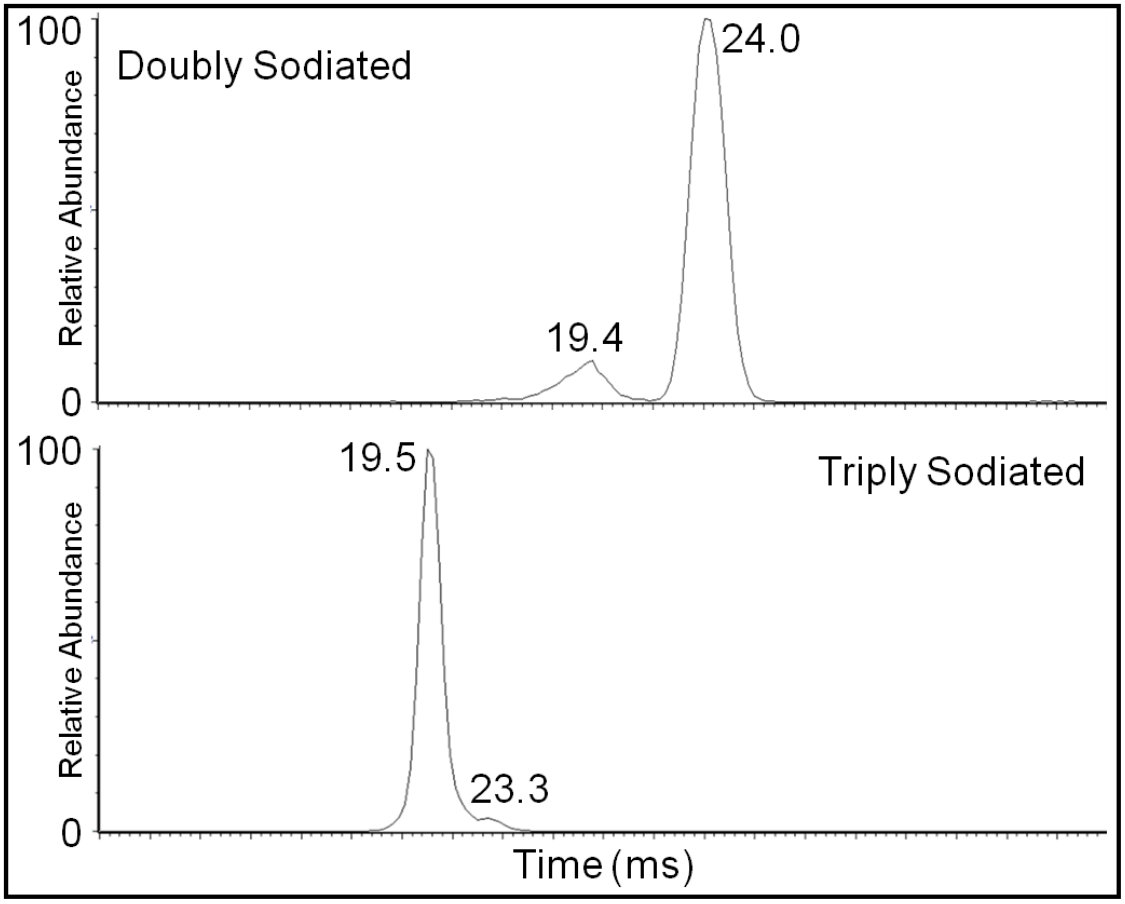


Figure 24. Charge state normalized ATD for the doubly (top) and triply (bottom) sodiated species of the 15-mer polymer. Both ATD indicate the presence of an abundant peak and a low abundant peak. The alignment in arrival time for the peaks observed between each species indicates that both species contains a population of each conformer. The difference in conformation preference between each species could indicate while a discrepancy was observed between the ETD analysis of each species.

Conclusion

The utility of ETD has been demonstrated for characterizing the regiochemistry of glucose derived polymers. ETD as well as low- and high-energy CID generated cross-ring cleavages of the glucose subunits. Interestingly, many of the same cross-ring cleavages are observed in the product ion spectra of all three dissociation

techniques. However, the majority of observed product ions do not indicate the subunit orientation making determination of regiochemistry difficult. Only the analysis of the triply sodiated polymer by ETD allows for the determination that all three regiochemistries (HH, TH, and TT) are present. The loss of 60 and 162 Da observed at every subunit allows for unambiguous assignment indicating that both orientations are present. Low-energy CID can only confirm the presence of the HT subunit orientation owing to the loss of 192 Da series whereas high-energy CID cannot confirm either subunit orientation. In summation, ETD of the triply sodiated polymer shows diagnostic ions for both HT and TH orientation for each subunit in the product ion spectrum indicating that both orientations are present in each subunit. The tandem MS analysis results indicate that the polymerization produces a randomization of orientation making all three regiochemistries (HH, HT, and TT) possible.

