ABSTRACT

Structural Analysis of Biomolecules Using Ion Mobility and Mass Spectrometry: Exploration of Ion Rearrangements and Conformations

Brett Daniel Harper, Ph.D.

Mentor: Touradj Solouki, Ph.D.

Early diagnosis of diseases is dependent on the development of reliable physiological and biochemical tests. The most precise medical tests employ analytical methods to identify specific molecules ("biomarkers") or molecular signatures ("biomarker panels") to monitor physiological responses to disease progression. It is conceivable that, one day, every disease could be classified by monitoring changes in biomolecular compositions. To realize this goal, however, it is necessary to develop highly sensitive tools that are capable of providing structural information about minute components of complex mixtures at the molecular and biologically relevant levels. Currently, mass spectrometry (MS) is the most suitable technique for routine and comprehensive molecular characterization of complex samples such as blood, saliva, urine, and tissue. Although MS is a promising tool for clinical diagnosis of diseases, ion rearrangements and presence of structural isomers can limit its utility. In this dissertation, results from investigations of ion rearrangements and biomolecular structures, using cutting-edge MS and ion mobility (IM)-MS techniques, are presented.

In Chapter One, the operating principles of various types of mass and IM spectrometers, and their applications to analysis of biomedical systems, are discussed. Next, results from isotope labeling, tandem mass spectrometry, and IM-MS are presented to demonstrate that losses of internal backbone carbonyls from y-type peptide ions are intermediate steps towards the formation of rearranged fragment ions (Chapter Two). In Chapter Three, MS, IM-MS, and theoretical modeling data are presented to demonstrate that collision-induced dissociation (CID) of 5′ phosphorylated DNA can generate rearranged $[phosphopurine]^-$ product ions.

Limited IM resolving powers may hinder adequate characterization of conformationally similar ions and reduce the accuracy of IM-based collision cross section (CCS) calculations. In Chapter Four, chemometric deconvolution of post-IM/CID MS data is used to extract IM drift times (DTs) of IM-unresolved isomers. Extracted DTs are used to calculate CCSs that are comparable to CCSs calculated from individual analysis of each isomer using traveling wave and drift tube IM-MS and the superposition approximation (PSA) molecular modeling. Finally, in Chapter Five, future directions and preliminary results from application of techniques discussed in Chapters One through Four to rapid analysis of tissue is presented and discussed.

Structural Analysis of Biomolecules Using Ion Mobility and Mass Spectrometry: Exploration of Ion Rearrangements and Conformations

by

Brett Daniel Harper, B.S.

A Dissertation

Approved by the Institute of Biomedical Studies

_____

Robert R. Kane, Ph.D., Chairperson

Submitted to the Graduate Faculty of
Baylor University in Partial Fulfillment of the
Requirements for the Degree
of

Doctor of Philosophy

Approved by the Dissertation Committee

_____

Touradj Solouki, Ph.D., Chairperson

_____

C. Kevin Chambliss, Ph.D.

_____

Christopher M. Kearney, Ph.D.

_____

Sascha Usenko, Ph.D.

_____

Sung Joon Kim, Ph.D.

Accepted by the Graduate School

May 2016

_____

J. Larry Lyon, Ph.D., Dean

Page bearing signatures is kept on file in the Graduate School.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF SCHEMES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	*Description*
AIMD	Automated Ion Mobility Deconvolution
AT	Arrival Time
$AT/\Delta AT_{50\%}$	Ion Mobility Resolving Power at Half-height (for DTIMS)
CCS	Collision Cross Section
$CCS/\Delta CCS_{50\%}$	Ion Mobility Resolving Power at Half-height (for TWIMS)
CEM	Chain Ejection Model
CRM	Charge Residue Model
CID	Collision-induced Dissociation
dα-MSH	Des-Acetylated-α-Melanocyte-Stimulating Hormone
dc	Direct Current
DNA	Deoxyribonucleic Acid
DT	Drift Time
DTIM	Drift Tube Ion Mobility
DTIMS	Drift Tube Ion Mobility Spectrometry
DTIM-MS	Drift Tube Ion Mobility-Mass Spectrometry
ECD	Electron Capture Dissociation
EI	Electron Impact Ionization
EIDTD	Extracted Ion Drift Time Distribution
EMT	Electron Multiplier Tube
ESI	Electrospray Ionization
FT	Fourier Transform
FT-ICR	Fourier Transform-Ion Cyclotron Resonance
FT-IR	Fourier Transform-Infrared Spectroscopy
FT-NMR	Fourier Transform-Nuclear Magnetic Resonance
GC	Gas Chromatography
HDX	Hydrogen/Deuterium Exchange
HPLC	High Performance Liquid Chromatography
ICR	Ion Cyclotron Resonance
I.D.	Inside Diameter
IEM	Ion Ejection Model
IM	Ion Mobility
IMS	Ion Mobility Spectrometry
IM-MS	Ion Mobility-Mass Spectrometry
IRMPD	Infrared Multiphoton Dissociation
LC	Liquid Chromatography
$M/\Delta M_{50\%}$	Mass Resolving Power at Half-height
MALDI	Matrix-Assisted Laser Desorption Ionization
MME	Mass Measurement Error

Abbreviation	*Description*
MS	Mass Spectrometry
MS/MS	Tandem Mass Spectrometry
MS^n	Multistage Tandem Mass Spectrometry
m/z	Mass-to-Charge Ratio
ppm	Parts-per-Million
PSA	Projected Superposition Approximation
PT	Proton Transfer
PTM	Post-translation Modification
rf	Radio Frequency
RFI	Radio Frequency Ionization
RMSD	Root Mean Squared Deviation
RNA	Ribonucleic Acid
SIM	Selected Ion Mobility
SIMPLISMA	SIMPLe-to-use Interactive Self-modeling Mixture Analysis
SPI	Spray Pump Ionization
SSI	Sonic Spray Ionization
TOF	Time-of-Flight
TWIG	Traveling Wave Ion Guide
TWIM	Traveling Wave Ion Mobility
TWIMS	Traveling Wave Ion Mobility Spectrometry
TWIM-MS	Traveling Wave Ion Mobility-Mass Spectrometry
UHV	Ultrahigh Vacuum
V-EASI	Venturi-Easy Ambient Sonic Spray Ionization

AMINO ACID ABBREVIATIONS

One Letter Abbreviation	*Three Letter Abbreviation*	*Amino Acid*
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartic Acid
E	Glu	Glutamic Acid
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

NUCLEOBASE ABBREVIATIONS

One Letter Abbreviation	*Nucleobase*
A	Adenine
C	Cytosine
G	Guanine
T	Thymine (DNA Only)
U	Uracil (RNA Only)

ACKNOWLEDGMENTS

First, I would like to thank my Lord and Savior, Jesus Christ, for his loving kindness and gentle guidance through the difficult process of earning a PhD. I would not have completed, or even started, this journey if it were not for you.

Next, I am deeply grateful for my doctoral advisor, and friend, Dr. Touradj Solouki. Through his constant feedback and encouragement, Touardj has pushed me to be a better scientist. I am thankful that I have had an advisor that I can speak with openly as a peer, and not just as an employee or student. Touradj accepted me into his group while I was an undergraduate student, despite the fact that I was a biology major who had little-to-no knowledge of mass spectrometry. Moreover, Touradj was kind enough to allow me to later join his group as a PhD student and train me to use state-of-the-art ion mobility and mass spectrometry instruments. I am confident that I would not have had the quality of opportunities that I have had under his direction at any other institution in the world. Thank you Touradj for everything you have done for me, both the things that I have seen as well as the things that have gone unseen.

I would like to extend my gratitude to my graduate committee members, Dr. Kevin Chambliss, Dr. Sascha Usenko, Dr. Christopher Kearney, and Dr. Sung-Joon Kim for their encouragement throughout my graduate studies. Moreover, I am grateful to all of my committee members for their valuable comments and questions which helped shape the direction of dissertation.

I am grateful for the support and friendship of past and present members of the Solouki research group including, Dr. Abayomi Daniel Olaitan, Dr. Behrooz Zekavat, Dr. Mahsan Miladi, Elizabeth Neumann, Matthew Brantley, Michael Pettit, Jackie Lochridge, Ian Anthony, Ting Wang, Brooke Brown, and others. I was truly blessed to work in a group who so strongly supports one another in and out of the lab setting. I am especially grateful for Daniel for being a constant encouragement, consistent friend, and a brother in Christ.

Furthermore, I would like to thank Dr. Alejandro Ramirez and Dr. Christopher Becker for their valuable assistance and providing training and support for instruments in the Baylor University mass spectrometry center. I am thankful for the faculty and staff of the Department of Chemistry and Biochemistry who offered me unencumbered access to their instruments and facilities in spite of me being a student in another department. On that note, I am also grateful for the generous support of the faculty and staff of the Institute of Biomedical Studies. Particularly, I would like to thank the director of the Institute of Biomedical Studies, Dr. Bob Kane, and the graduate program director for the biomedical studies program, Dr. Christopher Kearney, for their leadership of the institute as well as their personal encouragements to me. Likewise, I am thankful for Mrs. Rhonda Bellert for keeping me on track for graduation by always being available to help with administrative paperwork and processes.

In addition to my colleagues at Baylor University I am also thankful for my collaborators at Vanderbilt University, Dr. John McLean, Dr. Jody May, and Dr. Sarah Stow for their invaluable contributions to the ion mobility deconvolution/collision cross section project.

On a personal note, I am so thankful for my local church family at Antioch Community Church and my home church family in Fort Worth at Olive Place Baptist Church for their intentionality, genuineness, and investment in my life. Likewise, I am thankful for my parents, sister, and in-laws for their loving support. Finally, I must extend my deepest gratitude to my wife, Gina Harper. Gina has supported and stood by me through some of the toughest and most joyful times of my life thus far. I would not have made it to this point if it were not for her love, prayers, and support. Gina, words cannot express how much I love you. Thank you for always being there for me and, above all else, for honoring Jesus in everything that you do.

DEDICATION

To my wife, Gina Harper, for her continued prayers, love, encouragement, and support.

CHAPTER ONE

Introduction

Multidimensional Biomedical Mass Spectrometry and Allied Techniques

Mass spectrometry (MS) refers to a variety of analytical tools that are used to experimentally measure the mass-to-charge ratio (*m/z*) of ionized molecules and atoms. Because each element has a unique mass and isotopic distribution, mass spectra can be thought of as molecular "fingerprints" to identify analytes.[1] Moreover, using high-performance instruments and advanced MS techniques, information about molecular structure(s) and conformation(s) can be obtained. Although there are numerous types of commercial mass spectrometers available to address a variety of scientific inquiries, all MS systems are composed of three principal components that are generally housed in a vacuum or an ultrahigh vacuum (UHV) chamber: (1) ionization source, (2) mass analyzer, and (3) detector. In the following sections, various types and operating principles of these three major components of MS systems are presented and discussed.

Ionization Sources

The first step in MS analysis, after the analyte introduction, is ionization of the sample species. Ionization is the process of adding (or subtracting) a positive or negative charge to (or from) a neutral molecule or atom and is commonly achieved in the ionization source of the mass spectrometer. There are many types of ionization techniques employed in MS and these ionization types can each be broadly classified as one of the following two types: (1) "hard" ionization; (2) "soft" ionization.[2] In

experiments involving "hard ionization", the precursor molecules being analyzed may go through highly energetic and extensive ion fragmentation processes and yield a number of fragment ions.[2] Soft ionization techniques, on the other hand, are characterized by little to no fragmentation of the analyte molecules.[2] Although hard ionization techniques (*e.g.*, electron impact ionization (EI)[3-5]) are popular for small molecule (*e.g.*, < 1000 Da) analysis,[4] soft ionization techniques (*e.g.*, electrospray ionization (ESI)[6-8] and matrix assisted laser desorption ionization (MALDI)[9-11]), have become increasingly popular for analysis of larger biomolecules, such as proteins, deoxyribonucleic acids (DNAs), and lipids.[4] In the following section, a brief introduction of electrospray ionization (ESI) and closely related ionization techniques as well as their biomedical applications are provided.

Electrospray ionization (ESI). ESI was first described in the literature by Dole *et al.* in 1968.[6] However, the technique was not widely recognized until 1989 when Fenn *et al.* demonstrated that ESI could be used to analyze intact macromolecules *via* MS.[7] Fenn's discovery was a breakthrough in the MS community as it paved the way for broader biomedical MS applications and was duly recognized with a share of the 2002 Nobel prize in chemistry.[12]

Central to the success of ESI is the simplicity of its design and operation (Figure 1.1). To generate ions in ESI, a sample is first dissolved in a suitable solvent (*e.g.*, acidified methanol). The solution is then injected through an inert capillary (often with an inside diameter (I.D.) size of ~ 50 μm) and sprayed in front of the atmospheric pressure inlet of a mass spectrometer with an electric potential of ~1 kV to ~5 kV (relative to the MS inlet of the instrument, using positive or negative polarity depending on the negative-

or positive-mode of ionization) applied to the capillary needle. It is presumed that, after the sample nebulization, ensuing droplets undergo a continuous desolvation and this process forces charges on the surface of the droplet to come into closer proximity.[13] At a certain point (in time and/or space), the Coulombic repulsions of the like charges can overcome the surface tension of the droplet and yield two (or more) smaller droplets in a process called Coulomb fission (Figure 1.1, diverging arrows).[13] The final steps of charge deposition on analytes in ESI are still being investigated;[13-17] however, three models have been proposed and introduced in the literature to explain the phenomenon: (1) the charge residue model (CRM), (2) the ion evaporation model (IEM), and (3) the chain ejection model (CEM).[13] In the CRM, it is proposed that Coulomb fission keeps occurring until the solvent is completely evaporated and the charges that were on the surface of the solvent droplet are deposited on the still-folded analyte (*e.g.*, protein). In the IEM, it is presumed that analytes are "evaporated" from the surface of the solvent droplet, picking up a charge as they escape the solvation sphere. The IEM is generally used to explain ionization of smaller molecules (*e.g.*, pharmaceuticals) rather than larger molecules such as proteins. The CEM, is similar to the IEM but rather than the molecule escaping (or "evaporating") from the droplet at once, it is presumed that a portion or "tail" of the analyte escapes the droplet and pulls the rest of the analyte out (similar to pulling the string out of a wound up yo-yo toy). As the analyte is pulled out of the droplet it picks up charges until it is eventually ejected from the solvent droplet. The CEM can be used to explain the occasionally observed denaturation of ionized macromolecules.[13]

Several advancements and modifications to ESI have been described in the literature. For example, Cooks *et al.* developed a technique called desorption ESI (or

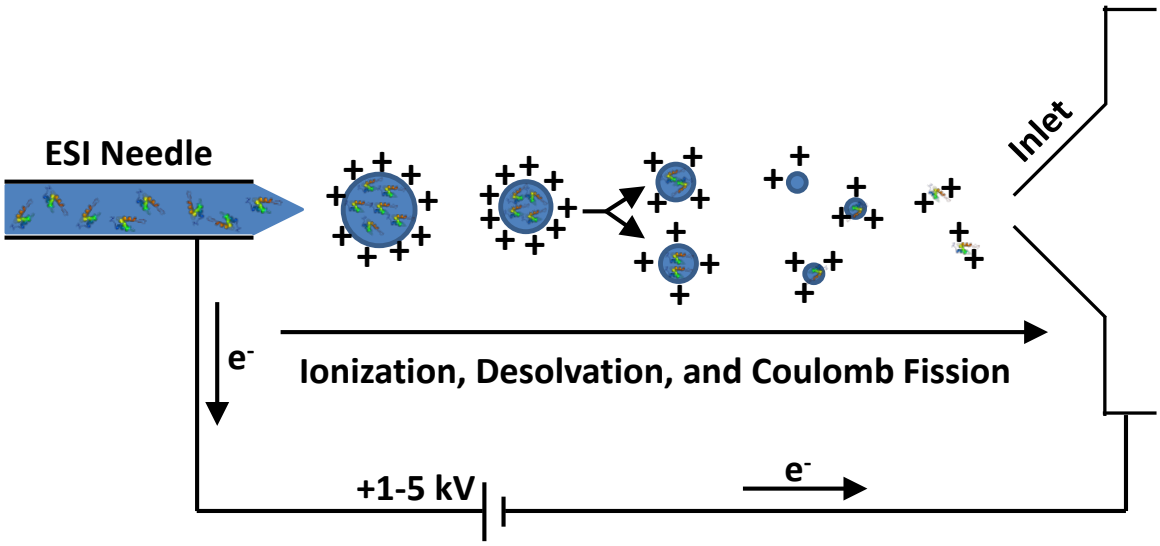


Figure 1.1. Schematic drawing of an electrospray ionization (ESI) source operating in positive-ionization mode. Detailed discussion of the mechanism and operating principles of ESI can be found in the main text. Please note: schematic is not drawn to scale.

DESI) which can be used to desorb and ionize molecules from various surfaces (*e.g.*, ink on paper).[18] The principal advantage of DESI is that samples do not have to be dissolved in an electrospray solvent, but can rather be analyzed in their natural environment.[18] One popular application of DESI is in imaging MS.[19] In DESI imaging MS, the electrospray needle is scanned along the surface of the sample (*e.g.*, tissue section) and the individual mass spectra at each spatial position of the sample are correlated to generate an image pixel. These individual pixels can then be re-plotted to generate a molecular image of the sample. For example, the distribution of a specific protein could be imaged in a tissue by monitoring the *m/z* value corresponding to that protein at each spatial position in the tissue.[19] The MS image can then be generated by assigning a color scale (*e.g.*, black for no signal and bright red for high signal intensity) corresponding to the MS signal intensity at each pixel across the surface of the tissue. DESI has shown promise in several biomedical applications and was recently used in the first MS guided human brain surgery.[20]

In addition to DESI, several other ionization techniques related to ESI have been developed. For example, in the mid-1990s, Hirabayashi *et al.* serendipitously developed sonic spray ionization (SSI) which uses a high-flow sheath gas around a sample capillary to nebulize aqueous analyte and generate ions, without the use of any high voltage.[21] Later, a self-pumping variant of SSI was developed (*i.e.,* Venturi-easy ambient sonic-spray ionization (V-EASI)),[22] which simplified the design of SSI by eliminating the need for a syringe pump. Although SSI and V-EASI were utilized for MS analysis in recent years, they are similar to early (*i.e.,* pre-ESI) nebulization-ionization experiments (*i.e.,* without any applied voltages).[23, 24] In these early experiments, however, singly-charge molecules (*e.g.,* sodium ions (Na^+)) were analyzed using ion mobility (IM) spectrometry[23, 24] (IM is a technique related to MS and will be discussed in-detail in later sections in Chapter One of this dissertation) rather than analyzing multiply-charged ions (*e.g.,* peptides) using MS (as with SSI and V-EASI experiments).[21, 22] Figure 1.2 shows mass spectra of singly-protonated sildenafil (Viagra) ions (*i.e.,* $[sildenafil + H]^+$ at *m/z* 475 (top row) and $[sildenafil - H]^-$ at *m/z* 473 (bottom row)) generated using sample nebulized using a spray pump bottle (herein referred to as spray pump ionization or SPI), SSI, and ESI. Experimental results indicate that SPI and SSI generate mass spectra that closely resemble those generated using ESI. However, in this particular example, SPI and SSI both showed lower signal intensity than ESI, presumably because fewer charges are deposited on the initial solvent droplets in SPI and SSI relative to the ESI approach. Unlike the earliest sample-nebulization ionization findings (with IM) that reported the generation of small singly-charged molecules,[23, 24] we found that large multiply-charged (*e.g.,* >20 charges) macromolecules could also be formed during the SPI (Figure 1.3).

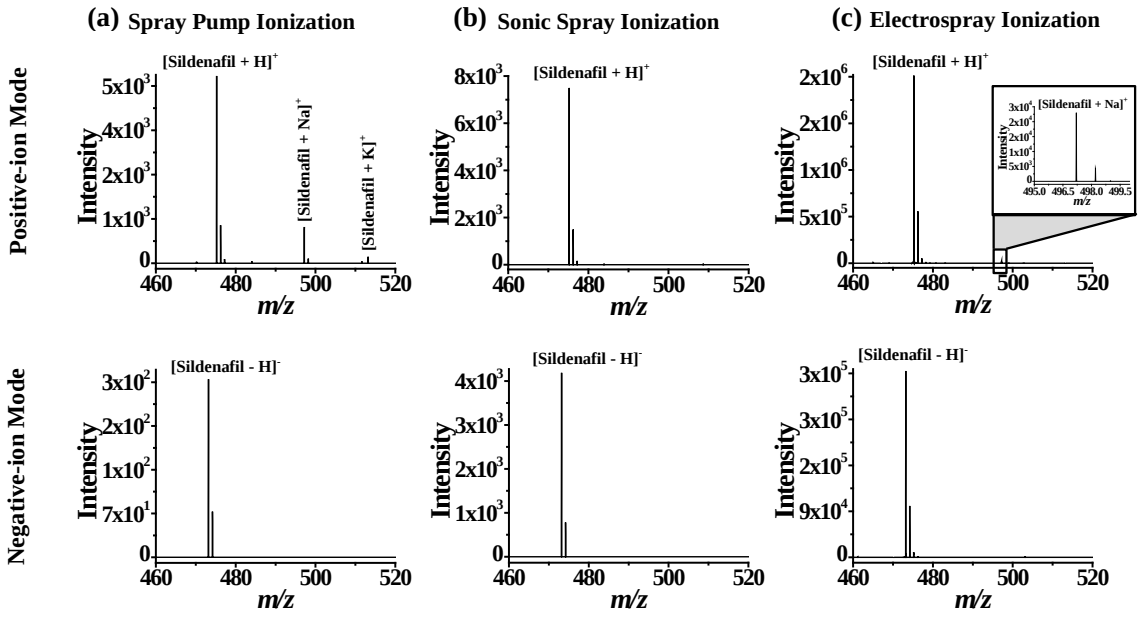


Figure 1.2. Mass spectra for 5 µg/mL sildenafil (a cGMP-specific phosphodiesterase type 5 (PDE5) inhibitor) generated from (a) spray pump ionization (SPI), (b) sonic spray ionization (SSI), and (c) electrospray ionization (ESI) in positive- (top row) and negative-ion (bottom row) modes collected using an Orbitrap Discovery mass spectrometer (Thermo Fisher Sci., Waltham, MA, USA). Inset (c, top row) shows the expanded mass range from *m/z* 495-500 which includes $[Sildenafil + Na]^+$ (*m/z* 497) generated from ESI.

traditional ESI, these newly introduced ionization methods may prove to be useful for unique applications such as field-based MS or in resource limited settings.[25, 26]

Mass Analyzers and Detectors

Similar to ionization sources, mass analyzers can also be broadly classified into several categories including electric field (E), magnetic field (B), and combined electric and/or magnetic field (*e.g.*,) analyzers. In addition, depending on the position of the mass analyzers with respect to other instrumental components (*e.g.*, ionization sources and detectors) mass spectrometers are often categorized as "spatial" or "temporal". For example, in spatially resolved mass spectrometers, the three major components of the mass spectrometers (*i.e.*, the ion source, analyzer, and detector) are physically located in

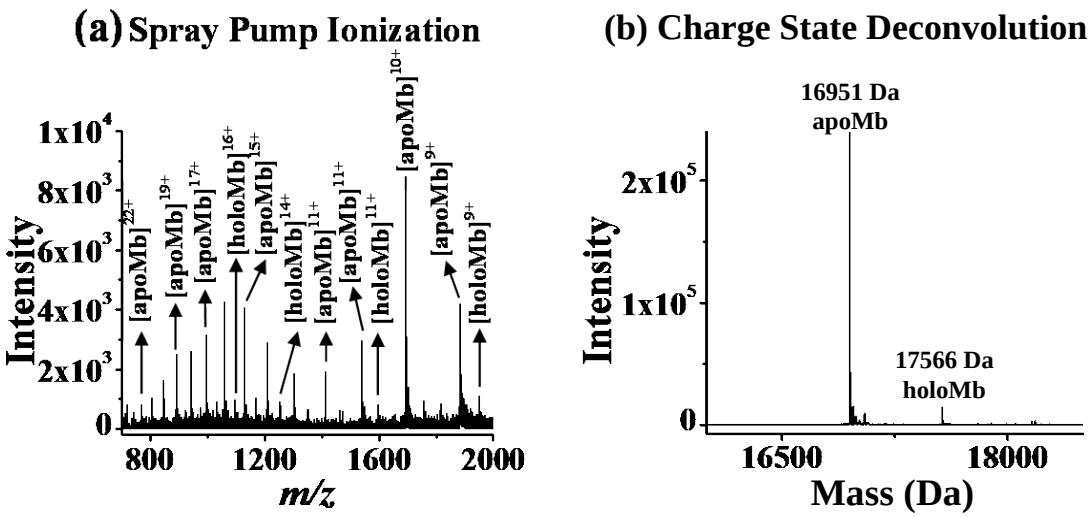


Figure 1.3. (a) SPI mass spectrum (average of 40 pumps/125 scans) of equine heart myoglobin (Mb) collected using a Synapt G2-S TOF mass spectrometer (Waters Corporation, Milford, MA, USA). (b) Charge state deconvolution (using MaxEnt1 tool in MassLynx Software (Version 4.1, Waters Corporation)) of the mass spectrum in (a) was used to confirm apoMB (16951 Da) and holoMB (17566 Da) identities.

different positions and hence ionization, separation, and detection events are separated in space. In other words, ions are first ionized, then sorted based on their unique *m/z* values in the mass analyzer, and finally detected at the detector (*i.e.*, distinct components are utilized for each task). Examples of mass spectrometers utilizing spatially resolved mass analysis approach include electric and/or magnetic sector instruments (where ions are separated based on their trajectories in electric and/or magnetic fields),[27, 28] time-of-flight (TOF) instruments (where ion separation occurs in a field free drift tube),[29] and quadrupole instruments (where ions are separated based on their flight-path stabilities through four rods with applied direct current (dc) and radio frequency (rf) voltages)[30]. In these spatial MS instruments (*i.e.*, sector, TOF, and quadrupoles), the *m/z* separated ions are first separated and then hit the detector (generally a variant of an electron multiplier tube (EMT;[31] discussed below)); detectors records the ions' *m/z*'s as a function either

magnetic and/or electric field or time to yield mass spectra. Conversely, in temporally resolved mass spectrometers, ionization, mass analysis, and detection events often take place in the same physical space but are separated in time. For example, in Fourier transform-ion cyclotron resonance (FT-ICR) MS[32, 33] and Orbitrap MS,[34] ions are trapped or confined, excited, and detected in the ICR or orbital trap cells (using magnetic and/or electric fields, respectively). For example, once the ions' cyclotron motions are radially excited in the ICR cell by applying appropriate range of resonance radiofrequencies to two "excitation plates", the two adjacent "detection plates" of the same ICR cell are used to detect the natural cyclotron frequencies of the trapped ions. Thus, detector electrodes in each (ICR or Orbitrap) cell are used to measure the cyclotron (ICR) or trapping (Orbitrap) frequencies of the trapped ions (by measuring the induced current on the detector electrodes as ions pass by the detection electrodes). These cyclotron (in an ICR cell) and trapping (in an Orbitrap cell) frequencies are inversely proportional to *m/z* values of the trapped ions and accurate characterization of these frequencies provide mass spectra with high mass measurement accuracies and high mass resolving powers.[33, 34] In the following sections, the operating principles of TOF and Orbitrap mass analyzers that were used for collection of the data presented in this dissertation are discussed. A brief discussion of EMT operation will be included at the end of the TOF section.

Time-of-flight (TOF) MS. In TOF MS, ions' *m/z* values are calculated based on their flight times through a field free drift tube (Figure 1.4).[29] The start time for TOF can be recorded by establishing a pulsed start event (*e.g.*, turning the laser pulse "on" in laser desorption TOF experiments or time regulation of a pusher electrode for ion acceleration with a high-voltage). A high-voltage (*e.g.*, ~20 kV) pulse (generally in the nanosecond

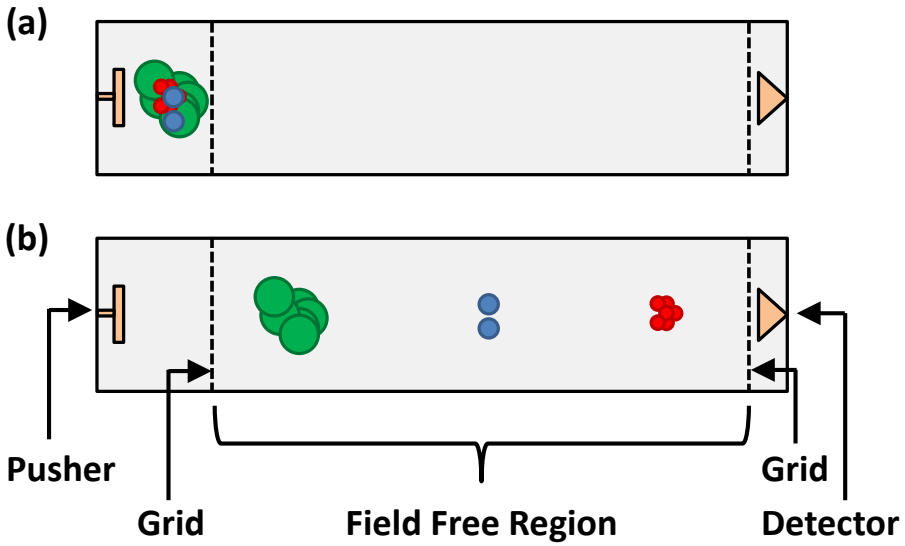


Figure 1.4. Schematic drawing of a linear TOF mass spectrometer. Ions (*i.e.*, colored circles) are shown before (a) and after (b) the TOF push. Lighter ions (*e.g.*, red circles) will have shorter flight times and hit the detector before more massive ions (e.g., green circles). Note: schematic is not drawn to scale.

scale peak widths) is applied to the pusher electrode to increase the ions' kinetic energies (E_k in joules) and accelerate the ions through the field free drift tube towards the ion detector. The field free drift region of the TOF is achieved by applying the same voltage (*i.e.*, equal potential) to two "grid" electrodes positioned at the beginning and end of the drift tube (Figure 1.4). Because all of the ions are pulsed into the field free region of the TOF at the same time (Figure 1.4a), they are considered to have the same start time. In reality, the assumption that all ions have the same start time is not accurate, as will be discussed in the description of reflectrons below. However, this assumption is sufficient for an initial description of the operating principles of TOF MS. Ions with smaller mass (m; in kilograms (kg)) (and/or more charges (z)) "feel" the push more strongly than the more massive ions (and/or ions with fewer charges) and therefore travel at greater velocities (v; in meters/second (m/s)) and hit the detector at an earlier time, t (in seconds),

than the heavier ions (which will arrive to the detector at a later time) (Figure 1.4b). Because the flight path (*d*; in meters) is constant, *m/z* values can be easily derived in TOF MS from fundamental Equations 1.1 through 1.4:[29]

Equation 1.1 $$E_k = \frac{mv^2}{2}$$

Equation 1.2 $$E_p = qV$$

Equation 1.3 $$q = ze$$

Equation 1.4 $$v = \frac{d}{t}$$

Units for E_k and potential energy (*Ep*) are both in joules; therefore, setting Equations 1.1 and 1.2 equal to one another, substituting *ze* for *q* (Equation 1.3, where "*e*" is the elementary charge in Coulombs), and substituting *d/t* for *v* (Equation 1.4), we get Equation 1.5:

Equation 1.5 $$\frac{md^2}{2t^2} = zeV$$

Equation 1.5 can be rearranged to yield the TOF equation for *m/z* (Equation 1.6):[29]

Equation 1.6 $$\frac{m}{z} = \frac{2eVt^2}{d^2}$$

It should be noted that all values in Equation 1.6 are in SI units and final mass (*m*) in *m/z* must be converted from kg to Daltons (Da) (or grams per mole) for meaningful utility.

As briefly mentioned above, Equation 1.6 assumes that the start time (and by extension position) of all ions are the same and that ions with the same *m/z* experience the same E_k from the pusher. In reality, some ions will be slightly further from or closer to the pusher during the pushing event. Ions that are closer to the pusher during the pushing event will "feel" a greater force, and will therefore have greater E_k, than ions with the same mass and charge positioned further away from the pusher during the pushing event.

To reduce the E_k spread of the ions, a reflectron can be placed in the ion path.[35] Accelerated ions with higher E_k can penetrate deeper into the electric field applied to the reflectron electrodes than ions with lower E_k. In effect, this can focus ions with the same *m/z* so that "slower" ions catch up with the "faster" ions and, ideally, hit the detector at the same time. A secondary effect of using a reflectron is an increased path length (*i.e.*, more time for similar *m/z* ions to be separated).[35] The more reflectrons used, the lower the kinetic energy spread, the longer the path length, and therefore the higher the mass resolving power (*i.e.*, $M/\Delta M_{50\%}$, where M and $\Delta M_{50\%}$ refer to the *m/z* of the ion and peak width at half height for this *m/z*, respectively) of the TOF.[35] However, because of ion losses to reflectron grids (and/or other electrodes), having more reflectrons often results in fewer ions making it to the detector and therefore, reduced sensitivity.[35]

The most common type of detector used in spatially resolved mass spectrometers, including TOF MS, is the EMT (or some variant thereof).[31] After ion separation in the mass analyzer, ions hit the EMT cathode, liberating electrons.[31] The liberated electrons then collide with a series of dynodes releasing even more electrons.[31] Eventually, all of the liberated electrons ($\sim 10^5$ or more liberated electrons per ion striking the cathode) strike the anode where the current can be measured and recorded as a detection event (*i.e.*, an ion hitting the detector).[31] In TOF MS, the pusher event corresponds to the "start" time whereas the detector response corresponds to the "end" time, or measured TOF. The measured TOF for a particular ionized molecule can then be plugged into Equation 1.6 to calculate the ion's *m/z* value. To improve the sensitivity and minimize ion saturation, multiple EMTs can be combined into a single detector called a microchannel plate detector.[31] Microchannel plate detectors can be stacked to create chevron detectors to

further improve detector performance and are often used in high-end TOF MS instruments.[31]

Orbitrap MS. Introduced in 2000, the Orbitrap MS is a relatively recent development.[34] The operating principles of Orbitrap MS are similar to those used in FT-ICR MS, a predecessor and direct influencer of the Orbitrap MS design. Therefore, it is useful to first describe the operating principles of FT-ICR MS before discussing Orbitrap MS.

In TOF MS (discussed above), only electric fields are used to manipulate ions, and mass analysis is achieved in an electric (and magnetic) field free region. In FT-ICR MS, on the other hand, both electric and magnetic fields are used to confine ions and determine their *m/z* values (Figure 1.5).[33] Figure 1.5 shows a schematic of a simple ICR cell consisting of six electrically isolated plates configured in a cube-shaped geometry.[33] The ICR cell is placed in the center of a doughnut shaped superconducting magnet (*e.g.*, 9 tesla), where the magnetic field lines run through the center of the magnet.[33] Trapping plates are placed perpendicular to the direction of the magnetic field lines and are used to confine ions in the "z" dimension (*i.e.*, along the magnetic field lines).[33]

The first step in FT-ICR mass analysis is to "quench" or eject any potentially pre-trapped ions within the ICR cell (*i.e.*, eject the left-over ions in the cell from the previous run). Ion quenching is achieved by applying a dc voltage to each of the trapping plates (with the opposite polarity on each plate) which pushes (or pulls, depending on the ions' polarity) the ions so that they hit one of the trapping plate electrodes and are neutralized and then evacuated by the vacuum pump(s). Once the ICR cell is cleaned, ions can either be generated in the ICR cell (*e.g.*, by EI or radio frequency ionization (RFI)[36]) or

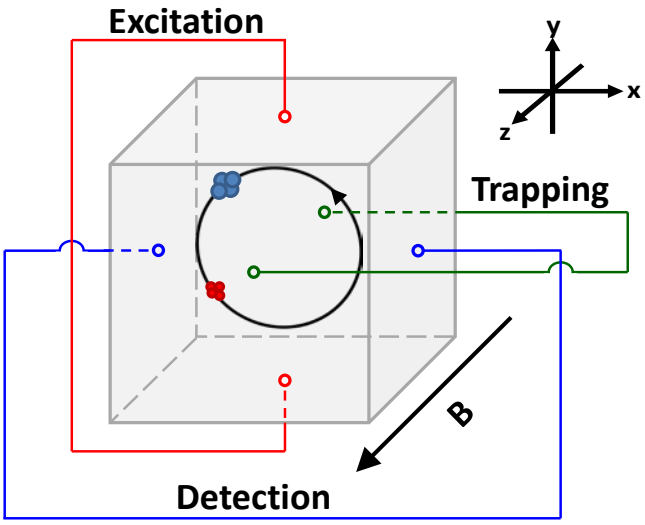


Figure 1.5. Schematic drawing of a cubic ICR cell. Orbital trajectory of two populations (*e.g.*, ions represented by filled red circles all have the same *m/z*) of positively charged ions where each ion population has different masses and/or charges (*i.e.*, ions represented by filled red and blue circles have different *m/z* values). Excitation, detection, and trapping plates are labeled. Open circles denote where wires are connected to each plate. Note: schematic is not drawn to scale.

externally (*e.g.*, by ESI, MALDI, SPI, *etc.*) and transferred into the ICR cell using transmission multipole rods (*e.g.*, quadrupole, hexapole, or octapole operated in "rf only" mode).[37] Once ions are in the ICR cell, they are confined radially by the high magnetic field and axially by applying symmetric dc potentials to the trapping plates. Next, ions are "excited" by applying an rf sweep (or "chirp") to the two excitation plates. During the ion excitation event, trapped ions' trajectories are displaced at 90˚ to the magnetic field line; however, because these trapped ions continuously encounter new field lines, they are continuously deflected by 90˚ and therefore follow orbital trajectories (Figure 1.5). The frequency of the ion's cyclotron motion can be determined using the two opposing detector plates. As ions pass the detector plates, they induce an image current on the plates (*i.e.*, by the attraction or repulsion [depending on the polarity of the ions' charges] of electrons on the surface of the detector plates) which can be amplified and recorded.

For isolated and excited ions with a single *m/z* value, the resulting time-domain signal resembles a sinusoidal (sin) wave. The frequency of this sin wave is inversely proportional to the ion's *m/z* as can be derived from the fundamental equations below:[33]

Equation 1.7 $$F_m = qvB \sin\theta$$

where "F_m" is the magnetic force on the ion (in newtons), "*B*" is the magnetic field strength in tesla, and "θ" is angle between the ion's trajectory and the magnetic field line ($\theta = 90^\circ$ in ICR). Because $\sin(90^\circ) = 1$, we can simplify Equation 1.7:

Equation 1.8 $$F_m = qvB$$

Substituting *ze* for *q* (Equation 1.3) we get:

Equation 1.9 $$F_m = zevB$$

The centripetal force (F_c) of an ion's cyclotron motion is equal to the product of the ion's mass (*m*; in kg) times its velocity (*v*; in m/s) squared, and divided by the radius of the ion's orbital trajectory (r):

Equation 1.10 $$F_c = \frac{mv^2}{r}$$

Given that:

Equation 1.11 $$\omega = \frac{v}{r}$$

and:

Equation 1.12 $$\omega = 2\pi f$$

where "ω" is the angular frequency and "f" is frequency (in Hz), Equation 1.10 can be rewritten as:

Equation 1.13 $$F_c = mv2\pi f$$

Setting Equations 1.9 and 1.13 equal to each other we get:

Equation 1.14 $$zevB = mv2\pi f$$

Simple rearrangement of Equation 1.14 yields the ICR formula for *m/z*:[33]

Equation 1.15 $$\frac{m}{z} = \frac{eB}{2\pi f}$$

Equation 1.15 shows that an ion's *m/z* is inversely proportional to its cyclotron frequency. However, this equation can only be used successfully if the individual frequencies for each ion are measurable. In other words, in the presence of two or more ion population types (*i.e.*, different *m/z* values and therefore different *f*), the time domain signal would be an additive (constructive and destructive) combination of the two or more individual sin waves. Therefore, much like other types of Fourier transform (FT) instruments (*e.g.*, FT-infrared spectroscopy (FT-IR) and FT-nuclear magnetic resonance spectroscopy (FT-NMR)), the resulting complex time domain signals would be difficult to interpret and might seem devoid of useful frequency information. Following the example of the NMR approach, in 1974, Alan Marshall and Melvin Comisarow solved the "messy" time domain issue in ICR MS using FT.[32] In their technique, dubbed FT-ICR MS, FT was used to convert the time-domain signal to a frequency-domain signal and allow accurate assignment of all frequency components (and therefore characterization of specific ICR frequencies of all ions and determination of their *m/z* values by using Equation 1.15).[32] Due to the multiplexing advantages of FT-ICR MS over the then conventional scanning method of sweeping through individual rf frequencies[38, 39] (or by changing the magnetic field[39-41]) to selectively excite individual *m/z*'s, the time necessary for mass spectra collection was reduced significantly.[32]

In 2000 Alexander Makarov reported on the development another FT-MS instrument based on electric field confinement of ions to orbital trajectories (Figure 1.6).[34] The operating principles of the so-called "Orbitrap" are similar to those described

for the Kingdon trap[42] and "wire ion guide"[43, 44] ICR cell.[45] In Makarov's design, as illustrated in Figure 1.6, a spindle shaped inner electrode is placed in the center of a barrel-shaped outer electrode consisting of two electrically isolated halves. Ions are injected off center into the orbital trap perpendicular to the axis of the center electrode (not shown in Figure 1.6), causing the ions to follow an orbital trajectory around the center electrode.

In orbital traps, a dc voltage is applied to the center electrode to attract the ions towards the center of the electrode. However, the ions' translational velocities cause them to constantly "overshoot" the center of the spindle shape (similar to how a satellite always falls towards the center of earth but never hits the earth due to its tangential velocity). Once the ions have passed the center line of the spindle they are re-attracted to the center of the spindle and therefore change directions. These cyclic trajectories continue in such a way that the translational motion of ions moving past the center electrode induces an image current on the two halves of the outer barrel-shaped electrode. This image current from the translational motion of ions in the cell (as opposed to the cyclotron motion in ICR MS) is recorded as a time-domain signal and converted to a frequency-domain signal using FT. Frequencies of the ions can then be converted to *m/z* values, similar to FT-ICR MS.[34] Orbitrap MS provides high mass-resolving-power (*i.e.*, reported $M/\Delta M_{50\%}$ of up to ~1,000,000 at *m/z* ~200 with routine $M/\Delta M_{50\%}$ of ~100,000 at *m/z* ~525)[46, 47] without the use of any magnetic fields, making it useful for a wide range of bioanalytical applications.

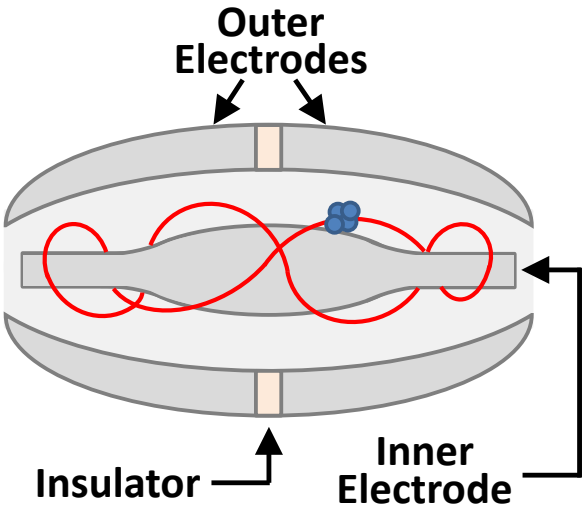


Figure 1.6. Schematic drawing of an Orbitrap mass analyzer's cross sectional cut. Ions orbit the spindle shaped inner electrode and induce an image current on each halve of the barrel shaped outer electrodes. The frequency of the induced currents on each of the outer electrodes is proportional to the ions' masses. The red line denotes the ions' (blue circles) path around the spindle after ion introduction into the trap (ion introduction is not depicted). Note: schematic is not drawn to scale.

Ion Mobility Spectrometry

Ion mobility spectrometry (IMS) is fundamentally similar to MS in many ways. Analogous to MS, IMS uses electric fields to control trajectories of ionized molecules in the gas-phase.[48-50] However, whereas MS is used to determine the mass of an ion, IMS is utilized to determine the collision cross-section (CCS) of an ion and/or as a separations technique for complex mixture analysis.[48-50] There are several different types of commercial IMS systems currently available.[49] Comparable to the different types of mass spectrometers discussed in preceding sections, the different types (and designs) of IM spectrometers also operate on slightly different principles.[49] Currently, the two most commonly used types of IMS instruments for biomedical research are drift time (DT) and traveling wave (TW) IM spectrometers.[49]

Drift time ion mobility spectrometry (DTIMS). The simplest type of IMS is DTIMS. In DTIMS, ions are confined radially by a homogenous dc electric-field applied to a series of ring electrodes.[49] Ions are gated into the DTIMS cell and a potential difference is applied between the entrance and exit of the cell to drive the ions through the drift tube.[49] The ions' flight through the DTIMS cell is impeded by a counter flow of buffer gas (*e.g.,* nitrogen (N_2) or helium (He) gas).[49] Ions with larger CCSs collide with more buffer gas molecules and therefore move at slower rates than ions with smaller CCSs which collide with fewer buffer gas molecules.[49] A detector (*e.g.,* EMT) is positioned at the end of the drift cell and is used to record ions' drift times (t_D's).[49] CCSs (Ω; in m^2 but generally converted to $Å^2$ for meaningful utility) can be simply calculated using Equation 1.16:

Equation 1.16 $$\Omega = \frac{Kt_D Fz\sqrt{T}}{NLP}\sqrt{\frac{1}{m_I}+\frac{1}{m_N}}$$

where "*K*" is a constant, "t_D" is the ions' drift times (in seconds), "*F*" is the electric field strength (in V/m), "*T*" is the temperature (in kelvin), "*N*" is the neutral buffer gas number density (in m^{-3}), "*L*" is the length of the drift region (in m), "*P*" is the pressure (in torr), and "m_N" and "m_I" are the masses (in kg) of the neutral buffer gas and ion, respectively.[51]

Traveling wave ion mobility spectrometery (TWIMS). In 2004, Giles *et al.* developed a fundamentally new type of IMS called TWIMS.[52] Figure 1.7 shows a schematic drawing of the TWIM-MS developed by Giles and coworkers. Analogous to DTIM spectrometers, TWIM spectrometers consist of a series of ring electrodes. However, in TWIMS ions are radially confined by applying an rf voltage to the stacked ring ion guide (where adjacent rings have opposite rf polarities).[52, 53] To move the ions from the entrance to the exit of the drift cell, a dc voltage is superimposed on the rf

electric field and scanned along each successive ring creating an electric field "wave" which propels ions through the drift cell.[52, 53] Rather than using a counter flow of buffer gas as in DTIMS, in TWIMS the IM cell is filled with a static buffer gas. Ions with smaller CCSs have fewer collisions with buffer gas (than ions with larger CCSs) and will "ride" the IM wave, similar to a surfer riding a wave in the ocean.[52, 53] Ions with larger CCSs undergo more collisions with the buffer gas (than ions with smaller CCSs) and fall over the back of wave and are slowly moved forward by successive waves.[52, 53] Like DTIMS, ions exiting the TWIM cell collide with a detector and the ions' drift times are recorded for CCS calculations.

Because of the complex motion of ions in the rf and dc electric fields in TWIMS,[52, 53] there is no simple fundamental formula relating TWIMS t_D's to CCSs (as there is for DTIMS [Equation 1.16]).[51] Although there is no fundamental formula relating t_D's to CCSs in TWIMS, ions' t_D's are proportional to their CCSs.[51] Using the method proposed by Ruotolo *et al.*,[51] ions' CCSs can be calculated from TWIMS t_D's by calibrating the analytes' measured t_D's with the t_D's of a set of ions with known CCSs. Although effective, the calibration method currently used for calculating CCSs in TWIMS is a time-intensive effort as it requires collection and analysis of data for both analyte *and* calibrant sets.

Regardless of the type of IM system used, two key advantages of IM are (a) rapid IM separation (*e.g.*, ms time-scale) and (b) compatibly of IM with MS systems (for straightforward construction of IM-MS systems).[49] The first commercially available TWIM-MS instrument was the Synapt G1 HDMS IM TOF MS produced by Waters Corporation. The Synapt G1 HDMS system consists of a modular ionization source (*e.g.*,

switchable to ESI, MALDI, *etc.*) for sample ionization, a quadrupole mass filter for ion isolation and/or transfer, a "TriWave" (*i.e.*, a traveling wave ion guide (TWIG) trap cell, IM cell, and transfer cell) for ion accumulation and IM, and a TOF mass spectrometer for *m/z* measurements (Figure 1.7). The reported IM resolving power (*i.e.*, $CCC/\Delta CCS_{50\%}$) for the Synapt G1 is ~11.[53] The instrument configuration of G1 system was later redesigned (Synapt G2-S HDMS; Figure 1.7) to improve instrumental performance characteristics. For example, the Synapt G2-S has an IM resolving power of ~44 (four times the Synapt G1).[53] Moreover, the Synapt G2-S has a helium (He) cell placed

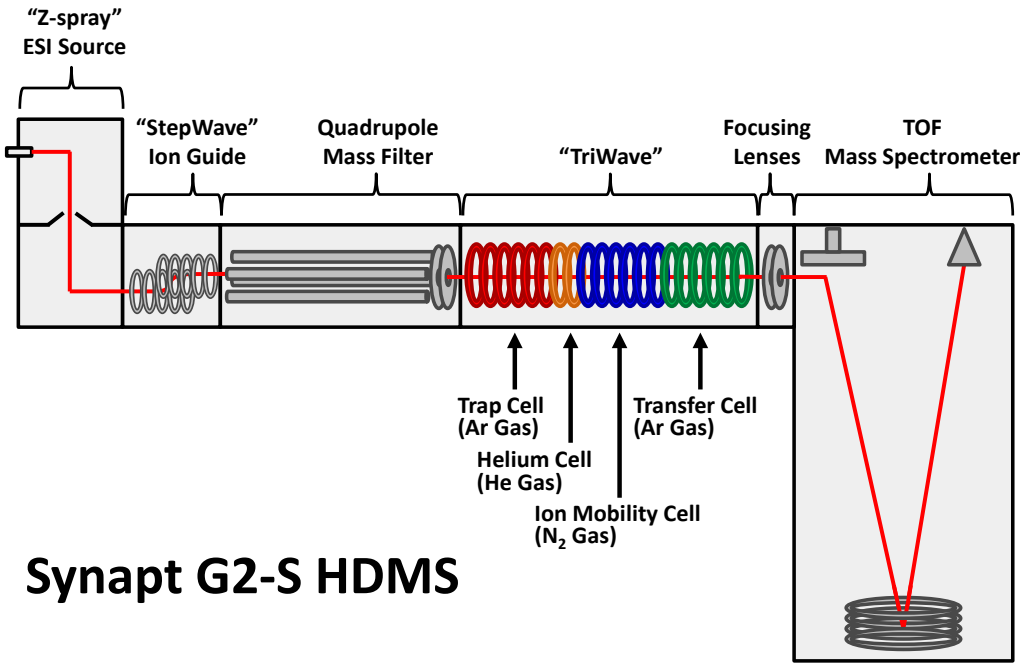


Figure 1.7. Schematic drawing of the Synapt G2-S HDMS IM TOF mass spectrometer (Waters Corporation). The red line shows the ion path through the instrument from the "z-spray" ion source to the detector. The red, orange, blue, and green rings correspond to the trap cell (for collision-induced dissociation (CID) and/or ion accumulation), helium cell (for collisional cooling), IM cell (for separation of ions based on their CSSs), and the transfer cell (for CID and/or guiding of ions to the TOF), respectively. Note: instrument not drawn to scale.

Despite the higher IM resolving power of the second generation TWIM-MS instrument (compared to the first generation instrument), isomers with very similar CCSs may still be unresolved by the Synapt G2-S.[54-56] To determine if an IM profile corresponds to two (or more) unresolved isomers, partially IM separated isomers (ions) can be fragmented to construct and analyze selected IM (SIM) profiles of informative individual fragment ions.[57-59] In this type of experiment, if a particular fragment ion is generated from a single isomer then the SIM profile for that "diagnostic" ion will correspond to the IM profile of the unresolved precursor ion.[57-59]

Although SIM profiles of diagnostic fragment ions may reveal the presence of IM unresolved isomers, this technique cannot be used to extract the individual CID mass spectra of IM unresolved isomers. Moreover, if fragmentation of the unresolved isomers generates similar product ions with slight differences in relative intensities, it can be difficult to determine if there are in-fact unresolved species present. To address these issues, Zekavat and Solouki used SIMPLe-to-use Interactive Self-modeling Mixture Analysis (SIMPLISMA),[60] and a similar approach to gas chromatography (GC)-MS deconvolution,[61] to deconvolute IM profiles *and* CID mass spectra of IM unresolved species.[62] Later, we improved the data preprocessing algorithm and integrated SIMPLISMA into an automated IM deconvolution (AIMD) software package for high-throughput data analysis.[54] Figure 1.8 shows a screen shot of the AIMD graphical user interface. Using this technique, we were able to analyze a variety of biomolecule mixtures of known[54, 56, 62] and unknown[55, 63, 64] sample compositions. Moreover, we recently demonstrated that chemometric deconvolution of IM-MS data is a CID

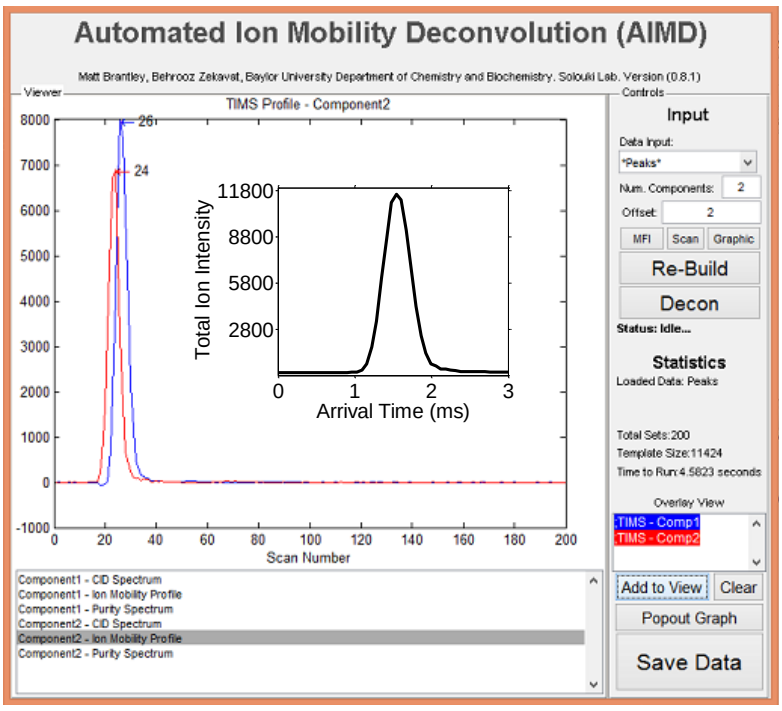


Figure 1.8. Screen capture of the automated IM deconvolution (AIMD) graphical user interface showing the deconvoluted IM profiles (red and blue traces) of two isomeric peptides. Inset shows the convoluted IM profile of the isomeric peptides before deconvolution.

energy dependent process and that experimental parameters can be optimized for IM deconvolution of complex mixtures.[56]

Mass Spectrometry Analysis of Biomolecules

Although early MS experiments were limited to analysis of small volatile molecules, [5, 12, 65] modern MS instruments are capable of routine analysis of large molecules and complex mixtures.[66] These new developments have facilitated high-throughput methods for elucidating biomolecule identities and structures. In the following sections, utilities of MS instruments for analysis of protein and oligonucleotide sequences are presented. In addition, potential challenges associated with structural analysis of macro molecules and their gas-phase sequencing are discussed.

Protein analysis. There are currently two primary methods for elucidating the identities and amino acid sequences of proteins *via* MS, namely, (1) bottom-up and (2) top-down proteomics approaches.[67-69] In bottom-up proteomics, unknown proteins are first enzymatically digested (*e.g.,* using trypsin) to generate smaller peptides from the intact protein.[67, 70] Digestion products are then (generally) injected onto a high-performance liquid chromatography (HPLC) column for separation prior to MS analysis.[67, 70] The MS data are then searched against libraries (*e.g.,* MASCOT) of proteins with known amino acid sequences to identify potential matches based off of scoring algorithms.[67, 71]

To enhance the sequence-coverage (and therefore match score[71]) of proteins, tandem MS (or "MS/MS") experiments can be performed.[67] In MS/MS experiments, ionized peptides are further fragmented in the gas phase to obtain additional sequence information by identifying contiguous sequences.[67] This process may be repeated several times where fragment ions can be isolated and further fragmented to yield even higher sequence-coverage. Multistage MS/MS experiments are differentiated from traditional tandem MS experiments by using the term "MS^n" (where "n" denotes the number of isolation/fragmentation stages).[72]

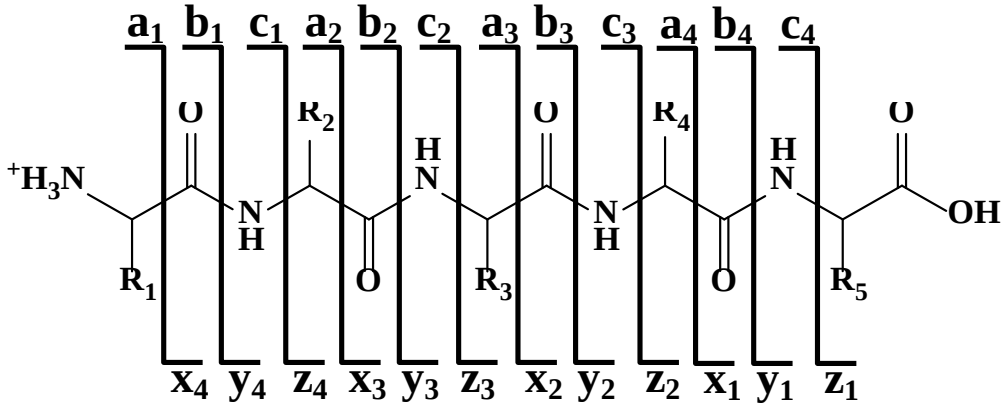
Top-down proteomics approaches are similar to bottom-up approaches; however, in top-down proteomics the digestion step is skipped and intact proteins are directly ionized and analyzed using MS.[67-69] Similar to the bottom-up approach, the protein is then isolated and fragmented (*viz.,* by MS/MS or MS^n) to generate product fragment ions which can be analyzed to determine the identity and amino acid sequence of the original protein. [67-69, 73]

The key advantage of top-down over bottom-up proteomics is the ability to preserve post-translational modifications (PTMs) and (potentially) achieve better sequence-coverage.[67-69, 73] In the bottom-up approach, enzymatic digestion can generate peptides with low proton affinities (*e.g.*, hydrophobic fragments may have low proton affinity values).[74] If this occurs and low proton affinity fragment ions are generated during the enzymatic digestion, then these sections of the protein cannot be identified (*i.e.*, because the peptides are not ionized) and therefore the sequence-coverage and matching score will be reduced.[68] Moreover, PTMs attached to these unidentified peptides, or PTMs that are completely removed from the protein during digestion, will also be unidentifiable.[67, 68] Therefore, even if the protein itself is correctly identified, the biologically relevant protein variant (*i.e.*, from modification) might not be identified. In the top-down approach, PTMs are preserved in the protein primary structure and, because the intact protein is ionized, the entire amino acid sequence is (potentially) available for analysis *via* MS^n.[67-69, 73]

Currently, the most ubiquitous gas-phase ion fragmentation technique is CID.[75-77] In CID, ions are generally accelerated in between two electrodes (using an electric potential gradient) and collided with neutral gas molecules (*e.g.*, nitrogen gas (N_2)). If there is sufficient collisional energy (*e.g.*, ~3.6-3.9 eV or ~83-90 kcal/mol for a C-C bond),[78] the ion will fragment. Because the collisional energy is randomized and distributed throughout the molecule,[79-82] fragmentation is most likely to occur at the weakest bond position.[83] In proteins, the weakest bonds are generally associated with the loss of small neutrals or dissociation of the amide bonds between amino acids. Therefore, CID of protein/peptide ions predominately yields peptide fragment ions *via* dissociation

of amide bonds.[70, 84] Because fragmentation occurs along the peptide backbone, product fragment ions can be interrogated to provide structural/sequence information.[71, 76, 77]

Although fragmentation of protein/peptide ions generally occurs at the amide bond positions during CID,[70, 84] there are several backbone and side chain positions that can also fragment (from CID or other fragmentation techniques). In 1984 Roepostorff and Fohlman proposed a simple nomenclature to name peptide fragment ions (Scheme 1.1).[85] In their nomenclature scheme (Scheme 1.1) fragment ions are named based off of the number of amino acids in the peptide (denoted by a subscript digit) as well as the location of the bond dissociation (denoted by the letters: a, b, c, x, y , and z; see Scheme 1.1). For example, if a peptide ion fragments at an amide bond the resulting fragment is called a "y-type" fragment ion if the charge is retained on the C-terminus or a "b-type" fragment ion if the charge is retained on the N-terminus. Moreover, if the peptide bond dissociates between the fifth and sixth amino acids (numbered from the N-terminus) and the charge is retained on the N-terminal fragment, the resulting fragment ion is called

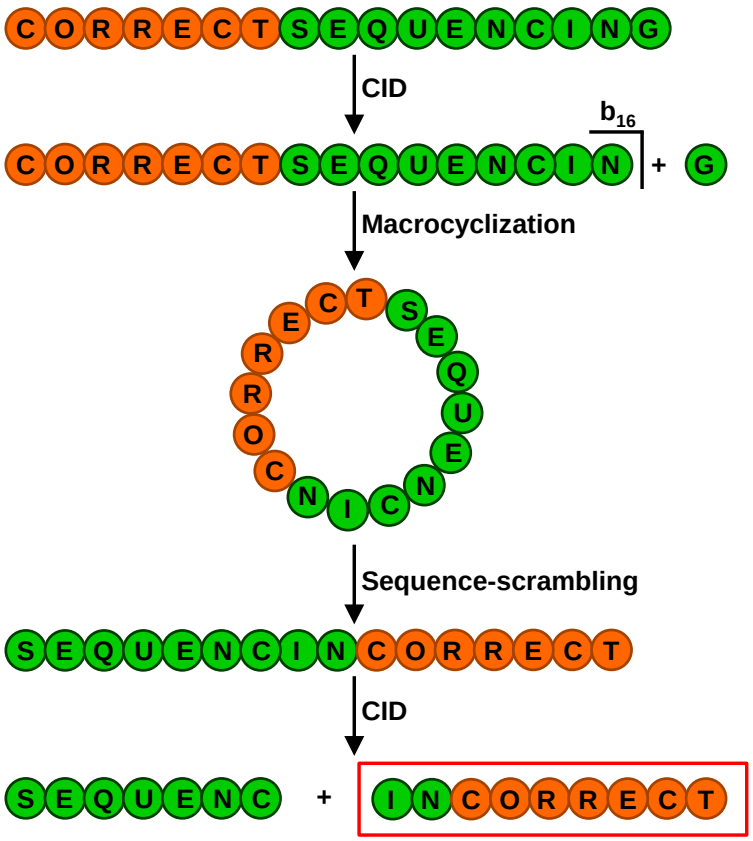


Scheme 1.1. Roepstorff and Fohlman's nomenclature for peptide ion fragmentation.[85] "b_5 fragment ion"; the subscript number (denoting the number of amino acids in the fragment ion sequence) is always counted from the terminus that keeps the charge (*e.g.*, y-type fragment ions are counted from the C-terminus).

a"b_5 fragment ion"; the subscript number (denoting the number of amino acids in the fragment ion sequence) is always counted from the terminus that keeps the charge (*e.g.*, y-type fragment ions are counted from the C-terminus).

Current sequence algorithms assume that all fragment ions are generated from "normal" Roepstorff-like fragmentation events, such as sequential loss of amino acids from the N- and/or C-termini.[71] However, recent evidence suggests that this may not always be a good assumption. For example, Solouki and coworkers showed that b-type peptide fragment ions can adopt (at least) two conformations corresponding to linear and macrocyclized structures.[86, 87] Extensive investigation of b-fragment ions has shown that, upon CID (*e.g.*, during MS^n), macro-cyclized b-fragment ions can reopen at internal amide bonds to generate peptides with new amino acid sequences that are different than the sequence of the original protein.[88-104]

Scheme 1.2 shows a hypothetical fragmentation pathway of a peptide with the amino acid sequence "CORRECTSEQUENCING". CID of the hypothetical peptide in Scheme 1.2 could generate a b_{16} fragment ion which can macro-cyclize and rearrange to generate a new linear sequence, "SEQUENCINCORRECT". The rearranged linear peptide ion (which has an identical *m/z* to the original b_{16} fragment ion) can then further fragment to generate a number of sequence-scrambled fragment ions such as "INCORRECT" (Scheme 1.2). If these new fragment ion sequences are misconstrued as accurately representing the original amino acid sequence, then the postulated protein identity may be incorrect. Misidentification of proteins is an undesired outcome in for any research study and particularly detrimental for those studies that are focused on identification of biomarkers for detection of a disease(s) or identification of drug targets.



Scheme 1.2. Simplified pictorial mechanism for sequence-scrambling of b-type fragment ions to generate re-arranged product fragment ions with amino acid sequences that are "incorrect" (*i.e.*, do not match the original amino acid sequence.

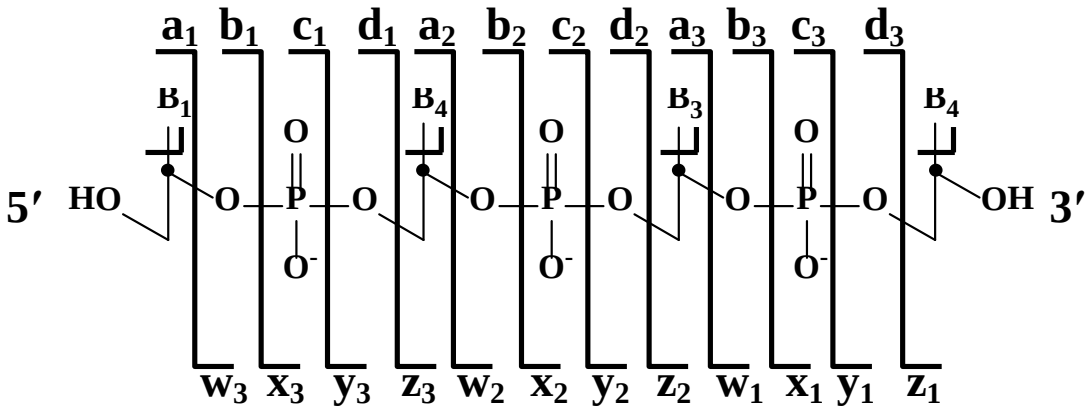
Sequence-scrambling of b-type fragment ions has generally been attributed to susceptibility of the C-terminal (imidazole-terminated) amino acid to nucleophilic attack by the lone pair of the N-terminal amino group.[88-104] On the other hand, y-type fragment ions are essentially truncated peptides (*i.e.*, with hydroxyl-terminated C-termini). Therefore, prior to a report from our lab, it was presumed that y-type fragment ions did not undergo sequence-scrambling.[105] However, in 2012 we showed, for the first time, that y-type fragment ions could also undergo sequence-scrambling, following a mechanism similar to b-type fragment ion scrambling.[106] Other examples of y-type fragment scrambling have since been reported.[55, 107] Detailed mechanism and experimental

evidence (including MS^n, IM-MS, and isotope labeling experiments) for y-type fragment ion sequence-scrambling will be discussed in Chapter Two of this dissertation.

Oligonucleotide analysis. As the resolution and throughput of MS has improved, analysis of oligonucleotides (*e.g.*, DNA and RNA) *via* MS has also increased.[108] Similar to sequencing the amino acid order of proteins, the nucleotide order of oligonucleotides can also be sequenced using MS.[109-113] Similar to CID of peptides, CID of an ionized DNA precursor ions also generates overlapping fragment ions whose identities can be used to determine the nucleotide sequence of the original precursor DNA strand.[114] Because of the complexity of DNA structures, there are more backbone and side chain cleavage sites in DNA than there are in peptides or proteins. Therefore, the CID mass spectra of DNA are generally very complex.[109, 114]

In 1992 McLuckey *et al.* proposed a new common nomenclature, based on Roepstorff and Fohlman's peptide fragmentation nomenclature,[85] for identifying oligonucleotide fragment ions (Scheme 1.3).[114] In McLuckey's nomenclature, fragment ions generated from the 5′ terminal (*i.e.*, retaining the charge on the 5′ terminal) are denoted as a-, b-, c-, or d-type fragment ions (analogous to a, b-, and c-type peptide fragment ions (Scheme 1.1)) whereas fragment ions generated from the 3′ terminal are denoted as w-, x-, y-, or z-type fragment ions (analogous to x-, y-, and z-type peptide fragment ions (Scheme 1.1)).[85, 114] Analogous to Roepstorff's peptide fragmentation nomenclature, in McLuckey's oligonucleotide fragmentation nomenclature, the subscript digit denotes the number of nucleotides in the fragment ion (also numbered from the terminus that keeps the charge). For example, a fragment ion denoted as "$[w_5]^{2-}$" refers to a doubly-deprotonated DNA fragment ion (where the charge is kept on the 3′ terminus)

containing five nucleotides with fragmentation occurring between the fifth and sixth nucleotides (number from the 3′ terminus) at the sixth nucleotide's 3′ carbon-oxygen bond. Side chain bases can be easily lost during fragmentation and are differentiated from other fragment ions by adding "-B" to the fragment ion name (*e.g.*, "a-B" denotes an a-type fragment which has lost a base).[114] The most commonly observed fragment ions generated from CID of oligonucleotides are w and a-B type ions which, analogous to b- and y-type peptide fragment ions, can be used to determine the nucleotide sequence of the original oligonucleotide.[114]



Scheme 1.3. McLuckey's nomenclature for oligonucleotide ion fragmentation. Structure is abbreviated for simplicity. Filled circles denote backbone sugars (deoxyribose or ribose).[114]

Justification for the Research Presented in this Dissertation

The goal of the research presented in this dissertation is to improve existing, and develop new, MS and IM-MS techniques for structural and conformational analysis of biomolecules. In Chapter Two we use a combination of isotope labeling, MS^n, IM-MS, and IM deconvolution to investigate the mechanism of sequence-scrambling in CID of y-type fragment ions.[55] Specifically, we present evidence suggesting that sequence-

scrambling of y-type ions proceeds through loss of internal backbone carbonyl groups prior to loss of internal amino acids.[55] Although sequence-scrambling has been of major interest to the protein/peptide MS community, few groups have investigated unusual fragmentation pathways for oligonucleotides. In Chapter Three we extend our results from Chapter Two to demonstrate, for the first time, that w-type DNA fragment ions can also rearrange upon CID.[115] We performed MS/MS experiments on 21 w-type fragment ions from three different single-stranded DNA sequences on three different mass spectrometers (*i.e.,* TOF MS, Orbitrap MS, and triple quadrupole MS) and demonstrated that purine bases can attack the 5′ phosphate group to generate "phosphopurine" rearranged fragment ions.[115] We also present results from IM-MS and theoretical modeling to support our proposed rearrangement mechanism.[115]

Chapters Two and Three focus on biomolecule primary structures (*e.g.,* sequence of amino acids and nucleotides) and in Chapter Four we demonstrate the utility of chemometric deconvolution of post-IM/CID MS data[54, 56, 62, 63] to determine accurate CCSs of IM unresolved peptide isomers. We report that experimentally determined CCSs for chemometrically deconvoluted IM profiles (using TWIMS) are not statistically different than CCSs calculated from individually analyzed peptides using both TWIMS and DTIMS.

In the final chapter of this dissertation, we discuss future directions and preliminary results from application of the techniques presented in Chapters Two through Four for the rapid analysis of human and animal tissues (*e.g.,* muscle). Specifically, we discuss how these techniques can be leveraged for biomarker discovery and investigation of disease pathophysiologies.

Author Contributions

All unpublished data and figures in the introduction and conclusion (Chapter One and Chapter Five, respectively) are the sole work of Brett Harper. Approximately seventy percent of the data in Chapter Two was collected by Brett Harper and approximately thirty percent was collected by Mahsan Miladi. The text in Chapter Two was contributed to by all authors (Brett Harper, Mahsan Miladi, and Touradj Solouki), with approximately eighty-five percent written by Brett Harper and equal contributions from Mahsan Miladi and Touardj Solouki. Approximately eighty percent of the data in Chapter Three was collected by Brett Harper and approximately twenty percent was collected by Elizabeth Neumann. Theoretical modeling was performed by Elizabeth Neumann. The text in Chapter Two was contributed to by all authors (Brett Harper, Elizabeth Neumann, and Touradj Solouki), with approximately ninety percent written by Brett Harper and equal contributions from Elizabeth Neumann and Touardj Solouki. Approximately seventy percent of the TWIM-MS data in Chapter Four was collected by Brett Harper and thirty percent collected by Elizabeth Neumann. DTIM-MS data was collected by Sarah Stow. All calculations in Chapter Four were done by Brett Harper. Theoretical modeling was performed Sarah Stow and Jody May. All authors (Brett Harper, Elizabeth Neumann, Sarah Stow, Jody May, John McLeann, and Touradj Solouki) contributed to the text in Chapter Four, with Brett Harper writing approximately eighty percent, Sara Stow writing five percent, and equal contributions from other authors.

CHAPTER TWO

Loss of Internal Backbone Carbonyls: Additional Evidence for Sequence-scrambling in Collision-induced Dissociation of y-Type Ions

This chapter published as: Harper, B.; Miladi, M.; Solouki, T., Loss of Internal Backbone Carbonyls: Additional Evidence for Sequence-Scrambling in Collision-induced Dissociation of y-Type Ions. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (10), 1716-1729.

Abstract

It is shown that y-type ions, after losing C-terminal H_2O or NH_3, can lose an internal backbone carbonyl (CO) from different peptide positions and yield structurally different product fragment ions upon collision-induced dissociation (CID). Such CO losses from internal peptide backbones of y-fragment ions are not unique to a single peptide and were observed in four of five model peptides studied herein. Experimental details on examples of CO losses from y-type fragment ions for an isotopically labeled AAAAHAA-NH_2 heptapeptide and des-acetylated-α-melanocyte-stimulating hormone (dα-MSH) (SYSMEHFRWGKPV-NH_2) are reported. Results from isotope labeling, tandem mass spectrometry (MS^n), and ion mobility-mass spectrometry (IM-MS) confirm that CO losses from different amino acids of *m/z*-isolated y-type ions yield structurally different ions. It is shown that losses of internal backbone carbonyls (as CID products of *m/z*-isolated y-type ions) are among intermediate steps towards formation of rearranged or permutated product fragment ions. Possible mechanisms for generation of the observed sequence-scrambled a-"like" ions, as intermediates in sequence-scrambling pathways of y-type ions, are proposed and discussed.

Introduction

Mechanistic understating of protein/peptide fragmentation in the gas-phase is an essential aspect of mass spectrometry-based proteomics and is gaining growing interest in the literature.[84, 116, 117] "Normal", or direct, product fragment ions are generated from losses of amino acids or amino acid segments from the N- (*e.g.*, x-, y-, and z-type ions [85]) and/or C-termini (*e.g.*, a-, b-, and c-type ions [85]) upon gas-phase dissociation.[84] Conversely, product fragment ion "rearrangement",[88-90] or "sequence-scrambling",[91-94] can occur when precursor ions lose internal residues to generate permutated product fragment ions (with "rearranged" amino acid orders) that do not match their respective precursor ions' original amino acid sequences. Presence of sequence-scrambled product fragment ions in collision-induced dissociation (CID) mass spectra complicates protein/peptide sequencing and may result in misidentification of proteins' primary structures.[71]

Generation of sequence-scrambled product fragment ions upon CID of b-type ions was first reported by Tang *et al.*.[88-90] Gas-phase structural probes (*e.g.*, hydrogen/deuterium exchange (HDX),[86, 87, 96, 97, 118, 119] ion mobility-mass spectrometry (IM-MS),[62, 120-122] infrared multiphoton dissociation (IRMPD),[96, 97, 104, 123-125] gas-phase proton transfer (PT) reactions,[126] and molecular modeling approaches [84, 127-129]) have been utilized to study the presence or absence of different structures (*i.e.*, oxazolone and macrocyclic) for b-type ions. Furthermore, mechanistic studies have suggested that sequence-scrambling in CID of b-type ions occurs through fragmentation of intermediate macrocyclic structures.[95, 96, 98-104]

Although a growing number of reports implicate b-type ions with sequence-scrambling, fewer studies have explored the possibility and extent of sequence-scrambling in CID of other types of product fragment ions (*i.e.*, a-, c-, x-, y-, and z-type ions [85]). However, there are indications that sequence-scrambling may not be exclusive to CID of b-type ions. For example, Riba-Garcia *et al.* demonstrated that $[a_n]^+$ (n = 4 and 5) product fragment ions generated from the C-terminal amidated peptide YAGFL-NH_2 could have multiple structures.[120] Additionally, IM-MS results by Badman *et al.* indicated multiple gas-phase structures/conformers for the $[y_{44}]^{4+}$ and $[y_n]^{5+}$ (n = 58-64) product fragment ions of ubiquitin.[130]

Recently, we reported that, similar to b ions, CID of *m/z*-isolated y-type ions could also yield sequence-scrambled product fragment ions.[106] Our experimental results suggested that sequence-scrambling in CID of y-type ions occurs through loss of H_2O or NH_3 (for C-terminal amidated peptides) from the C-terminus followed by formation of macrocyclic b-"like" ions.[106] Such macrocyclic structures can reopen at internal amide bond locations and yield sequence-scrambled product fragment ions.[106]

In this chapter, we utilize isotope labeling, MS^n, and ion mobility to demonstrate that CID of isolated y-type ions can generate sequence-scrambled product fragment ions through a-type ion chemistry, *via* loss of different carbonyl groups from various backbone positions. Moreover, based on the observed fragmentation patterns, we propose and discuss possible mechanisms for generation of the observed sequence-scrambled a-"like" ions as intermediates in sequence-scrambling pathways.

Materials and Methods

Sample Preparation

Optima grade water, methanol, and acetic acid were purchased from Fisher Scientific (Pittsburgh, PA, USA). Hexaglycine (G_6) (used as a lock mass in Orbitrap experiments), pentapeptide (SDGRG), and des-acetylated-α-melanocyte-stimulating hormone (dα-MSH) (SYSMEHFRWGKPV-NH_2) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Angiotensin II antipeptide (EGVYVHPV), angiotensin II (DRVYIHPF), and isotopically (^{13}C) labeled heptapaptide (AAAAH<u>A</u>A-NH_2) were purchased from CPC Scientific Incorporated (Sunnyvale, CA, USA). The isotopically labeled (^{13}C) control peptide AAAAH***A***A-NH_2 (where "***A***" indicates a carbon thirteen (^{13}C) isotope on the alanine *side chain* (*i.e.,* R group)) was synthesized in-house [131] using a solid-phase peptide synthesis protocol.[132] Electrospray ionization (ESI) solvent (methanol (~49.95 %): water (~49.95 %): acetic acid (0.1 %)) and peptide samples were used without further purification. Micromolar concentrations of analytes (*e.g.*, AAAH<u>A</u>A-NH_2 in ESI solvent) were used for all MS^n and IM-MS experiments.

Instrumentation

Ion Mobility-Mass Spectrometry. IM-MS experiments were performed using a Synapt G2-S HDMS (Waters Corporation, Manchester, UK) operated in positive-ion mode ESI. The time-of-flight (TOF) was operated in sensitivity mode (*i.e.*, "V" mode) for enhanced detection of low abundance ions. The sample flow rate for ESI was set at 0.2 μL/min (AAAAH<u>A</u>A-NH_2) or 0.5 μL/min (dα-MSH) using a Harvard Pump 11 Plus dual syringe pump (Harvard Apparatus, Holliston, MA, USA). The MS scan time was set

to 1 s. Synapt G2-S HDMS IM-MS data reported in here correspond to an average of 180 scans (AAAAHAA-NH_2) or 1800 scans (dα-MSH) which provided signal-to-noise ratios of >3 for all product fragment ions of interest (deconvolution of IM-unresolved dα-MSH product fragment ions [Figure 5] necessitated a large number [*i.e.*, 1800] of scans). For the isotopically labeled heptapeptide (*i.e.*, AAAAHAA-NH_2), the capillary, sampling cone, and source offset voltages were set at +2.0 kV, +85.0 V, and +35.0 V, respectively. Harsher source conditions for dα-MSH were necessary to generate $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664), and thus the capillary, sampling cone, and source offset voltages were set at +4.0 kV, +100.0 V, and +65.0 V, respectively. Source temperature was set to 100 ˚C for all ESI TOF experiments. Isolation parameters (*i.e.*, quadrupole RF settings, ion optics, and *m/z*-isolation window) were optimized to isolate the first isotope ($^{13}C_1$) peak of $[y_6]^+$ (*m/z* 511) generated from in-source fragmentation of AAAAHAA-NH_2 and the complete isotopic packet (due to mass isolation restrictions) of $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) from dα-MSH. Argon gas pressures in the trap- and transfer-cells were both set to ~2 x 10^{-2} mbar. Helium (He) gas pressure in the He cell was ~1 bar. Nitrogen (N_2) gas pressure in the IM cell was ~3.5 mbar. Pressure in the TOF region of the instrument was ~1.00 x 10^{-6} mbar. In all of the IM experiments, a wave height of +40 V and optimized wave velocities of ~1250 m/s (AAAAHAA-NH_2) and ~1000 m/s (dα-MSH) were applied to the IM stacked ring ion guide assembly.

Product fragment ions from the isotopically labeled peptide (*i.e.*, AAAAHAA-NH_2), were generated prior to IM separation (pre-IM CID) in the trap cell using collision energies corresponding to a 25 V potential difference between the trap cell exit and entrance of the helium cell. The potential difference between the exit of the IM cell and

entrance of the transfer cell was set at 2.0 V to prevent post-IM fragmentation of ions of interest, except in post-IM CID experiments.

Product fragment ions from dα-MSH were generated after IM separation (post-IM CID) of *m/z*-isolated $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) in the transfer cell using collision energies corresponding to a 40 V potential difference between the exit of the IM cell and entrance of the transfer cell. The potential difference between the trap cell exit and entrance of the helium cell was set at 1.2 V to prevent pre-IM fragmentation.

Default manufacture settings were used for all other instrument parameters (*e.g.*, trap cell wave velocity (400 m/s) and height (+3.0 V), transfer cell wave velocity (665 m/s) and height (+3.0 V), pusher frequency (~14.5 kHz), and all other ion optic parameters).

IM-MS data were analyzed using MassLynx software (version 4.1) (Waters Corporation, Manchester, UK). Chemometric analysis of post-IM CID results for dα-MSH were performed in Matlab 7.0 (The MathWorks Inc., Natick, MA, USA) using a previously published method [62]. Mass corrected IM drift times (for $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) from AAAAH<u>A</u>A-NH_2) were calculated using the method proposed by Ruotolo *et al.*[133] Reported IM profiles for $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) from AAAAH<u>A</u>A-NH_2 correspond to the centroid *m/z* ± 0.02 mass range, extracted from the CID mass spectra of mono-isotopically isolated $[y_6]^+$.

Ion trap tandem mass spectrometry. A linear trap quadrupole (LTQ) Orbitrap Discovery mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) operated in the positive ion-mode was used to acquire all MS^n CID mass spectra. The ESI sample

flow rate was set to 5.0 μL/min. The mass resolving power ($M/\Delta M_{50\%}$) of the Orbitrap was set to ~30,000 (at *m/z* 400). All LTQ Orbitrap mass spectra reported in here correspond to an average of 30 scans (where each scan is an average of 5 microscans); these optimized experimental conditions provided signal-to-noise ratios of >3 for all product fragment ions of interest. The ESI source, metal capillary, and tube lens voltages were set at +4.55 kV, +50 V, and +145 V, respectively. The LTQ metal capillary temperature was set at 275 ˚C. LTQ parameters (*i.e.*, quadrupole RF settings and isolation window) were optimized to isolate the first isotopic peak of $[y_6]^+$ (*m/z* 511) generated from in-source fragmentation of AAAAHAA-NH_2. CID of *m/z*-isolated ions was carried out in the LTQ at optimized normalized collision energies [134] of 15% to 30%. He gas (partial pressure of ~1 x 10^{-5} torr) was used as the collision gas for LTQ Orbitrap CID experiments. Protonated hexaglycine (*m/z* 361) was used as a lock mass for acquisition of all LTQ Orbitrap mass spectra. Mass measurement errors (MMEs) for all LTQ Orbitrap experiments were better than ~1.5 ppm. Orbitrap mass spectra were analyzed using Xcalibur software (version 2.1) (Thermo Fisher Scientific, Waltham, MA, USA). Instrumental parameters (for the Orbitrap mass spectrometer) for generation of the mass spectra shown in Figures 2.1 and 2.2, as well as all Appendix A information (*i.e.*, Figures A.1, A.2, A.3, and Table A.1) were similar to those discussed above and have been previously reported in detail.[106]

All presented IM-MS and MS^n experiments were performed a minimum of three times to ensure the reproducibility of acquired data. Reported IM drift times are the average of three (n = 3) replicates and confidence intervals are reported at the 95% confidence level.