CHAPTER VI

CONCLUSION

Summary

An alternative approach for top-down proteomics was developed. Inherent to top-down proteomics is the limitation in dynamic range associated for product ions owning to the large number of fragmentation channels that are available for large ionic systems, *i.e.* proteins, and the resulting spectral congestion. This approach alleviates these limitations. Incorporation of ion mobility to top-down proteomics, the complex mixture of protein CID product ions are dispersed in two-dimensions, specifically size-to-charge (IM) and mass-to-charge (MS), and the resulting 2-D data display greatly facilitates the top-down information contents; (i) charge state specific trend lines, (ii) increased dynamic range, and (iii) separation of overlapping ion signals. The increase in peak capacity allows for the detection of low abundant fragment ions which increases the likelihood of identifying all the possible sequence variations and modifications in the protein of interest as seen with the methyl ester derivative of ubiquitin and the single point mutation of human iron sulfur cluster U. Detection of low abundant ions as well as separation of overlapping ion signals increases the primary sequence coverage and the confidence of sequence assignments.

The information provided by this approach, specifically, the 2-D data display as well as the conformational information contained in the ATD, has been applied to screening of PTM, specifically, disulfide bonds. Intra-molecular disulfide bonds create

cyclic peptides and loops in proteins constraining the molecule, limiting the degrees of freedom. MS-CID-IM-MS analysis of disulfide containing ions demonstrates the formation of a peak at a later ATD for the collisionally activated parent ion. This later ATD peak is consistent with an elongated structure and analysis of small model systems suggest that this conformational change is the result of backbone cleavages and not ion heating. Monitoring the ATD of the parent pre- and post- CID has been demonstrated as a qualitative technique for the detection of disulfide bonds and holds potential as a rapid screening technique.

Analysis of disulfide bond containing molecules by MS-CID-IM-MS demonstrated the conformational content of collisionally activated molecules gained by this approach. Therefore, investigations into whether specific secondary structural elements are preserved in CID product ions were performed. This approach allows CID product ions to be separated on the basis of charge state specific trend lines. As product ions are separated in conformational space, observed deviations from charge state specific trend lines provide additional information about the product ion conformation (random coil, helical/extended coil, and collapsed globule) and in some cases result in distinct trend lines for *b*- and *y*-type ions. Identifying the onset in divergence from the charge state specific trend line allows the change in secondary structure to be localized to the peptide sequence.

The separation of product ions in conformational space afford by this approach allows the rapid and accurate determination of CID product ion CCS. Incorporation of CCS data has several advantages, most notably an increased dispersion of product ions

compared with the other mobility-mass data. The CCS suggests that the helical character of the B-chain of insulin is retained upon dissociation indicating that some elements of solution phase structure can be successful are retained by the gas phase ions. The combination of contributions of the non-covalent interactions that stabilize the helical conformation of the gas phase ion exceeds the energy required to dissociate a covalent bond allowing the helical character to be detected in the gas phase by IM. Moreover, when the CCS-molecular weight plots are correlated to product ion identity and abundance, a better understanding of the fragmentation mechanism and conformational changes is gained. The *b*-and *y*-type product ions which retain helical character contains a positive charge aligned with the macrodipole of the peptide backbone. However, *a*-type fragment ions, formed either by direct bond cleavage of the C_α-keto carbon bond or by initial cleavage of the amide bond followed by loss of CO, likely have a delocalized charge which acts to destabilize the helical conformation and promote formation of a random coil conformer.

The MS-CID-IM-MS approach has applications beyond biological samples and has been used to aid in the determination of the regiochemistry of glucose derived polycarbonates. This approach was used to detect low abundant ion signals corresponding to cross-ring cleavages. However, the CID data indicated the presence of only one orientation, contrary to ETD analysis and NMR results.

REFERENCES

(1) Venter, J. C.; Adams, M. D.; Myers, E. W.; Li, P. W.; Mural, R. J.; Sutton, G. G.; Smith, H. O.; Yandell, M.; Evans, C. A.; Holt, R. A.; Gocayne, J. D.; Amanatides, P.; Ballew, R. M.; Huson, D. H.; Wortman, J. R.; Zhang, Q.; Kodira, C. D.; Zheng, X. H.; Chen, L.; Skupski, M.; Subramanian, G.; Thomas, P. D.; Zhang, J.; Gabor Miklos, G. L.; Nelson, C.; Broder, S.; Clark, A. G.; Nadeau, J.; McKusick, V. A.; Zinder, N.; Levine, A. J.; Roberts, R. J.; Simon, M.; Slayman, C.; Hunkapiller, M.; Bolanos, R.; Delcher, A.; Dew, I.; Fasulo, D.; Flanigan, M.; Florea, L.; Halpern, A.; Hannenhalli, S.; Kravitz, S.; Levy, S.; Mobarry, C.; Reinert, K.; Remington, K.; Abu-Threideh, J.; Beasley, E.; Biddick, K.; Bonazzi, V.; Brandon, R.; Cargill, M.; Chandramouliswaran, I.; Charlab, R.; Chaturvedi, K.; Deng, Z.; Di Francesco, V.; Dunn, P.; Eilbeck, K.; Evangelista, C.; Gabrielian, A. E.; Gan, W.; Ge, W.; Gong, F.; Gu, Z.; Guan, P.; Heiman, T. J.; Higgins, M. E.; Ji, R. R.; Ke, Z.; Ketchum, K. A.; Lai, Z.; Lei, Y.; Li, Z.; Li, J.; Liang, Y.; Lin, X.; Lu, F.; Merkulov, G. V.; Milshina, N.; Moore, H. M.; Naik, A. K.; Narayan, V. A.; Neelam, B.; Nusskern, D.; Rusch, D. B.; Salzberg, S.; Shao, W.; Shue, B.; Sun, J.; Wang, Z.; Wang, A.; Wang, X.; Wang, J.; Wei, M.; Wides, R.; Xiao, C.; Yan, C. *Science* **2001**, *291*, 1304.

(2) Fields, S. *Science* **2001**, *291*, 1221.

(3) Wilkins, M. R.; Sanchez, J. C.; Williams, K. L.; Hochstrasser, D. F. *Electrophoresis* **1996**, *17*, 830.

(4) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198.

(5) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. *J Am Chem Soc* **1999**, *121*, 806.

(6) Mann, M.; Wilm, M. *Anal Chem* **1994**, *66*, 4390.

(7) Senko, M. W.; Speir, J. P.; Mclafferty, F. W. *Analytical Chemistry* **1994**, *66*, 2801.

(8) Chi, A.; Huttenhower, C.; Geer, L. Y.; Coon, J. J.; Syka, J. E. P.; Bai, D. L.; Shabanowitz, J.; Burke, D. J.; Troyanskaya, O. G.; Hunt, D. F. *P Natl Acad Sci USA* **2007**, *104*, 2193.

(9) Molina, H.; Horn, D. M.; Tang, N.; Mathivanan, S.; Pandey, A. *P Natl Acad Sci USA* **2007**, *104*, 2199.

(10) Kelleher, N. L.; Zubarev, R. A.; Bush, K.; Furie, B.; Furie, B. C.; McLafferty, F. W.; Walsh, C. T. *Anal Chem* **1999**, *71*, 4250.

(11) Kjeldsen, F.; Haselmann, K. F.; Budnik, B. A.; Sorensen, E. S.; Zubarev, R. A. *Anal Chem* **2003**, *75*, 2355.

(12) Gorshkov, M. V.; Masselon, C. D.; Nikolaev, E. N.; Udseth, H. R.; Pasa-Tolic, L.; Smith, R. D. *Int J Mass Spectrom* **2004**, *234*, 131.

(13) Patrie, S. M.; Ferguson, J. T.; Robinson, D. E.; Whipple, D.; Rother, M.; Metcalf, W. W.; Kelleher, N. L. *Mol Cell Proteomics* **2006**, *5*, 14.

(14) Forbes, A. J.; Patrie, S. M.; Taylor, G. K.; Kim, Y. B.; Jiang, L.; Kelleher, N. L. *Proc Natl Acad Sci U S A* **2004**, *101*, 2678.

(15) Cooper, H. J.; Hakansson, K.; Marshall, A. G. *Mass Spectrometry Reviews* **2005**, *24*, 201.

(16) 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.

(17) Marshall, A. G.; Guan, S. H. *Rapid Commun Mass Sp* **1996**, *10*, 1819.

(18) Anderson, N. G.; Ahmad, T. *Front Biosci* **2002**, *7*, d1926.

(19) Ofarrell, P. H. *J Biol Chem* **1975**, *250*, 4007.

(20) Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P. *Analytical Biochemistry* **1992**, *203*, 173.

(21) Bjellqvist, B.; Ek, K.; Righetti, P. G.; Gianazza, E.; Gorg, A.; Westermeier, R.; Postel, W. *J Biochem Bioph Meth* **1982**, *6*, 317.

(22) Anderson, N. L.; Anderson, N. G. *Electrophoresis* **1991**, *12*, 883.

(23) Lim, P.; North, R.; Vigh, G. *Electrophoresis* **2007**, *28*, 1851.

(24) Cologna, S. M.; Russell, W. K.; Lim, P. J.; Vigh, G.; Russell, D. H. *J Am Soc Mass Spectrom* **2010**, *21*, 1612.

(25) Righetti, P. G.; Castagna, A.; Herbert, B.; Reymond, F.; Rossier, J. S. *Proteomics* **2003**, *3*, 1397.

(26) Makinen, M. A.; Anttalainen, O. A.; Sillanpaa, M. E. *Anal Chem* **2010**, *82*, 9594.

(27) Fenn, L. S.; Kliman, M.; Mahsut, A.; Zhao, S. R.; McLean, J. A. *Anal Bioanal Chem* **2009**, *394*, 235.

(28) Becker, C.; Qian, K.; Russell, D. H. *Anal Chem* **2008**, *80*, 8592.

(29) Becker, C.; Fernandez-Lima, F. A.; Russell, D. H. *Spectroscopy* **2009**, *24*, 38.