Nomenclature

Product fragment ion assignments reported in this text reflect the origin of product fragment ions (*i.e.,* losses from isolated y-ions) and should not be confused with conventional a, b, and other product fragment ions from protonated molecular ions. According to current recommendations by the International Union of Pure and Applied Chemistry (IUPAC), an a-type ion is defined as a "fragment ion containing the peptide N-terminus formed upon dissociation of a peptide ion at the peptide backbone C-C bond".[135] Because y-type ions are formed *via* losses of original N-termini from their respective precursor peptides (*e.g.,* loss of one or more amino acids(s)), subsequent losses of H_2O/NH_3 and CO from y-type ions should not be classified as a-type product ions. To acknowledge the structural similarity of these product ions (formed by neutral losses from y-type ions) to a- and b-type ions, we refer to them as a- and b-"like" ions throughout the text. Further justification and discussion of the nomenclature used in here can be found in our recent publication.[106]

Results and Discussion

In the following sections we present MS and IM-MS data that show loss of internal carbonyl groups during the CID of *m/z*-isolated $[y_6]^+$ generated from protonated AAAAH<u>A</u>A-NH_2 heptapeptide. First, we provide MS^n CID evidence confirming the loss of an internal, isotopically labeled backbone carbonyl group (*i.e.,* $[y_6 - (NH_3 + {}^{13}CO)]^+$) following CID of *m/z*-isolated $[y_6]^+$ (generated from CID of protonated AAAAH<u>A</u>A-NH_2). Second, we utilize IM-MS to compare the gas-phase structures of $[y_6 - (NH_3 + {}^{12}CO)]^+$ and $[y_6 - (NH_3 + {}^{13}CO)]^+$, and to probe the potential for $[y_6 - (NH_3 + {}^{12}CO)]^+$ to be a mixture of different product fragment ions that lose carbonyls from multiple

locations. Next, we perform MS^n using an LTQ Orbitrap mass spectrometer to map different positions of $[y_6]^+$ (from AAAAH<u>A</u>A-NH_2) that lose carbonyl groups. Finally, to test the effects of backbone carbonyl losses from y-fragments of other peptides, we perform post-IM CID on $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ from dα-MSH and discuss our findings.

Evidence for Loss of Internal Backbone Carbonyls Following CID of $[y_6]^+$ from Protonated AAAAH<u>A</u>A-NH_2

During a recent investigation assessing sequence-scrambling in subsequent CID of *m/z*-isolated y-type ions,[106] we noticed the generation of unusual product fragment ions corresponding to potential loss of a labeled backbone carbonyl group (from an internal amino acid) from $[y_n]^+$ (where n = 4-6). These $[y_n]^+$ product fragment ions were generated from CID of a C-terminal amidated heptapeptide AAAAH<u>A</u>A-NH_2. For brevity, only results from CID of *m/z*-isolated $[y_6]^+$ are presented in this chapter. Representative CID mass spectra for *m/z*-isolated $[y_5]^+$ and $[y_4]^+$ that indicate losses of internal ^{13}C labeled carbonyl groups can be found in Appendix A (Figures A.1 and A.2, respectively). Figure 2.1 shows the CID mass spectrum of mono-isotopically isolated $[y_6]^+$ (*m/z* 511) generated from in-source fragmentation of AAAAH<u>A</u>A-NH_2 (*m/z* 582) in the LTQ Orbitrap. Product fragment ions at *m/z* 465.25674 and 466.26019 (Figure 2.1 inset) correspond to $[y_6 - (NH_3 + {}^{13}CO)]^+$ and $[y_6 - (NH_3 + {}^{12}CO)]^+$, with calculated mass measurement errors (MMEs) of approximately 0.21 ppm and 0.02 ppm, respectively.

Product fragment ion assignment of $[y_6 - (H_2O + {}^{12}CO)]^+$ (theoretical *m/z* 465.27618) was also considered for the ion at *m/z* 465. However, calculated MME for $[y_6 - (H_2O + {}^{12}CO)]^+$ product fragment ion assignment is approximately 41.48 ppm, which is far greater than the instrument's expected MME of ≤ 2 ppm when using an appropriate

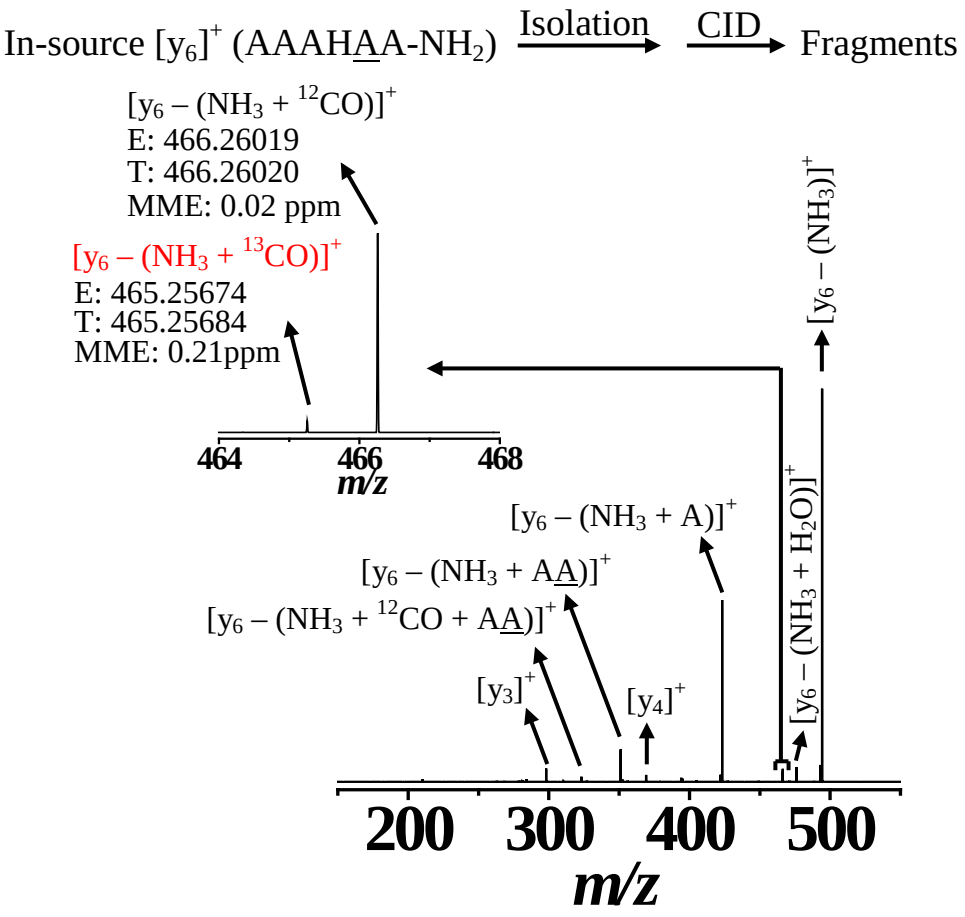


Figure 2.1. LTQ Orbitrap CID mass spectrum of *m/z*-isolated $[y_6]^+$ (*m/z* 511) (generated from in-source fragmentation of protonated AAAAHAA-NH_2). Inset shows the expanded view of the *m/z* range from 464 to 468 corresponding to $[y_6 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) and scrambled product fragment ion $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) (red font). Experimental (E) and theoretical (T) *m/z* values as well mass measurement error (MME) values for select postulated product fragment ion assignments are indicated.

lock mass.[47] Mass measurement error calculations suggest that observed product fragment ion at *m/z* 465 should be assigned as $[y_6 – (NH_3 + {}^{13}CO)]^+$ (0.21 ppm) and not $[y_6 – (H_2O + {}^{12}CO)]^+$ (41.48 ppm). Furthermore, to confirm that no $[y_6 – (H_2O + {}^{12}CO)]^+$ is generated, and ^{13}CO is lost from an internally labeled backbone carbonyl group, we performed CID on a control counterpart protonated peptide, AAAAH***A***A-NH_2 in the LTQ

Orbitrap (where "***A***" indicates a carbon thirteen (^{13}C) isotope on the alanine *side chain* (*i.e.,* R group) rather than its backbone).

Figure 2.2 shows the CID mass spectrum of mono-isotopically isolated $[y_6]^+$ (*m/z* 511) generated from in-source fragmentation of protonated AAAAH***A***A-NH_2 (*m/z* 582). Fragmentation patterns for the control peptide were similar to those observed for CID of $[y_6]^+$ (*m/z* 511) generated from in-source fragmentation of AAAAH<u>A</u>A-NH_2 (Figure 2.1). Although the product fragment ion at *m/z* 466 (*i.e.,* $[y_6 - (NH_3 + {}^{12}CO)]^+$) was observed in CID mass spectrum of *m/z*-isolated $[y_6]^+$ (*m/z* 511) generated from AAAAH***A***A-NH_2 (*i.e.,* control peptide), there was no peak present at *m/z* 465. We can draw two conclusions from CID results of this control peptide (Figure 2.2). First, under our experimental conditions, AAAAH***A***A-NH_2 does not appear to generate a product fragment ion at *m/z* 465 through $[y_6 - (H_2O + {}^{12}CO)]^+$ channel; this suggests that AAAAH<u>A</u>A-NH_2 should not generate $[y_6 - (H_2O + {}^{12}CO)]^+$ either. Second, $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465.25674) was not present in the CID mass spectrum of $[y_6]^+$ (*m/z* 511) generated from AAAAH***A***A-NH_2 (Figure 2.2). Because the control peptide (*i.e.,* AAAAH***A***A-NH_2) does not have a ^{13}C label on the Ala carbonyl group, absence of an ion at *m/z* 465 is expected and confirms that CO loss is from a backbone carbonyl (rather than the side chain of Ala). Results from MME calculations and CID of control AAAAH***A***A-NH_2 peptide are in agreement with one another and consistent with the assignment of $[y_6 - (NH_3 + {}^{13}CO)]^+$ for ions at *m/z* 465.25674 (Figure 2.1).

To further confirm our findings on CO loss assignment, we examined the CID mass spectra of y-type ions from other C-terminus free and amidated peptides. For

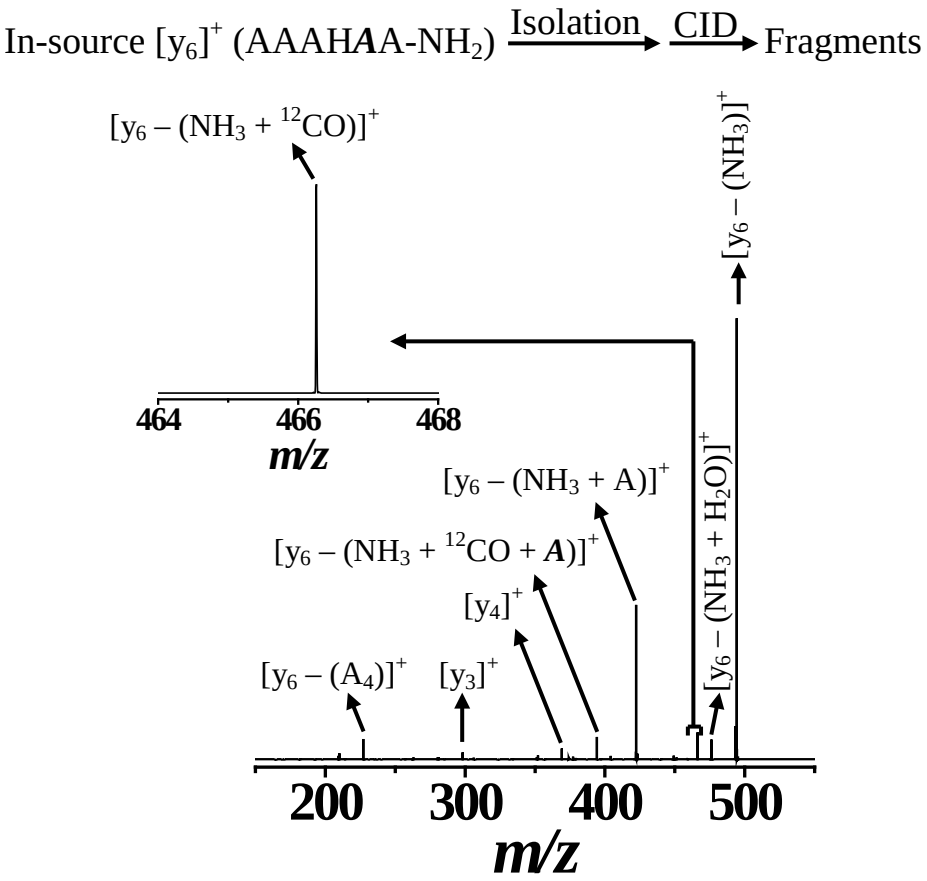


Figure 2.2. LTQ Orbitrap CID mass spectrum of *m/z*-isolated $[y_6]^+$ (*m/z* 511) (generated from in-source fragmentation of protonated control peptide, AAAAH***A***A-NH_2). Inset shows the expanded view of the *m/z* range from 464 to 468 corresponding to $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466). No product fragment ions were observed at *m/z* 465.

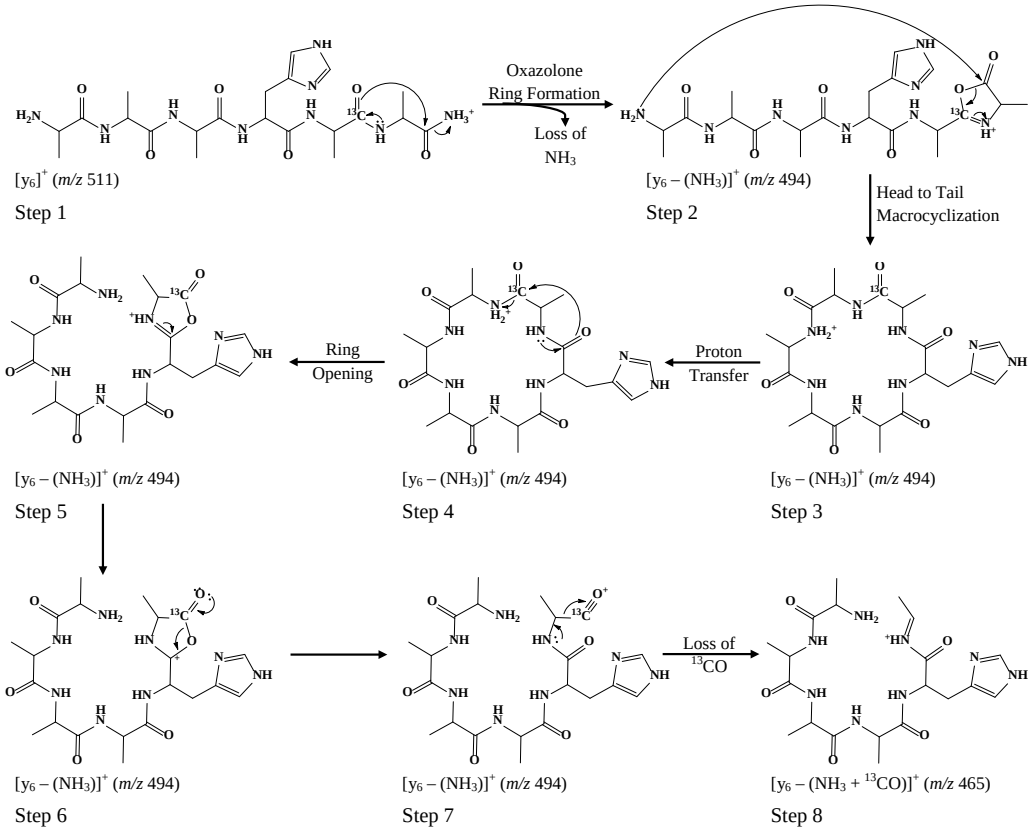
acid sequence: SYSMEHFRWGKPV-NH_2) and $[y_5]^{2+}$ and $[y_6]^+$ species of angiotensin II antipeptide (amino acid sequence: EGVYVHPV). Results from these experiments are summarized in Appendix A (Table A.1). CID of *m/z*-isolated y-type ions of different lengths and charge states from dα-MSH and angiotensin II antipeptide generated a-"like" ions, similar to $[y_6 - (NH_3 + {}^{12/13}CO)]^+$ of AAAAH<u>A</u>A-NH_2. Generation of these product fragment ions suggests that loss of C-terminal H_2O or NH_3 (from amidated peptides) and ^{12}CO from backbone positions may be a common loss from y-type ions upon CID.

Consequently, sequence-scrambling following CID of *m/z*-isolated y-type ions may be more common than previously assumed. Furthermore, we did not observe any $[y_n - (H_2O + {}^{12}CO)]^{m+}$ (n = 6-12; m = 1, 2, and/or 3) in the CID mass spectra of *m/z*-isolated y-type ions from amidated dα-MSH. These results support our conclusions from CID of the control peptide (*i.e.*, protonated AAAAH***A***A-NH_2; Figure 2.2) and reconfirm that the product fragment ion at *m/z* 465 in Figure 2.1 is $[y_6 - (NH_3 + {}^{13}CO)]^+$ and not $[y_6 - (H_2O + {}^{12}CO)]^+$.

The observed product fragment ion at *m/z* 465 (*i.e.*, $[y_6 - (NH_3 + {}^{13}CO)]^+$ in Figure 2.1) can only be formed by a loss of an internal ^{13}C labeled carbonyl group. Loss of an internal ^{13}CO suggests potential sequence-scrambling. The loss of ammonia (NH_3) (in addition to ^{13}CO in $[y_6 - (NH_3 + {}^{13}CO)]^+$) is consistent with our previous results which suggested that sequence-scrambling in CID of isolated y-type ions occurred through dissociation of intermediate macrocyclic b-"like" ions.[106] Likewise, recent results by Harrison *et al.* suggested that protonated C-terminal amidated peptides readily lose NH_3 to form macrocyclic b-type ions which can undergo further fragmentation and result in sequence-scrambled ions following a similar mechanism.[136] There are five non-labeled carbonyl (^{12}CO) groups in $[y_6]^+$ of AAAAH<u>A</u>A-NH_2, and without a systematic isotopic labeling, the ion at *m/z* 466 (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO)]^+$) cannot strictly be assigned as a sequence-scrambled product fragment ion. In other words, loss of carbonyl groups from five different locations (*e.g.*, see discussion of Scheme 2.2) could yield product fragment ions at *m/z* 466 (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO)]^+$); hence ions at *m/z* 466 could be a mixture of different sequence-scrambled and sequence-intact product fragment ion isomers. Based on the number of possible carbonyl groups in AAAAH<u>A</u>A-NH_2, CO losses should

predominately be of the ^{12}CO type with a lesser probability for ^{13}CO losses; as shown in Figure 2.1 (inset), the observed lower relative abundance for *m/z* 465 (as compared to the peak at *m/z* 466) is consistent with this view. However, reaction thermochemistry, chemical kinetics for metastable decay, and other variable parameters might also influence ion dissociation and mass spectral patterns.

Previous studies have demonstrated that imine terminated a-type ions (CO loss from C-terminus) can undergo macrocyclization and generate sequence-scrambled product ions.[137-139] In this study, however, the ion at *m/z* 465, *i.e.*, $[y_6 - (NH_3 + {}^{13}CO)]^+$ (Figure 2.1) is generated from the loss of internal ^{13}CO (suggesting that scrambling occurs prior to carbonyl expulsion). Scheme 2.1 shows a proposed mechanism for the loss of ^{13}C labeled internal backbone carbonyl group from CID of $[y_6]^+$ (generated from CID of protonated AAAAHAA-NH_2). Loss of NH_3 could be initiated *via* oxazolone ring formation at the C-terminus of $[y_6]^+$ (Scheme 2.1, Step 1).[104, 123] The oxazolone ring structure may subsequently reopen following a head-to-tail macrocyclization reaction (isobaric to macrocyclic b-type ions) (Scheme 2.1, Step 2). Following an "intramolecular" proton transfer (Scheme 2.1, Step 3), a second oxazolone ring can be formed at the labeled alanine residue (A) cleaving the A-A amide bond (Scheme 2.1, Step 4). Similar to generation of sequence-scrambled and normal a-type ions,[94, 140] the second oxazolone ring may then rearrange to lose ^{13}CO (Scheme 2.1, Step 5-8) and generate the N-terminal imine sequence-scrambled product fragment ion at *m/z* 465 (*i.e.*, $[y_6 - (NH_3 + {}^{13}CO)]^+$ (Scheme 2.1, Step 6).



Scheme 2.1. Proposed mechanism for generation of sequence-scrambled $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) product fragment ion from CID of $[y_6]^+$ from protonated AAAAHAA-NH_2.

CID of isolated protonated molecular ion of AAAAHAA-NH_2 (*i.e.*, $[M + H]^+$) should generate doublet (*i.e.*, loss of ^{12}CO and ^{13}CO) a-type ions.[121, 141] Results from CID of *m/z*-isolated $[M + H]^+$ (*m/z* 582) are provided in Appendix A, Figure A.3. CID of *m/z*-isolated $[M + H]^+$ (*m/z* 582) precursor ions generated $[a_n]^+$ ions (where n = 5-7) corresponding to losses of ^{12}CO and ^{13}CO. Generation of doublet peaks corresponding to losses of ^{12}CO and ^{13}CO (and C-terminal NH_3) from $[M + H]^+$ supports our proposed mechanism (Scheme 2.1) for generation of a-"like" product fragment ions at *m/z* 465 and 466 (Figure 2.1).

To further investigate the potential for sequence-scrambling *via* expulsion of internal carbonyl groups from CID of y-type ions, we performed IM-MS on product

fragment ions $[y_6 – (NH_3 + {}^{13}CO)]^+$ and $[y_6 – (NH_3 + {}^{12}CO)]^+$ to probe for the presence of different gas-phase conformers. Results from IM-MS studies of $[y_6 – (NH_3 + CO)]^+$ product fragment ions are presented in the following section.

IM-MS Analysis of $[y_6 – (NH_3 + {}^{13}CO)]^+$ and $[y_6 – (NH_3 + {}^{12}CO)]^+$ from Protonated AAAAHAA-NH_2

Previous studies have suggested that sequence-scrambling in CID of b-type ions occur when product fragment ions cyclize and reopen at various existing internal amide bonds.[95, 96, 98-104] Due to changes in the sequence-scrambled product fragment ion's primary structure, the ion's secondary and tertiary structures may also be affected. Assuming that similar mechanisms for b-type ion scrambling are involved in y-type ion scrambling, it would be reasonable to predict that $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) might have different structures and conformations and hence distinguishable ion mobility profiles.

The pre-IM CID mass spectrum of *m/z*-isolated $[y_6]^+$ (first isotopic peak at *m/z* 511) generated from in-source (Synapt G2-S HDMS) fragmentation of AAAAHAA-NH_2 is shown in Figure 2.3. Because of the differences in energetics of CID (*e.g.*, collision gas types and partial pressures) involved, CID of *m/z*-isolated $[y_6]^+$ in the Synapt G2-S HDMS (Figure 2.3a) generated more product fragment ions than in the LTQ Orbitrap (Figure 2.1). Differences in ESI source conditions (*e.g.*, temperatures and ESI spray voltage) between the LTQ Orbitrap and Synapt G2-S HDMS could lead to formation of ions with different initial internal energies. However, collisional and, to a lesser extent, radiative cooling during ion isolation and ion trapping events should be sufficient to

thermalize and reduce the internal energy differences of surviving ions and allow reasonable comparisons between the two experiments.

Under our experimental conditions, single-stage/direct ESI mass spectra or MS^1 of AAAAHAA-NH_2 obtained using either the LTQ Orbitrap or the Synapt G2-S HDMS system, and in the absence of ion fragmentation (*e.g.*, no applied CID for collisional excitation) showed similar patterns; this suggested that ions formed in both systems had comparable initial internal energy contents. For a quantitative comparison of fragment ions generated in the LTQ Orbitrap and the Synapt G2-S HDMS, it would be necessary to operate both instruments under identical experimental conditions (source temperature, spray conditions, collision gas type [*i.e.,* reduced mass], and pressure [*i.e.,* number of collisions]). Therefore, presence of different isomers under varied experimental conditions cannot be excluded. However, in IM-MS experiments, analogous to the LTQ Orbitrap CID results and consistent with the presented data in Figure 2.1, pre-IM CID MS of mono-isotopically *m/z*-isolated $[y_6]^+$ (*m/z* 511) showed the presence of product fragment ions at *m/z* 465 ($[y_6 - (NH_3 + {}^{13}CO)]^+$) and 466 (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO)]^+$) (Figure 2.3a inset).

Figure 2.3b shows the IM profiles for $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466). Under our experimental conditions, the IM profile of $[y_6 - (NH_3 + {}^{13}CO)]^+$ (Figure 2.3b top) showed one major gas-phase conformer (drift time (DT) = 4.47 (± 0.05) ms) and at least two partially resolved minor conformers at shorter DTs. The presence of the two minor conformers may be due to differential product fragment ion folding and/or the existence of alternative fragmentation pathways for

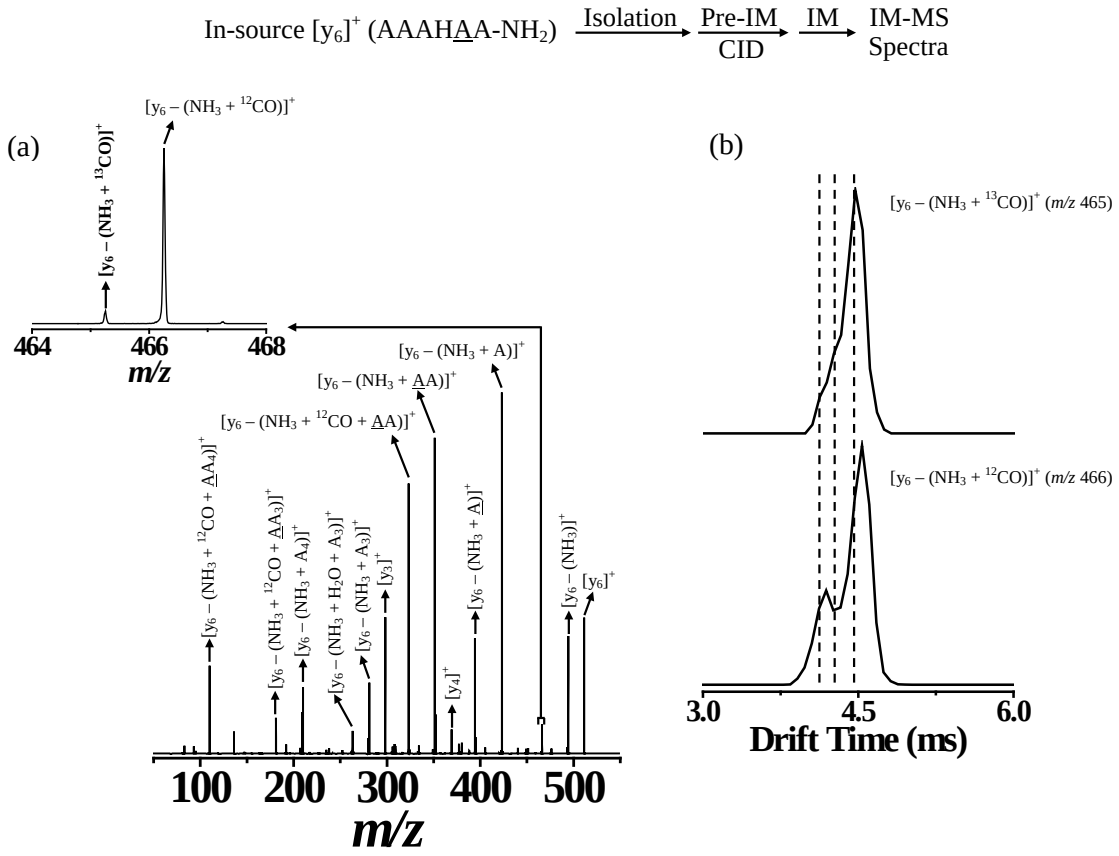


Figure 2.3. (a) Pre-ion mobility collision-induced dissociation (pre-IM CID) mass spectrum of m/z-isolated $[y_6]^+$ (*m/z* 511) (generated from in-source fragmentation of protonated AAAAHAA-NH_2). Inset shows the expanded view of the *m/z* range from 464 to 468 corresponding to $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) and scrambled product fragment ion $[y_6 - (NH_3 + {}^{13}CO)]^+$ (m/z 465). (b) Mobility profiles of $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) (top panel) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) (bottom panel) generated from pre-IM CID of *m/z*-isolated $[y_6]^+$ (*m/z* 511).

generation of $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465). Additionally, observed partially resolved conformers may be the result of multiple protonation sites for $[y_6 - (NH_3 + {}^{13}CO)]^+$, which has been shown to affect the mobility of ions in drift time and traveling wave IM experiments [142, 143]. Although we expected to see five conformers for $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) (due to potential loss of ^{12}CO from the five non-labeled backbone carbonyl groups in $[y_6]^+$), the IM profile of $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) (Figure 2.3b

bottom) showed the presence of two overlapping ion populations (DT = 4.19 (± 0.00[1]) ms and 4.54 (± 0.00) ms). Please, note that the Synapt G2-S HDMS system reports DTs to only two decimal places and all replicates yielded the same DT resulting in calculated standard deviations of ± 0.00 ms. It is possible that five (or more) conformers exist but are clustered into two groups with similar collision cross sections, which are unresolved under our experimental conditions. The two populations in the IM profile of $[y_6 – (NH_3 + {}^{12}CO)]^+$ suggest the presence of at least two structural conformers (or ion populations) with different collision cross sections. The observed IM profile (in Figure 2.3b, bottom panel) is likely the result of losses of non-labeled carbonyl groups from independent locations (*e.g.,* C-terminal carbonyl and internal carbonyl(s)). Differences in DTs of $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) are statistically significant at the 95% confidence level ($t_{experimental}$ = 6.062, and t_{table} = 4.303, calculated using the case two Student's t test for unpaired samples [144, 145]) indicating that, as a consequence of sequence-scrambling, these two ions have different gas-phase structures.

Post-IM CID of $[y_6 – (NH_3 + {}^{13}CO)]^+$ and $[y_6 – (NH_3 + {}^{12}CO)]^+$ from Protonated AAAAHAA-NH_2

Ideally, we would have liked to isolate $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) prior to IM separation to perform post-IM CID (*i.e.,* CID in the transfer cell) and search for possible conformation specific product fragment ions.[57-59, 62] However, with the existing Synapt G2-S setup, we were limited to one isolation event per scan prior to the TriWave assembly (*i.e.,* the trap, IM, and transfer cells collectively).[53, 146] Hence, isolation of $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 – (NH_3 + {}^{12}CO)]^+$ (*m/z*

[1] The Synapt G2-S HDMS system reports DTs to only two decimal places and all replicates yielded the same DT resulting in calculated standard deviations of ± 0.00 ms.

466) between the trap and IM cells was not possible. Additionally, in-source generation of $[y_6 – (NH_3 + {}^{13}CO)]^+$ and $[y_6 – (NH_3 + {}^{12}CO)]^+$ followed by isolation in the quadrupole could not be performed due to potential in-source generation of $[b_6 – ({}^{13}CO)]^+$ and $[b_6 – ({}^{12}CO)]^+$ (*e.g.*, Appendix A, Figure A.3, inset b), which have identical masses to $[y_6 – (NH_3 + {}^{13}CO)]^+$ and $[y_6 – (NH_3 + {}^{12}CO)]^+$, respectively. Therefore, we performed CID on isolated $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) in the LTQ Orbitrap mass spectrometer; these CID results are presented in the following section.

CID of m/z-Isolated $[y_6 – (NH_3 + {}^{13}CO)]^+$ and $[y_6 – (NH_3 + {}^{12}CO)]^+$ from Protonated AAAAHAA-NH_2

The CID mass spectrum of *m/z*-isolated $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) (*i.e.*, in-source CID of $[M + H]^+$ → isolation of $[y_6]^+$ (*m/z* 511) → CID of $[y_6]^+$ → isolation of $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) → CID of $[y_6 – (NH_3 + {}^{13}CO)]^+$) is shown in Figure 2.4a. Due to presence of potential rearrangements in the original sequence of $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465), we expected to observe a major portion of its CID product fragment ions to be scrambled. As expected, all product fragment ions generated from CID of isolated $[y_6 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) can be assigned as rearranged (with respect to the peptide's original amino acid sequence: AAAAHAA-NH_2) product fragment ions.

Next, we performed CID on *m/z*-isolated $[y_6 – (NH_3 + {}^{12}CO)]^+$ (Figure 2.4b) (*i.e.*, in-source CID on $[M + H]^+$ → isolation of $[y_6]^+$ (*m/z* 511) → CID on $[y_6]^+$ → isolation of $[y_6 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) → CID on $[y_6 – (NH_3 + {}^{12}CO)]^+$ to yield the mass spectrum in Figure 2.4b). The fragmentation pattern from CID of *m/z*-isolated $[y_6 – (NH_3 + {}^{12}CO)]^+$ (Figure 2.4b) is considerably different than the fragmentation pattern from CID of *m/z*-isolated $[y_6 – (NH_3 + {}^{13}CO)]^+$ (Figure 2.4a). The product fragment ion at *m/z* 378

(*i.e.*, $[y_6 - (2NH_3 + {}^{12}CO + A)]^+$; red font in Figure 2.4b) is unique to the CID of $[y_6 - (NH_3 + {}^{12}CO)]^+$. The CID mass spectrum of $[y_6 - (NH_3 + {}^{12}CO)]^+$ contains two abundant sequence-scrambled product fragment ions (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3 + \underline{A})]^+$ (*m/z* 285) and $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3)]^+$ (*m/z* 357)). Results from CID of isolated $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) confirm the IM-MS data (in Figure 3) and suggest that these product fragment ions have different gas-phase structures.

Although ^{12}CO can be lost from multiple backbone positions, postulated identities of sequence-scrambled product fragment ions generated from CID of $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3 + \underline{A})]^+$ (*m/z* 285) and $[y_6 - (NH_3 + {}^{12}CO +$

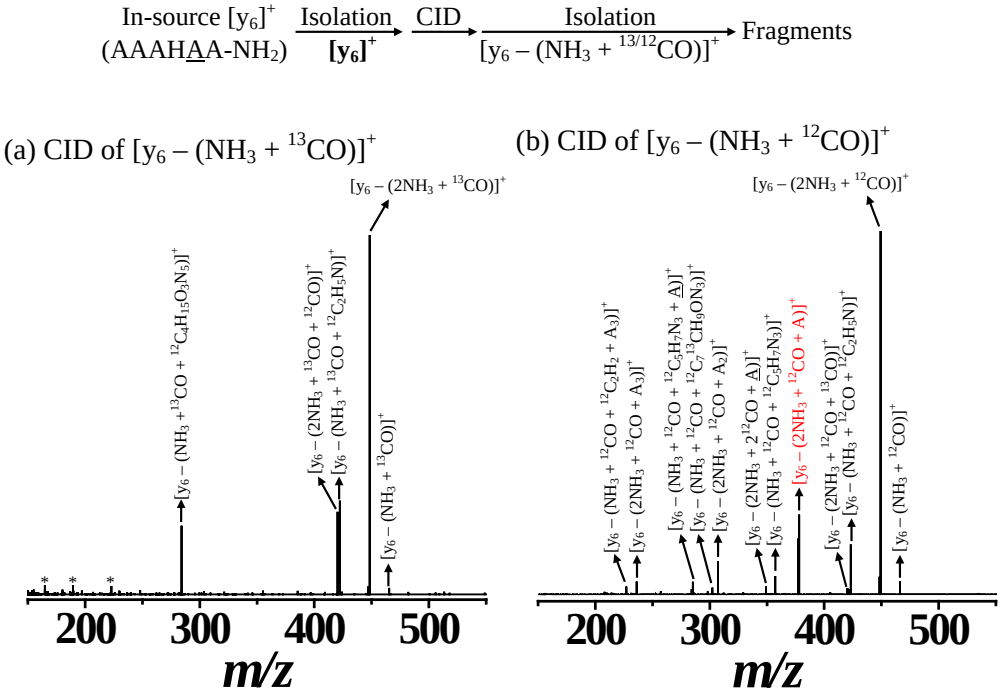
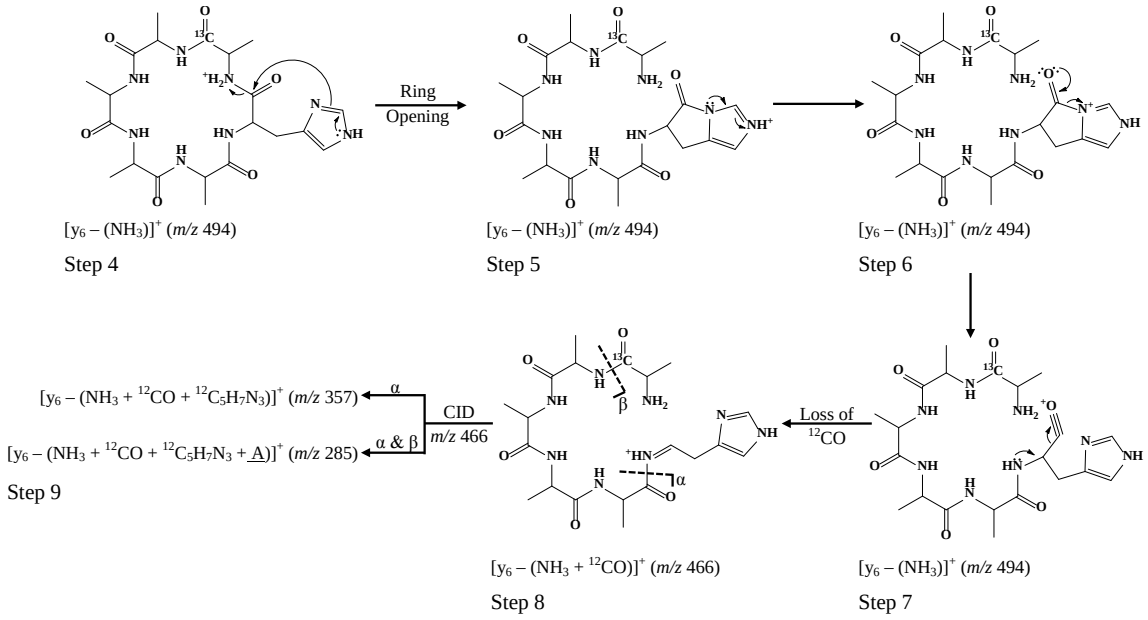


Figure 2.4. LTQ Orbitrap collision-induced dissociation (CID) mass spectra of *m/z*-isolated (a) $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and (b) $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) at 15% normalized collision energy. Low intensity peaks marked with an asterisk (*) indicate electronic noise.

$^{12}C_5H_7N_3)]^+$ (*m/z* 357)) suggest that one of the backbone carbonyls can be lost from the histidine residue. Adding the chemical compositions of ^{12}CO and $^{12}C_5H_7N_3$ yields $^{12}C_6H_7N_3O$ (*i.e.*, the chemical composition for histidine) suggesting the initial ^{12}CO loss was likely from the histidine residue. Because isotopically labeled alanine (A) does not have a ^{12}CO in its structure, we can conclude that the His carbonyl group (^{12}CO) was lost *prior* to CID of $[y_6 - (NH_3 + {}^{12}CO)]^+$ and the remainder of His (*i.e.*, $^{12}C_5H_7N_3$) was lost *after* CID of $[y_6 - (NH_3 + {}^{12}CO)]^+$. These results agree with those reported by Tsaprailis *et al.* where it was demonstrated that the His side chain could direct fragmentation at its C-terminus.[147]

Scheme 2.2 shows a possible mechanism for loss of the carbonyl group from His residue (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466)) generated from CID of $[y_6]^+$ of the



Scheme 2.2. Proposed mechanism for loss of a carbonyl (^{12}CO) group from histidine residue, that is, $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 465) in CID of $[y_6]^+$ from protonated AAAAHAA-NH_2, and generation of subsequent scrambled product fragment ions, that is, $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3 + A)]^+$ (*m/z* 285) and $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3)]^+$ (*m/z* 357).

isotopically labeled peptide AAAAH<u>A</u>A-NH_2. The first three steps (*i.e.*, loss of C-terminal NH_3, head-to-tail macrocyclization, and proton transfer) can occur following a similar pathway as proposed for the generation of $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) (Scheme 2.1, Steps 1-3) and, for brevity, are not shown in Scheme 2.2. Following macrocyclization of $[y_6 - (NH_3)]^+$ (*m/z* 494), the histidine side chain can form a bicyclic ring and direct cleavage at the His-<u>Ala</u> amide bond (Scheme 2.2, Step 4) (consistent with previous reports by Tsaprailis *et al.* [147]). The bicyclic ring can then internally rearrange (Scheme 2.2, Step 5-7) to expel the carbonyl group (similar to expulsion of CO from imidazole ring [94, 140]) and generate the product fragment ion at *m/z* 466 (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO)]^+$) (Scheme 2.2, Step 8) containing an imine derived from histidine (denoted as His′). Upon mass isolation and CID of $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466), cleavage of the Ala-His′ bond (designated as "α" in Scheme 2.2, Step 8) results in formation of the observed sequence-scrambled product fragment ion $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3)]^+$ (*m/z* 357) (Scheme 2.2, Step 9). Subsequent cleavage of the Ala-<u>Ala</u> amide bond (designated as "β" in Scheme 2.2, Step 8) generates the observed sequence-scrambled product fragment ion $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3 + \underline{A})]^+$ (*m/z* 285) (Scheme 2.2, Step 9).

Although presence of sequence-scrambled product fragment ions (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3 + \underline{A})]^+$ (*m/z* 285) and $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_5H_7N_3)]^+$ (*m/z* 357)) generated from CID of isolated $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) suggest that a carbonyl is lost from the His residue, there are four Ala residues which could also lose ^{12}CO and generate product fragment ions with the same mass as $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466). Identities of the product fragment ions in Figure 2.4b can be used to determine if ^{12}CO is

lost from Ala residues. For example, the product fragment ion at *m/z* 423 (Figure 2.4b) is generated *via* neutral losses of NH_3, ^{12}CO, and $^{12}C_2H_5N$ from $[y_6]^+$ (*i.e.*, $[y_6 - (NH_3 + {}^{12}CO + {}^{12}C_2H_5N)]^+$). In this case, addition of ^{12}CO and $^{12}C_2H_5N$ yields the chemical composition of alanine (*i.e.*, $^{12}C_3H_5NO$), suggesting that ^{12}CO is also lost from a "non-labeled" alanine(s). Due to the presence of four Ala residues in $[y_6]^+$ of AAAAHAA-NH_2 it is not possible to assign this particular ^{12}CO loss to a specific alanine. However, presence of the product fragment ions at *m/z* 236 (*i.e.*, $[y_6 - (2NH_3 + {}^{12}CO + A_3)]^+$), *m/z* 307 (*i.e.*, $[y_6 - (2NH_3 + {}^{12}CO + A_2)]^+$), and *m/z* 378 (*i.e.*, $[y_6 - (2NH_3 + {}^{12}CO + A)]^+$) suggest that ^{12}CO is likely lost from the C-terminal Ala. Sequential loss of NH_3 and ^{12}CO from the C-terminus of y-type ions is analogous to formation of general a-type ions in CID of protonated peptides/proteins. Therefore, the fragmentation mechanism(s) for generation of $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) loosing C-terminal NH_3 and ^{12}CO is likely similar to those described for formation of a-type ions.[140, 148]

Based on our MS^n results we can conclude that $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) is a rearranged product fragment ion. Additionally, CID of *m/z*-isolated $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) suggest that ^{12}CO is lost from His and C-terminal Ala residues. Carbonyl group losses from His and Ala may be representative of the two major gas-phase conformers observed in our IM-MS results of $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) (Figure 2.3b bottom). In the following section we provide evidence that loss of carbonyl groups from multiple locations is not limited to y-type ions generated from AAAAHAA-NH_3.