(30) Karas, M.; Hillenkamp, F. *Anal Chem* **1988**, *60*, 2299.

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

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

(33) Loo, J. A.; Edmonds, C. G.; Udseth, H. R.; Smith, R. D. *Analytical Chemistry* **1990**, *62*, 693.

(34) Sowell, R. A.; Koeniger, S. L.; Valentine, S. J.; Moon, M. H.; Clemmer, D. E. *J Am Soc Mass Spectr* **2004**, *15*, 1341.

(35) Stone, E.; Gillig, K. J.; Ruotolo, B.; Fuhrer, K.; Gonin, M.; Schultz, A.; Russell, D. H. *Anal Chem* **2001**, *73*, 2233.

(36) Stone, E. G.; Gillig, K. J.; Ruotolo, B. T.; Russell, D. H. *Int J Mass Spectrom* **2001**, *212*, 519.

(37) Ptacek, J.; Devgan, G.; Michaud, G.; Zhu, H.; Zhu, X.; Fasolo, J.; Guo, H.; Jona, G.; Breitkreutz, A.; Sopko, R.; McCartney, R. R.; Schmidt, M. C.; Rachidi, N.; Lee, S. J.; Mah, A. S.; Meng, L.; Stark, M. J.; Stern, D. F.; De Virgilio, C.; Tyers, M.; Andrews, B.; Gerstein, M.; Schweitzer, B.; Predki, P. F.; Snyder, M. *Nature* **2005**, *438*, 679.

(38) Cohen, P. *Nat Cell Biol* **2002**, *4*, E127.

(39) Lee, N. Y.; Koland, J. G. *Protein Sci* **2005**, *14*, 2793.

(40) Taylor, K. A.; Feig, M.; Brooks, C. L., 3rd; Fagnant, P. M.; Lowey, S.; Trybus, K. M. *J Struct Biol* **2013**.

(41) Jonsson, T. J.; Tsang, A. W.; Lowther, W. T.; Furdui, C. M. *J Biol Chem* **2008**, *283*, 22890.

(42) Anfinsen, C. B. *Science* **1973**, *181*, 223.

(43) Givol, D.; Delorenz.F; Goldberg.Rf; Anfinsen, C. B. *P Natl Acad Sci USA* **1965**, *53*, 676.

(44) Yano, H.; Kuroda, S.; Buchanan, B. B. *Proteomics* **2002**, *2*, 1090.

(45) Mansfeld, J.; Vriend, G.; Dijkstra, B. W.; Veltman, O. R.; VandenBurg, B.; Venema, G.; UlbrichHofmann, R.; Eijsink, V. G. H. *J Biol Chem* **1997**, *272*, 11152.

(46) Shimaoka, M.; Lu, C.; Palframan, R. T.; von Andrian, U. H.; McCormack, A.; Takagi, J.; Springer, T. A. *Proc Natl Acad Sci U S A* **2001**, *98*, 6009.

(47) Narasimhan, L.; Singh, J.; Humblet, C.; Guruprasad, K.; Blundell, T. *Nat Struct Biol* **1994**, *1*, 850.

(48) Renaud, L. P.; Bourque, C. W. *Prog Neurobiol* **1991**, *36*, 131.

(49) Lehrer, R. I.; Ganz, T. *Curr Opin Immunol* **1999**, *11*, 23.

(50) Patterson, S. D.; Katta, V. *Anal Chem* **1994**, *66*, 3727.

(51) Janecki, D. J.; Nemeth, J. F. *J Mass Spectrom* **2011**, *46*, 677.

(52) Fukuyama, Y.; Iwamoto, S.; Tanaka, K. *J Mass Spectrom* **2006**, *41*, 191.

(53) Quinton, L.; Demeure, K.; Dobson, R.; Gilles, N.; Gabelica, V.; De Pauw, E. *Journal of Proteome Research* **2007**, *6*, 3216.

(54) Thakur, S. S.; Balaram, P. *J Am Soc Mass Spectrom* **2008**, *19*, 358.

(55) Williams, B. J.; Russell, W. K.; Russell, D. H. *J Mass Spectrom* **2010**, *45*, 157.

(56) Gunawardena, H. P.; O'Hair, R. A.; McLuckey, S. A. *Journal of Proteome Research* **2006**, *5*, 2087.

(57) Lioe, H.; Duan, M.; O'Hair, R. A. *Rapid Commun Mass Spectrom* **2007**, *21*, 2727.

(58) Hirs, C. H. *J Biol Chem* **1956**, *219*, 611.

(59) King, G. J.; Jones, A.; Kobe, B.; Huber, T.; Mouvadov, D.; Hume, D. L.; Ross, I. L. *Anal Chem* **2008**, *80*, 5036.

(60) Soderblom, E. J.; Goshe, M. B. *Anal Chem* **2006**, *78*, 8059.

(61) Chrisman, P. A.; McLuckey, S. A. *Journal of Proteome Research* **2002**, *1*, 549.