IM-MS and Post-IM CID Analysis of $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ from Protonated Des-Acetylated-α-Melanocyte-Stimulating Hormone

In our previous study on sequence-scrambling of y-type ions,[106] we identified several a-"like" ions generated from CID of various size and charge state y-type ions from C-terminus free and amidated peptides (results are summarized in Appendix A, Table A.1). For example, CID of *m/z*-isolated y-type ions of different lengths and charge states from dα-MSH (SYSMEHFRWGKPV-NH_2) and angiotensis II antipeptide (EGVYVHPV) generated a-"like" ions, similar to $[y_6 - (NH_3 + {}^{12/13}CO)]^+$ of AAAAHAA-NH_2. These results suggest that backbone CO losses from y-type ions can occur in peptides of different sizes and amino acid compositions. To further investigate carbonyl losses from internal backbone positions of y-type ions, we performed IM-MS analysis on a pentapeptide (SDGRG), angiotensin II antipeptide (EGVYVHPV), angiotensin II (DRVYIHPF), a pentapeptide (SDGRG), and dα-MSH (SYSMEHFRWGKPV-NH_2).

No $[y_n - (H_2O + {}^{12}CO)]^+$ fragment ions were generated from the SDGRG pentapeptide (data not shown), suggesting that a-"like" fragments (and subsequent sequence-scrambling) are not formed in CID of all peptides. Based on the evidence from CID of $[y_4]^+$ ions from AAAAHAA-NH_2 (Appendix A, Figure A.2), it is likely that generation of sequence-scrambled a-"like" ions is not limited to larger size peptides. Alternatively, generation of a-"like" ions may be mitigated in CID of the SDGRG pentapeptide due to residue directed (i.e., non-random) fragmentation pathways.[106, 119, 128, 147, 149, 150] For example, peptides with C-terminal arginine residues (similar to SDGRG pentapeptide) have been shown to follow different fragmentation pathways than related

peptides without arginine due to proton immobilization by the basic arginine side chain.[151, 152]

IM-MS results from angiotensin II antipeptide and angiotensin II indicated multiple conformers for $[y_n - (H_2O + {}^{12}CO)]^+$ product fragment ions (Appendix A, Figures A.4 and A.5, respectively), similar to $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) from AAAAHAA-NH_2 (Figure 2.3b, bottom panel). Intriguingly, in-source CID of dα-MSH generated $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) with a single Gaussian-like IM arrival time distribution (Figure 2.5a). To investigate the potential for multiple IM unresolved conformers for $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) generated from dα-MSH, we performed post-IM CID and present our findings in this section.

Figure 2.5a shows the post-IM CID (Synapt G2-S HDMS) arrival time distribution of *m/z*-isolated $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) generated from in-source CID of dα-MSH (amino acid sequence: SYSMEHFRWGKPV-NH_2). There are eleven backbone (and two side chain (*i.e.,* glutamic acid)) carbonyl groups in the structure of $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) and we expected to observe the presence of several conformational/structural isomers at different arrival times. However, the IM profile of $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) showed a single Gaussian-like distribution (Figure 2.5a). The observation of single Gaussian-like distribution for $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) may suggest: (i) ^{12}CO is lost from a single position yielding one major conformation/structure or (ii) ^{12}CO is lost from multiple locations yielding IM unresolved structures/conformers. To determine if the ion population at *m/z* 664 (*i.e.,* $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$) represented multiple structures/conformers, we applied a previously reported

chemometric deconvolution approach [62, 63] and assessed post-IM CID mass spectra across the IM profile shown in Figure 2.5a.

Results from IM profile deconvolution of post-IM CID of *m/z*-isolated $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) are shown in Figure 2.5b, bottom panel. Post-IM CID deconvolution results (Figure 2.5b) suggest that ^{12}CO is lost from multiple locations (in $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ of dα-MSH) yielding IM overlapping structures/conformers. Although deconvolution of post-IM CID data revealed two distinct structural components for $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664), existence of additional isomers should not be disregarded. For example, the deconvoluted IM profile of the early arrival time conformer (Figure 2.5b, solid line) has a small shoulder at a longer arrival time (AT ~ 6.8 ms), suggesting the presence of an additional conformer(s) with similar fragmentation pattern(s).

Insets in Figure 2.5b show the deconvoluted CID mass spectra corresponding to the two deconvoluted IM profiles for $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ of dα-MSH (Figure 2.5b, bottom panel). Deconvoluted mass spectrum for the early arrival time conformer (Figure 2.5b, solid line) contained sequence-scrambled product fragment ions (denoted in red font). For example, ions at *m/z* 434 ($[y_8 - (2NH_3 + {}^{12}CO + P)]^{2+}$), *m/z* 607 ($[y_{11} - (2NH_3 + {}^{12}CO + P)]^{2+}$), *m/z* 729 ($[y_7 - (2NH_3 + {}^{12}CO + P)]^{+}$), *m/z* 866 ($[y_8 - (2NH_3 + {}^{12}CO + P)]^{+}$), and *m/z* 1214 ($[y_{11} - (2NH_3 + {}^{12}CO + P)]^{+}$) are all assigned as sequence-scrambled product fragment ions from the original $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ or SMEHFRWGKPV-NH_2 sequence (Figure 2.5b, top left panel). Identities of sequence-scrambled product fragment ions in Figure 2.5b, top left panel suggest that the proline residue is directing the fragmentation process. Site directed fragmentation in the presence of proline, also

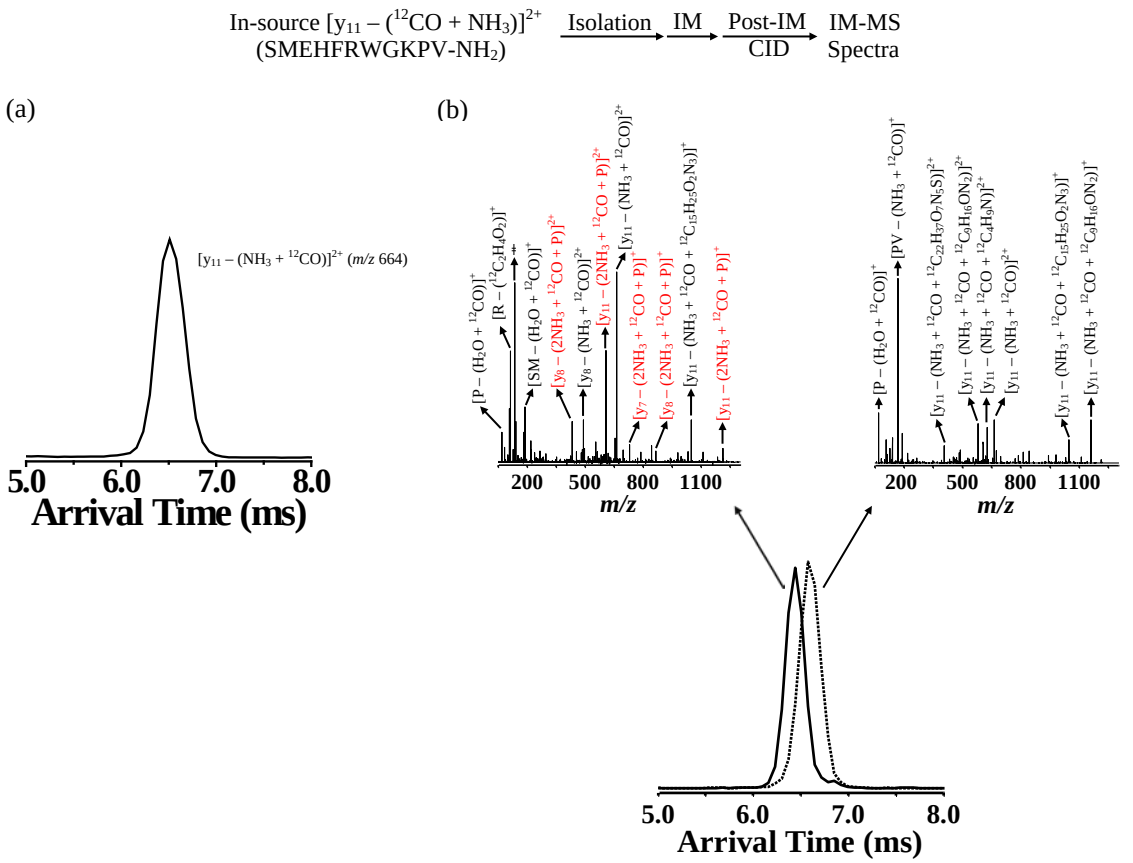


Figure 2.5. (a) Post-IM CID mobility profile of $[y_{11} – (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) (generated from in-source CID of des-acetylated-α-melanocyte-stimulating hormone). (b) Deconvoluted IM (bottom panel) and mass (top panels) spectra generated *via* chemometric analysis of post-IM CID data (see main text for details). The peak labeled with a double cross bar (ǂ) at experimental *m/z* 138.1334 (in mass spectrum of the shorter arrival time isomer shown in left panel) may correspond to an unusual fragment ion with formula $[{}^{12}C_3H_{16}N_5O]^+$ (theoretical *m/z* 138.1349); however, due to large MME (11 ppm) we did not label this peak in the figure.

called the "proline effect", has been well studied and may play a role in sequence-scrambling of proline containing peptides.[106, 149, 150] Likewise, the proline effect may explain why we only observed two major conformers for the deconvoluted post-IM CID data for $[y_{11} – (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) (Figure 2.5b). In other words, if proline is involved in directed fragmentation of dα-MSH then ^{12}CO groups at some backbone positions should have a higher probability of being lost than at other backbone positions.

Residue directed fragmentation mechanism may result in fewer numbers of gas-phase conformers (*e.g.*, eleven sites for backbone ^{12}CO loss) for $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) than initially expected.

The post-IM CID mass spectrum of the longer arrival time conformer for $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) (Figure 2.5b, top right panel) showed product fragment ions corresponding to normal losses (*i.e.*, sequence-intact product fragment ions) as major fragment ions. Postulated product fragment ion identities in Figure 2.5b (top right panel), suggest that a backbone carbonyl can be lost from the C-terminal valine residue. For example, the doubly charged product fragment ion at *m/z* 629 (Figure 2.5b, top right panel) is generated *via* neutral losses of NH_3, ^{12}CO, and $^{12}C_4H_{10}N$ from $[y_{11}]^{2+}$ (*i.e.*, $[y_{11} - (NH_3 + {}^{12}CO + {}^{12}C_4H_9N)]^{2+}$). Similar to our discussion of losses from the isotopically labeled peptide (*e.g.*, Figure 2.4), when we add the chemical composition of ^{12}CO and $^{12}C_4H_9N$ we arrive at the chemical composition of valine (*i.e.*, $^{12}C_5H_9ON$). Because there is only one valine (at the C-terminus) in the amino acid sequence of dα-MSH, we can conclude that one backbone ^{12}CO must be lost from the C-terminal position to yield a sequence-intact a-"like" product fragment ion, similar to generation of normal a-type ions.[140, 148]

Chemometric deconvolution results from post-IM CID of $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) from dα-MSH (Figure 2.5) indicate that IM resolution of currently available commercial instruments may be insufficient for resolving some isomer populations. Based on IM results for $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ from dα-MSH (Figure 2.5) and complementary results from IM analysis of b-type ions,[63] we suspect that IM analysis of a-"like" ions from other peptides systems will similarly show unresolved structures,

which may be misinterpreted as having a single, pure, and sequence-intact isomer. As previously reported,[119] and is clear from the IM profiles in Figures 2.3, 2.5, and Appendix A, Figures A.4 and A.5, presence of multiple IM unresolved isomers results in broadened AT distributions. Therefore, peak width analysis,[119] along with post-IM CID,[62, 63] can be used as a rapid screen for IM unresolved structural isomers and/or sequence-scrambling.

Results from post-IM CID of $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) from dα-MSH (Figure 2.5) are in line with our observations from isotopically labeled heptapeptide, AAAAH<u>A</u>A-NH_2 (Figures 2.1 to 2.4), and suggest that backbone carbonyl groups can be lost from multiple backbone positions. Furthermore, CID results from both peptides suggest that loss of internal backbone carbonyl groups is an important step in sequence-scrambling of y-type ions.

Conclusions

Presented results are in agreement with previous findings on the possibility of generating sequence-scrambled peptide product fragments in CID of *m/z*-isolated a-, b-, and y-type ions [106, 120, 153]. The loss of internal carbonyl groups in CID of y-type ions from the isotopically labeled heptapeptide AAAAH<u>A</u>A-NH_2 provides an additional evidence for sequence-scrambling in CID of y-type ions. Sequence-scrambled $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) product fragment ions were observed in CID of $[y_6]^+$ (generated from CID of protonated AAAAH<u>A</u>A-NH_2). IM-MS results suggest the existence of different gas-phase conformers for $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466). Furthermore, MS^n CID results on *m/z*-isolated $[y_6 - (NH_3 + {}^{13}CO)]^+$ (*m/z* 465) and $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) reveal unique fragmentation patterns for each product fragment ion. Moreover, product fragment ions

generated from CID of *m/z*-isolated $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) indicate loss of carbonyl groups from different amino acid residues (*i.e.*, His and non-labeled Ala residues). Independent carbonyl group losses from His and Ala may be representative of the two major gas-phase conformers of $[y_6 - (NH_3 + {}^{12}CO)]^+$ (*m/z* 466) observed in our IM-MS results (Figure 2.3b bottom). Likewise, post-IM CID results for *m/z*-isolated $[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$ (*m/z* 664) from dα-MSH suggest that carbonyl groups are lost from both C-terminal (yielding direct product fragment ions) and internal (yielding sequence-scrambled product fragment ion) positions. Additionally, IM-MS analysis of select y-type ions generated from dα-MSH and angiotensin II antipeptide suggest that loss of internal carbonyl groups, and consequently sequence-scrambling, may be a common feature in CID of y-type ions.

MS^n and IM-MS results suggest that loss of internal backbone CO may be a crucial step in formation of sequence-scrambled product fragment ions from CID of *m/z*-isolated y-type ions. For example, sequence-scrambled product fragment ions may be produced *via* following pathways: Step 1. Formation of y-type ions → Step 2. CID to form b-"like" ions → Step 3. Formation of cyclic b-"like" ions → *Step 4. Loss of a neutral CO (from available positions)* → Step 5. Loss of different internal amino acids or fragments (*e.g.*, see Scheme 2.2). Although steps 1-3 and 5 have been previously reported,[106] the proposed intermediate step 4 (*i.e.*, loss of internal backbone carbonyl group(s)) has not been reported before. Comprehensive characterization of CO losses and their inclusion as possible steps in structural rearrangements may abate existing knowledge-gaps in mechanistic understanding of peptide rearrangement pathways.

Using isotope labeling, ion mobility-mass spectrometry (IM-MS), and tandem mass spectrometry (MS^n) we have demonstrated that y-type ions can rearrange to lose internal backbone carbonyl(s) and yield ions that can serve as intermediate structures in the sequence-scrambling pathway. Further experiments must be performed to assess effects of product fragment ion size, charge, and amino acid composition on the extent of expulsion of internal backbone carbonyls of y-type ions.

Acknowledgements

The authors would like to thank Baylor University for financial support. The authors also thank Dr. Behrooz Zekavat for helpful discussions of IM-MS and MS^n data reported in this manuscript.

CHAPTER THREE

DNA Oligonucleotide Fragment Ion Rearrangements Upon Collision-induced Dissociation

This chapter published as: Harper, B.; Neumann, E. K.; Solouki, T., DNA Oligonucleotide Fragment Ion Rearrangements Upon Collision-induced Dissociation. *J. Am. Soc. Mass Spectrom.* **2015,** *26* (8), 1404-1413.

Abstract

Collision-induced dissociation (CID) of *m/z*-isolated w type fragment ions and an intact 5′ phosphorylated DNA oligonucleotide generated rearranged product ions. Of the 21 studied w ions of various nucleotide sequences, fragment ion sizes, and charge states, 18 (~86%) generated rearranged product ions upon CID in a Synapt G2-S HDMS (Waters Corporation, Manchester, England, UK) ion mobility-mass spectrometer. Mass spectrometry (MS), ion mobility spectrometry (IMS), and theoretical modeling data suggest that purine bases can attack the free 5′ phosphate group in w type ions and 5′ phosphorylated DNA to generate sequence permuted $[phosphopurine]^-$ fragment ions. We propose and discuss a potential mechanism for generation of rearranged $[phosphopurine]^-$ and complementary y-B type product ions.

Introduction

The emergence of commercially available electrospray ionization (ESI) sources and high resolving power mass spectrometers has allowed for routine sequencing of oligonucleotides *via* mass spectrometry (MS) based techniques.[108] Similar to MS sequencing of peptides and proteins,[71] to sequence oligonucleotides, precursor molecular

ions are generally first isolated and then fragmented using gas-phase dissociation techniques (*e.g.*, CID,[114, 154] infrared multiphoton dissociation (IRMPD),[155, 156] electron capture dissociation (ECD),[155] *etc.*). The resulting product fragment ion mass spectra are then analyzed and interpreted to determine the original nucleotide sequences.[109-113]

A potential drawback of oligonucleotide sequencing by tandem MS (MS^n) techniques is the data analysis challenges associated with generation of complex mass spectra composed of fragment ions from the 3′ and 5′ ends, internal fragments (generated from sequential fragmentation), and base losses.[109] Although CID of oligonucleotides predominately generates a-B and w type fragment ions (McLuckey's nomenclature [114]), generation of unidentifiable product ions and alternative fragmentation pathways (in CID and other fragmentation techniques) have also been reported.[157-160] For example, Juhasz *et al.* reported generation of six unusual/unidentifiable high intensity product ion peaks from the DNA oligonucleotide, CAC ACG CCA GT;[157] these authors purported that presence of unidentifiable fragment ions may make MS based *de novo* sequencing of some oligonucleotides impractical.[157] Likewise, Nyakas *et al.* recently described an unusual fragmentation pathway involving loss of internal $[PO_3]^-$ through formation of a cyclic intermediate in CID of highly charged DNA oligonucleotides.[160]

Because of their potential impact on "top-down" proteomics [67, 68] and high throughput biomolecule sequencing, amino acid "sequence-scrambling" and ion rearrangements in the gas-phase have drawn much attention.[89, 106, 121] Sequence-scramblings in peptides and proteins are believed to occur through macrocyclization and subsequent reopening at internal amide bonds to generate product ions with rearranged amino acid sequences.[101] Though the impact of rearrangements on sequence coverage of

peptides in amino acid sequencing is uncertain,[105, 161] positive identification of permuted fragment ions and mechanistic understandings of their generation could help improve sequencing algorithms. Despite the number of investigations focused on rearrangement pathways of peptide ions, there has been a paucity of examples addressing unusual fragmentation of oligonucleotide ions.

We hypothesized that, similar to CID of peptide fragment ions,[89, 101, 106, 121] CID of DNA fragment ions could also result in generation of secondary fragment ions with rearranged atomic structures (relative to precursor ion structures). For this study, we chose to focus on CID of w type fragment ions for two reasons: (a) w type ions (along with a-B type ions) are the most commonly observed fragment ions in CID of DNA ions,[114] and are thus often used for sequencing,[109-113] and (b) w type ions retain all of the bases in the precursor ion series (unlike a-B type ions which lack nucleobases at various positions [114]).

In this chapter, we report the first experimental evidence for generation of rearranged fragment ions during CID of isolated w type oligonucleotide ions and 5′ phosphorylated DNA. We performed CID on twenty-one $[w_n]^{z-}$ (where n = 1-8 and z = 1-3) fragment ions from three different DNA oligonucleotides. Of the twenty-one studied w type fragments, ~86 % showed evidence for subsequence rearrangement upon CID. We also compared our results to CID of a 5′ phosphorylated DNA oligonucleotide which showed similar results.

Materials and Methods

Sample Preparation

Imidazole and angiotensin II antipeptide (amino acid sequence: EGVYVHPV; used as a lock mass) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Optima grade methanol and water were purchased from Fisher Scientific (Pittsburgh, PA, USA). All DNA oligonucleotides (Oligo-1 [ATG CGT CCG GCG TAG A], Oligo-2 [p-GTA GA; where "p" denotes a 5′ phosphorylation], Oligo-3 [TCG AGG TCG ACG GTA TC], and Oligo-4 [AGA GTT TGA TCC TGG CTC AG]) were purchased from Integrated DNA Technologies (Coralville, IA, USA). Oligo-1, 3, and 4 were chosen based on their small sizes (*i.e.*, $\leq$ 20 nucleotides), presence of all four normally occurring DNA bases (*i.e.*, A, G, T, and C) in their sequences, and occurrence of different, non-repeating nucleotide sequences. Oligo-2 was custom synthesized to be homologous with the sequence of the $[w_5]^{z-}$ (where "z" denotes the charge state(s)) fragment ions from Oligo-1. Micromolar concentrations of oligonucleotides were prepared in a spray solvent containing 1.0 mM imidazole in a 1:1 mixture of water:methanol. All chemicals and analytes were used "as is" and without any additional purification or modification.

Instrumentation

Ion mobility-mass spectrometry. IM-MS experiments were performed using a Synapt G2-S HDMS ion mobility-mass spectrometer (IM-MS) (Waters Corporation, Manchester, England, UK) operated in negative-ion mode ESI. The time-of-flight (TOF) mass analyzer was operated either in high resolution mode (*i.e.*, "W" mode for IM experiments in Figure 3 and to distinguish between potential isobaric species) or

sensitivity mode (*i.e.*, "V" mode for all other experiments to detect low abundance ions). The experimental scan time for data acquisition was set at 1 second.

Sample flow rate was set at 5 μL/min (Figure 3.3) or 1.5 μL/min (Figures 3.1 and 3.2, and Tables 3.1 and 3.2) using a Harvard Pump 11 Plus dual syringe pump (Harvard Apparatus, Holliston, MA, USA). Source temperature was set at 100 °C. The source voltage was set at -3.0 kV for all IM-MS experiments. The software defined sampling cone voltage was set at -150.0 V (Figures 3.1a and 3.2, and Tables 3.1 and 3.2), -35.0 V (Figure 3.1b), or -120.0 V (Figure 3.3). The source offset voltage was set at -35.0 V (Figure 3.1b) or -150.0 V (all other Figures/Tables). Please note that source conditions were "softer" for Figure 1b to prevent in-source fragmentation. Ion isolation parameters (*i.e.*, quadrupole rf/dc) were optimized to isolate the isotopic distribution of the ions of interest.

Post-IM/CID [162] was performed by setting a 40.0 to 50.0 V (Figure 3.1 and Tables 3.1 and 3.2) or 20.0 V (Figure 3.2) potential difference between the IM and transfer cells. Pre-IM/CID (Figure 3.3) was performed by setting a 30.0 V potential difference between the trap and helium cells. IM cell wave height and velocity were set at 40.0 V and 750 m/s, respectively, and nitrogen gas (N_2) pressure in the IM cell was ~3.5×10^{-3} bar (N_2 flow rate of 90.0 mL/min). Helium gas (used for collisional cooling prior to IM [53]) pressure in the helium cell was ~1.4 bar (helium gas flow rate of 180 mL/min). Argon gas pressures in the trap and transfer cells were both kept at ~2.4×10^{-5} bar (argon gas flow rate of 2.0 mL/min). All gas pressures are reported as direct instrument readouts and have not been corrected for geometry [64, 163] or sensitivity factors.[164]

Ion trap tandem mass spectrometry. A Linear Trap Quadrupole (LTQ) Orbitrap MS (Thermo Fisher Scientific, Waltham, MA, USA) operated in negative-ion mode ESI was used to collect the MS^n data in Figure 3.4 and Appendix B, Figure B.2. Mass resolving power ($M/\Delta M_{50\%}$) of the Orbitrap MS was set to the highest resolving power for the instrument (*i.e.*, ~30,000 at *m/*z 400). Sample flow rate was set to 20 μl/min. The sheath and auxiliary gas (N_2) flow rates were set to 60 AU and 5 AU, respectively. The source, capillary, and tube lens voltages were optimized to -5.0 kV, -42.0 V, and -133.7 V, respectively. The metal capillary temperature was set to 275 °C. The entire isotopic distribution of the ions of interest were isolated and fragmented in the LTQ at normalized collision energies [134] of 30% (Figure 3.4) or 20% (Appendix B, Figure B.2). Wideband excitation [134] was turned off to minimize re-activation of fragment ions. Helium gas was used as the collision gas for CID (partial pressure of ~1.3 x 10^{-8} bar). Due to low abundance of y-B type ions from Oligo-2, sufficiently high sample concentration (100 μM) was used to collect the data displayed in Figure 3.4. Singly and doubly deprotonated angiotensin II antipeptide ions (*m/z* 897 and *m/z* 448, respectively) were used as lock masses for all LTQ Orbitrap experiments.

Data Analysis

IM-MS data were analyzed using MassLynx software (version 4.1) (Waters Corporation, Manchester, England, UK). To obtain signal-to-noise ratios of >3 for all ions of interest, acquired IM-MS spectra were averaged for 5 minutes (Figure 3.1 and Tables 3.1 and 3.2), 10 minutes (Figure 3.2), or 20 minutes (Figure 3.3). IM profiles displayed in Figure 3.3 correspond to the centroid *m/z* ± 0.02 Th, extracted from the associated mass spectra.

LTQ Orbitrap MS data were analyzed using Xcalibur (version 2.2) (Thermo Fisher Scientific, Waltham, MA, USA). The data in Figures 3.4 and Appendix B, Figure B.2 correspond to the average of 150 scans and 50 scans, respectively, which provided signal-to-noise ratios of >3 for all ions of interest. Mass measurement error (MME) for all ions identified using the LTQ Orbitrap were ≤ 2 ppm.

Fragment ions corresponding to "typical" losses were identified using Mongo Oligo Mass Calculator (version 2.06).[165] All fragment ion assignments were based on McLuckey's nomenclature.[114] Unless otherwise indicated (*viz.*, figure caption for Figure 3.1a), fragment ion identities were assigned with respect to the original oligonucleotide's sequence.

Theoretical Calculations

Theoretical calculations were performed using density functional theory (DFT) at B3LYP/6-31G(d) level of theory and basis set [166] with Gaussian 09 suite of programs (Gaussian Inc., Wallingford, CT, USA).[167] Thermochemical data are reported as thermally corrected changes in Gibbs free energies at 298 K (ΔG_{298}) and are expressed in kcal/mol.

Results and Discussion

Fragment Ion Rearrangements in CID of w Type and Intact 5′ Phosphorylated Oligonucleotide Ions

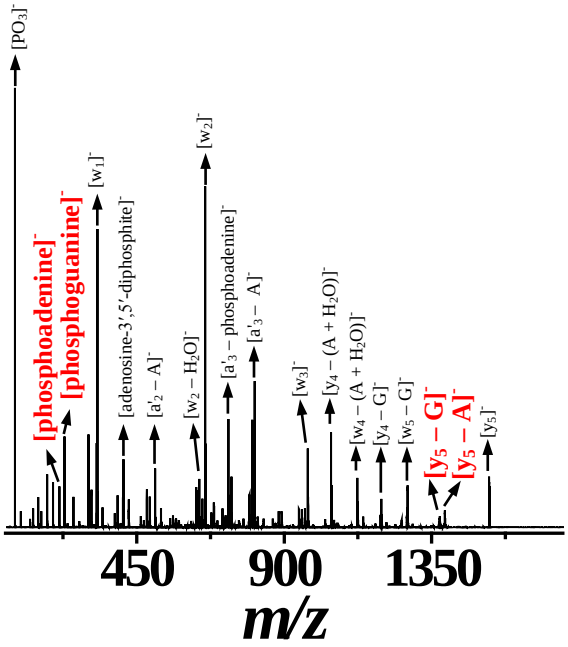
Figure 3.1a shows the MS^n CID mass spectrum of *m/z*-isolated $[w_5]^{2-}$ (*m/z* 802) generated from in-source CID of Oligo-1 in a Synapt G2-S HDMS system. Under our experimental conditions, CID of $[w_5]^{2-}$ generated two unusual product ions at *m/z* 214 and

m/z 230 (denoted in red font, Figure 3.1a) which could not be assigned to a known fragment ion type using conventional nomenclature or fragmentation mechanisms. Calculated low MMEs for the assigned peaks suggested that chemical compositions of these observed ions at *m/z* 214 and *m/z* 230 might correspond to metaphosphate ($[PO_3]^-$) ion [160] plus adenine (*i.e.*, [phosphoadenine]$^-$; ~5.84 ppm) and $[PO_3]^-$ plus guanine (*i.e.*, [phosphoguanine]$^-$; ~5.96 ppm), respectively. Fragment ion assignments of [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230) are further supported by the presence of complementary fragment ions $[y_5 - A]^-$ (*m/z* 1390) and $[y_5 - G]^-$ (*m/z* 1374) in the CID mass spectrum of $[w_5]^{2-}$ (denoted in red font, Figure 3.1a).

Numerous investigations of peptide ion scrambling indicate that fragment ions (*e.g.*, b-type ions; Roepstorff's nomenclature [168]) generated during CID of peptide ions are themselves rearranged, and that the evidence of such rearrangements (*i.e.*, "scrambled" fragment ions) can be observed upon subsequent fragmentation.[63, 64, 169-173] If a similar "scrambling" trend holds true for generation of [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230), then the precursor w type ion (*i.e.*, $[w_5]^{2-}$) may also be structurally rearranged. To test this hypothesis, and to determine if [phosphopurines]$^-$ (*i.e.*, [phosphoadenine]$^-$ and [phosphoguanine]$^-$) could be produced from CID of intact 5′ phosphorylated DNA, we repeated an analogous experiment to that presented in Figure 3.1a on a 5′ phosphorylated oligonucleotide (*i.e.*, Oligo-2) with the same sequence as the last five nucleotides of Oligo-1 (*i.e.*, p-GTA GA). Figure 3.1b shows the CID mass spectrum of *m/z*-isolated doubly-charged Oligo-2 (*i.e.*, [Oligo-2 – 2H]2- at *m/z* 802).

The CID mass spectrum of $[\text{Oligo-2} - 2H]^{2-}$ (Figure 3.1b) is essentially identical to the CID mass spectrum of *m/z*-isolated $[w_5]^{2-}$ fragment ion (*m/z* 802) generated from

(a) **ESI Oligo-1 → In-source CID → Isolate $[w_5]^{2-}$ → Post-IM/CID → Mass Spectrum**



(b) **ESI Oligo-2 → Isolate $[Oligo\text{-}2 - 2H]^{2-}$ → Post-IM/CID → Mass Spectrum**

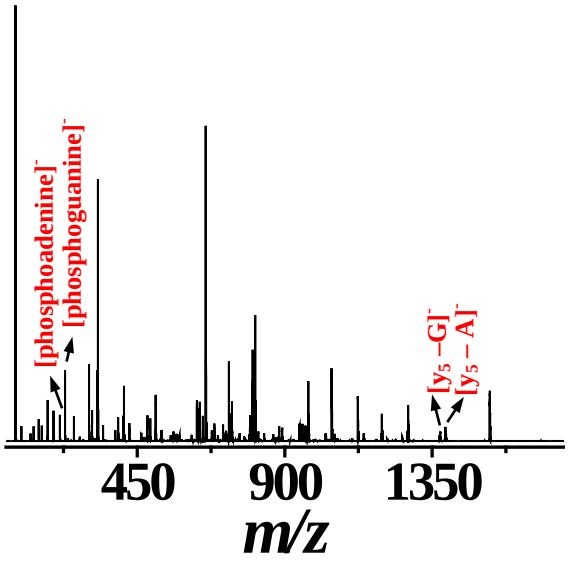


Figure 3.1. CID mass spectra of *m/z*-isolated (a) $[w_5]^{2-}$ (*m/z* 802) generated from in-source CID of Oligo-1 and (b) doubly-charged intact Oligo-2 ($[Oligo\text{-}2 - 2H]^{2-}$ at *m/z* 802). Sequence for both ions is p-GTA GA. Rearranged $[phosphopurine]^-$ and complementary y-B type ions (see discussion in main text) are denoted in red font. Peaks labeled with a prime sign (*e.g.*, $[a'_3 - A]^-$ at *m/z* 810) denote a-B type fragment ions which were generated from a precursor w type ion.

in-source CID of Oligo-1 (Figure 3.1a). It should be noted that both [phosphopurines]$^-$ and complementary y-B fragment ions (denoted in red font, Figures 3.1a and 3.1b) were present in both mass spectra, and with comparable fragment ion relative abundances. Although we cannot eliminate the possibility that Oligo-2 cyclizes/rearranges prior to fragmentation (*e.g.*, during ionization), results in Figure 3.1b show that the observed [phosphopurines]$^-$ can be generated from the originally linear structures. Therefore, it is likely that $[w_5]^{2-}$ (Figure 3.1a) is also a linear structure (*i.e.*, not cyclized). This is especially important for sequencing efforts as rearranged w type ions could undermine the validity of MS-based DNA sequencing. Furthermore, results from Figure 1b indicate that, in addition to w type ions, 5′ phosphorylated oligonucleotides can also generate [phosphopurines]$^-$.

To confirm the presumed identities of [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230), MS^n experiments (Figures 3.2a and 3.2b, respectively) were carried out. As expected, CID of [phosphoadenine]$^-$ species generated a high intensity peak at *m/z* 79 (corresponding to $[PO_3]^-$) and a low intensity peak for [adenine]$^-$ (at *m/z* value of 134) (Figure 3.2a). Similarly, CID of [phosphoguanine]$^-$ (*m/z* 230) also generated a high intensity peak for $[PO_3]^-$ (at *m/z* 79) and a low intensity peak for [guanine]$^-$ (at *m/z* value of 150) (Figure 3.2b). Data from the MS/MS mass spectra show that the relative abundance of [adenine]$^-$ with respect to $[PO_3]^-$ (*i.e.*, [adenine]$^-$/$[PO_3]^-$ in Figure 3.2a) is higher than that observed for [guanine]$^-$ (*i.e.*, [guanine]$^-$/$[PO_3]^-$ in Figure3.2b). These results are in agreement with findings by Pan *et al.* who demonstrated that [adenine]$^-$ is more stable than [guanine]$^-$ and therefore more likely to be generated during the CID of deprotonated DNA oligonucleotides.[174]

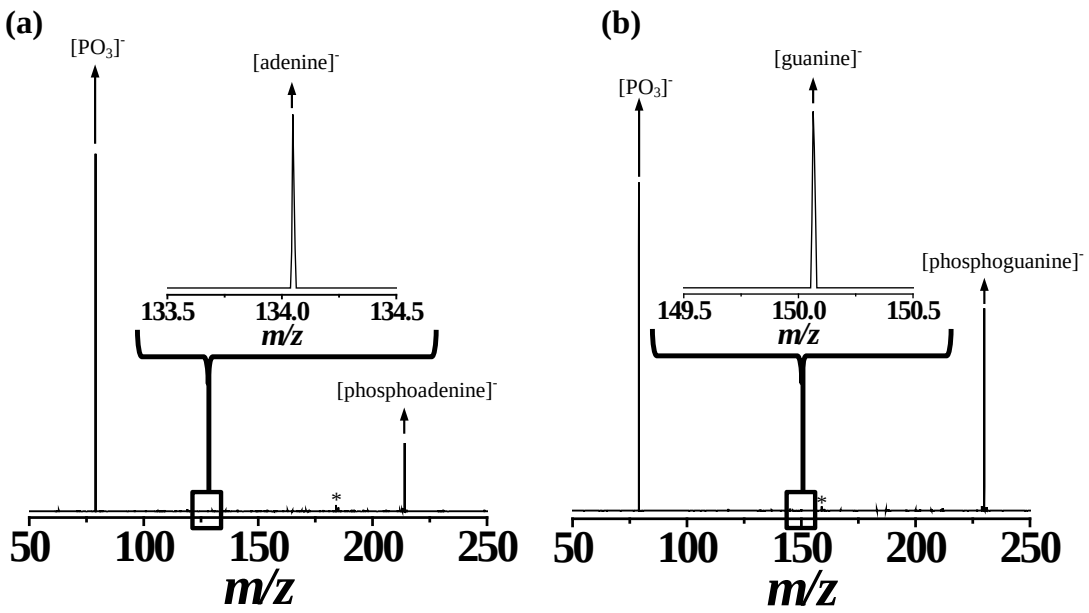


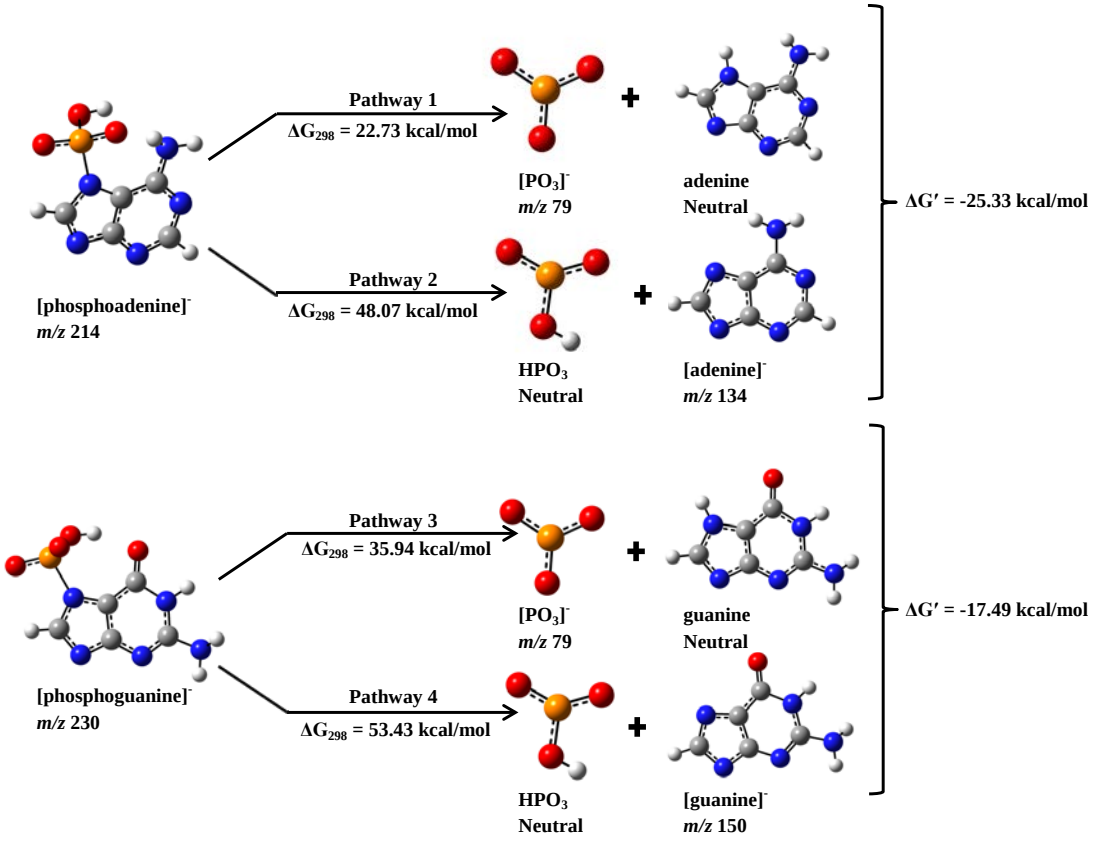
Figure 3.2. Post-IM/CID mass spectra of (a) [phosphoadenine]$^-$ (*m/z* 214) and (b) [phosphoguanine]$^-$ (*m/z* 230). Insets are shown at 150x and 250x expansion, respectively. Peaks marked with an asterisk (*) denote background noise.

Experimental results in Figure 3.2 are further supported by predictions from theoretical calculations (Scheme 3.1) which suggest that formation of $[PO_3]^-$ (*m/z* 79) and a corresponding neutral base (*i.e.*, adenine or guanine; Scheme 3.1, pathways 1 and 3, respectively) are lower energy dissociation pathways for [phospho*purines*]$^-$ than for the formation of neutral HPO_3 and purine ions (*i.e.*, [adenine]$^-$ (*m/z* 134) or [guanine]$^-$ (*m/z* 150); Scheme 1, pathways 2 and 4, respectively). For example, the ΔG_{298} for generation of neutral HPO_3 and [adenine]$^-$ from [phosphoadenine]$^-$ (Scheme 3.1, pathway 2) is 48.07 kcal/mol, whereas, the ΔG_{298} for generation of $[PO_3]^-$ and neutral adenine (Scheme 3.1, pathway 1) is 22.73 kcal/mol (*i.e.*, $\Delta G' = -25.33$ kcal/mol, where "$\Delta G'$" denotes the difference in ΔG_{298} between the two pathways). Similar results were observed for [phosphoguanine]$^-$ pathways (Scheme 3.1, pathways 3 and 4; $\Delta G' = -17.49$ kcal/mol), suggesting that generation of $[PO_3]^-$ is a lower energy pathway than formation of

[guanine]$^-$. Although the general trends between theoretical (Scheme 3.1) and experimental (Figure 3.2) evidences are in agreement, an in-depth study of reaction kinetics (*e.g.*, activation barrier(s) to formation and depletion of reactants/products, charge affinities, stability of transition species, *etc.*) would be necessary to draw concrete conclusions. Although useful, these types of calculations are outside the scope of this dissertation, but may be worth exploring in future studies.

The DNA oligonucleotides used in this study do not contain any non-natural bases or modified groups. Therefore, presence of peaks corresponding to [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230), in Figures 3.1 and 3.2, indicates the existence of reaction channels involving intramolecular rearrangements. In other words, based on the high mass measurement accuracy (or MME < 6.00 ppm), observed peaks at *m/z* 214.0137 and 230.0060 (compared to theoretical *m/z* values of 214.0125 and 230.0074, respectively) can be assigned to negatively charged [phosphoadenine]$^-$ and [phosphoguanine]$^-$ (*i.e.*, exclusively as rearranged fragment ions). Presence of such rearranged phosphopurine fragment ions may be more prevalent than previously believed. For example, Schulten *et al.*, observed analogous "phosphourus acid + nucleobase" ions in a positive-ion mode pyrolysis field desorption (Py-FD)-MS study.[175] In their discussion, Schulten *et al.* confer that they could not determine how "phosphourus acid + nucleobase" ions were generated, but assumed it was due to pyrolysis of starting materials.[175] In the present study, however, negative ions were generated *via* CID of ESI generated species; hence, no pyrolysis products were formed.

Similar to CID of $[w_5]^{2-}$ (Figure 3.1), CID of $[w_n]^{z-}$ (where n = 2-4, 6, and 7; z = 1 or 2) also generated [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230)



Scheme 3.1. Theoretical pathways for dissociation of $[phosphoadenine]^-$ (pathways 1 and 2) and $[phosphoguanine]^-$ (pathways 3 and 4).

isotopically labeled peptide AAAAHAA-NH_2. The first three steps (*i.e.*, loss of C-species (Table 3.1). Of the studied w type ions from Oligo-1, only $[w_1]^-$ (*m/z* 330) did not generate a [phospho*purine*]$^-$ (Table 3.1). Although structural rigidity of the smaller $[w_1]^-$ (as compared to large w type ions) may restrict rearrangement reactions, the absence of *m/z* 214 ([phosphoadenine]$^-$) in the CID mass spectrum of $[w_1]^-$ is surprising for two reasons: (a) $[w_1]^-$ contains an adenine in its structure and (b) CID of $[w_2]^-$ from Oligo-1 (Table 3.1, row 4) generates [phosphoadenine]$^-$ (*m/z* 214) *and* [phosphoguanine]$^-$ (*m/z* 230), suggesting that losses are possible from both positions. These unusual results may be explained by the existence of other lower energy dissociation pathways for $[w_1]^-$, such

Table 3.1. Summary of Precursor w Ion Sizes, Charge States, and Presence or Absence of [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230) Generated from Oligo-1

w_n	*z*	*m/z* 214	*m/z* 230 *
Oligo-1: ATG CGT CCG GCG TAG A			
1	1-	✗	–
2	1-	✓	✓
3	1-	✓	✓
4	2-	✓	✓
5	2-	✓	✓
6	2-	✓	✓
7	2-	✓	✓

* "–" Denotes that [phosphoguanine]$^-$ generation was not possible due to lack of guanine in the precursor w type fragment ion sequence.

as those resulting from charge localization.[174, 176-178] For example, Pan *et al.* demonstrated that deprotonating a phosphate group enhances the proton affinity, and therefore cleavage, of the neighboring nucleobase.[174] In the cases of $[w_n]^{z-}$ fragment ions where $n \geq 2$, the negative charge(s) can be distributed to multiple sites in the oligonucleotide, therefore decreasing the chance of a negative charge being at the 5′ terminal phosphate group. The $[w_1]^-$ fragment ion, however, has only one phosphate group that can be deprotonated. Thus, according to previously reported findings,[174, 176] charge localization to the 5′ terminal phosphate group could increase the proton affinity of the neighboring adenine and direct fragmentation through loss of adenine, rather than generation of [phosphoadenine]$^-$.

To study the conformation(s) of rearranged [phosphopurines]$^-$ and check the possibility of the phosphorus containing group being attached to multiple purine ring positions (*i.e.*, structural isomers), we acquired ion mobility spectra of [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230) (Figure 3.3). Presence of multiple structural

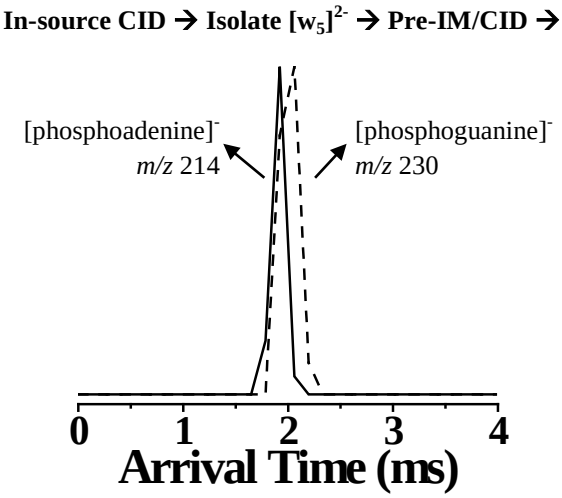


Figure 3.3. IM profiles for [phosphoadenine]- (*m/z* 214; solid line) and [phosphoguanine]$^-$ (*m/z* 230; dashed line) generated from pre-IM CID of $[w_5]^{2-}$ (*m/z* 802) from Oligo-1.

isomers for each [phosphopurine]$^-$ would be indicated by presence of multiple IM arrival time (AT) peaks. IM profiles for [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230) (Figure 3.3) each showed single IM peaks, suggesting the presence of a single structure (type) for each ion population. However, it is possible for [phosphopurines]$^-$ to have multiple isomers or structures that require higher mobility resolving powers for their successful separation. Hence, under our current experimental conditions, the IM-MS results (Figure 3.3) suggest that the phosphorus containing group is attached to purine bases at only one position.

In addition to experiments in the Synapt G2-S HDMS system (*i.e.*, Figures 3.1-3), rearranged [phospho*purines*]$^-$ were also observed in the CID mass spectra of $[w_5]^{2-}$ from Oligo-1 in a triple quadrupole MS system (Appendix B, Figure B.1) and $[\text{Oligo-2} - 2H]^{2-}$ in an LTQ Orbitrap MS (Appendix B, Figure B.2). Results from these three different mass spectrometer types suggest that the [phosphopurine]$^-$ fragment ion rearrangement pathway can be accessed in different MS system types and collisional activation regimes.

For example, the CID results in Figures 3.1-3 and Appendix B, Figure B.1 were collected on instruments where collision energy was applied through a voltage potential difference between two electrodes.[52, 53] In this type of non-selective collision regime, product fragment ions continue to undergo collisional activation and fragmentation. Conversely, in the LTQ Orbitrap (Appendix B, Figure B.2), precursor ions were selectively (*i.e.*, on-resonance) excited and fragmented in the LTQ which can greatly reduce reactivation (and further fragmentation) of the product fragment ions.[134, 179] Moreover, LTQ Orbitrap results further support the conclusion that the $[y_5 - G]^-$ (*m/z* 1374) and $[y_5 - A]^-$ (*m/z* 1390) ions in Figure 3.1 and Appendix B, Figure B.2 are complementary fragments to $[phosphoguanine]^-$ and $[phosphoadenine]^-$ respectively. In other words, these ions might be generated from a common fragmentation pathway (*i.e.*, $[w_5]^{2-} \rightarrow [y_5 - B]^-$) rather than from independent losses (*e.g.*, $[w_5]^{2-} \rightarrow [y_5]^- \rightarrow [y_5 - B]^-$).

To further investigate the fragmentation mechanism for complementary y-B type ions, we performed MS^n experiments in the LTQ Orbitrap. Figures 3.4a and 3.4b show the MS^3 CID mass spectra of $[y_5 - G]^-$ (*m/z* 1374) and $[y_5 - A]^-$ (*m/z* 1390), respectively, generated from Oligo-2 (*i.e.*, isolate $[Oligo\text{-}2 - 2H]^{2-}$ (*m/z* 802) → CID → isolate $[y_5 - G$ or $A]^-$ (*m/z* 1374 or *m/z* 1390) → CID to yield the mass spectra in Figure 4). Observed fragments, generated from CID of the ions at *m/z* 1374 (Figure 3.4a) and *m/z* 1390 (Figure 3.4b), support the fragment ion assignments for $[y_5 - G]^-$ and $[y_5 - A]^-$. Furthermore, none of the product fragment ions generated from CID of isolated $[y_5 - G]^-$ or $[y_5 - A]^-$ precursors suggest the presence of atomic rearrangements or sequence-scrambling, further supporting the conclusion that macrocyclization may not be involved in product ion formation.

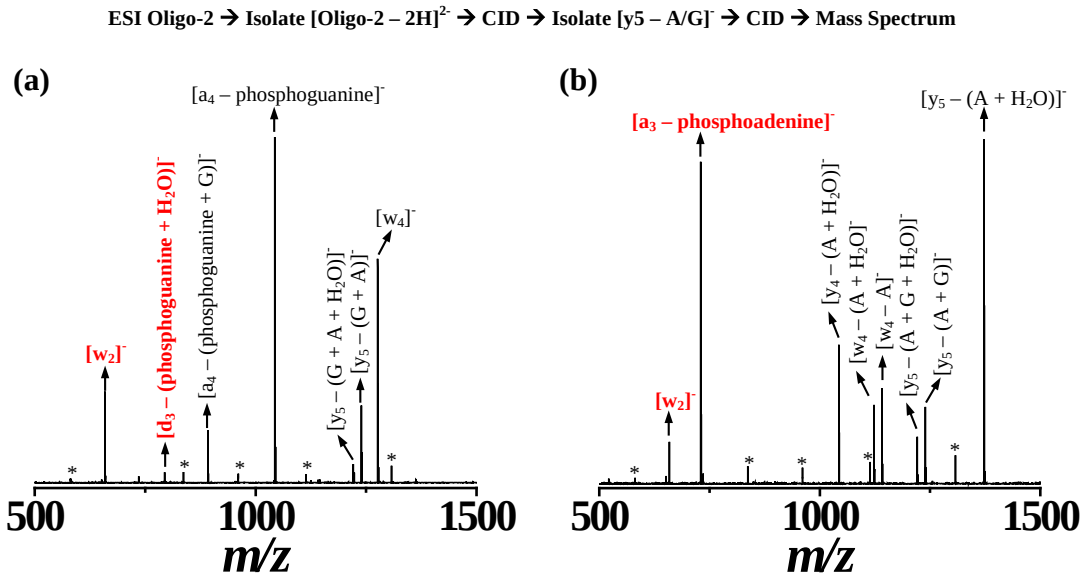


Figure 3.4. CID mass spectra of *m/z*-isolated (a) $[y - G]^-$ (*m/z* 1374) and (b) $[y - A]^-$ (*m/z* 1390) generated from Oligo-2 (p-GTA GA) at 30 % normalized collision energy. Fragment ions labeled in red font denote peaks which can be used to determine positions of base losses. Peaks marked with an asterisk (*) denote electronic noise.

Beyond confirming the fragment ion identities, the CID mass spectra in Figure 3.4 can also be probed to determine sites of guanine and adenine losses. For example, CID of $[y_5 - G]^-$ (Figure 3.4a) generated two ions (denoted in red font) at *m/z* 659 and *m/z* 794, which correspond to $[w_2]^-$ and $[d_3 - (\text{phosphoguanine} + H_2O)]^-$, respectively. Presence of the $[w_2]^-$ (*m/z* 659) fragment ion suggests that the guanine at position four (from the 5′ end) is not lost. Moreover, presence of the ion at *m/z* 794 (*i.e.*, $[d_3 - (\text{phosphoguanine} + H_2O)]^-$) suggests that the guanine at the first position (from the 5′ end) is lost. These two pieces of information together can be used to conclude that at least a portion of the [phosphoguanine]- ions observed at *m/z* 230 in Figures 3.1-3.3, and Appendix B, Figures B.1 and B.2, are generated from losses of the guanines at the first positions from the 5′ ends of $[w_5]^{2-}$ and $[\text{Oligo-2} - 2H]^{2-}$. Although the results in Figure 3.4a suggest that a guanine is lost from the first position, they do not negate the possibility for presence of

multiple pathways to form a [phosphoguanine]$^-$ (*i.e.*, from the guanines at positions one and/or four); in other words, observed ions at *m/z* 230 could be a mixture of these two [phosphoguanine]$^-$ populations. Similar conclusions can be drawn from the CID mass spectra of [y_5 - A]$^-$ (*m/z* 1390) in Figure 3.4b. CID of [y_5 - A]$^-$ (*m/z* 1390) generated two fragment ions (denoted in red font) at *m/z* 659 (again corresponding to the [w_2]$^-$ ion) and *m/z* 730, corresponding to [a_3 – phosphoadenine]$^-$ (Figure 3.4b). Identities of these two ions suggest that adenine is lost from position three; similarly this finding alone cannot be used to conclusively establish that adenine is not lost from position five as well.

Our MS^n and IM-MS results suggest that [phosphopurines]$^-$ are generated from atomic rearrangement of Oligo-1 (ATG CGT CCG GCG TAG A) w type ions and intact 5′ phosphorylated Oligo-2 (p-GTA GA). In the following section we provide additional experimental data that confirm the prevalence of [phosphopurine]$^-$ fragment ion formation (*i.e.*, we show that formation of rearranged [phosphopurine]$^-$ fragments are not limited to CID of w type ions from Oligo-1 and intact Oligo-2).

Fragment Ion Rearrangements in CID of w Type Ions from Additional DNA Oligonucleotides

Based on the results from CID of [w_n]$^{z-}$ (where n = 1-7; *z* = 1 or 2) from Oligo-1 (Table 3.1), we hypothesized that CID of w type ions from other oligonucleotides should also generate [phosphoadenine]$^-$ (*m/z* 214) and/or [phosphoguanine]$^-$ (*m/z* 230), assuming there is an adenine and/or guanine in the nucleotide sequence. To test this hypothesis, we performed additional CID experiments on fourteen [w_n]$^{z-}$ (where n = 1-8; *z* = 1-3) ions (validation set) generated from two DNA oligonucleotides: Oligo-3 (TCG AGG TCG ACG GTA TC) and Oligo-4 (AGA GTT TGA TCC TGG CTC AG). These CID results

are summarized in Table 3.2. Please note that no $[w_3]^{z-}$ ions were generated from in-source CID of Oligo-4 and are thus absent in Table 3.2.

Results from CID of $[w_n]^{z-}$ (where n = 1-8; *z* = 1-3) ions from Oligo-3 and Oligo-4 confirm our hypothesis that [phosphopurine]$^-$ ions can be generated from oligonucleotides with different sequences. Based on our experimental results, the precursor ions' charge state (where *z* = 1-3) did not have any noticeable effect on generation of [phosphopurine]$^-$ ions; however, potential effects of the precursor ion's charge state on [phosphopurine]$^-$ generation when $z > 3$ (which were not tested here) should not be negated. The precursor ions' sizes did have some influence on [phospho*purine*]$^-$ generation. Notably, [phosphopurine]$^-$ ions were only formed from precursor $[w_n]^{z-}$ ions when $n \geq 2$ (*i.e.*, w type ions containing more than one nucleotide in their sequences), which agree with results from CID of $[w_1]^-$ from Oligo-1 (Table 3.1).
In total, we performed tandem MS experiments on twenty-one w type ions (Tables 3.1 and 3.2) and no [phosphopyrimidine]$^-$ product ions were observed in any of the MS^n experiments. Preferential generation of [phosphopurine]$^-$ over [phosphopyrimidine]$^-$ suggests that the bicyclic nitrogenous ring structure of purines might be essential for formation of intermediate and/or final product ions and their stability. Based on a valuable reviewer comment, we also checked the mass spectra for possibilities of purine interactions with phosphate groups at internal positions, which would be indicated by presence of y+B type ions; however, under our experimental conditions, these types of ions were not observed. It is possible that steric hindrance from the next sequential deoxyribose group may prevent this type of rearrangement (*i.e.*, the rearrangement must involve a terminal phosphate group). These results confirm our IM and MS^n data and are

Table 3.2. Summary of Precursor w Ion Sizes, Charge States, and Presence or Absence of [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230) Generated from Validation Set

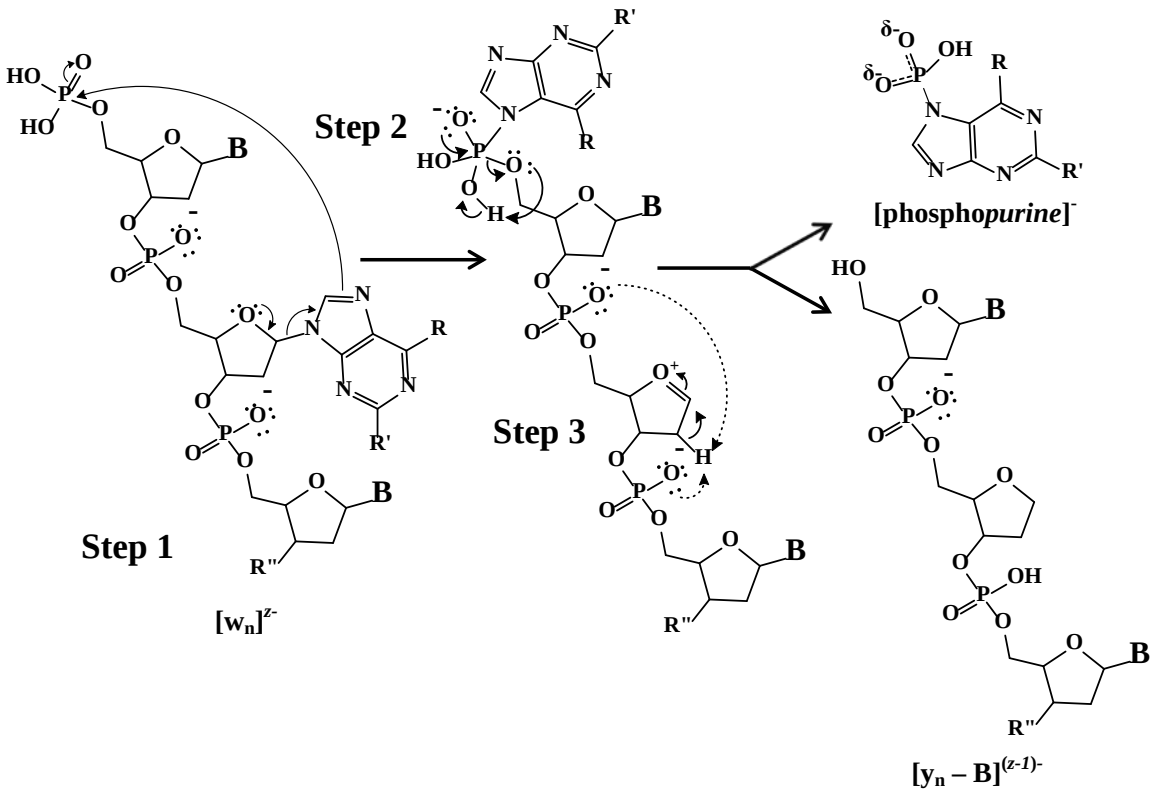
w_n	z	*m/z* 214*	*m/z* 230*
Oligo-3: TCG AGG TCG ACG GTA TC			
1	1-	–	–
2	1-	–	–
3	1-	✓	–
4	2-	✓	–
5	2-	✓	✓
6	2-	✓	✓
7	2-	✓	✓
8	3-	✓	✓
Oligo-4: AGA GTT TGA TCC TGG CTC AG			
1	1-	–	✗
2	1-	✓	✓
4	2-	✓	✓
5	2-	✓	✓
6	2-	✓	✓
7	2-	✓	✓

* "–" Denotes that [phospho*purine*]$^-$ generation was not possible due to lack of guanine and/or adenine in the precursor w type fragment ion sequence.

in agreement with the theoretical evidence and allow elucidation of a potential mechanism for formation of [phosphopurine]$^-$ ions.

Scheme 3.2 shows a proposed mechanism for generation of [phospho*purine*]$^-$ (*i.e.*, [phosphoadenine]$^-$ (*m/z* 214) and [phosphoguanine]$^-$ (*m/z* 230)) and complementary y-B type ions (*i.e.*, $[y_n - A]^{(z-1)-}$ and $[y_n - G]^{(z-1)-}$) from CID of *m/z*-isolated w type ions. Please note that the purine base depicted in Scheme 3.2 could either be guanine or adenine and that nucleobases denoted by a capital "B" could be any nucleobase (*i.e.*, A, G, C, or T).

For succinctness, the oligonucleotide structure in Scheme 3.2 is abbreviated to three nucleotides and we only show the interaction of a purine at nucleobase position two (from the 5′ terminus) with the 5′ phosphate group, although purines at other positions can also follow the same mechanism. The first step in formation of [phosphopurines]$^-$



Scheme 3.2. Proposed mechanism for generation of rearranged [phospho*purine*]$^-$ and y-B type ions.

may proceed *via* a mechanism similar to one recently proposed by Stucki *et al.* in a study of homo-DNA oligonucleotides.[180] First, a lone pair of electrons from the ether oxygen (of the deoxyribose sugar) initiates cleavage of the purine base from the sugar backbone and subsequent nucleophilic attack of the 5′ phosphate group to form a phosphoamide tetrahedral transition species (Scheme 3.2, Step 1). Next, the phosphoamide tetrahedral transition species can rearrange to generate [phosphopurine]$^-$ (*m/z* 214 and 230) (Scheme 3.2, Step 2), which may be structurally similar to phosphoimidazole.[181-183] Finally, neutralization of the oxonium ion may be achieved by β-elimination of a proton from the deoxyribose sugar by one of the adjacent phosphate groups (dashed arrows in Scheme

3.2, Step 3), moving the pi electrons back to the valence shell of the ether oxygen and generating complementary y-B type ions (Scheme 3.2, Step 3).

Conclusions

Our experimental and theoretical results indicate that CID of *m/z*-isolated 5′ phosphorylated DNA oligonucleotide ions can generate rearranged product ions. Furthermore, CID results suggest that, unlike peptide ion rearrangements,[89, 106, 121] oligonucleotide ions do not necessarily need to go through a macrocyclic intermediate to generate [phosphopurines]$^-$. The proposed mechanism for generation of [phosphopurine]$^-$ ions also indicates a previously undefined pathway for formation of y-B type fragment ions.

The predictability of [phosphopurine]$^-$ formation may be leveraged as an additional probe to determine if an unknown DNA sequence (or fragment ion) contains a 5′ phosphate group, assuming its sequence contains an adenine and/or guanine nucleobase. Moreover, the predictable nature of [phosphopurine]$^-$ formation and incorporation of these fragmentation pathways into existing nucleotide sequencing software/algorithms could improve data analysis.

Although presence of the two reported [phosphopurine]$^-$ ions may be inconsequential to sequence analysis of oligonucleotides, their presence in the CID mass spectra of oligonucleotide ions opens up the possibility that other, more consequential, rearrangements or fragmentation pathways may also exist. The current report gives a small snapshot of how oligonucleotide ions rearrange upon CID. Additional systematic studies are necessary to determine the full extent of oligonucleotide ion rearrangements

(*e.g.*, from CID of intact, a, b, c, d, w, x, y, and z type ions [114]) and their effects on sequencing efforts.

Acknowledgements

The authors wish to thank Baylor University for financial support. The authors also thank Dr. Patrick Farmer for providing the imidazole used in this study. Furthermore, the authors thank Dr. Clinton George for helpful discussions of Scheme 2 and Dr. Abayomi Olaitan for assistance with theoretical calculations. Lastly, the authors are grateful to the reviewers of this manuscript whose excellent suggestions and comments helped shape the scientific inquiries and conclusions discussed in here.

CHAPTER FOUR

Calculating Accurate Collision Cross-sections of Ion Mobility Unresolved Isomers Using Tandem Mass Spectrometry and Chemometric Deconvolution

Abstract

Ion mobility (IM) is an important analytical technique for determining gas-phase collision cross-section (CCS) and gaining insight into molecular structures and conformations. However, limited instrumental resolving powers for IM may restrict adequate characterization of conformationally similar ions, such as structural isomers, and reduce the accuracy of IM-based CCS calculations. Recently, we introduced an automated technique for extracting "pure" IM and collision-induced dissociation (CID) mass spectra of IM overlapping species using chemometric deconvolution of post-IM/CID mass spectrometry (MS) data [*J. Am. Soc. Mass Spectrom.*, 2014, **23**, 1873-1884]. Here we extend those capabilities to demonstrate how extracted IM profiles can be used to calculate accurate CCS values of peptide isomers which are not fully resolved by IM. We show that CCS values obtained from deconvoluted IM spectra match with CCS values measured from the individually analyzed corresponding peptides. We introduce an approach that utilizes experimentally determined IM arrival time (AT) "shift factors" to compensate for ion acceleration variations during post-IM/CID and significantly improve the accuracy of the calculated CCS values. We discuss details of the proposed IM deconvolution approach and compare empirical CCS values from traveling wave

(TW)IM-MS and drift tube (DT)IM-MS with theoretically calculated CCS values using the projected superposition approximation (PSA).

Introduction

Ion mobility-mass spectrometry (IM-MS) is a valuable analytical technique for analyzing the structure(s) and conformation(s) of gas-phase ions.[48, 49] IM-MS can be used to rapidly separate ions based on their gas-phase size and shape, while collision cross-sections (CCSs) determined from the IM drift time measurement can aid molecular identification, and can be used to gain insight into the size and folding of molecules. There is compelling evidence that the structural insights obtained from IM-MS studies may be relevant to solution-phase structures and chemistry.[184-190] Several different commercial IMS and IM-MS systems are available which are capable of determining CCSs, namely the most prevalent are based on drift tube (DT)IM-MS and traveling wave (TW)IM-MS technologies.[48, 49] CCS measurements in DTIM-MS are relatively straightforward because ion drift times (t_D) are directly proportional to their CCS. [48, 191] However, in TWIM-MS where a non-uniform field electric field is used, ion CCS values are proportional to t_D^x, where "*x*" is a parameter that must be calculated from analysis of calibrants with known CCSs.[51]

The capability of IM instruments to resolve isomeric species is governed by the IM resolving power (*i.e.,* arrival time (AT)/$\Delta AT_{50\%}$ or CCS/$\Delta CCS_{50\%}$).[191-195] Continuous improvements in the resolving power of contemporary instrumentation have extended the ability to differentiate isomeric systems which exhibit similar CCS values. To improve IM resolving power, several groups have focused on different strategies including using alternative drift gasses or gas modifiers,[196-201] developing new instruments,[162, 202-206] or

customizing existing systems.[53] Although effective, these approaches aimed at improving the IM resolving power can be complicated and costly to implement. To circumvent instrumental limitations, some groups have used post-IM fragmentation and assigned IM profiles of diagnostic fragment ions to IM unresolved precursor ions.[57, 59, 207-211] For example, similar to spectral deconvolution of overlapping gas chromatography (GC)/MS peaks,[61] Clemmer's group demonstrated that isomers with unresolved drift time distributions could be differentiated by analyzing extracted ion drift time distributions (EIDTDs) of unique fragment ions generated from post-DTIM collision induced dissociation (CID).[207, 208] Recently, Khakinejad *et al.* demonstrated that EIDTDs from post-DTIM/electron transfer dissociation data could be used to calculate CCSs of IM unresolved peptide conformers.[210] These post-IM/CID approaches demonstrated that fragment ions could be used to differentiate unresolved isomers present in a mixed AT, but to date there have not been any attempts at reconstructing pure AT profiles from convoluted IM spectra which are suitable for determining CCS values.

Previously, we introduced a chemometric data deconvolution technique[62] based on SIMPLe-to-use Interactive Self-modeling Mixture Analysis (SIMPLISMA)[60] which allowed the extraction of "pure" IM and post-IM/CID mass spectra for IM unresolved species. Subsequently, we integrated SIMPLISMA into an automated IM deconvolution (AIMD) software, which improved speed and accuracy of IM-MS analyses.[54] Moreover, we showed that chemical rank determination techniques, such as principal component analysis and Malinowski's factor indicator,[212, 213] could be used to identify the number of IM unresolved species without prior knowledge of the sample composition.[54, 55, 62-64]

In this manuscript, we demonstrate another unique utility of chemometric deconvolution of post-IM/CID MS; specifically, we show that CCSs of IM unresolved species, such as reverse pentapeptide isomers (*i.e.*, SDGRG and GRGDS) and enkephalin hexapeptide isomers (*i.e.*, RYGGFM, RMFGYG, MFRYGG, and FRMYGG), can be accurately calculated. A method is developed which utilizes an AT "shift factor" to compensate for effects of electric field variations during post-IM/CID, which significantly reduces differences between mean calculated CCSs for deconvoluted profiles and CCSs obtained from individually measured analytes. Finally, we compare our deconvoluted TWIM-MS CCSs for the enkephalin hexapeptide isomers to values obtained using DTIM-MS and values calculated using projected superposition approximation (PSA), a contemporary computational approach used to determine theoretical CCSs.

Materials and Methods

Sample Preparation

Reverse pentapeptide isomers (amino acid sequences: SDGRG and GRGDS) and poly-Alanine (used as the CCS calibrant)[51, 214] were purchased from Sigma-Aldrich (St. Louis, MO, USA). [Arg0] Met-enkephalin (amino acid sequence: RYGGFM) was purchased from American Peptide (Sunnyvale, CA, USA). Scrambled sequence enkephalin hexapeptides (amino acid sequences: RMFGYG, MFRYGG, and FRMYGG) were custom synthesized by Peptide 2.0, Inc. (Chantilly, VA, USA) and their identities were confirmed by CID MS analyses (Appendix C, Figure C.1). The reverse pentapeptide mixture was chosen based on its previous use to characterize the IM resolving powers of

the Synapt HDMS TWIM-MS instrument.[53] The enkephalin hexapeptide isomers were chosen because they have overlapping IM profiles and their CCSs have not, to the best of our knowledge, been previously characterized. Optima LC/MS grade acetic acid and methanol were purchased from Fisher-Scientific (Waltham, MA, USA). Water was purified in-house using a Direct-Q 3 UV water purification system (EMD Millipore Corporation, Billerica, MA, USA). Peptides were prepared by reconstituting the lyophilized powder using a water: methanol: acetic acid (49.95: 49.95: 0.1) solvent composition at final concentrations of ~1 μM. Peptide mixtures contained approximately equimolar concentrations of each isomer.

Instrumentation

TWIM-MS experiments were performed as previously described[54, 55, 62-64, 215] using both a Synapt G1 (for the reverse pentapeptide binary mixture) and a G2-S (for the enkephalin hexapeptide quaternary mixture) system (Waters Corporation, Milford, MA, USA). Details of the experimental conditions used to collect the TWIM-MS data reported in this manuscript can be found in Appendix C, Table C.1. Drift tube CCS measurements were obtained on a commercial ESI-IM-QTOF Agilent 6560 ion mobility-mass spectrometer (Agilent Technologies, Santa Clara, CA, USA). The method used has been previously described,[216, 217] but briefly, the buffer gas (*i.e.*, nitrogen (N_2)) was maintained at a pressure of *ca.* 4 Torr and the drift voltages were varied in order to correct for the non-IM flight time of the ions through the interfacing ion optics. Agilent tune mix was run with each enkephalin hexapeptide sample as an internal CCS standard.

Data Analysis

Post-IM/CID MS data were deconvoluted using in-house developed automated IM deconvolution (AIMD) software.[54] All IM arrival times for original raw data (i.e., poly-Alanine, individually run peptides, and peptide mixtures) correspond to the experimental values for the centroid m/z of the monoisotopic (viz., $^{12}C_{all}$) peak ±0.03 Th mass range (e.g., m/z 491.22 ± 0.03 for singly-protonated SDGRG ($[SDGRG + H]^+$)). Post-IM/CID MS data analyzed with AIMD software were the average of 10 minutes (588 scans; reverse pentapeptides) or 1 minute acquisition times (58 scans; enkephalin hexapeptides) which were sufficient to provide signal-to-noise ratios of >3 for all deconvoluted ions of interest. All other TWIM data were collected for 1 minute acquisition time (58 scans), unless otherwise noted. DTIM-MS experiments were collected for 2 minutes acquisition time (108 scans). All experiments were conducted a minimum of three times; reported errors are at the 95% confidence level. Experimental TWIM-MS CCSs were calculated using the method proposed by Ruotolo and coworkers.[51] The Mason-Schamp equation [218] was used to determine experimental CCS values from ion drift times for all DTIM-MS experiments.

Theoretical Modeling

Theoretical CCSs for enkephalin hexapeptide isomers (*i.e.*, RYGGFM, RMFGYG, MFRYGG, and FRMYGG) were calculated using a simulated annealing protocol implemented with the AMBER software.[219] The doubly-protonated enkephalin isomers were constructed using standard amino acid templates in the xLEaP module such that protonation sites are at the arginine and N-terminus. Because xLEaP contains amino acid templates that do not incorporate hydrogens on the C-terminal residues, these were

constructed manually and subsequently geometry optimized at the Hartree-Fock level with a 6-31G* basis set using Gaussian 09.[220] The enkephalin isomers were then energy minimized with the sander module, and subsequently heated to 700-800 K over a 10 ps molecular dynamics simulation. An extended molecular dynamic simulation was run at 700-800 K for 9,000 ps and structural snapshots were saved every 3 ps throughout the simulation, resulting in 3,000 unique structures. These structures were cooled to 300 K during 15 ps molecular dynamics simulations resulting in low-energy conformations that were used for theoretical CCS determination.[221, 222]

A combination of MOBCAL[223-226] and projected superposition approximation (PSA)[227-230] were used to calculate the theoretical CCS values. Helium (He) values were obtained in MOBCAL using the projection approximation. PSA was then used to calculate nitrogen (N_2) CCS values for a series of low-energy structures (~20) spanning the entire CCS range for which the structures were obtained. A linear function was determined from these values to convert the remaining He CCS values to N_2 CCS values as previously demonstrated.[221] Briefly, it was found that, to a first approximation, theoretical He and N_2 CCS values could be directly correlated by a simple linear function which helps offset the time requirement for generating a large pool of values.[221] PSA CCS values were calculated using the online service hosted by the University of California, Santa Barbara.[231]

Results and Discussion

Figure 4.1 (black) shows the IM profile of an equimolar mixture of two isomeric reverse pentapeptides (*i.e.,* $[SDGRG + H]^+$ and $[GRGDS + H]^+$ at *m/z* 491.22). As previously reported, the IM resolving power of the Synapt G1 is insufficient to separate

these two isomeric peptides, while the Synapt G2 is only capable of differentiating the doubly-protonated forms of this isomeric mixture.[53, 54] Characterization of unknown samples with overlapping IM profiles (such as those shown in Figure 4.1, black IM profile) would be challenging, as unresolved IM peaks could erroneously be assigned to single analytes. Particularly in this example, the mixture exhibits a symmetric Gaussian-like IM profile indicative of what would be obtained from a single analyte. Resultantly, this would lead to the measurement of a CCS corresponding to a mixture of analytes. In the following sections, we show that chemometric deconvolution of post-IM/CID MS data can be used to extract the ATs of IM unresolved isomers which can then be used to calculate accurate CCS values. Briefly, because of ion acceleration during post-IM/CID, ATs of deconvoluted IM profiles are shorter than ATs of the individually analyzed species without post-IM/CID (*i.e.*, obtained under low-energy conditions). Here, we demonstrate a method to correct, or "zero", ATs from deconvoluted IM spectra to match low-energy individually analyzed IM spectra for the purpose of determining the CCS.

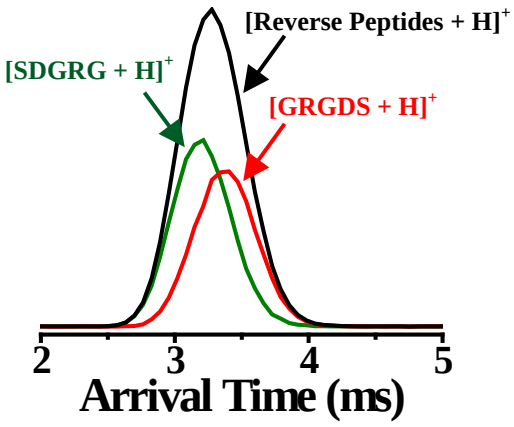


Figure 4.1. IM profiles of $[SDGRG + H]^+$ (green) and $[GRGDS + H]^+$ (red) before (black IM profile) and after (colored IM profiles) IM deconvolution at an IM wave height of 7.6 V and 40 V post-IM/CID acceleration voltage.

Chemometric Deconvolution

Using our previously reported AIMD software,[54] the unresolved IM profile in Figure 4.1 (black IM profile) was deconvoluted to yield extracted IM profiles for $[SDGRG + H]^+$ and $[GRGDS + H]^+$ (green and red IM spectra in Figure 4.1, respectively). To check the validity of the deconvolution, we compared constructed CID mass spectra for each peptide from deconvolution of the data in Figure 4.1 to CID mass spectra of individually analyzed peptides (Appendix C, Figure C.2). The degree of similarity between deconvoluted and pure CID mass spectra were assessed using a previously described matching score (*R*) algorithm:[215]

Equation 4.1
$$R = 1 - \frac{\sum_{m=1}^{n} |x_m - x'_m|}{\sum_{m=1}^{n} (x_m - x'_m)}$$

where n is the total number of fragment ion m/z values, m, and xm and x'm are the fragment ion intensities (for a given value of m) for pure and deconvoluted spectra, respectively. R values of 0 and 1 indicate no correlation and a perfect match, respectively.[215] An R value of $\geq$ 0.75 was considered a "successful" threshold for deconvolution.[215] Calculated R values for all deconvoluted reverse peptide isomer CID mass spectra were > 0.86. Moreover, cross-validation matching scores (R') were calculated by comparing the deconvoluted CID mass spectra for each peptide isomer with the reference spectrum for the non-matching isomer. The highest calculated R' value was ~0.46. Validation and cross-validation results suggested that, for each peptide isomer, deconvolution was successful both in terms of the fragment ions present in each mass spectrum, as well as relative intensities of those fragment ions.

Arrival Time Correction Factor and CCS Calculations

Although the match scores (*i.e.*, R values) for all deconvoluted CID mass spectra (Appendix C, Figure C.2) for the reverse peptides suggest that MS deconvolution is successful, the degree of match between the ion mobility ATs of deconvoluted and pure isomers cannot be discerned or quantified from the MS data alone. In-fact, as a result of electric potential difference used to induce post-IM/CID, ATs extracted from the deconvoluted post-IM/CID data in Figure 1 did not match the ATs of individually run peptides (*i.e.*, without post-IM/CID). The higher potential difference between the IM and transfer cells in post-IM/CID experiments results in faster acceleration of the ions as they exit the IM cell and pass through the transfer cell and thus leads to faster measured ATs than what is obtained in low-energy (non-fragmentation) experiments. For example, the AT of deconvoluted (high-energy) $[SDGRG + H]^+$ (labeled at the peak maxima, 3.18 ±0.06 ms, of green IM profile in Figure 4.2a) was approximately 22 ms shorter than the AT for the individually analyzed (low-energy) pure $[SDGRG + H]^+$ (labeled at the peak

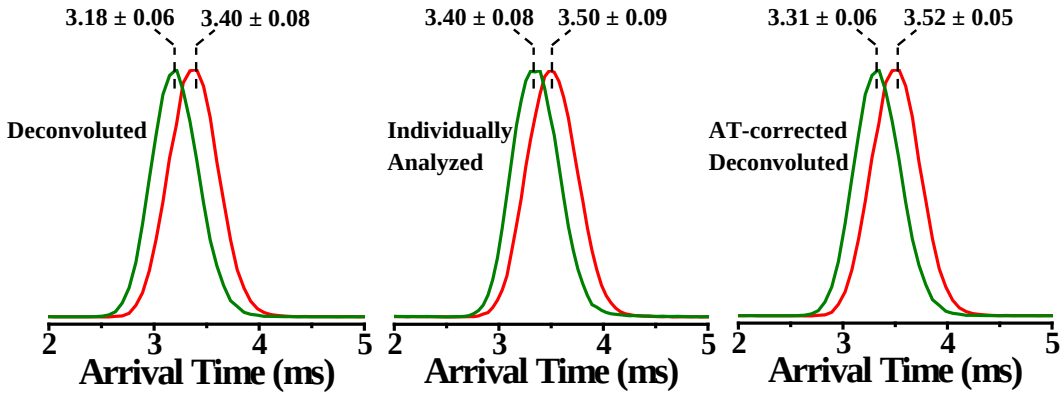


Figure 4.2. Normalized IM profiles for (a) deconvoluted, (b) individually analyzed pure, and (c) AT-corrected deconvoluted $[SDGRG +H]^+$ (green) and $[GRGDS + H]^+$ (red) using IM wave height voltage, wave velocity, and post-IM/CID acceleration voltage of 7.6 V, 300 m/s, and 40 V, respectively. AT errors are reported at the 95% confidence level (n_{pure} = 5 and n_{decon} = 4).

maxima, 3.40 ±0.08 ms, of green IM profile in Figure 4.2b); these AT values are statistically different at the 95% confidence level (data summarized in Appendix C) and suggest that deconvoluted ATs should only be used for CCS calculations after applying an appropriate AT correction factor to the deconvoluted IM spectra.

One solution to avoid AT variations would be to run the CCS calibration mixture at the same post-IM/CID energy as used for deconvolution. However, in cases where the CCS calibrant is a polymer (such as poly-Alanine used here),[63, 214, 232, 233] post-IM/CID could generate several isomers with different and/or overlapping IM profiles (*e.g.*, CID of Ala_5 through Ala_n [where n > 5] can generate Ala_{1-5}), or completely fragment the precursor ion (*e.g.,* Ala_5 is depleted while CID of Ala_n [where n > 5] still generates the Ala_5 fragment ion), and complicate extraction of pure IM ATs. Moreover, if post-IM/CID were used for the calibration mixture, then the calibration plots would only be valid for that specific collision-energy. Therefore, multiple calibration profiles would have to be generated in order to deconvolute several different IM unresolved species at different collision-energies.[215] Hence, it would be advantageous to correct ATs of deconvoluted IM profiles and fit them to a "conventional" low-energy CCS calibration plot.

To correct for the change in ATs caused by post-IM/CID acceleration voltage, we ran the reverse pentapeptide mixture at low collision energy (*i.e.,* 4 V; which is insufficient to induce CID) and high collision energy (*i.e.*, 40 V) and calculated the difference in AT (herein referred to as the AT "shift factor"). We then added this shift factor (of ~0.13 ms) to ATs of deconvoluted IM profiles to generate corrected-IM distributions (Figure 4.2c, discussed in-detail below); note that similar shift factors can be applied to all ions exiting the IM cell. Garmón-Lobato *et al.* presented a similar strategy

for aligning low- and high-energy IM profiles;[234] however, their approach was not focused on characterizing ions with similar CCSs, such as those presented in this manuscript, or calculating CCSs from the resulting aligned spectra.[234]

The magnitude of the AT correction can be estimated from known parameters of the experiment. Experimentally measured ATs in the Synapt G1 (and G2-S) instruments are the summation of four ion transit times which includes: (1) IM drift time, (2) time spent in the transfer cell, (3) TOF MS flight time, and (4) time spent traveling across the interface between each of these instrument regions.[51, 53, 146] IM drift times should not change significantly between low- and high-energy experiments because the electric fields (*i.e.,* traveling wave, entrance and exit lenses, and radio frequency confinement voltages) and IM pressure, which govern ions' mobilities through the IM cell, are the same in both sets of experiments.[53, 235] Differences in TOF (in the field free vacuum region of the MS) for precursor and product ions generated between the IM and transfer cells (*i.e.*, prior to TOF) is on the order of 1 to 30 μs,[236] which is one to two orders of magnitude smaller than the AT difference (*e.g.*, ~0.13 ms for the reverse pentapeptides) between deconvoluted IM profiles and individually analyzed pentapeptides in Figures 4.2a and 4.2b, respectively. Nevertheless, the TOF difference constitutes a lower limit for expected AT shifts between the high- and low-energy experiments.

In addition to different TOF, the time spent in the interface region between IM and transfer cells is also decreased during post-IM/CID (as compared to low-energy experiments) due to a larger potential difference between these two cells. Moreover, increased kinetic energy of fragment ions (as a result of post-IM acceleration) entering the transfer cell may result in a shorter time to traverse the transfer cell. In the Synapt G1,

the center-to-center spacing between ring electrodes (*e.g.*, between the IM and transfer cells) is ~1.5 mm and the length of the transfer cell is ~100 mm.[146, 235] Based on the physical geometry and electric fields of the instrument,[146, 235] ions will spend a considerably longer time in the transfer cell than in the region between the IM and transfer cells or in the TOF mass spectrometer. Therefore, we can approximate a maximum AT shift factor to be equal to the maximum amount of time an ion can spend in the transfer cell. The transfer cell is operated at lower pressure than the IM cell and ions therefore travel with the traveling wave rather than falling over the wave as they do in IM.[146, 194, 201, 235] Therefore, assuming an ion entering the transfer cell had insufficient kinetic energy to overtake the traveling wave in the transfer cell, and ignoring phase effects of the traveling wave, the maximum amount of time an ion could spend in the transfer cell ($t_{transfer}$) would be equal to the length of the cell (d_t) divided by the transfer cell wave velocity (v_t):[51]

Equation 4.2
$$t_{transfer} = \frac{d_t}{v_t}$$

For a cell length of 100 mm (0.1 m) and transfer cell wave velocity of 248 m/s (as used in our experiments), the maximum time an ion should spend in the transfer cell (under our experimental conditions) is ~0.40 ms. Based on these simple estimations, under our experimental conditions we should expect shift factors between ~0.10 x 10^{-2} ms (minimum TOF variation for fragment and precursor ions)[236] and ~0.40 ms (maximum transfer cell time).