(62) Brgles, M.; Kurtovic, T.; Halassy, B.; Allmaier, G.; Marchetti-Deschmann, M. *J Mass Spectrom* **2011**, *46*, 153.

(63) Gilbert, H. F. *Methods Enzymol* **1995**, *251*, 8.

(64) Morris, H. R.; Pucci, P.; Panico, M.; Marino, G. *Biochem J* **1990**, *268*, 803.

(65) Gorman, J. J.; Wallis, T. P.; Pitt, J. J. *Mass Spectrom Rev* **2002**, *21*, 183.

(66) Xu, H.; Zhang, L.; Freitas, M. A. *J Proteome Res* **2008**, *7*, 138.

(67) Zubarev, R. A.; Kruger, N. A.; Fridriksson, E. K.; Lewis, M. A.; Horn, D. M.; Carpenter, B. K.; McLafferty, F. W. *Journal of the American Chemical Society* **1999**, *121*, 2857.

(68) Chowdhury, S. M.; Munske, G. R.; Ronald, R. C.; Bruce, J. E. *J Am Soc Mass Spectrom* **2007**, *18*, 493.

(69) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *Journal of the American Chemical Society* **1998**, *120*, 3265.

(70) Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F. *Proc Natl Acad Sci U S A* **2004**, *101*, 9528.

(71) Liu, J.; Gunawardena, H. P.; Huang, T. Y.; McLuckey, S. A. *Int J Mass Spectrom* **2008**, *276*, 160.

(72) Han, X.; Jin, M.; Breuker, K.; McLafferty, F. W. *Science* **2006**, *314*, 109.

(73) Ryan, C. M.; Souda, P.; Halgand, F.; Wong, D. T.; Loo, J. A.; Faull, K. F.; Whitelegge, J. P. *J Am Soc Mass Spectrom* **2010**, *21*, 908.

(74) Ren, D.; Pipes, G. D.; Liu, D.; Shih, L. Y.; Nichols, A. C.; Treuheit, M. J.; Brems, D. N.; Bondarenko, P. V. *Analytical Biochemistry* **2009**, *392*, 12.

(75) Chait, B. T. *Science* **2006**, *314*, 65.

(76) Hoaglund-Hyzer, C. S.; Counterman, A. E.; Clemmer, D. E. *Chem Rev* **1999**, *99*, 3037.

(77) Jarrold, M. F. *Annu Rev Phys Chem* **2000**, *51*, 179.

(78) Ruotolo, B. T.; Tate, C. C.; Russell, D. H. *J Am Soc Mass Spectrom* **2004**, *15*, 870.

(79) Ruotolo, B. T.; Gillig, K. J.; Woods, A. S.; Egan, T. F.; Ugarov, M. V.; Schultz, J. A.; Russell, D. H. *Anal Chem* **2004**, *76*, 6727.

(80) Carstensen, L.; Zoldak, G.; Schmid, F. X.; Sterner, R. *Biochemistry* **2012**, *51*, 3420.

(81) Walter, P.; Ron, D. *Science* **2011**, *334*, 1081.

(82) Honig, B.; Ray, A.; Levinthal, C. *Proc Natl Acad Sci U S A* **1976**, *73*, 1974.

(83) Guinn, E. J.; Kontur, W. S.; Tsodikov, O. V.; Shkel, I.; Record, M. T., Jr. *Proc Natl Acad Sci U S A* **2013**, *110*, 16784.

(84) Baker, D. *Nature* **2000**, *405*, 39.

(85) Gianni, S.; Guydosh, N. R.; Khan, F.; Caldas, T. D.; Mayor, U.; White, G. W.; DeMarco, M. L.; Daggett, V.; Fersht, A. R. *Proc Natl Acad Sci U S A* **2003**, *100*, 13286.

(86) White, S. H.; Wimley, W. C. *Annu Rev Biophys Biomol Struct* **1999**, *28*, 319.

(87) Shi, P.; Li, D.; Lai, C.; Zhang, L.; Tian, C. *Biochem Biophys Res Commun* **2013**, *437*, 408.

(88) Levitt, M.; Chothia, C. *Nature* **1976**, *261*, 552.

(89) Gao, J.; Chen, J. *J Phys Chem B* **2013**, *117*, 8330.

(90) Mattos, C.; Ringe, D. *Curr Opin Struct Biol* **2001**, *11*, 761.

(91) Chen, L.; Chen, S. H.; Russell, D. H. *Anal Chem* **2013**, *85*, 7826.

(92) Ruotolo, B. T.; Gillig, K. J.; Stone, E. G.; Russell, D. H. *J Chromatogr B* **2002**, *782*, 385.

(93) Ruotolo, B. T.; Gillig, K. J.; Woods, A. S.; Egan, T. F.; Ugarov, M. V.; Schultz, J. A.; Russell, D. H. *Anal Chem* **2004**, *76*, 6727.

(94) Kohtani, M.; Jones, T. C.; Schneider, J. E.; Jarrold, M. F. *J Am Chem Soc* **2004**, *126*, 7420.