In this study, we introduce a simple approach to calculate AT shift factors that can be used to correct the experimentally measured AT values obtained under high (post-IM) collision energy conditions. We show that the proposed approach provides CCS values

for individual components of convoluted mixtures (at high post-IM collision energy conditions) that are statistically indistinguishable from individually measured CCS values (under low energy ion transfer conditions) from pure samples of the same species. Future studies could improve the current experimental approach by deriving a formula to express the impact of the electric potential differences between the IM and transfer cells on ions' drift times in the transfer cell. These values could then be subtracted from the measured ATs and therefore eliminate the need for measuring the ΔAT between low- and high-energy experiments. However, we note here that while it would be desirable to determine corrected AT values *a priori*, such efforts would be challenging as the traveling wave complicates mathematical derivation by introducing new variables including wave phase, height, and velocity.[235]

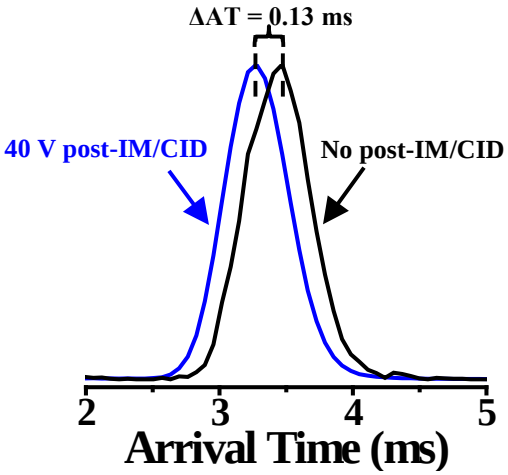


Figure 4.3. Unresolved IM profiles of an equimolar mixture of $[SDGRG + H]^+$ and $[GRGDS + H]^+$ (*m/z* 491.22) using IM wave height and velocity of 7.6 V and 300 m/s, respectively, without post-IM/CID (black IM profile) and with post-IM/CID at 40 V acceleration voltage (blue IM profile). The arrival time shifts (ΔAT = 0.13 ms) to a shorter value, due to post-IM/CID acceleration voltage.

Figure 4.3 shows the observed shift in IM AT for the binary reverse pentapeptide mixture at 4 V (black IM profile) and 40 V (blue IM profile) post-IM/CID collision-energies. The measured shift factor for 7.6 V IM wave height was consistently 0.13 ms across eight trials (*i.e.*, four low-energy (4 V) and four high-energy (40 V)). The calculated shift factors vary as a function of wave height ranging from 0.12 ms to 0.18 ms for wave heights of 7.0 V to 8.0 V. Experimentally determined AT shift factors (*i.e.*, 0.12 ms to 0.18 ms) are within the estimated ~0.10 x 10^{-2} ms to ~0.40 ms range discussed above. Deconvoluted IM profiles were then corrected by adding AT shift factor time (*i.e.*, 0.13 ms) as shown in Figure 4.2a to yield the time-corrected IM profiles in Figure 4.2c. ATs for individually measured peptide isomers and AT-corrected deconvoluted IM profiles at wave height of 7.6 V were statistically indistinguishable at the 95% confidence level (Appendix C, summary of statistical tests). Because of the approximations involved in the deconvolution process (*e.g.*, peak centroid identification)[54] and CCS dependent post-IM ion velocity, adding a single "linear" correction factor is not expected to improve all AT values identically. For example, the average corrected-AT value for GRGDS (3.52 ms, from the red IM profile in Figure 4.2c) is closer to its individually measured value (3.50 ms, from the red IM profile in Figure 4.2b) than the corrected-AT for SDGRG (3.31 ms, from the green IM profile in Figure 4.2c) is to its individually measured value (3.40 ms, from the green IM profile in Figure 4.2b). However, based on the aforementioned statistical justifications, corrected-deconvoluted ATs (as labeled in Figure 4.2c) were sufficient for use in CCS calculations (*e.g.*, CCS results from AT corrected deconvolution and pure IM peak analyses agreed at the 95% confidence level).

Table 4.1 shows a summary of CCS calculations for the reverse pentapeptide mixture. The calculated CCS of the reverse pentapeptide mixture (*i.e.,* $210._3 \pm 3._2$ Å^2; where the subscript digit is the first insignificant figure) is comparable to the mathematical average CCS (~211 Å^2) of individually measured $[SDGRG + H]^+$ ($209._1 \pm 3._1$ Å^2) and $[GRGDS + H]^+$ ($212._7 \pm 3._3$ Å^2) (Table 4.1). Our calculated CCSs for individual $[SDGRG + H]^+$ and $[GRGDS + H]^+$ (from pure samples) are within the 3% error range of the previously reported values for the same peptide samples calculated using DTIM-MS.[237] Slight differences in CCS measurements may be due to several factors including (but not limited to) differences in IM cell humidity,[238, 239] mixture of gasses between the trap/transfer and IM cells (*i.e.,* changing the reduced mass and average drift gas polarizability),[214, 237, 240-242] accuracy and precision of calibrants' CCSs,[237] and/or other experimental condition variations.[191] CCSs of deconvoluted $[SDGRG + H]^+$ and $[GRGDS + H]^+$ without AT corrections were $204._9 \pm 3._0$ Å^2 and $209._5 \pm 3._1$ Å^2, respectively (Table 4.1) Uncorrected CCSs were not statistically different than the individually measured values for both peptides at the 95% confidence interval (Appendix C, summary of statistical tests); however, at the 90% confidence interval, the uncorrected CCS for $[SDGRG + H]^+$ was significantly different than the corresponding pure CCS (supporting information, summary of statistical tests). By comparison,

Table 4.1. Experimental CCSs for Two Reverse Pentapeptide Isomers

Identity	Mixture CCS (Å^2)*	Pure CCS (Å^2)*	Decon. CCS (Å^2)*	Corrected Decon. CCS (Å^2)*
$[Reverse\ pentapeptides + H]^+$ (mixture)	$210._3 \pm 3._2$	—	—	—
$[GRGDS + H]^+$	—	$212._7 \pm 3._3$	$209._5 \pm 3._1$	$212._7 \pm 4._6$
$[SDGRG + H]^+$	—	$209._1 \pm 3._1$	$204._9 \pm 3._0$	$208._3 \pm 4._4$

* Errors are reported at the 95% confidence level ($n_{mixtures} = 24$, $n_{pure} = 25$)

calculated CCSs for the AT-corrected deconvoluted IM profiles for $[SDGRG + H]^+$ and the AT correction approach) are not statistically different than the CCSs of individually $[GRGDS + H]^+$ were $208._3 \pm 4._4$ Å^2 and $212._7 \pm 4._6$ Å^2, respectively (Table 4.1); these calculated CCSs (after applying measured peptide isomers at the 90% and 95% confidence levels (Appendix C, summary of statistical tests).

Our results indicate that CCSs of IM-unresolved reverse sequence peptide isomers (*i.e.*, SDGRG and GRGDS) can be accurately calculated using chemometric deconvolution and AT correction using a Synapt G1 system. In the next section, we extend those results to demonstrate the utility of IM-unresolved CCS calculations on a four-component IM-unresolved enkephalin isomer mixture (*i.e.*, RYGGFM, RMFGYG, MFRYGG, and FRMYGG) using a higher IM resolving power Synapt G2-S system.

Application of IM-unresolved CCS Calculations Using a Synapt G2-S System

Clemmer and coworkers demonstrated that by using LC-IM-MS they could identify ~82% of isomers in an approximately 4000-component combinatorial peptide library.[243] For complex mixture analyses, such as performed by Clemmer,[243] it is unlikely (although not impossible) to have more than two LC/IM unresolved isomers. However, to test our approach in a "worst case scenario" we analyzed a mixture of four constitutional isomeric enkephalin hexapeptides with unreported CCSs and limited our data collection times to one minute (as opposed to 10 minutes as used for the reverse pentapeptides or 45 minutes as used for our previously reported four-component mixture results[215]). Results were compared to calculated theoretical CCS values (*i.e.*, using PSA) and experimentally derived DTIM-MS CCS values.

Figure 4.4 shows the IM profile for a doubly-protonated (*m/z* 365.67) four-component enkephalin hexapeptide mixture before (black IM profile) and after (colored IM profiles where the orange, purple, cyan, and gray IM profiles correspond to RYGGFM, RMFGYG, MFRYGG, and FRMYGG, respectively) IM deconvolution at IM wave height of 22 V and wave velocity of 1300 m/s. The enkephalin isomer mixture (Figure 4.4, black IM profile) yielded a bimodal IM distribution which could erroneously be characterized as corresponding to two components. Deconvolution, *R* and *R'* match scoring, and AT corrections were performed as discussed in the preceding sections. Calculated *R* and *R'* values for all deconvoluted CID mass spectra for the enkephalin hexapeptides were ≥ 0.83 and ≤ 0.37, respectively. For brevity, only final CCS values for the enkephalin hexapeptides are discussed in this section. Examples of deconvoluted and pure CID mass spectra for enkephalin hexapeptide isomers can be found in Appendix C, Figure C.1.

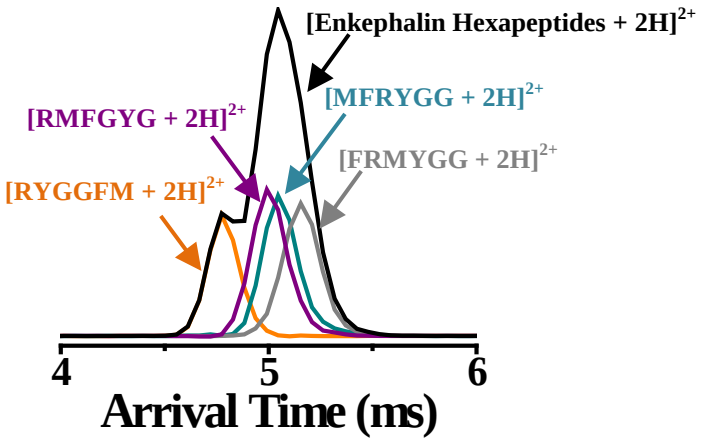


Figure 4.4. IM profiles of $[RYGGFM + 2H]^{2+}$ (orange), $[RMFGYG + 2H]^{2+}$ (purple), $[MFRYGG + 2H]^{2+}$ (cyan), and $[FRMYGG + 2H]^{2+}$ (gray) before (black IM profile) and after (colored IM profiles) IM deconvolution at IM wave height voltage, wave velocity, and post-IM/CID acceleration voltage of 40 V, 1300 m/s, and 22 V, respectively.

Table 4.2 shows a summary of CCS calculations for the enkephalin isomer mixture. Similar to the reverse pentapeptide mixture (Table 4.1), the calculated CCS for the late-AT profile in the enkephalin hexapeptide mixture (*i.e.,* without deconvolution; $296._9 \pm 7._0$ Å^2) is close to the mathematical average (*i.e.*, ~298 Å^2) of the CCSs of the four individually run enkephalin isomers (*i.e.,* $[RYGGFM + 2H]^{2+}$ ($292._7 \pm 6._9$ Å^2), $[RMFGYG +2H]^{2+}$ ($299._6 \pm 7._1$ Å^2), $[MFRYGG + 2H]^{2+}$ ($299._2 \pm 7._1$ Å^2), and $[FRMYGG + 2H]^{2+}$ ($301._5 \pm 7._0$ Å^2)) (Table 4.2). In other words, if the late-AT IM profile in Figure 4.4 were (incorrectly) identified as a single component, then the calculated CCS (*i.e.*, $296._9 \pm 7._0$ Å^2) would be close to the average of the CCSs of the individual IM unresolved constitutional isomers that contribute to the late-AT IM peak. Likewise, the calculated CCS for the early-AT profile in the enkephalin hexapeptide mixture (*i.e.*, without deconvolution; $288._8 \pm 6._8$ Å^2) at the 95% confidence level (Appendix C, summary of statistical tests) is statistically indistinguishable from the CCS of the first eluting enkephalin hexapeptide (*i.e.*, $[RYGGFM + 2H]^{2+}$; $292._7 \pm 6._9$ Å^2).

Table 4.2. Experimental CCSs for Four Enkephalin Hexapeptide Isomers

Identity	Mixture CCS (Å^2)*	Pure CCS (Å^2)*	Decon. CCS (Å^2)*	Corr. Decon. CCS (Å^2)*	Drift Tube CCS (Å^2)*
$[Enkephalins + 2H]^{2+}$ (mixture; early AT Peak)	$288._8 \pm 6._8$	—	—	—	—
$[Enkephalins + 2H]^{2+}$ (mixture; late AT Peak)	$296._9 \pm 7._0$	—	—	—	—
$[RYGGFM + 2H]^{2+}$	—	$292._7 \pm 6._9$	$285._7 \pm 6._7$	$288._8 \pm 9._6$	$285.2_3 \pm 0.5_3$
$[RMFGYG + 2H]^{2+}$	—	$299._6 \pm 7._1$	$292._0 \pm 6._9$	$295._1 \pm 9._9$	$296.1_1 \pm 0.3_8$
$[MFRYGG + 2H]^{2+}$	—	$299._2 \pm 7._1$	$293._8 \pm 6._9$	$296._8 \pm 10._0$	$298.9_7 \pm 0.4_8$
$[FRMYGG + 2H]^{2+}$	—	$301._5 \pm 7._0$	$297._1 \pm 7._1$	$300._1 \pm 10._1$	$303.7_5 \pm 0.4_0$

*Subscript digits denote the first insignificant figure. Errors are reported at the 95% confidence level (TWIMS: $n_{mixtures} = 24$, $n_{pure} = 23$ (FRMYGG) or $n_{pure} = 24$ (all others); DTIMS: n = 24 (for each peptide))

Similar to IM results from analyses of the reverse pentapeptides, CCSs calculated using the AT-corrected deconvoluted IM profiles were statistically indistinguishable from the corresponding individually analyzed peptides (Table 4.2). Although results for neither raw nor AT-corrected deconvoluted IM data were statistically different from the results obtained for the individual peptides, average CCSs for the corrected-AT deconvolution results were closer to the average CCSs for the individually run peptides. For example, calculated CCS for the individually measured $[FRMYGG +2H]^{2+}$ was $301._5 \pm 7._0$ Å^2, which is closer to the calculated CCS for the AT-corrected deconvoluted data of $300._1 \pm 10._1$ Å^2 than to the estimated CCS value of $297._1 \pm 7._1$ Å^2 for deconvoluted raw data. These results suggest that AT-correction improves CCS calculations (both for data from Synapt G1 and G2-S instruments) and can be applied to data from complex mixtures containing more than two IM-overlapping components.

To further demonstrate the accuracy of CCS values calculated from AT-corrected deconvoluted IM profiles, we ran the same enkephalin peptide isomers on an Agilent Technologies DTIM-MS and measured their individual experimental CCSs. Results from DTIM-MS measurements are summarized in Table 4.2, column six. CCS values calculated from traveling wave and drift tube type instruments are in close agreement. Moreover, CCSs derived from AT-corrected post-TWIM/CID MS deconvoluted data are not significantly different than CCSs derived from the individually analyzed peptides run on the DTIM-MS instrument (95% confidence interval using the Type 2 Student's t-test for unequal variance).

The number of significant figures in CCSs calculated using TWIM-MS (Tables 4.1 and 4.2) were limited by the number of significant figures (*viz.* accuracy) in the

literature (calibrant) CCSs. Therefore, TWIM-MS CCSs are reported to three significant figures where as DTIM-MS CCSs (which are not limited by calibrant CCSs) are reported to four significant figures. It should be noted that all experimental errors reported in this manuscript were calculated by propagating error through each step of CCS calculations. For DTIM-MS measurements there is only error in the recorded DTs and in the DT correction (*i.e.,* to compensate for time outside of the IM portion of the instrument). For TWIM-MS measurements, on the other hand, there is error in calibrant and analyte DTs (and error in "shift factors" used for AT-correction) as well as error in fitting parameters and exponential calibration factors derived from calibrant DTs. TWIM-MS calibration errors were calculated using the "LINEST" (*i.e.,* linear least squares fit and errors) function in Microsoft Excel. While propagation of error calculations can be time consuming, they more accurately represent the precision of IM measurements and underscore the importance of developing a calibration independent TWIM-MS CCS equation. Nevertheless, our results suggest that CCSs of IM unresolved isomers can be calculated using deconvolution of post-IM/CID MS data that are not statistically different than individually analyzed peptides on both TW and DTIMS.

In addition to the drift tube CCS values obtained for the enkephalin isomers, theoretical CCS values were determined for an ensemble of conformations resulting from molecular dynamics simulations. Although PSA is considered preliminary for nitrogen CCS determination, good agreement is observed between experimentally measured (DTIM-MS) CCS values (vertical lines) and the conformational space plots in Figure 5. Root mean squared deviation (RMSD) clustering representative conformations for each of the isomers that fall within the experimental range (experimentally determined CCS

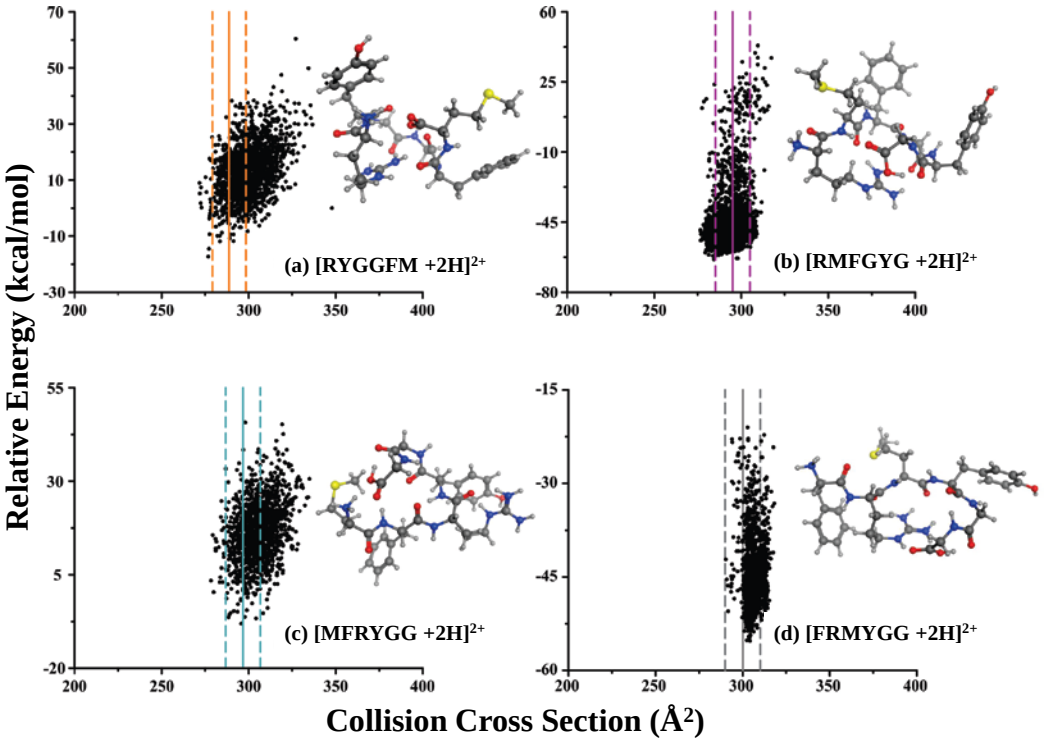


Figure 4.5. Computational results for calculating theoretical CCS values of the four enkephalin hexapeptides including (a) $[RYGGFM + 2H]^{2+}$, (b) $[RMFGYG + 2H]^{2+}$, (c) $[MFRYGG + 2H]^{2+}$, and (d) $[FRMYGG + 2H]^{2+}$. Solid vertical lines on each plot represent the AT-corrected deconvoluted TWIM-MS CCS values and the associated 95% confidence intervals are indicated with dashed lines; line colors match the corresponding deconvoluted IM profiles for Enkephalin hexapeptides in Figure 4.4. RMSD clustering representative conformations for each isomer are also included.

value plus or minus the 95% confidence interval) are also shown in Figure 4.5.[230] The structure of $[RYGGFM + 2H]^{2+}$ in Figure 4.5a shows increased flexibility with the two glycine amino acids in the middle of the peptide chain. The other three isomers have more similar CCS values as seen in the ion mobility traces in Figure 4.4. The one internal glycine gives $[RMFGYG + 2H]^{2+}$ in Figure 4.5b some flexibility. Both $[MFRYGG + 2H]^{2+}$ in Figure 4.5c and $[FRMYGG + 2H]^{2+}$ in Figure 4.5d have the two glycine amino acids on the end of the peptide chain reducing flexibility in the other four bulkier amino acids. The interaction of the positive arginine residue with the c-terminus carboxylate

group in Figure 4.5d extends the structure by pushing out the tryptophan and phenylalanine benzene rings of $[FRMYGG + 2H]^{2+}$ resulting in the largest conformation. Additional RMSD clustering representative conformations from the experimental range of the four isomers are shown in Appendix C, Figure C.3-C.6.

Conclusion

We presented a method to calculate the CCSs of IM unresolved isomers using a modified version of the protocol proposed by Ruotolo *et al.*[51] To accurately measure CCSs of unresolved isomers, it was necessary to calculate IM AT "shift factors" that could be added to the ATs of chemometrically deconvoluted IM profiles. Use of these shift factors improved the measured CCS values as compared to CCSs obtained from pure analytical standards run on both TWIM-MS and DTIM-MS instruments. Moreover, CCSs calculated using IM-deconvoluted data matched closely with theoretical space plots constructed using PSA. Although we utilized TWIM-MS instruments in this report for convolution experiments, the proposed time-corrected deconvolution technique is not limited to TWIM-MS and AT shift factors can be calculated for any IM-MS instrument capable of post-IM/CID. Moreover, the reported protocol could also be utilized in other MS techniques that exploit post-IM/CID[209] to determine the presence of unresolved species and extract CCS information from subsequently deconvoluted spectra.

Acknowledgements

The authors thank Baylor University and National Science Foundation (NSF-IDBR, award number 1455668) for financial support. Support for Vanderbilt authors was provided by the Defense Threat Reduction Agency (Grant CBMXCEL-XL1=2=0001)

and the National Institutes of Health (Grants 2R01GM092218, 5UH3TR000491-04, and 3UH3TR000491-0451) and the National Science Foundation (CHE-1229341). The authors also thank Matthew Brantley for assistance with calculating CID mass spectra matching scores and Brooke Brown for assistance with running experimental replicates.

CHAPTER FIVE

Conclusion and Future Directions

Overview of Dissertation

Chapter One

Chapter One introduced the operating principles and applications of mass spectrometers and allied technologies (*e.g.*, ion mobility spectrometry (IMS)). Briefly, mass spectrometry (MS) is a technique which uses magnetic and/or electric fields to manipulate ionized gas-phase molecules' trajectories to measure their masses. Furthermore, we discussed how collision-induced dissociation (CID) and IMS can be leveraged to gain insight into the structures and conformations of biomolecules. Moreover, we introduced peptide fragment ion sequence-scrambling and chemometric data deconvolution of IM-unresolved isomers.

Chapter One also introduced a simple-to-use ionization technique, spray pump ionization (SPI), which can be likened to early nebulization-ionization IMS experiments.[23, 24] Nebulization-type ionization techniques, such as SPI, have fallen out of favor because of the availability of more efficient ionization methods[7] (*e.g.*, electrospray ionization (ESI)); however, the simplicity of nebulization-type ionization techniques suggest that these simple-to-use techniques might be useful in resource limited settings (*e.g.*, testing pharmaceutical integrity in the developing-world).[25] Our preliminary results suggest that SPI could be used to rapidly test for presence or absence of some pesticides in food products.[25] For example, Figure 5.1 shows the SPI mass spectra of corn husk with

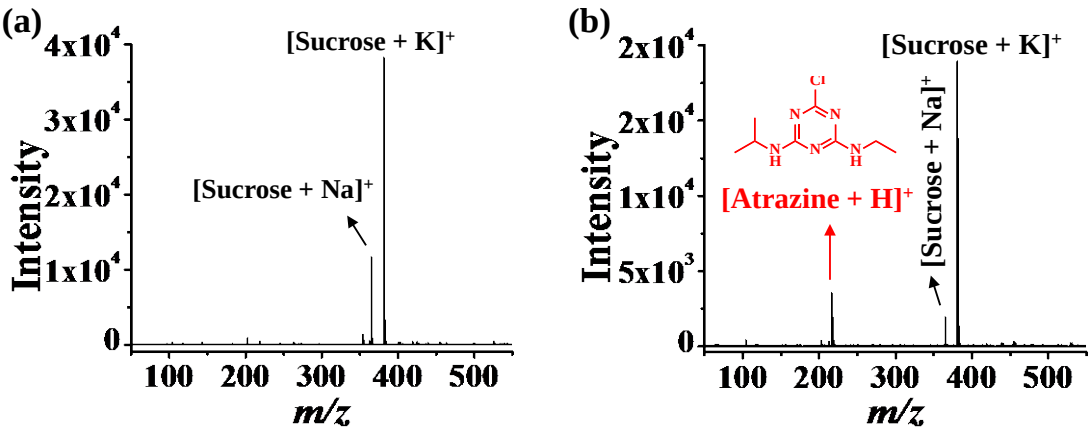


Figure 5.1. SPI mass spectra of corn husk (a) without and (b) with atrazine pesticide treatment. Corn husks were directly placed in the spray pump bottle which contained 0.1% acetic acid in methanol spray solution. Data was collected using a Synapt G2-S HDMS TOF mass spectrometer.

and without treatment with the pesticide atrazine. The spectra in Figure 5.1 were generated by placing the corn husk directly in spray pump bottles containing acidified methanol and nebulizing the mixture in front of a mass spectrometer. Although these results are promising, the primary disadvantage of nebulization-type ionization techniques is poor ionization efficiencies (*e.g.*, see discussion of Figure 1.2 above). Future studies could focus on improving the sensitivity of SPI, such as using an atmospheric pressure ion funnel[244-246] like the one shown in Figure 5.2.

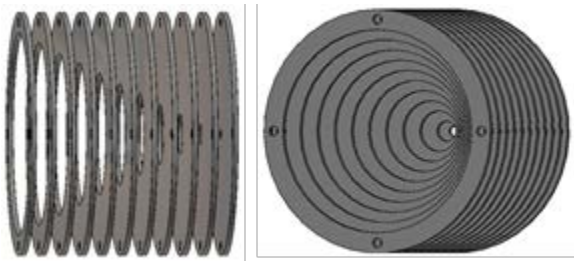


Figure 5.2. CAD rendering of an atmospheric pressure ion funnel[244-246] for enhancing the sensitivity of SPI MS.

Chapter Two

Sequence-scrambling of peptide fragment ions is a well-documented phenomenon that occurs as a result of gas-phase cyclization and reopening at internal amide bonds during CID.[71, 86-104] In 2012, we reported experimental evidence suggesting that, akin to b-type fragment ions, y-type fragment ions can undergo sequence-scrambling during CID.[119] In Chapter Two of this dissertation we presented evidence further confirming the propensity for y-type fragment ions to sequence-scramble, and identified a previously unknown intermediate step in the sequence-scrambling pathway.[55] Specifically, we showed that, upon CID, y-type peptide fragment ions, can lose internal backbone carbonyls (COs) from different positions and yield sequence-scrambled product fragment ions. We supported our proposed mechanism using results from isotope labeling, tandem mass spectrometry (MS^n), and IM-MS experiments. Moreover, we utilized chemometric deconvolution[62] of IM-unresolved des-acetylated-α-melanocyte-stimulating hormone (dα-MSH) (SYSMEHFRWGKPV-NH_2) fragment ions to demonstrate that some ions with a single-*m/z* and Gaussian-like IM distribution may in-fact correspond to mixtures of rearranged and sequence-intact fragment ions. Future studies may focus on the effects of amino acid sequence, charge state, and fragment ion size on the extent of internal carbonyl loss and potential implications for protein/peptide sequencing.

Chapter Three

Whereas Chapter Two focused on unusual rearrangement pathways in CID of peptide fragment ions,[55] Chapter Three extends those results to investigate a novel rearrangement pathway in CID of deoxyribonucleic acid (DNA) fragment ions.[115] We demonstrated that CID of *m/z*-isolated w type fragment ions and an intact 5′

phosphorylated DNA oligonucleotide could generate rearranged product ions. Eighteen out of twenty-one (~86%) studied w type ions of various nucleotide sizes, sequences, and charge states, generated rearranged product ions upon CID. Our experimental results and theoretical modeling data indicated that purine bases (*i.e.*, adenine and guanine) could attack the free 5′ phosphate group in w type ions (and 5′ phosphorylated DNA) to generate rearranged [phosphopurine]$^-$ fragment ions. Analysis of unusual oligonucleotide fragmentation pathways has not been frequently investigated and there are many unanswered questions that should be addressed. For example, the data presented in Chapter Three of this dissertation only focused on analysis of w-type fragment ions and 5′ phosphorylated DNA and did not address potential rearrangement pathways of other fragment ion types (*e.g.*, a-B, c, d, w, x, y, and z type fragment ions[114]) or ribonucleic acid (RNA) ions. Moreover, potential effects (positive or negative) of rearrangements on MS-based oligonucleotide sequencing efforts are currently unknown. Therefore, identification of DNA rearrangements could potentially open up a previously unrecognized field of oligonucleotide analysis.

Chapter Four

Chapters Two and Three discussed MS analysis of primary structures (*i.e.*, sequences of amino acids (Chapter Two) and nucleic acids (Chapter Three). In Chapter Four we progressed beyond the characterization of primary structures and developed a technique which could be used to investigate higher order (*i.e.*, secondary through quaternary) structures. Specifically, we used in-house developed automated IM deconvolution (AIMD) software[54] (see discussion of Figure 1.8 above) to extract IM profiles of IM-unresolved constitutional isomers (similar to results discussed in Chapter

Two[55] for analysis of IM-unresolved "a-like" peptide fragment ions (Figure 2.5)). Arrival times (ATs) of extracted IM profiles were then used to calculate accurate collision cross-section (CCS) values. CCS values obtained from deconvoluted IM spectra were not statistically different than CCS values measured from the individually analyzed peptides using both traveling wave (TW)IM-MS and drift tube (DT)IM-MS. Moreover, we introduced an approach to determine IM AT "shift factors" to compensate for post-IM/CID ion acceleration variations and further improve CCS calculations. To date, we have utilized TWIM-MS for all of our IM-MS deconvolution studies.[54-56, 62-64] Future studies could utilize DTIM-MS and chemometric deconvolution for analysis of IM-unresolved data to improve throughput of CCS calculations by eliminating the need to collect and analyze calibration data.[49]

Future Applications

All of the experiments discussed in this dissertation have focused on addressing fundamental MS questions, such as mechanisms and existence of rearrangement pathways and development of improved CCS calculations. These research endeavors were the necessary and required steps for successful characterization of complex sample mixtures and addressing biomedical research needs. Therefore, the next step in this research area is to apply the knowledge we have learned from these studies, and the techniques developed, to answer biomedically significant questions. Arguably, one of the most significant biomedical applications of MS is biomarker discovery.[247] Traditionally, MS-based biomarker discovery has been relegated to gas chromatography (GC) or liquid chromatography (LC)-MS experiments.[248] These approaches, however, have two primary limitations. First, traditional GC and LC runs can be time consuming (*e.g.*, >30 minutes

in many cases), which limits the throughput of analysis.[249] Second, for large biomolecules, LC-MS cannot be used to analyze secondary through quaternary structures. Therefore, if the biomarker has the same primary structure as a normal biomolecule but a different folding pattern (*e.g.*, prions), then it may not be identified. Using the techniques discussed in Chapters Two through Four, specifically using IM-MS, we should be able to identify biomarkers without the need for GC or LC, even if they correspond to conformational variations. To demonstrate the utility of these techniques for rapid biomarker analysis we analyzed the total cellular extract of fresh bovine cricoarytenoid muscle.

Bovine cricoarytenoid muscle cellular extracts were prepared by first flash freezing the tissue with liquid nitrogen (LN_2) and crushing the frozen tissue using a ceramic mortar and pestle to lyse the cells. Next, the defrosted-liquefied tissue was centrifuged at 20xG for 5 minutes at 4˚C to pellet connective tissue, unlysed cells, and other large material. The supernatant was then removed and placed in 50 kDa centrifuge cutoff filters (EMD Millipore, Darmstadt, Germany) and centrifuged at 10xG for 10 minutes at 4˚C to remove large proteins, residual connective tissue, and unlysed cells. Following filtration, samples were desalted using C_{18} ZipTips (EMD Millipore). Finally, purified biomolecules were dissolved in ESI solvent (*i.e.*, water (49.95): methanol (49.95): acetic acid (0.1)) and directly analyzed using a Synapt G2-S HDMS IM-mass spectrometer (Waters Corporation, Milford, MA, USA).

Figure 5.3a and b, show the total summed (five minutes data collection) direct infusion ESI mass spectra of bovine cricoarytenoid muscle extract before and after desalting with C_{18} ZipTips, respectively. Use of ZipTips increased signal intensity by

about an order of magnitude for some proteins (*e.g.*, Figure 5.3c and d). The spectra in Figure 5.3 correspond to the total cellular extract of bovine cricoarytenoid muscle. Without separation (*e.g.*, LC) to reduce sample complexity, presence of high abundance proteins in muscle (*e.g.*, myoglobin [Mb]) suppresses signal intensity of lower abundance biologically relevant molecules. Rather than using LC to fractionate the cellular extract, we utilized IM which separates molecules in < 14 ms. Figure 5.4 shows a heat map of IM arrival time versus *m/z* for the same (desalted) sample in Figure 5.3b, generated using

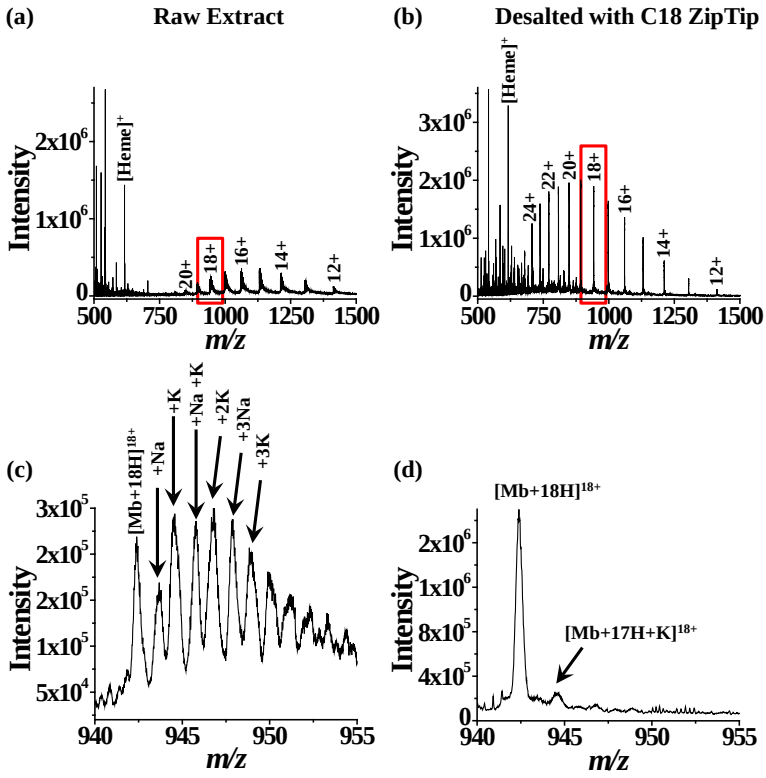


Figure 5.3. Total summed mass spectra before (a and c) and after (b and d) desalting and protein concentration using C_{18} ziptips. Panels (c) and (d) show the zoomed in mass spectra (from *m/z* 940 to 955) of the area highlighted by red boxes in (a) and (b), respectively corresponding to the 18+ charge state for bovine myoglobin extracted from cricoarytenoid muscle.

DriftScope software (Waters Corporation). From the spectra in Figure 5.4 we can identify the presence of numerous unique ions from the cricoarytenoid muscle extract.

Using the data analysis tools in DriftScope we can remove high intensity proteins or contaminants to identify low intensity biomolecule ions. For example, Figure 5.5 shows the ESI summed mass spectra of desalted cricoarytenoid muscle extract before (a and c) and after (c and d) removing the polyethylene glycol (PEG) and Mb series highlighted in Figure 5.4. Removal of high intensity PEG and Mb signals revealed the presence of multiple proteins which could not be previously identified. For example, Figure 3c and d show the zoomed in mass spectra from *m/z* 938 to 952 before and after PEG and Mb removal in DriftScope. After removal of the Mb peak in Figure 5.5c, the presence of several, otherwise unobservable, proteins can be identified in the mass spectrum (Figure 5.5d).

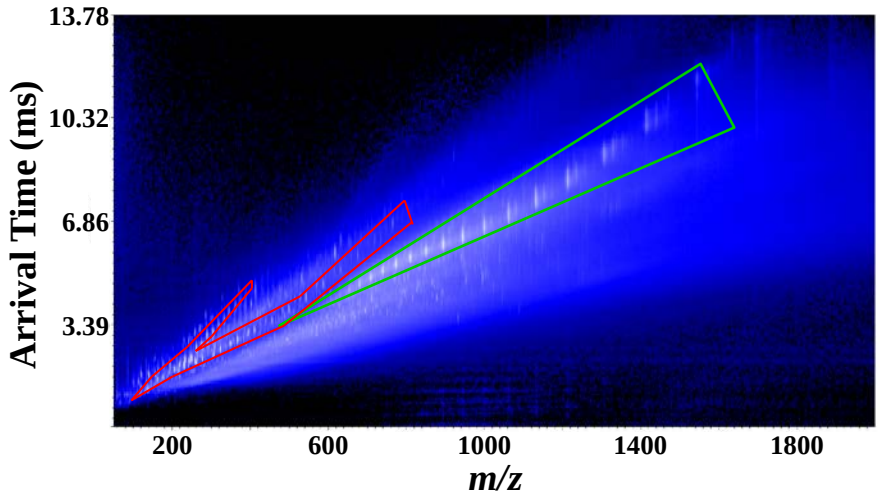


Figure 5.4. DriftScope (Waters Corporation) heat map of IM arrival time versus *m/z* for total bovine cricoarytenoid muscle extract. Color scale: black → blue → white; where black is the lowest (*i.e.*, zero) ion intensity and white is the highest ion intensity. Area outline in red and green correspond to polyethylene glycol (PEG; a ubiquitous polymer) and myoglobin (Mb) series, respectively.

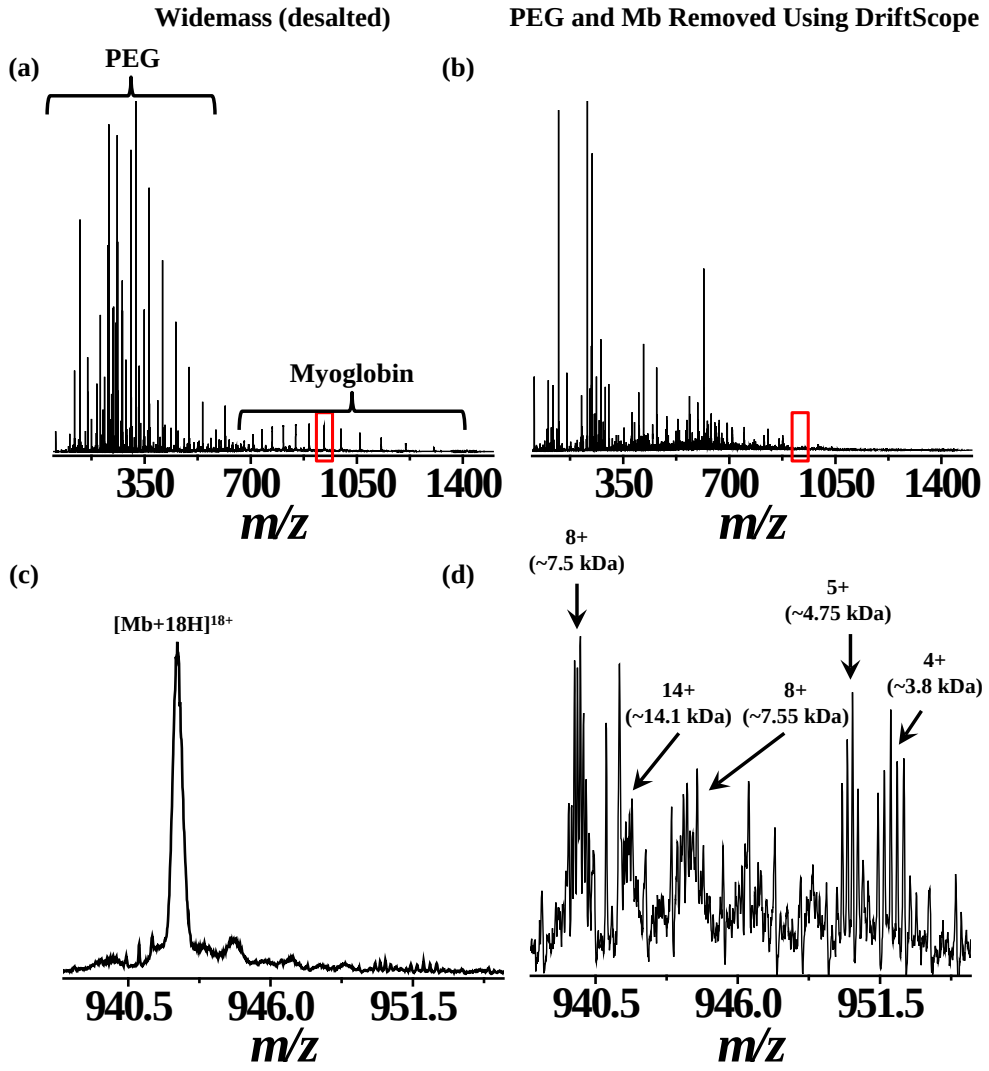


Figure 5.5. ESI mass spectra of bovine cricoarytenoid muscle extract before (a and c) and after (c and d) removing the PEG and Mb series highlighted in Figure 2. Panels (c) and (d) show the zoomed in mass spectra (from *m/z* 938 to 952) of the area highlighted by red boxes in (a) and (b), respectively showing the presence of numerous proteins that were suppressed by the high intensity Mb peak.