(95) Kohtani, M.; Jones, T. C.; Sudha, R.; Jarrold, M. F. *J Am Chem Soc* **2006**, *128*, 7193.

(96) Kohtani, M.; Schneider, J. E.; Jones, T. C.; Jarrold, M. F. *J Am Chem Soc* **2004**, *126*, 16981.

(97) Koenig, J. L. *Physical Properties of Polymers*; Cambridge University Press: New York, USA, 2004.

(98) Chang, T. Y. *Adv Polym Sci* **2003**, *163*, 1.

(99) Messaud, F. A.; Sanderson, R. D.; Runyon, J. R.; Otte, T.; Pasch, H.; Williams, S. K. R. *Prog Polym Sci* **2009**, *34*, 351.

(100) Heigl, N.; Petter, C. H.; Rainer, M.; Najam-Ul-Haq, M.; Vallant, R. M.; Bakry, R.; Bonn, G. K.; Huck, C. W. *J near Infrared Spec* **2007**, *15*, 269.

(101) Ngamna, O.; Morrin, A.; Killard, A. J.; Moulton, S. E.; Smyth, M. R.; Wallace, G. G. *Langmuir* **2007**, *23*, 8569.

(102) Medhioub, H.; Zerrouki, C.; Fourati, N.; Smaoui, H.; Guermazi, H.; Bonnet, J. J. *J Appl Phys* **2007**, *101*.

(103) Brostow, W.; Gorman, B. P.; Olea-Mejia, O. *Mater Lett* **2007**, *61*, 1333.

(104) Montaudo, G.; Samperi, F.; Montaudo, M. S. *Prog Polym Sci* **2006**, *31*, 277.

(105) Weidner, S. M.; Trimpin, S. *Anal Chem* **2010**, *82*, 4811.

(106) Hilton, G. R.; Jackson, A. T.; Thalassinos, K.; Scrivens, J. H. *Analytical Chemistry* **2008**, *80*, 9720.

(107) Jackson, A. T.; Green, M. R.; Bateman, R. H. *Rapid Communications in Mass Spectrometry* **2006**, *20*, 3542.

(108) Morris, H. R.; Paxton, T.; Dell, A.; Langhorne, J.; Berg, M.; Bordoli, R. S.; Hoyes, J.; Bateman, R. H. *Rapid Commun Mass Spectrom* **1996**, *10*, 889.

(109) Polce, M. J.; Ocampo, M.; Quirk, R. P.; Leigh, A. M.; Wesdemiotis, C. *Anal Chem* **2008**, *80*, 355.

(110) Mikami, K.; Lonnecker, A. T.; Gustafson, T. P.; Zinnel, N. F.; Pai, P. J.; Russell, D. H.; Wooley, K. L. *J Am Chem Soc* **2013**.

(111) Ashline, D.; Singh, S.; Hanneman, A.; Reinhold, V. *Anal Chem* **2005**, *77*, 6250.

(112) Sandra, K.; Devreese, B.; Van Beeumen, J.; Stals, I.; Claeyssens, M. *J Am Soc Mass Spectrom* **2004**, *15*, 413.

(113) Jiang, H.; Desaire, H.; Butnev, V. Y.; Bousfield, G. R. *J Am Soc Mass Spectrom* **2004**, *15*, 750.

(114) Reinhold, V. N.; Reinhold, B. B.; Costello, C. E. *Analytical Chemistry* **1995**, *67*, 1772.

(115) Lemoine, J.; Fournet, B.; Despeyroux, D.; Jennings, K. R.; Rosenberg, R.; Dehoffmann, E. *J Am Soc Mass Spectr* **1993**, *4*, 197.

(116) Adamson, J. T.; Hakansson, K. *Anal Chem* **2007**, *79*, 2901.

(117) Han, L.; Costello, C. E. *J Am Soc Mass Spectrom* **2011**, *22*, 997.

(118) Zhao, C.; Xie, B.; Chan, S. Y.; Costello, C. E.; O'Connor, P. B. *J Am Soc Mass Spectr* **2008**, *19*, 138.

(119) Bogdanov, B.; Smith, R. D. *Mass Spectrom Rev* **2005**, *24*, 168.

(120) Han, X. M.; Aslanian, A.; Yates, J. R. *Curr Opin Chem Biol* **2008**, *12*, 483.

(121) Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J. H.; Begley, T. P.; McLafferty, F. W. *Journal of the American Chemical Society* **2002**, *124*, 672.

(122) Tran, J. C.; Zamdborg, L.; Ahlf, D. R.; Lee, J. E.; Catherman, A. D.; Durbin, K. R.; Tipton, J. D.; Vellaichamy, A.; Kellie, J. F.; Li, M.; Wu, C.; Sweet, S. M.; Early, B. P.; Siuti, N.; LeDuc, R. D.; Compton, P. D.; Thomas, P. M.; Kelleher, N. L. *Nature* **2011**, *480*, 254.

(123) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Anal Chem* **2001**, *73*, 3274.

(124) Stephenson, J. L.; McLuckey, S. A. *Anal Chem* **1998**, *70*, 3533.

(125) Halgand, F.; Habchi, J.; Cravello, L.; Martinho, M.; Guigliarelli, B.; Longhi, S. *Anal Chem* **2011**, *83*, 7306.

(126) Reid, G. E.; Simpson, R. J.; O'Hair, R. A. J. *J Am Soc Mass Spectr* **1998**, *9*, 945.

(127) Tsai, C. L.; Bridwell-Rabb, J.; Barondeau, D. P. *Biochemistry* **2011**, *50*, 6478.

(128) Giles, K.; Williams, J. P.; Campuzano, I. *Rapid Commun Mass Sp* **2011**, *25*, 1559.

(129) Chanthamontri, C.; Liu, J.; McLuckey, S. A. *Int J Mass Spectrom* **2009**, *283*, 9.

(130) Ruotolo, B. T.; Verbeck, G. F.; Thomson, L. M.; Gillig, K. J.; Russell, D. H. *J Am Chem Soc* **2002**, *124*, 4214.

(131) Clemmer, D. E.; Jarrold, M. F. *Journal of Mass Spectrometry* **1997**, *32*, 577.

(132) McLean, J. A.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. *Int J Mass Spectrom* **2005**, *240*, 301.

(133) Terwilliger, T. C.; Eisenberg, D. *J Biol Chem* **1982**, *257*, 6010.

(134) Bazzo, R.; Tappin, M. J.; Pastore, A.; Harvey, T. S.; Carver, J. A.; Campbell, I. D. *Eur J Biochem* **1988**, *173*, 139.

(135) Vanberkel, G. J.; Glish, G. L.; Mcluckey, S. A.; Tuinman, A. A. *Anal Chem* **1990**, *62*, 786.

(136) Meotner, M.; Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *Rapid Commun Mass Sp* **1995**, *9*, 829.

(137) Little, D. P.; Speir, J. P.; Senko, M. W.; Oconnor, P. B.; Mclafferty, F. W. *Anal Chem* **1994**, *66*, 2809.

(138) Vaisar, T.; Urban, J. *Journal of Mass Spectrometry* **1996**, *31*, 1185.

(139) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J Am Chem Soc* **1996**, *118*, 8365.

(140) Loo, J. A.; Edmonds, C. G.; Smith, R. D. *Analytical Chemistry* **1993**, *65*, 425.

(141) Huang, Y.; Tseng, G. C.; Yuan, S.; Pasa-Tolic, L.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Journal of Proteome Research* **2008**, *7*, 70.

(142) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc Natl Acad Sci U S A* **2000**, *97*, 10313.

(143) 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.

(144) Engel, B. J.; Pan, P.; Reid, G. E.; Wells, J. M.; McLuckey, S. A. *Int J Mass Spectrom* **2002**, *219*, 171.

(145) Cobb, J. S.; Easterling, M. L.; Agar, J. N. *J Am Soc Mass Spectrom* **2010**, *21*, 949.

(146) Wiesner, J.; Premsler, T.; Sickmann, A. *Proteomics* **2008**, *8*, 4466.

(147) DeSouza, L.; Diehl, G.; Rodrigues, M. J.; Guo, J. Z.; Romaschin, A. D.; Colgan, T. J.; Siu, K. W. M. *Journal of Proteome Research* **2005**, *4*, 377.

(148) Apuy, J. L.; Chen, X. H.; Russell, D. H.; Baldwin, T. O.; Giedroc, D. P. *Biochemistry* **2001**, *40*, 15164.

(149) Apuy, J. L.; Park, Z. Y.; Swartz, P. D.; Dangott, L. J.; Russell, D. H.; Baldwin, T. O. *Biochemistry* **2001**, *40*, 15153.

(150) Zinnel, N. F.; Pai, P. J.; Russell, D. H. *Anal Chem* **2012**, *84*, 3390.

(151) Park, Z. Y.; Russell, D. H. *Anal Chem* **2000**, *72*, 2667.

(152) Morris, H. R.; Pucci, P. *Biochem Bioph Res Co* **1985**, *126*, 1122.

(153) Yamada, Y.; Post, S. R.; Wang, K.; Tager, H. S.; Bell, G. I.; Seino, S. *Proc Natl Acad Sci U S A* **1992**, *89*, 251.

(154) Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; Li, Q.; Imura, H.; Seino, S.; Seino, Y. *Biochem Bioph Res Co* **1993**, *195*, 844.

(155) Valentine, S. J.; Counterman, A. E.; Hoaglund-Hyzer, C. S.; Clemmer, D. E. *J Phys Chem B* **1999**, *103*, 1203.

(156) Gross, D. S.; Schnier, P. D.; RodriguezCruz, S. E.; Fagerquist, C. K.; Williams, E. R. *P Natl Acad Sci USA* **1996**, *93*, 3143.

(157) Valentine, S. J.; Anderson, J. G.; Ellington, A. D.; Clemmer, D. E. *J Phys Chem B* **1997**, *101*, 3891.