Preliminary results from analysis of bovine cricoarytenoid muscle extract suggest that we can successfully extract, purify, concentrate, and analyze abundant proteins using the proposed technique. Using IM we are able to significantly reduce the amount of time necessary for sample characterization to a few minutes. Rapid tissue processing and examination using the proposed technique will enable high-throughput clinical

pathological analysis of normal and diseased thyroids, which could potentially be used to direct treatment and save lives. Moreover, using our in-house developed and commercial software, we are able to identify isomeric and low abundance biomarkers that cannot otherwise be identified.

APPENDICES

APPENDIX A

Loss of Internal Backbone Carbonyls: Additional Evidence for Sequence-scrambling in Collision-induced Dissociation of y-Type Ions

Table A.1. Summary of experimental *m/z* values, mass measurement errors (ppm), and identities of a-"like" product fragment ions generated from CID of *m/z*-isolated $[y_n]^{m+}$ (n = 6-12; m = 1, 2, and/or 3) species of des-acetylated-α-melanocyte-stimulating hormone and $[y_5]^{2+}$ and $[y_6]^{+}$ species of angiotensin II antipeptide in the LTQ Orbitrap from (JASMS, 2013, 24, 1755-1766)

Isolated y Fragment	Identity of Fragment Ion	Experimental *m/z*	Error (ppm)
Des-acetylated-α-melanocyte-stimulating Hormone (SYSMEHFRWGKPV-amidated)			
$[y_{12}]^{3+}$	$[y_{12} - (NH_3 + {}^{12}CO)]^{3+}$	497.584	0.26
$[y_{11}]^{3+}$	$[y_{11} - (NH_3 + {}^{12}CO)]^{3+}$	443.229	0.09
$[y_{11}]^{2+}$	$[y_{11} - (NH_3 + {}^{12}CO)]^{2+}$	664.340	0.42
$[y_{10}]^{2+}$	$[y_{10} - (NH_3 + {}^{12}CO)]^{2+}$	620.824	0.45
$[y_9]^{2+}$	$[y_9 - (NH_3 + {}^{12}CO)]^{2+}$	555.304	0.14
$[y_8]^{3+}$	$[y_8 - (NH_3 + {}^{12}CO)]^{3+}$	327.524	0.09
$[y_7]^{+}$	$[y_7 - (NH_3 + {}^{12}CO)]^{+}$	422.253	0.00
$[y_6]^{2+}$	$[y_6 - (NH_3 + {}^{12}CO)]^{2+}$	348.719	0.17
Angiotensisn II Antipeptide (EGVYVHPV)			
$[y_7]^{+}$	$[y_7 - (NH_3 + {}^{12}CO)]^{+}$	724.414	0.52
$[y_5]^{2+}$	$[y_5 - (NH_3 + {}^{12}CO)]^{2+}$	284.666	0.18
$[y_5]^{2+}$	$[y_5 - (NH_3 + {}^{12}CO)]^{2+}$	285.158	0.04

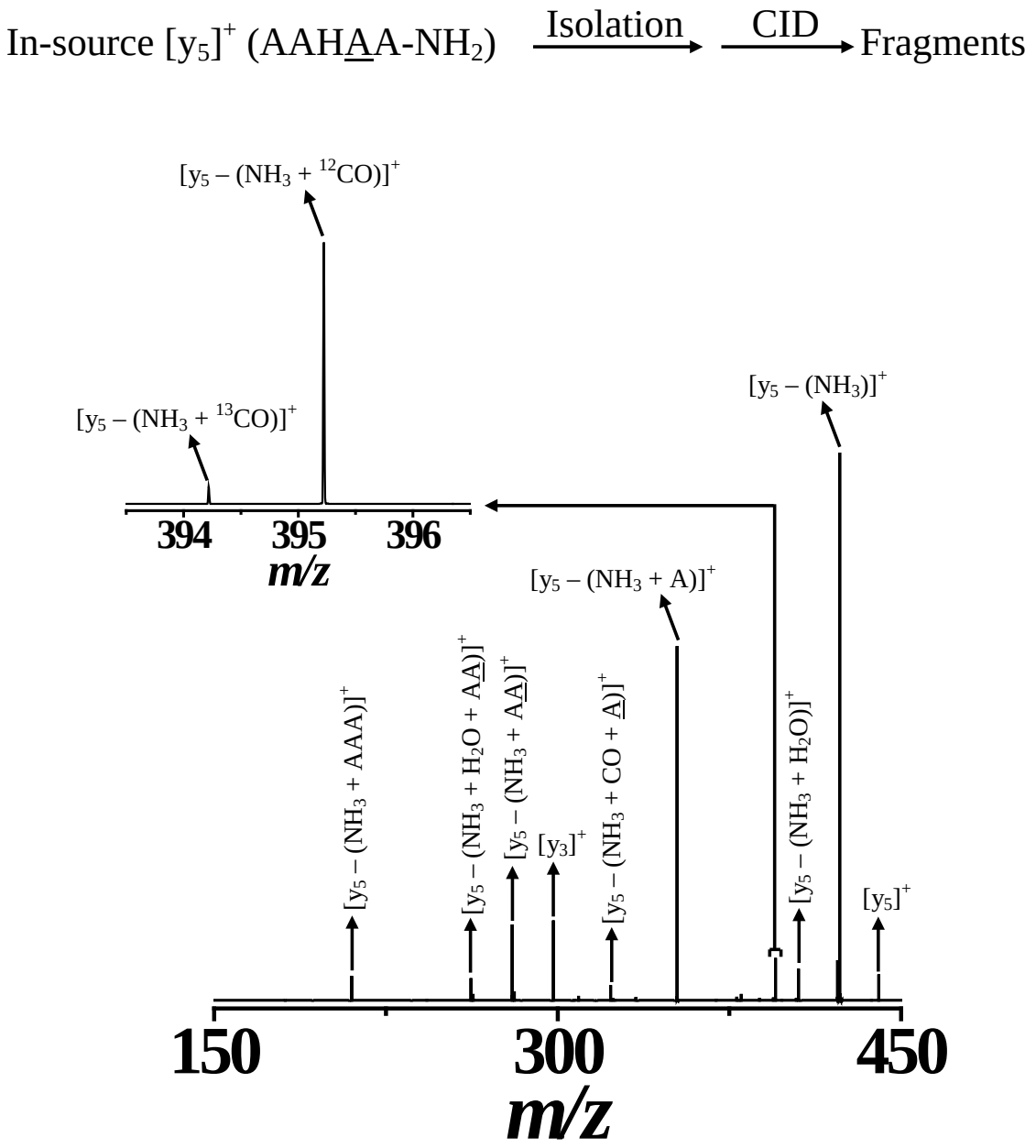


Figure A.1. LTQ Orbitrap collision-induced dissociation (CID) mass spectrum of *m/z*-isolated $[y_5]^+$ (*m/z* 440) (generated from in-source fragmentation of protonated AAAAHAA-NH_2). Inset shows the expanded view of the *m/z* range from 393.5 to 396.5 corresponding to $[y_5 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 395) and rearranged product fragment ion $[y_5 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 394).

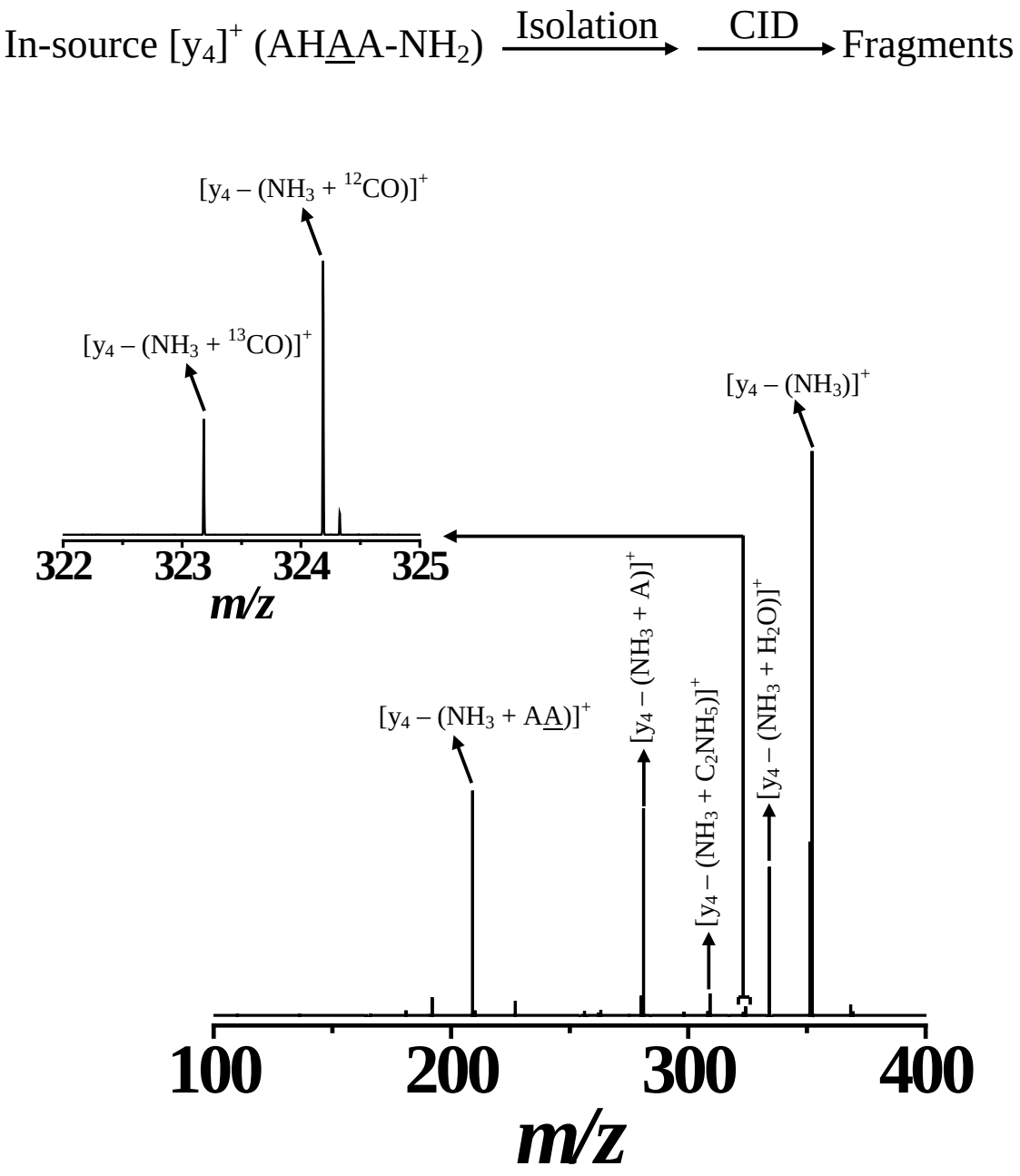


Figure A.2. LTQ Orbitrap collision-induced dissociation (CID) mass spectrum of *m/z*-isolated $[y_4]^+$ (*m/z* 369) (generated from in-source fragmentation of protonated AAAAHAA-NH_2). Inset shows the expanded view of the *m/z* range from 322 to 325 corresponding to $[y_4 – (NH_3 + {}^{12}CO)]^+$ (*m/z* 324) and rearranged product fragment ion $[y_4 – (NH_3 + {}^{13}CO)]^+$ (*m/z* 323).

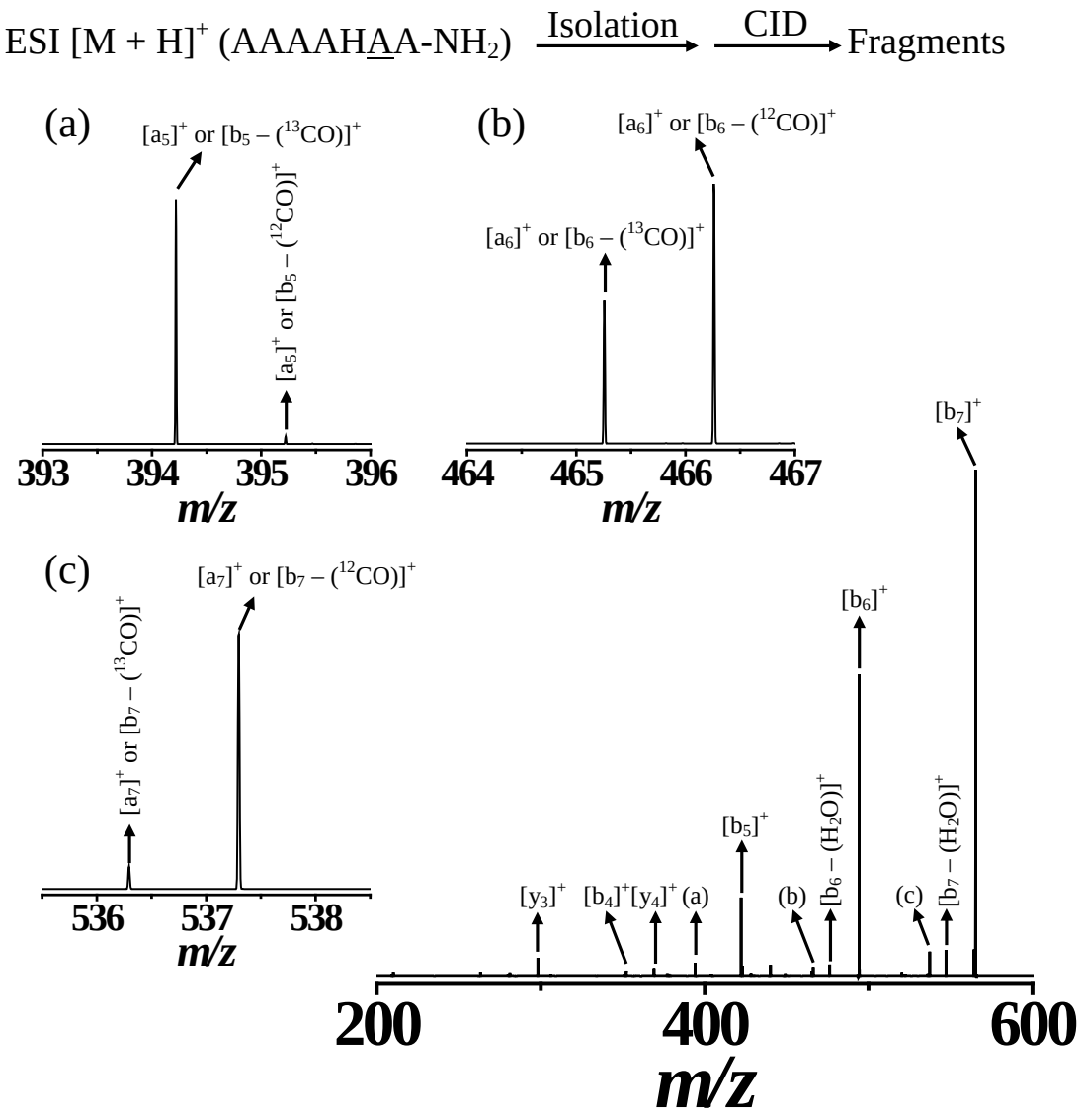


Figure A.3. LTQ Orbitrap collision-induced dissociation (CID) mass spectrum of *m/z*-isolated protonated molecular ion of AAAAHAA-NH_2 (*m/z* 582). Insets show the expanded views of the *m/z* ranges from (a) 393 to 396 corresponding to $[a_5]^+$ ($[b_5 - (^{12}CO)]^+$) (*m/z* 395) and rearranged product fragment ion $[a_5]^+$ ($[b_5 - (^{13}CO)]^+$) (*m/z* 394), (b) 464 to 467 corresponding to $[a_6]^+$ ($[b_6 - (^{12}CO)]^+$) (*m/z* 466) and rearranged product fragment ion $[a_6]^+$ ($[b_6 - (^{13}CO)]^+$) (*m/z* 465), and (c) 535.5 to 538.5 corresponding to $[a_7]^+$ ($[b_7 - (^{12}CO)]^+$) (*m/z* 537) and rearranged product fragment ion $[a_7]^+$ ($[b_7 - (^{13}CO)]^+$) (*m/z* 536).

Please note the change in ratios of ^{13}CO to ^{12}CO losses as a function of fragment ion size (Figure A.3, insets). In smaller a-type ions, there is a higher probability of losing

the ^{13}CO (or the entire labeled alanine (*i.e.*, A)) because fewer unlabeled backbone carbonyl groups are available. Because the precursor peptide contains five isobaric alanines, there are (at least) two possible pathways for generation of the a-type ions in Figure A.3. Pathway I: Isolation of $[M + H]^+$ (*m/z* 582) → CID → Generation of Normal a-Type Ions. Pathway II: Isolation of $[M + H]^+$ (*m/z* 582) → CID → Generation of y-Type Ions → Gas-phase Rearrangement → Generation of a-"Like" Ions *via* main text Scheme 1 and/or 2. Results from CID of *m/z*-isolated y-type ions (Figure 2.1, A.1, and A.2) suggest that pathway II may contribute to the CID mass spectrum of protonated molecular ions. Therefore, sequence-scrambling through gas-phase dissociation of y-type ions should not be negated.

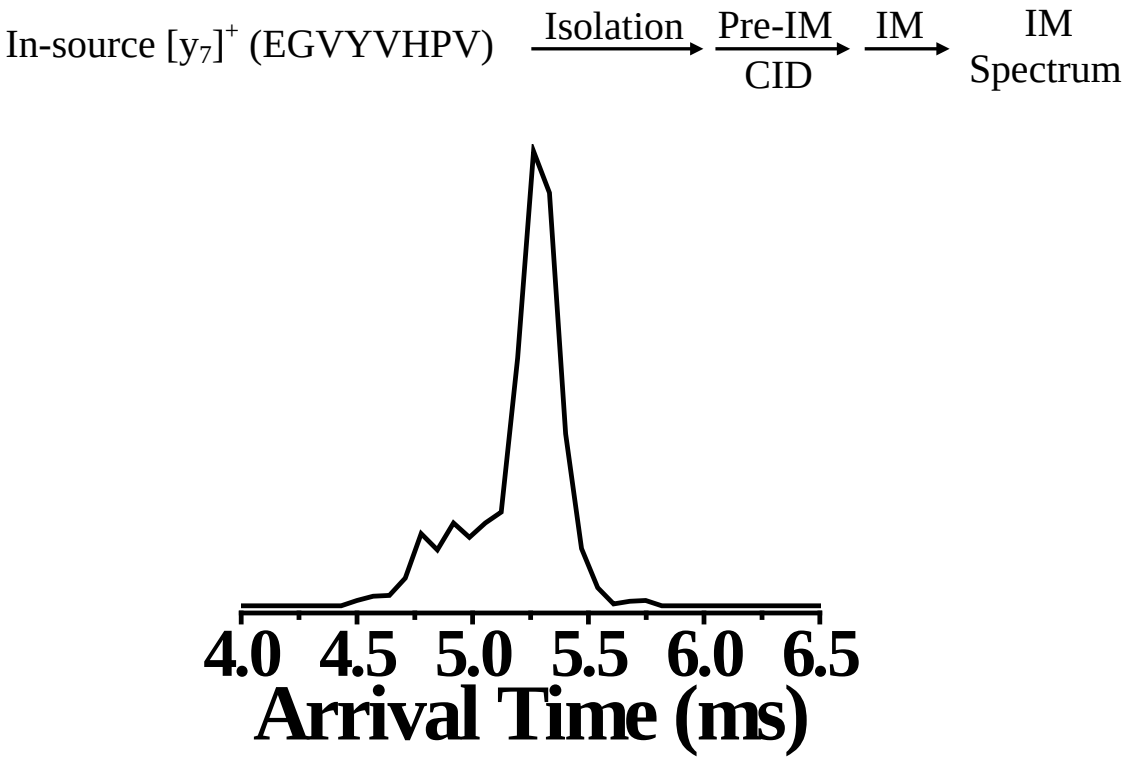


Figure A.4. Synapt G2-S HDMS ion mobility (IM) profile showing multiple structural isomers for $[y_7 - (H_2O + {}^{12}CO)]^+$ (*m/z* 724) generated from pre-IM CID of $[y_7]^+$ (*m/z* 770) from angiotensin II antipeptide (amino acid sequence: EGVYYVHPV).

Figure A.4 shows the selected IM profile for *m/z* 724 generated from pre-IM CID of *m/z*-isolated $[y_7]^+$ (*m/z* 770) from angiotensin II antipeptide. Specifically, product fragment ions from $[y_7]^+$ of angiotensin II antipeptide were generated prior to IM separation (pre-IM CID) in the trap cell using collision energies corresponding to a 38 V potential difference between the trap cell exit and entrance of the helium cell. The potential difference between the exit of the IM cell and entrance of the transfer cell was set at 2.0 V to prevent post-IM fragmentation of ions of interest. The IM profile in Figure A.4 corresponds to the average of 2400 scans. All other instrument parameters were kept similar to those used for IM-MS analysis of des-acetylated-α-melanocyte-stimulating hormone (see the experimental section of Chapter Two for details).

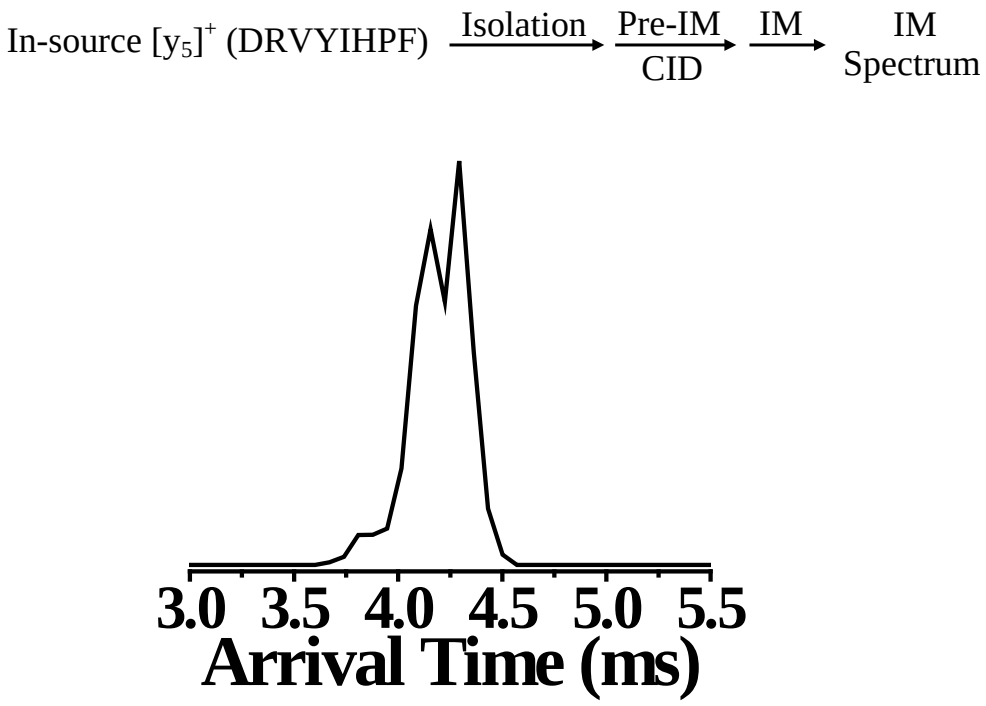


Figure A.5. Synapt G2-S HDMS ion mobility (IM) profile showing multiple structural isomers for $[y_5 - (H_2O + {}^{12}CO)]^+$ (*m/z* 630) generated from pre-IM CID of $[y_5]^+$ (*m/z* 676) from angiotensin II, human (amino acid sequence: DRVYIHPF).

To generate the IM profile in Figure A.5, product fragment ions from $[y_5]^+$ of angiotensin II were generated prior to IM separation (pre-IM CID) in the trap cell using collision energies corresponding to a 30 V potential difference between the trap cell exit and entrance of the helium cell. The potential difference between the exit of the IM cell and entrance of the transfer cell was set at 2.0 V to prevent post-IM fragmentation of ions of interest. IM profile corresponds to the average of 1200 scans. Source offset was set to 85 V. Nitrogen (N_2) gas pressure in the IM cell was ~4.0 mbar. IM wave velocity was set to 500 m/s. All other instrument parameters were kept similar to those used for IM-MS analysis of des-acetylated-α-melanocyte-stimulating hormone (see the experimental section of Chapter Two for details).

APPENDIX B

DNA Oligonucleotide Fragment Ion Rearrangements Upon Collision-induced Dissociation

Appendix B Experimental

Triple Quadrupole Mass Spectrometer

Triple quadrupole mass spectrometry (MS) experiments (Figure B.1) were conducted using a Waters Xevo TQ-S (Waters Corporation, Manchester, UK). Sample flow rate was set at 10 μL/min. Cone and desolvation gas (nitrogen [N_2]) flows were set at 150 L/min and 1000 L/min, respectively. Nebulizer gas (N_2) pressure was set at 7.00 bar. Desolvation and source temperatures were set at 500 °C and 150 °C, respectively. The software defined capillary, sampling cone, and source offset voltages were set at -2.5 kV, -80.0 V, and -50.0 V, respectively. Ion isolation parameters (*i.e.*, quadrupole rf/dc) were optimized to isolate the isotopic packet of $[w_5]^{2-}$ (*m/z* 802) generated from in-source CID of pET Upstream DNA. Ions of interest were mass isolated in the first quadrupole (Q1) and fragmented in the ScanWave™ collision cell by setting a 25.0 V potential difference between Q1 and the analyzing quadrupole (Q3). Collision gas (argon) flow was set at 0.18 mL/min. The mass spectrum in Figure S1 corresponds to the average of 299 scans (5 minutes run time). Data were analyzed using MassLynx software (version 4.1) (Waters Corporation, Manchester, UK).

Linear Trap Quadrupole Orbitrap Mass Spectrometer

Orbital trap MS experiments in Figure B.2 were conducted using a Linear Trap Quadrupole (LTQ) Orbitrap Discovery MS (Thermo Fisher Scientific, Waltham, MA, USA). Sample concentration was 1 x 10^{-6} M. Doubly-charged Oligo-2 (*i.e.*, [Oligo-2 – 2H]$^{-}$) at *m/z* 802 was mass isolated in the LTQ and collisionally excited using a normalized collision energy [134] of 20%. Fragment ions generated in the LTQ were then sent to the Orbitrap for high mass accuracy analysis. All other instrument conditions were the same as those reported for Chapter Three, Figure 3.4.

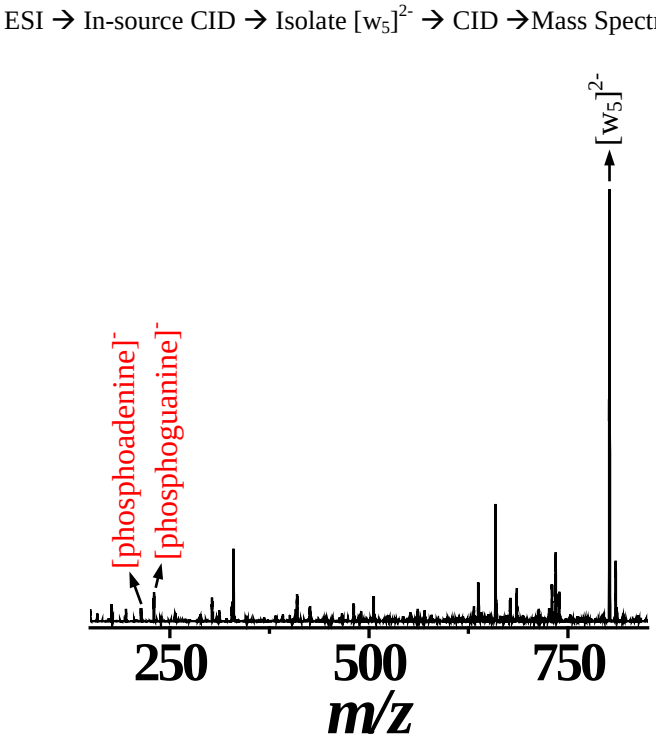


Figure B.1. CID mass spectrum of *m/z*-isolated $[w_5]^{2-}$ (*m/z* 802) generated from in-source CID of Oligo-1 using a Waters Xevo TQ-S (Waters Corporation, Manchester, England, UK). Only [phosphopurine]$^{-}$ ions (red font) and precursor $[w_5]^{2-}$ ions are labeled.

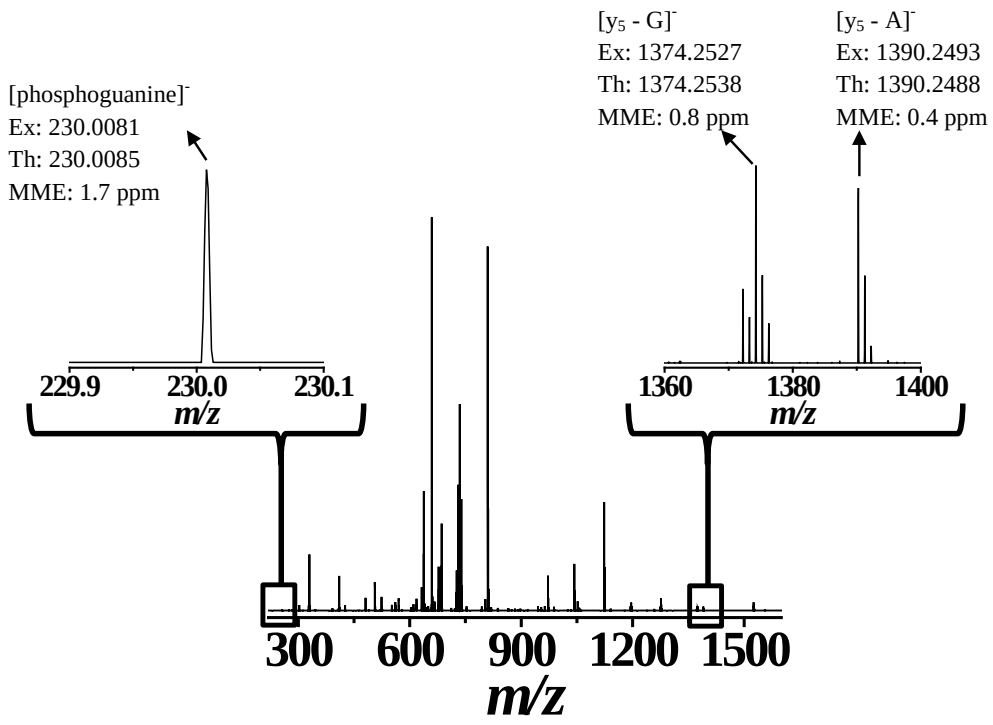


Figure B.2. CID mass spectrum of *m/z*-isolated $[Oligo\text{-}2 - 2H]^{2-}$ (*m/z* 802; sequence: p-GTAGA) using a Linear Trap Quadrupole (LTQ) Orbitrap Discovery MS (Thermo Fisher Scientific, Waltham, MA, USA). Instrument low mass limit of *m/z* 220 (for MS^n experiments) prevented the inclusion of $[phosphoadenine]^-$ (*m/z* 214) in the CID mass spectrum; however, complementary $[y_5 – A]^-$ was observed at *m/z* 1390, suggesting that $[phosphoadenine]^-$ was also formed but not detected. Experimental (Ex) and Theoretical (Th) masses as well as mass measurement errors (MME, reported in ppm) for $[phosphoguanine]^-$ and complementary y-B ions are labeled. Insets for *m/z* ranges corresponding to $[phosphoguanine]^-$ (*m/z*: 299.9 to 230.1) and complementary y-B ions (*m/z*: 1360 to 1400) are shown at ~1000x and ~100x y-scale expansions, respectively.

APPENDIX C

Calculating Accurate Collision Cross-sections of Ion Mobility Unresolved Isomers Using Tandem Mass Spectrometry and Chemometric Deconvolution

Table C.1. Instrumental Operating Conditions for Ion Mobility-Mass Spectrometry (IM-MS) Experiments

Experimental Parameter	Synapt G1 (Reverse Pentapeptides)	Synapt G2-S (Enkephalin Hexapeptides)
Time-of-Flight (TOF)	"V" Mode	"V" Mode
IM Wave Velocity	300 m/s	1200–1700 m/s (100 m/s steps)
IM Wave Height	7.0–8.0 V (0.2 V steps)	40.0 V
Nitrogen (N_2) Pressure: Trap Cell*	~1.8×10^{-5} bar	~1.6×10^{-5} bar
Nitrogen (N_2) Pressure: Transfer Cell*	~1.8×10^{-5} bar	~1.9×10^{-5} bar
Nitrogen (N_2) Pressure: IM Cell*	~5.3×10^{-4} bar	~3.8×10^{-3} bar
Helium (He) Pressure: He Cell*	—	~1.4 bar
Post-IM Acceleration Voltage (Low-energy Experiments)	4.0 V	4.0 V
Post-IM Acceleration Voltage (High-energy Experiments)	40.0 V	22.0 V

*All gas-pressures are reported from direct instrument readouts without correcting for sensitivity[164] or geometry factors[64, 163]

Summary of Statistical Tests Discussed in Chapter Four

ATs for pure and deconvoluted IM profiles (Figure 4.2):

SDGRG: type 2 Student's t test for equal variance, $n_{pure} = 5$, $n_{decon} = 4$, $t_{calc} = 5.17$, $t_{95\%} = 2.365$

GRGDS: type 2 Student's t test for equal variance, $n_{pure} = 5$, $n_{decon} = 4$, $t_{calc} = 2.81$, $t_{95\%} = 2.365$

ATs for pure and AT-corrected deconvoluted IM profiles (Figure 4.2):

SDGRG: type 2 Student's t test for equal variance, $n_{pure} = 5$, $n_{decon} = 4$, $t_{calc} = 2.14$, $t_{95\%} = 2.365$

GRGDS: type 2 Student's t test for equal variance, $n_{pure} = 5$, $n_{decon} = 4$, $t_{calc} = 0.542$, $t_{95\%} = 2.365$

TWIM-MS CCSs calculated using pure and deconvoluted ATs (Table 4.1):

SDGRG: type 2 Student's t test for equal variance, $n_{pure} = 25$, $n_{decon} = 24$, $t_{calc} = 1.986$, $t_{95\%} = 2.012$, $t_{90\%} = 1.678$

GRGDS: type 2 Student's t test for equal variance, $n_{pure} = 25$, $n_{decon} = 24$, $t_{calc} = 1.436$, $t_{95\%} = 2.012$, $t_{90\%} = 1.678$

TWIM-MS CCSs calculated using pure and corrected-deconvoluted ATs (Table 1):

SDGRG: type 2 Student's t test for equal variance, $n_{decon} = 24$, $n_{pure} = 25$, $t_{calc} = 0.331$, $t_{95\%} = 2.012$, $t_{90\%} = 1.678$

GRGDS: type 2 Student's t test for equal variance, $n_{decon} = 24$, $n_{pure} = 25$, $t_{calc} = 0.00546$, $t_{95\%} = 2.012$, $t_{90\%} = 1.678$

TWIM-MS CCS of early AT conformer and pure $[RYGGFM + 2H]^{2+}$ (Table 2):

type 2 Student's t test for equal variance, $n_{earlyAT}$ = 24, n_{RYGGFM} = 24, t_{calc} = 0.841, $t_{95\%}$ = 2.013, $t_{90\%}$ = 1.679

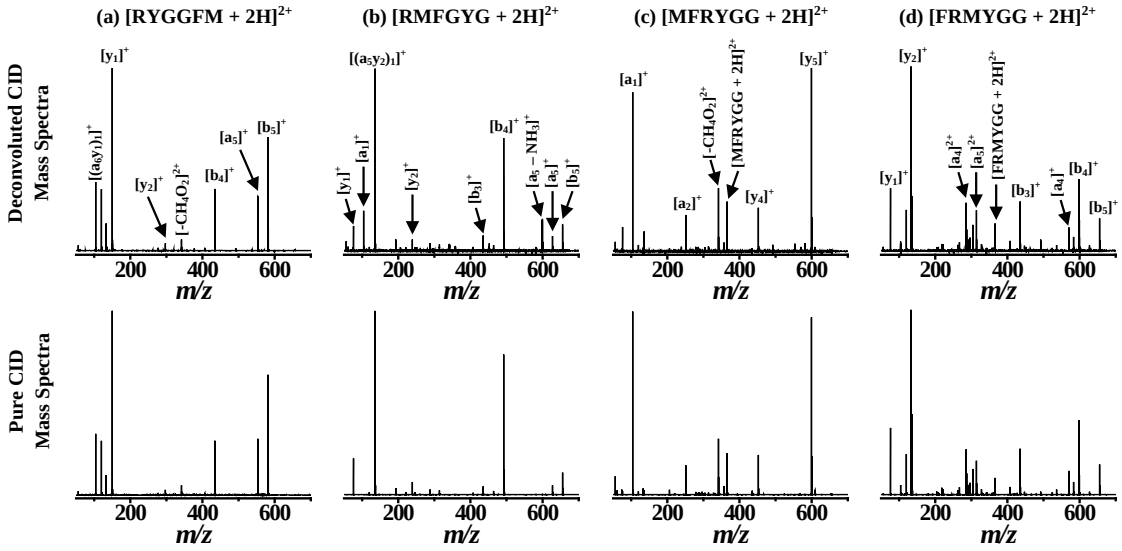


Figure C.1. Deconvoluted (top row) and pure (bottom row) CID mass spectra of (a) $[RYGGFM + 2H]^{2+}$, (b) $[RMFGYG + 2H]^{2+}$, (c) $[MFRYGG + 2H]^{2+}$ and (d) $[FRMYGG + 2H]^{2+}$. Deconvoluted CID mass spectra (top row) correspond to the deconvoluted IM profiles in Fig. 4 of the main text. Select fragment ions are labeled. Please note that fragment ion intensities (*e.g.*, $[a_1]^+$ from $[RMFGYG + 2H]^{2+}$ in (b)) for some deconvoluted *m/z* values (top row) are different than the respective fragment ion intensities in the pure CID mass spectra (bottom row). Slight intensity differences between the deconvoluted (top row) and pure (bottom row) CID mass spectra can be observed for fragment ions that belong to two or more IM unresolved peptide isomers.[215]

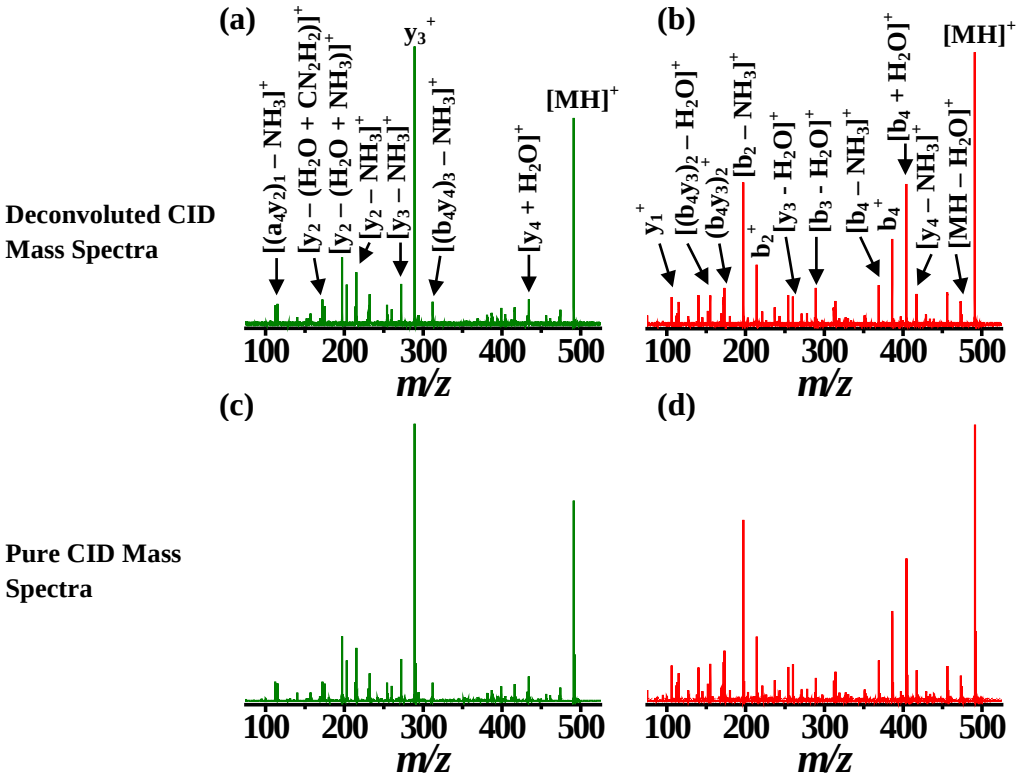


Figure C.2. Deconvoluted post-IM/CID mass spectra from a mixture of (a) [SDGRG + H]$^+$ (green) and (b) [GRGDS + H]$^+$ (red) (corresponding to the deconvoluted IM profiles in main text Fig. 1) compared to the individually analyzed post-IM/CID mass spectra for (c) pure [SDGRG + H]$^+$ and (d) pure [GRGDS + H]$^+$. Select high relative intensity ions are labeled in (a) and (b).

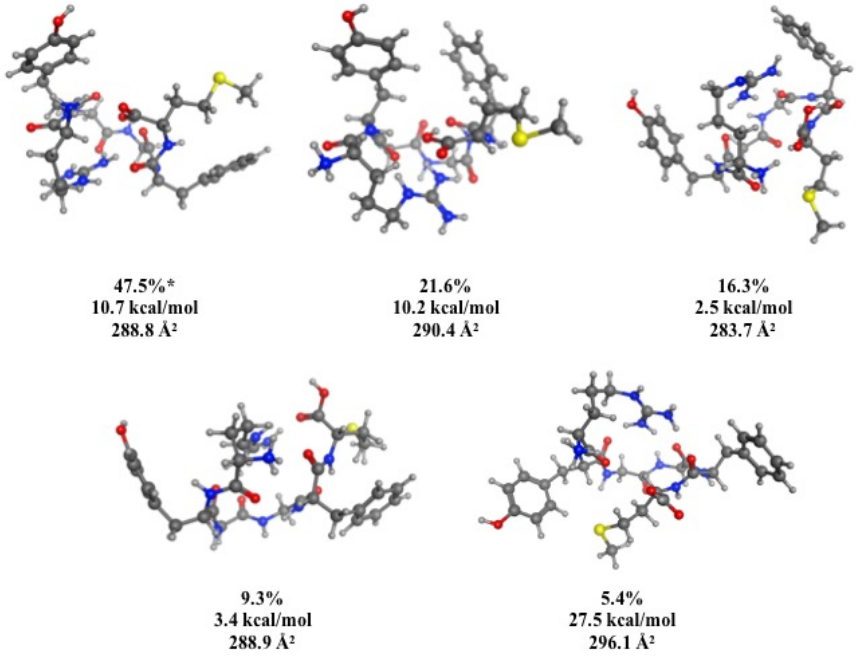


Figure C.3. Theoretical structures from within the experimental CCS range for $[RYGGFM + 2H]^{2+}$. Each structure is annotated with the percentage of conformations it represents from RMSD clustering analysis, its theoretical relative energy, and theoretical CCS value. The most populous conformation was shown in Figure 5 in the manuscript.