(158) Desaire, H.; Clark, D. F.; Go, E. P.; Toumi, M. L. *J Am Soc Mass Spectr* **2011**, *22*, 492.

(159) McLuckey, S. A.; Liu, J.; Gunawardena, H. P.; Huang, T. Y. *Int J Mass Spectrom* **2008**, *276*, 160.

(160) Shaffer, J.; Adams, J. A. *Biochemistry* **1999**, *38*, 5572.

(161) Ruotolo, B. T.; Russell, D. H. *J Phys Chem B* **2004**, *108*, 15321.

(162) Lin, K.; Yang, H.; Gao, Z.; Li, F.; Yu, S. *Eur Biophys J* **2013**, *42*, 455.

(163) Gerlits, O.; Waltman, M. J.; Taylor, S.; Langan, P.; Kovalevsky, A. *Biochemistry* **2013**.

(164) Clery, A.; Schubert, M.; Allain, F. H. *Chimia (Aarau)* **2012**, *66*, 741.

(165) Chatterjee, A.; Abeydeera, N. D.; Bale, S.; Pai, P. J.; Dorrestein, P. C.; Russell, D. H.; Ealick, S. E.; Begley, T. P. *Nature* **2011**, *478*, 542.

(166) Han, X.; Aslanian, A.; Yates, J. R., 3rd *Curr Opin Chem Biol* **2008**, *12*, 483.

(167) O'Brien, J. P.; Pruet, J. M.; Brodbelt, J. S. *Anal Chem* **2013**.

(168) Skinner, O. S.; McLafferty, F. W.; Breuker, K. *J Am Soc Mass Spectrom* **2012**, *23*, 1011.

(169) Merkley, E. D.; Baker, E. S.; Crowell, K. L.; Orton, D. J.; Taverner, T.; Ansong, C.; Ibrahim, Y. M.; Burnet, M. C.; Cort, J. R.; Anderson, G. A.; Smith, R. D.; Adkins, J. N. *J Am Soc Mass Spectrom* **2013**, *24*, 444.

(170) Lee, J.; Park, H.; Kwon, H.; Kwon, G.; Jeon, A.; Kim, H. I.; Sung, B. J.; Moon, B.; Oh, H. B. *Anal Chem* **2013**.

(171) Zhang, Q.; Chen, J.; Kuwajima, K.; Zhang, H. M.; Xian, F.; Young, N. L.; Marshall, A. G. *Sci Rep* **2013**, *3*, 1247.

(172) Sawyer, H. A.; Marini, J. T.; Stone, E. G.; Ruotolo, B. T.; Gillig, K. J.; Russell, D. H. *J Am Soc Mass Spectrom* **2005**, *16*, 893.

(173) Liu, F.; Goshe, M. B. *Analytical Chemistry* **2010**, *82*, 6215.

(174) Zhong, Y.; Hyung, S. J.; Ruotolo, B. T. *Expert Rev Proteomics* **2012**, *9*, 47.

(175) Zekavat, B.; Miladi, M.; Becker, C.; Munisamy, S. M.; Solouki, T. *J Am Soc Mass Spectrom* **2013**, *24*, 1355.

(176) Riba-Garcia, I.; Giles, K.; Bateman, R. H.; Gaskell, S. J. *J Am Soc Mass Spectrom* **2008**, *19*, 609.

(177) McLean, J. R.; McLean, J. A.; Wu, Z.; Becker, C.; Perez, L. M.; Pace, C. N.; Scholtz, J. M.; Russell, D. H. *J Phys Chem B* **2010**, *114*, 809.

(178) Clemmer, D. E.; Hudgins, R. R.; Jarrold, M. F. *J Am Chem Soc* **1995**, *117*, 10141.

(179) Sowell, R. A.; Koeniger, S. L.; Valentine, S. J.; Moon, M. H.; Clemmer, D. E. *J Am Soc Mass Spectrom* **2004**, *15*, 1341.

(180) Ross, P. L.; Huang, Y. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol Cell Proteomics* **2004**, *3*, 1154.

(181) Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V. *Science* **2005**, *310*, 1658.

(182) Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S. J.; Robinson, C. V. *Nat Protoc* **2008**, *3*, 1139.

(183) Chen, S. H.; Russell, W. K.; Russell, D. H. *Anal Chem* **2013**, *85*, 3229.

(184) Brange, J.; Dodson, G. G.; Edwards, D. J.; Holden, P. H.; Whittingham, J. L. *Proteins* **1997**, *27*, 507.

(185) Alessandri, S.; Sancho, A.; Vieths, S.; Mills, C. E.; Wal, J. M.; Shewry, P. R.; Rigby, N.; Hoffmann-Sommergruber, K. *PLoS One* **2012**, *7*, e39785.

(186) Orlando, R.; Bush, C. A.; Fenselau, C. *Biomed Environ Mass* **1990**, *19*, 747.

(187) Robinson, E. W.; Leib, R. D.; Williams, E. R. *J Am Soc Mass Spectrom* **2006**, *17*, 1469.