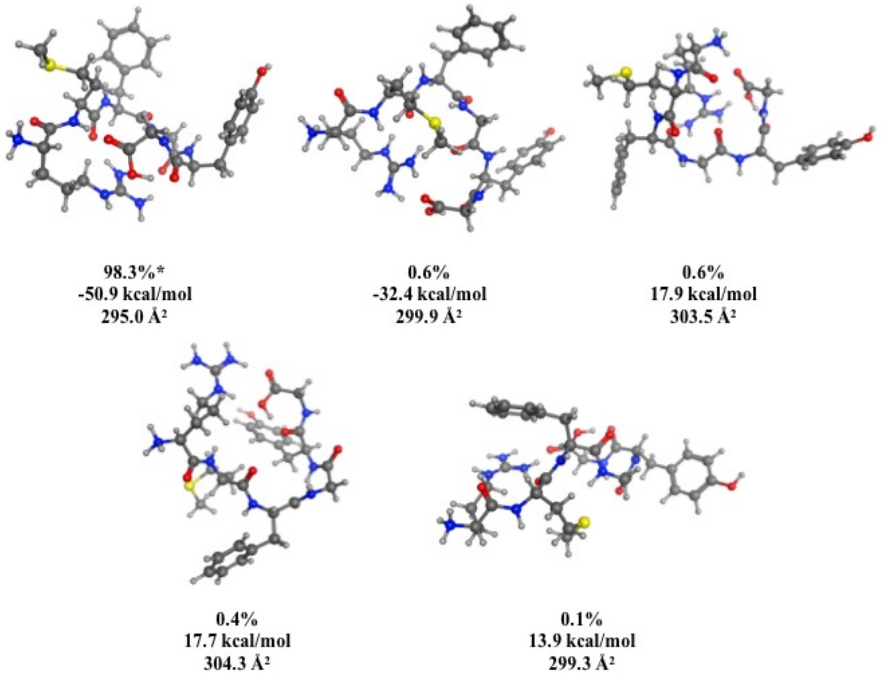


Figure C.4. Theoretical structures from within the experimental CCS range for $[RMFGYG + 2H]^{2+}$. Each structure is annotated with the percentage of conformations it represents from RMSD clustering analysis, its theoretical relative energy, and theoretical CCS value. The most populous conformation was shown in Figure 5 in the manuscript.

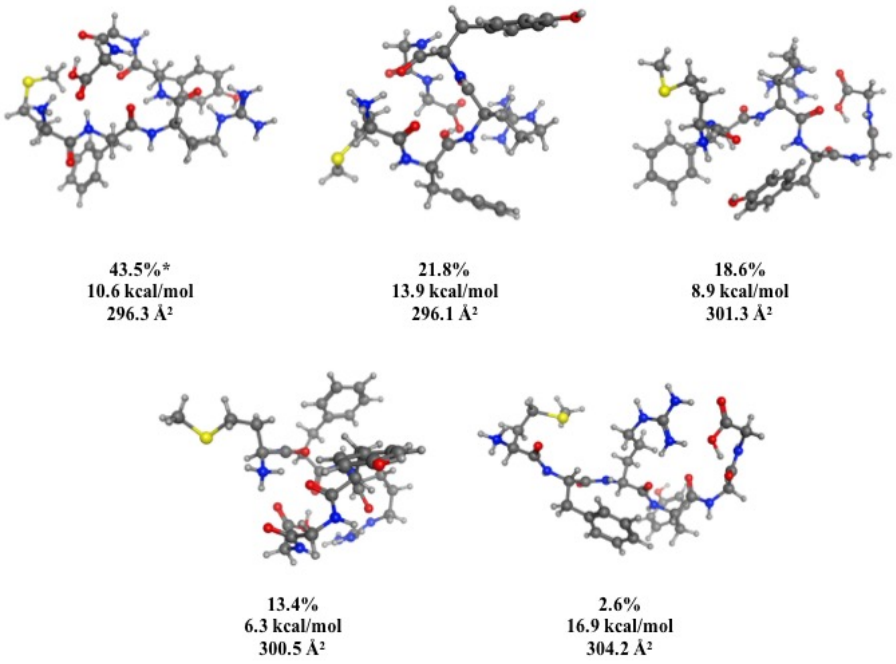


Figure C.5. Theoretical structures from within the experimental CCS range for $[MFRYGG + 2H]^{2+}$. Each structure is annotated with the percentage of conformations it represents from RMSD clustering analysis, its theoretical relative energy, and theoretical CCS value. The most populous conformation was shown in Figure 5 in the manuscript.

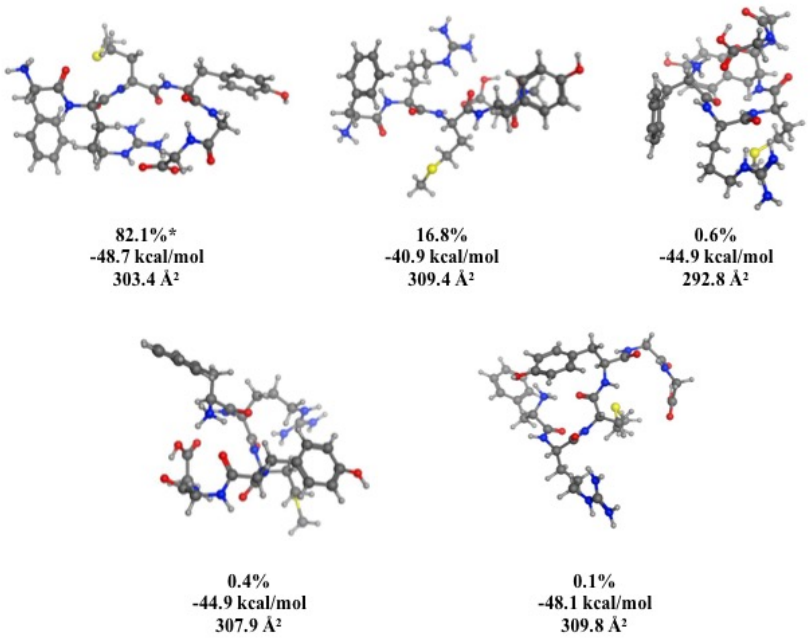


Figure C.6. Theoretical structures from within the experimental CCS range for $[FRMYGG + 2H]^{2+}$. Each structure is annotated with the percentage of conformations it represents from RMSD clustering analysis, its theoretical relative energy, and theoretical CCS value. The most populous conformation was shown in Figure 5 in the manuscript.

BIBLIOGRAPHY

1. Szulejko, J.; Solouki, T., Potential analytical applications of interfacing a gc to an FT-ICR MS: Fingerprinting complex sample matrixes. *Anal. Chem.* **2002,** *74* (14), 3434-3442.

2. Vékey, K., Internal energy effects in mass spectrometry. *J. Mass Spectrom.* **1996,** *31* (5), 445-463.

3. Dempster, A. J., The properties of slow canal rays. *Phys. Rev.* **1916,** *8* (6), 651-662.

4. Kind, T.; Fiehn, O., Advances in structure elucidation of small molecules using mass spectrometry. *Bioanal. Rev.* **2010,** *2* (1-4), 23-60.

5. Nier, A. O., A mass spectrometer for routine isotope abundance measurements. *Rev. Sci. Instrum.* **1940,** *11* (7), 212-216.

6. Dole, M.; Mack, L. L.; Hines, R. L.; Mobley, R. C.; Ferguson, L. D.; Alice, M. B., Molecular beams of macroions. *J. Chem. Phy.* **1968,** *49* (5), 2240-2249.

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

8. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization–principles and practice. *Mass Spectrom. Rev.* **1990,** *9* (1), 37-70.

9. Yoshida, Y.; Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, T. Development of laser ionization time of flight mass spectrometer-curved electric field ion reflector, *12th Procedings of the Japanese Society for Medical Mass Spectrometry* **1987,** Kyoto, Japan.

10. Tanaka, K.; Waki, H.; Ido, Y.; Akita, S.; Yoshida, Y.; Yoshida, T.; Matsuo, T., Protein and polymer analyses up to *m/z* 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Commun. Mass Spectrom.* **1988,** *2* (8), 151-153.

11. Karas, M.; Bachmann, D.; Bahr, U. e.; Hillenkamp, F., Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int. J. Mass Spectrom.* **1987,** *78*, 53-68.

12. Griffiths, J., A brief history of mass spectrometry. *Anal. Chem.* **2008,** *80* (15), 5678-5683.

13. Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S., Unraveling the mechanism of electrospray ionization. *Anal. Chem.* **2013,** *85* (1), 2-9.

14. Konermann, L.; Rodriguez, A. D.; Liu, J., On the formation of highly charged gaseous ions from unfolded proteins by electrospray ionization. *Anal. Chem.* **2012,** *84* (15), 6798-6804.

15. Ahadi, E.; Konermann, L., Modeling the behavior of coarse-grained polymer chains in charged water droplets: Implications for the mechanism of electrospray ionization. *J. Phys. Chem. B* **2012,** *116* (1), 104-112.

16. Kebarle, P.; Verkerk, U. H., Electrospray: From ions in solution to ions in the gas phase, what we know now. *Mass Spectrom. Rev.* **2009,** *28* (6), 898-917.

17. Iavarone, A. T.; Williams, E. R., Mechanism of charging and supercharging molecules in electrospray ionization. *J. Am. Chem. Soc.* **2003,** *125* (8), 2319-2327.

18. Takats, Z.; Wiseman, J. M.; Gologan, B.; Cooks, R. G., Mass spectrometry sampling under ambient conditions with desorption electrospray ionization. *Science* **2004,** *306* (5695), 471-473.

19. Wiseman, J. M.; Ifa, D. R.; Venter, A.; Cooks, R. G., Ambient molecular imaging by desorption electrospray ionization mass spectrometry. *Nat. Protocols* **2008,** *3* (3), 517-524.

20. Santagata, S.; Eberlin, L. S.; Norton, I.; Calligaris, D.; Feldman, D. R.; Ide, J. L.; Liu, X.; Wiley, J. S.; Vestal, M. L.; Ramkissoon, S. H.; Orringer, D. A.; Gill, K. K.; Dunn, I. F.; Dias-Santagata, D.; Ligon, K. L.; Jolesz, F. A.; Golby, A. J.; Cooks, R. G.; Agar, N. Y. R., Intraoperative mass spectrometry mapping of an onco-metabolite to guide brain tumor surgery. *Proc. Natl. Acad. Sci.* **2014,** *111* (30), 11121-11126.

21. Hirabayashi, A.; Sakairi, M.; Koizumi, H., Sonic spray ionization method for atmospheric pressure ionization mass spectrometry. *Anal.Chem.***1994,** *66* (24), 4557-4559.

22. Santos, V. G.; Regiani, T. s.; Dias, F. F. G.; Romão, W.; Jara, J. L. P.; Klitzke, C. c. F.; Coelho, F.; Eberlin, M. N., Venturi easy ambient sonic-spray ionization. *Anal.Chem.* **2011,** *83* (4), 1375-1380.

23. Chapman, S., Carrier mobility spectra of spray electrified liquids. *Phys. Rev.***1937,** *52* (3), 184-190.

24. Iribarne, J.; Thomson, B., On the evaporation of small ions from charged droplets. *J. Chem. Phy.* **1976,** *64* (6), 2287-2294.

25. Harper, B.; Lochridge, J.; Solouki, T. Spray pump ionization: A simple-to-use and low cost ionization technique for regulatory and formulation mass spectrometry analysis of pharmaceuticals. *Pan Africa Chemistry Network Congress 2015: Healthcare: From Discovery to Delivery* **2015,** Nairobi, Kenya.

26. Schwab, N. V.; Porcari, A. M.; Coelho, M. B.; Schmidt, E. M.; Jara, J. L.; Visentainer, J. V.; Eberlin, M. N., Easy dual-mode ambient mass spectrometry with venturi self-pumping, canned air, disposable parts and voltage-free sonic-spray ionization. *Analyst* **2012,** *137* (11), 2537-2540.

27. Nier, A. O.; Schlutter, D. J., High-performance double-focusing mass spectrometer. *Rev. Sci. Inst.* **1985,** *56* (2), 214-219.

28. Cody, R. B.; Tamura, J.; Musselman, B. D., Electrospray ionization/magnetic sector mass spectrometry: Calibration, resolution, and accurate mass measurements. *Anal.Chem.***1992,** *64* (14), 1561-1570.

29. Guilhaus, M., Principles and instrumentation in time-of-ight mass spectrometry. *J. Mass Spectrom.* **1995,** *30* (1519), 1532.

30. Douglas, D. J.; Frank, A. J.; Mao, D., Linear ion traps in mass spectrometry. *Mass Spectrom. Rev.* **2005,** *24* (1), 1-29.

31. Gys, T., Micro-channel plates and vacuum detectors. *Nucl. Instrum. Meth. A* **2015,** *787*, 254-260.

32. Comisarow, M. B.; Marshall, A. G., Fourier transform ion cyclotron resonance spectroscopy. *Chem. Phys. Lett.* **1974,** *25* (2), 282-283.

33. Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S., Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrom. Rev.* **1998,** *17* (1), 1-35.

34. Makarov, A., Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis. *Anal. Chem.* **2000,** *72* (6), 1156-1162.

35. Mamyrin, B.; Karataev, V.; Shmikk, D.; Zagulin, V., The mass-reflectron, a new non-magnetic time-of-flight mass spectrometer with high resolution. *Zh. Eksp. Teor. Fiz* **1973,** *64*, 82-89.

36. Zekavat, B.; Solouki, T., Radio-frequency ionization of organic compounds for mass spectrometry analysis. *Angew. Chem. Int. Ed.* **2013,** *52* (9), 2426-2429.

37. McIver Jr, R. T., Trajectory calculations for axial injection of ions into a magnetic field: Overcoming the magnetic mirror effect with an r.f. quadrupole lens. *Int. J. Mass Spectrom.* **1990,** *98* (1), 35-50.

38. Beauchamp, J. L., Ion cyclotron resonance spectroscopy. *Annu. Rev. Phys. Chem.* **1971,** *22* (1), 527-561.

39. Marshall, A. G., Milestones in fourier transform ion cyclotron resonance mass spectrometry technique development. *Int. J. Mass Spectrom.***2000,** *200* (1–3), 331-356.

40. Sommer, H.; Thomas, H. A.; Hipple, J. A., The measurement of e/m by cyclotron resonance. *Phys. Rev.* **1951,** *82* (5), 697-702.

41. Wobschall, D., Ion cyclotron resonance spectrometer. *Rev. Sci. Inst.* **1965,** *36* (4), 466-475.

42. Kingdon, K. H., A method for the neutralization of electron space charge by positive ionization at very low gas pressures. *Phys. Rev.* **1923,** *21* (4), 408-418.

43. Solouki, T.; Gillig, K. J.; Russell, D. H., Detection of high-mass biomolecules in Fourier transform ion cyclotron resonance mass spectrometry: Theoretical and experimental investigations. *Anal. Chem.* **1994,** *66* (9), 1583-1587.

44. Gillig, K. J.; Bluhm, B. K.; Russell, D. H., Ion motion in a Fourier transform ion cyclotron resonance wire ion guide cell. *Int. J. Mass Spectrom.* **1996,** *157–158*, 129-147.

45. Perry, R. H.; Cooks, R. G.; Noll, R. J., Orbitrap mass spectrometry: Instrumentation, ion motion and applications. *Mass Spectrom. Rev.* **2008,** *27* (6), 661-699.

46. Denisov, E.; Damoc, E.; Lange, O.; Makarov, A., Orbitrap mass spectrometry with resolving powers above 1,000,000. *Int. J. Mass Spectrom.* **2012,** *325–327*, 80-85.

47. Makarov, A.; Denisov, E.; Kholomeev, A.; Balschun, W.; Lange, O.; Strupat, K.; Horning, S., Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal. Chem.* **2006,** *78* (7), 2113-2120.

48. May, J. C.; McLean, J. A., Ion mobility-mass spectrometry: Time-dispersive instrumentation. *Anal. Chem.* **2015,** *87* (3), 1422-1436.

49. Cumeras, R.; Figueras, E.; Davis, C. E.; Baumbach, J. I.; Gracia, I., Review on ion mobility spectrometry. Part 1: Current instrumentation. *Analyst* **2015,** *140* (5), 1376-1390.

50. Cumeras, R.; Figueras, E.; Davis, C.; Baumbach, J.; Gràcia, I., Review on ion mobility spectrometry. Part 2: Hyphenated methods and effects of experimental parameters. *Analyst* **2015,** *140* (5), 1391-1410.

51. Ruotolo, B. T.; Benesch, J. L. P.; Sandercock, A. M.; Hyung, S.-J.; Robinson, C. V., Ion mobility-mass spectrometry analysis of large protein complexes. *Nat. Protocols* **2008,** *3* (7), 1139-1152.

52. Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H., Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.* **2004,** *18* (20), 2401-2414.

53. Giles, K.; Williams, J. P.; Campuzano, I., Enhancements in travelling wave ion mobility resolution. *Rapid Commun. Mass Spectrom.* **2011,** *25* (11), 1559-1566.

54. Brantley, M.; Zekavat, B.; Harper, B.; Mason, R.; Solouki, T., Automated deconvolution of overlapped ion mobility profiles. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (10), 1810-1819.

55. Harper, B.; Miladi, M.; Solouki, T., Loss of internal backbone carbonyls: Additional evidence for sequence-scrambling in collision-induced dissociation of y-type ions. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (10), 1716-1729.

56. Pettit, M.; Brantley, M.; Harper, B.; Solouki, T., Collision-energy resolved ion mobility characterization of isomeric mixtures. *Analyst* **2015,** *140* (20), 6886-6896.

57. Creese, A. J.; Cooper, H. J., Separation and identification of isomeric glycopeptides by high field asymmetric waveform ion mobility spectrometry. *Anal. Chem.* **2012,** *84* (5), 2597-2601.

58. Lee, S.; Li, Z.; Valentine, S. J.; Zucker, S. M.; Webber, N.; Reilly, J. P.; Clemmer, D. E., Extracted fragment ion mobility distributions: A new method for complex mixture analysis. *Int. J. Mass Spectrom.* **2012,** *309* (0), 154-160.

59. Xuan, Y.; Creese, A. J.; Horner, J. A.; Cooper, H. J., High-field asymmetric waveform ion mobility spectrometry (FAIMS) coupled with high-resolution electron transfer dissociation mass spectrometry for the analysis of isobaric phosphopeptides. *Rapid Commun. Mass Spectrom.* **2009,** *23* (13), 1963-1969.

60. Windig, W.; Guilment, J., Interactive self-modeling mixture analysis. *Anal. Chem.* **1991,** *63* (14), 1425-1432.

61. Colby, B. N., Spectral deconvolution for overlapping gc/ms components. *J. Am. Soc. Mass Spectrom.* **1992,** *3* (5), 558-562.

62. Zekavat, B.; Solouki, T., Chemometric data analysis for deconvolution of overlapped ion mobility profiles. *J. Am. Soc. Mass Spectrom.* **2012,** *23* (11), 1873-1884.

63. Zekavat, B.; Miladi, M.; Becker, C.; Munisamy, S.; Solouki, T., Combined use of post-ion mobility/collision-induced dissociation and chemometrics for b fragment ion analysis. *J. Am. Soc. Mass Spectrom.***2013,** *24* (9), 1355-1365.

64. Zekavat, B.; Miladi, M.; Al-Fdeilat, A.; Somogyi, A.; Solouki, T., Evidence for sequence scrambling and divergent h/d exchange reactions of doubly-charged isobaric b-type fragment ions. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (2), 226-236.

65. Rudge, M. R. H., Theory of the ionization of atoms by electron impact. *Rev.Mod. Phys.* **1968,** *40* (3), 564-590.

66. Aebersold, R.; Mann, M., Mass spectrometry-based proteomics. *Nature* **2003,** *422* (6928), 198-207.

67. Bogdanov, B.; Smith, R. D., Proteomics by FTICR mass spectrometry: Top down and bottom up. *Mass Spectrom. Rev.* **2005,** *24* (2), 168-200.

68. Siuti, N.; Kelleher, N. L., Decoding protein modifications using top-down mass spectrometry. *Nat. Methods* **2007,** *4* (10), 817-821.

69. Scherperel, G.; Reid, G. E., Emerging methods in proteomics: Top-down protein characterization by multistage tandem mass spectrometry. *Analyst* **2007,** *132* (6), 500-506.

70. Hunt, D. F.; Yates, J. R.; Shabanowitz, J.; Winston, S.; Hauer, C. R., Protein sequencing by tandem mass spectrometry. *Proc. Natl. Acad. Sci.* **1986,** *83* (17), 6233-6237.

71. Sadygov, R. G.; Cociorva, D.; Yates, J. R., 3rd, Large-scale database searching using tandem mass spectra: Looking up the answer in the back of the book. *Nat. Methods* **2004,** *1*, 195-202.

72. Solouki, T.; Paša-Tolic, L.; Jackson, G. S.; Guan, S.; Marshall, A. G., High-resolution multistage MS, MS^2, and MS^3 matrix-assisted laser desorption/ionization FT-ICR mass spectra of peptides from a single laser shot. *Anal. Chem.* **1996,** *68* (21), 3718-3725.

73. Scherperel, G.; Yan, H.; Wang, Y.; Reid, G. E., 'Top-down'characterization of site-directed mutagenesis products of staphylococcus aureus dihydroneopterin aldolase by multistage tandem mass spectrometry in a linear quadrupole ion trap. *Analyst* **2006,** *131* (2), 291-302.

74. Pesavento, J. J.; Mizzen, C. A.; Kelleher, N. L., Quantitative analysis of modified proteins and their positional isomers by tandem mass spectrometry: Human histone H4. *Anal. Chem.* **2006,** *78* (13), 4271-4280.

75. McLafferty, F. W.; Bente, P. F.; Kornfeld, R.; Tsai, S.-C.; Howe, I., Metastable ion characteristics. XXII. Collisional activation spectra of organic ions. *J. Am. Chem.Soc.* **1973,** *95* (7), 2120-2129.

76. Levsen, K.; Wipf, H. K.; McLafferty, F., Mass spectrometric studies of peptides—VI:. Applications of metastable ion and collisional activation spectra. *J. Mass Spectrom.* **1974,** *8* (1), 117-128.

77. Hunt, D. F.; Buko, A. M.; Ballard, J. M.; Shabanowitz, J.; Giordani, A. B., Sequence analysis of polypeptides by collision activated dissociation on a triple quadrupole mass spectrometer. *Biol. Mass Spectrom.* **1981,** *8* (9), 397-408.

78. Blanksby, S. J.; Ellison, G. B., Bond dissociation energies of organic molecules. *Acc. Chem. Res.* **2003,** *36* (4), 255-263.

79. Rice, O. K.; Ramsperger, H. C., Theories of unimolecular gas reactions at low pressures. *J. Am. Chem. Soc.* **1927,** *49* (7), 1617-1629.

80. Kassel, L. S., Studies in homogeneous gas reactions. I. *J. Phys. Chem.* **1927,** *32* (2), 225-242.

81. Kassel, L. S., Studies in homogeneous gas reactions. II. Introduction of quantum theory. *J. Phys. Chem.* **1927,** *32* (7), 1065-1079.

82. Marcus, R. A., Unimolecular dissociations and free radical recombination reactions. *J. Chem. Phys.* **1952,** *20* (3), 359-364.

83. Sleno, L.; Volmer, D. A., Ion activation methods for tandem mass spectrometry. *J. Mass Spectrom.* **2004,** *39* (10), 1091-1112.

84. Paizs, B.; Suhai, S., Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* **2005,** *24*, 508-548.

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

86. Fattahi, A.; Solouki, T., Conformational analysis of metal complexed model peptides and their fragment ions using ft-icr ms and gas-phase h/d exchange reactions. *Proceedings of the 49th ASMS Conference on Mass Spectrometry and Allied Topics* **2001,** Chicago, IL.

87. Fattahi, A.; Zekavat, B.; Solouki, T., H/d exchange kinetics: Experimental evidence for formation of different b fragment ion conformers/isomers during the gas-phase peptide sequencing. *J. Am. Soc. Mass Spectrom.* **2010,** *21*, 358-369.

88. Tang, X. J.; Boyd, R. K., An investigation of fragmentation mechanisms of doubly protonated tryptic peptides. *Rapid. Commun. Mass. Spectrom.* **1992,** *6*, 651-657.

89. Tang, X. J.; Thibault, P.; Boyd, R. K., Fragmentation reactions of multiply-protonated peptides and implications for sequencing by tandem mass spectrometry with low-energy collision-induced dissociation. *Anal. Chem.* **1993,** *65*, 2824-2834.

90. Tang, X. J.; Boyd, R. K., Rearrangements of doubly charged acylium ions from lysyl and ornithyl peptides. *Rapid. Commun. Mass. Spectrom.* **1994,** *8*, 678-686.

91. Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B., Scrambling of sequence information in collision-induced dissociation of peptides. *J. Am. Chem. Soc.* **2006,** *128*, 10364-10365.

92. Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B., Infrared spectroscopy and theoretical studies on gas-phase protonated leu-enkephalin and its fragments: Direct experimental evidence for the mobile proton. *J. Am. Chem. Soc.* **2007,** *129*, 5887-5897.

93. Polfer, N. C.; Oomens, J., Reaction products in mass spectrometry elucidated with infrared spectroscopy. *Phys. Chem. Chem. Phys.* **2007,** *9*, 3804-3817.

94. Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S. n.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. l., Sequence-scrambling fragmentation pathways of protonated peptides. *J. Am. Chem. Soc.* **2008,** *130* (52), 17774-17789.

95. Bythell, B. J.; Erlekam, U.; Paizs, B.; Maitre, P., Infrared spectroscopy of fragments from doubly protonated tryptic peptides. *Chem. Phys. Chem.* **2009,** *10*, 883-885.

96. Chen, X.; Steill, J. D.; Oomens, J.; Polfer, N. C., Oxazolone versus macrocycle structures for leu-enkephalin b_2-b_4: Insights from infrared multiple-photon dissociation spectroscopy and gas-phase hydrogen/deuterium exchange. *J. Am. Soc. Mass Spectrom.* **2010,** *21*, 1313-1321.

97. Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C., Effect of peptide fragment size on the propensity of cyclization in collision-induced dissociation: Oligoglycine b_2-b_8. *J. Am. Chem. Soc.* **2009,** *131*, 18272-18282.

98. Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M., Spectroscopic evidence for an oxazolone structure of the b_2 fragment ion from protonated tri-alanine. *J. Am. Soc. Mass Spectrom.* **2009,** *20*, 334-339.

99. Paizs, B.; Simonyi, M., Ring inversion barrier of diazepam and derivatives: An ab initio study. *Chirality* **1999,** *11* (8), 651-658.

100. Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B., Spectroscopic and theoretical evidence for oxazolone ring formation in collision-induced dissociation of peptides. *J. Am. Chem. Soc.* **2005,** *127*, 17154-17155.

101. Tirado, M.; Polfer, N. C., Defying entropy: Forming large head-to-tail macrocycles in the gas phase. *Angew. Chem. Int. Ed. Engl.* **2012,** *124*, 1-4.

102. Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G., The structure and fragmentation of b_n ($n \geq 3$) ions in peptide spectra. *J. Am. Soc. Mass Spectrom.* **1996,** *7*, 233-242.

103. Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G., Why are b ions stable species in peptide spectra? *J. Am. Soc. Mass Spectrom.* **1995,** *6*, 1165-1174.

104. Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H., IRMPD spectroscopy shows that AGG forms an oxazolone b_2^+ ion. *J. Am. Chem. Soc.* **2008,** *130*, 17644-17645.

105. Goloborodko, A. A.; Gorshkov, M. V.; Good, D. M.; Zubarev, R. A., Sequence scrambling in shotgun proteomics is negligible. *J. Am. Soc. Mass Spectrom.* **2011,** *22*, 1121-1124.

106. Miladi, M.; Harper, B.; Solouki, T., Evidence for sequence scrambling in collision-induced dissociation of y-type fragment ions. *J. Am. Soc. Mass Spectrom.* **2013,** *24* (11), 1755-1766.

107. Gonzalez-Sanchez, M.-B.; Lanucara, F.; Hardman, G. E.; Eyers, C. E., Gas-phase intermolecular phosphate transfer within a phosphohistidine phosphopeptide dimer. *Int. J. Mass Spectrom.* **2014,** *367*, 28-34.

108. Meng, Z.; Simmons-Willis, T. A.; Limbach, P. A., The use of mass spectrometry in genomics. *Biomolecular Engineering* **2004,** *21* (1), 1-13.

109. Castleberry, C. M.; Rodicio, L. P.; Limbach, P. A., Electrospray ionization mass spectrometry of oligonucleotides. *Curr. Prot. Nucleic Acid Chem.* **2001,** John Wiley & Sons, Inc..

110. Murray, K. K., DNA sequencing by mass spectrometry. *J. Mass Spectrom.* **1996,** *31* (11), 1203-1215.

111. Nakayama, H.; Akiyama, M.; Taoka, M.; Yamauchi, Y.; Nobe, Y.; Ishikawa, H.; Takahashi, N.; Isobe, T., Ariadne: A database search engine for identification and chemical analysis of rna using tandem mass spectrometry data. *Nucleic Acids Res.* **2009,** *37* (6), e47.

112. Ni, J.; Pomerantz, S. C.; Rozenski, J.; Zhang, Y.; McCloskey, J. A., Interpretation of oligonucleotide mass spectra for determination of sequence using electrospray ionization and tandem mass spectrometry. *Anal. Chem.* **1996,** *68* (13), 1989-1999.

113. Rozenski, J.; McCloskey, J. A., Sos: A simple interactive program for ab initio oligonucleotide sequencing by mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2002,** *13* (3), 200-203.

114. McLuckey, S. A.; Van Berkel, G. J.; Glish, G. L., Tandem mass spectrometry of small, multiply charged oligonucleotides. *J. Am. Soc. Mass Spectrom.* **1992,** *3* (1), 60-70.

115. Harper, B.; Neumann, E. K.; Solouki, T., DNA oligonucleotide fragment ion rearrangements upon collision-induced dissociation. *J. Am. Soc. Mass Spectrom.* **2015,** *26* (8), 1404-1413.

116. Zhurov, K. O.; Fornelli, L.; Wodrich, M. D.; Laskay, Ü. A.; Tsybin, Y. O., Principles of electron capture and transfer dissociation mass spectrometry applied to peptide and protein structure analysis. *Chem. Soc. Rev.* **2013,** *42* (12), 5014-5030.

117. Asakawa, D.; Smargiasso, N.; Quinton, L.; Pauw, E. D., Peptide backbone fragmentation initiated by side-chain loss at cysteine residue in matrix-assisted laser desorption/ionization in-source decay mass spectrometry. *J. Mass Spectrom.* **2013,** *48* (3), 352-360.

118. Somogyi, A., Probing peptide fragment ion structures by combining sustained off-resonance collision-induced dissociation and gas-phase H/D exchange (SORI-HDX) in fourier transform ion-cyclotron resonance (ft-icr) instruments. *J. Am. Soc. Mass Spectrom.* **2008,** *19*, 1771-1775.

119. Miladi, M.; Zekavat, B.; Munisamy, S. M.; Solouki, T., A systematic study on the effect of histidine position and fragment ion size on the formation of b_n ions. *Int. J. Mass Spectrom.* **2012,** *316-318*, 164-173.

120. Riba-Garcia, I.; Giles, K.; Bateman, R. H.; Gaskell, S. J., Evidence for structural variants of a- and b-type peptide fragment ions using combined ion mobility/mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2008,** *19*, 609-613.

121. Riba Garcia, I.; Giles, K.; Bateman, R. H.; Gaskell, S. J., Studies of peptide a- and b-type fragment ions using stable isotope labeling and integrated ion mobility/tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2008,** *19*, 1781-1787.

122. Polfer, N. C.; Bohrer, B. C.; Plasencia, M. D.; Paizs, B.; Clemmer, D. E., On the dynamics of fragment isomerization in collision-induced dissociation of peptides. *J. Phys. Chem. A.* **2008,** *112*, 1286-1293.

123. Perkins, B. R.; Chamot-Rooke, J.; Yoon, S. H.; Gucinski, A. C.; Somogyi, A.; Wysocki, V. H., Evidence of diketopiperazine and oxazolone structures for HA b_2^+ ion. *J. Am. Chem. Soc.* **2009,** *131*, 17528-17529.

124. Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maitre, P., Infrared spectroscopy of fragments of protonated peptides: Direct evidence for macrocyclic structures of b_5 ions. *J. Am. Chem. Soc.* **2009,** *131*, 11503-11508.

125. Bianco, G.; Labella, C.; Pepe, A.; Cataldi, T. I., Scrambling of autoinducing precursor peptides investigated by infrared multiphoton dissociation with electrospray ionization and fourier transform ion cyclotron resonance mass spectrometry. *Anal Bioanal. Chem.* **2013,** *405* (5), 1721-1732.

126. Solouki, T.; Zekavat, B.; Miladi, M., On the existence of structurally different isobaric b_n fragment ions. *Proceedings of the 58th ASMS Conference on Mass Spectrometry and Allied Topics* **2010,** Salt Lake City, UT.

127. Bythell, B. J.; Somogyi, A.; Paizs, B., What is the structure of b_2 ions generated from doubly protonated tryptic peptides? *J. Am. Soc. Mass Spectrom.* **2009,** *20*, 618-624.

128. Bythell, B. J.; Knapp-Mohammady, M.; Paizs, B.; Harrison, A. G., Effect of the his residue on the cyclization of b ions. *J. Am. Soc. Mass Spectrom.* **2010,** *21*, 1352-1363.

129. Bythell, B. J.; Csonka, I. P.; Suhai, S.; Barofsky, D. F.; Paizs, B., Gas-phase structure and fragmentation pathways of singly protonated peptides with n-terminal arginine. *J. Phys. Chem. B.* **2010,** *114*, 15092-15105.

130. Badman, E. R.; Myung, S.; Clemmer, D. E., Gas-phase separations of protein and peptide ion fragments generated by collision-induced dissociation in an ion trap. *Anal. Chem.* **2002,** *74*, 4889-4894.

131. Badger, J. Implication of basic amino acid residues on peptide fragmentation: An H/D exchange and SORI-CID study. *Undergraduate Thesis* **2011,** University of Maine, Orono, ME.

132. Amblard, M.; Fehrentz, J. A.; Martinez, J.; Subra, G., Methods and protocols of modern solid phase peptide synthesis. *Mol. Biotechnol.* **2006,** *33*, 239-254.

133. Brandon, T. R.; Justin, L. P. B.; Alan, M. S.; Suk-Joon, H.; Carol, V. R., Ion mobility–mass spectrometry analysis of large protein complexes. *Nat. Protocols* **2008,** *3* (7), 1139-1152.

134. Lopez, L. L.; Tiller, P. R.; Senko, M. W.; Schwartz, J. C., Automated strategies for obtaining standardized collisionally induced dissociation spectra on a benchtop ion trap mass spectrometer. *Rapid Commun. Mass Spectrom.* **1999,** *13* (8), 663-668.

135. Murray, K. K.; Boyd, R. K.; Eberlin, M. N.; Langley, G. J.; Li, L.; Naito, Y., Definitions of terms relating to mass spectrometry (IUPAC recommendations 2013). *Pure Appl. Chem.* **2013,** *85* (7), 1515-1609.

136. Harrison, A.; Tasoglu, C.; Yalcin, T., Non-direct sequence ions in the tandem mass spectrometry of protonated peptide amides—an energy-resolved study. *J. Am. Soc. Mass Spectrom.* **2013,** *24* (10), 1565-1572.

137. Bythell, B. J.; Maître, P.; Paizs, B., Cyclization and rearrangement reactions of a_n fragment ions of protonated peptides. *J. Am. Chem. Soc.* **2010,** *132* (42), 14766-14779.

138. Paizs, B.; Bythell, B.; Maître, P., Rearrangement pathways of the a_4 ion of protonated YGGFL characterized by IR spectroscopy and modeling. *J. Am. Soc. Mass Spectrom.* **2012,** *23* (4), 664-675.

139. Wassermann, T.; Boyarkin, O.; Paizs, B.; Rizzo, T., Conformation-specific spectroscopy of peptide fragment ions in a low-temperature ion trap. *J. Am. Soc. Mass Spectrom.* **2012,** *23* (6), 1029-1045.

140. Ambihapathy, K.; Yalcin, T.; Leung, H.-W.; Harrison, A. G., Pathways to immonium ions in the fragmentation of protonated peptides. *J. Mass Spectrom.* **1997,** *32* (2), 209-215.

141. Cooper, T.; Talaty, E.; Grove, J.; Van Stipdonk, M.; Suhai, S.; Paizs, B., Isotope labeling and theoretical study of the formation of a_3* ions from protonated tetraglycine. *J. Am. Soc. Mass Spectrom.* **2006,** *17*, 1654-1664.

142. Karpas, Z.; Berant, Z.; Stimac, R., An ion mobility spectrometry/mass spectrometry (IMS/MS) study of the site of protonation in anilines. *Struct. Chem.* **1990,** *1* (2-3), 201-204.

143. Lalli, P. M.; Iglesias, B. A.; Toma, H. E.; de Sa, G. F.; Daroda, R. J.; Silva Filho, J. C.; Szulejko, J. E.; Araki, K.; Eberlin, M. N., Protomers: Formation, separation and characterization via travelling wave ion mobility mass spectrometry. *J. Mass Spectrom.* **2012,** *47* (6), 712-719.

144. Student, Probable error of a correlation coefficient. *Biometrika* **1908,** *6* (2/3), 302-310.

145. Student, The probable error of a mean. *Biometrika* **1908,** *6* (1), 1-25.

146. Pringle, S. D.; Giles, K.; Wildgoose, J. L.; Williams, J. P.; Slade, S. E.; Thalassinos, K.; Bateman, R. H.; Bowers, M. T.; Scrivens, J. H., An investigation of the mobility separation of some peptide and protein ions using a new hybrid quadrupole/travelling wave ims/oa-tof instrument. *Int. J. Mass Spectrom.* **2007,** *261* (1), 1-12.

147. Tsaprailis, G.; Nair, H.; Zhong, W.; Kuppannan, K.; Futrell, J. H.; Wysocki, V. H., A mechanistic investigation of the enhanced cleavage at histidine in the gas-phase dissociation of protonated peptides. *Anal. Chem.* **2004,** *76*, 2083-2094.

148. Paizs, B.; Szlávik, Z.; Lendvay, G.; Vékey, K.; Suhai, S., Formation of a_2^+ ions of protonated peptides. An ab initio study. *Rapid Commun. Mass Spectrom.* **2000,** *14* (9), 746-755.

149. Breci, L. A.; Tabb, D. L.; Yates, J. R.; Wysocki, V. H., Cleavage N-terminal to proline: Analysis of a database of peptide tandem mass spectra. *Anal. Chem.* **2003,** *75*, 1963-1971.

150. Kish, M. M.; Wesdemiotis, C., Selective cleavage at internal lysine residues in protonated vs. metalated peptides. *Int. J. Mass Spectrom.* **2003,** *227*, 191–203.

151. Bythell, B. J.; Suhai, S.; Somogyi, Á.; Paizs, B., Proton-driven amide bond-cleavage pathways of gas-phase peptide ions lacking mobile protons. *J. Am. Chem. Soc.* **2009,** *131* (39), 14057-14065.

152. Chawner, R.; Eyers, C. E.; Gaskell, S. J., The influence of a C-terminal basic residue on peptide fragmentation pathways. *Int. J. Mass Spectrom.* **2012,** *316–318*, 284-291.

153. Paizs, B.; Suhai, S., Fragmentation pathways of protonated peptides. *Mass Spectrom. Rev.* **2005,** *24* (4), 508-548.

154. McLuckey, S. A.; Habibi-Goudarzi, S., Decompositions of multiply charged oligonucleotide anions. *J. Am. Chem. Soc.* **1993,** *115* (25), 12085-12095.

155. Håkansson, K.; Hudgins, R.; Marshall, A.; O'Hair, R. J., Electron capture dissociation and infrared multiphoton dissociation of oligodeoxynucleotide dications. *J. Am. Soc. Mass Spectrom.* **2003,** *14* (1), 23-41.

156. Keller, K. M.; Brodbelt, J. S., Collisionally activated dissociation and infrared multiphoton dissociation of oligonucleotides in a quadrupole ion trap. *Anal. Biochem.* **2004,** *326* (2), 200-210.

157. Juhasz, P.; Roskey, M. T.; Smirnov, I. P.; Haff, L. A.; Vestal, M. L.; Martin, S. A., Applications of delayed extraction matrix-assisted laser desorption ionization time-of-flight mass spectrometry to oligonucleotide analysis. *Anal. Chem.* **1996,** *68* (6), 941-946.

158. Premstaller, A.; Huber, C. G., Factors determining the performance of triple quadrupole, quadrupole ion trap and sector field mass spectrometer in electrospray ionization mass spectrometry. 2. Suitability for de novo sequencing. *Rapid Commun. Mass Spectrom.* **2001,** *15* (13), 1053-1060.

159. Yang, J.; Mo, J.; Adamson, J. T.; Håkansson, K., Characterization of oligodeoxynucleotides by electron detachment dissociation fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **2005,** *77* (6), 1876-1882.

160. Nyakas, A.; Eberle, R.; Stucki, S.; Schürch, S., More than charged base loss — revisiting the fragmentation of highly charged oligonucleotides. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (7), 1155-1166.

161. Saminathan, I. S.; Wang, X. S.; Guo, Y.; Krakovska, O.; Voisin, S.; Hopkinson, A. C.; Siu, K. W., The extent and effects of peptide sequence scrambling via formation of macrocyclic b ions in model proteins. *J. Am. Soc. Mass Spectrom.* **2010,** *21*, 2085-2094.

162. Hoaglund-Hyzer, C. S.; Li, J.; Clemmer, D. E., Mobility labeling for parallel CID of ion mixtures. *Anal. Chem.* **2000,** *72* (13), 2737-2740.

163. Solouki, T.; Szulejko, J., Bimolecular and unimolecular contributions to the disparate self-chemical ionizations of α-pinene and camphene isomers. *J. Am.Soc. Mass Spectrom.* **2007,** *18* (11), 2026-2039.

164. Bartmess, J. E.; Georgiadis, R. M., Empirical methods for determination of ionization gauge relative sensitivities for different gases. *Vacuum* **1983,** *33* (3), 149-153.

165. Rozenski, J., Mongo oligo mass calculator v2.06 http://mods.Rna.Albany.Edu/masspec/mongo-oligo. *The RNA Institute, College of Arts and Sciences, State University of New York at Albany,* **1999,** On-line Interface.

166. Francl, M. M.; Pietro, W. J.; Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A., Self-consistent molecular orbital methods. XXIII. A polarization-type basis set for second-row elements. *J. Chem. Phys.* **1982,** *77* (7), 3654-3665.

167. Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G., Gaussian 09, rev. C. 01, Gaussian. *Inc.* **2009,** Wallingford, CT.

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

169. Fattahi, A.; Zekavat, B.; Solouki, T., H/d exchange kinetics: Experimental evidence for formation of different b fragment ion conformers/isomers during the gas-phase peptide sequencing. *J. Am. Soc. Mass Spectrom.* **2010,** *21* (3), 358-369.

170. Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maître, P., Infrared spectroscopy of fragments of protonated peptides: Direct evidence for macrocyclic structures of b_5 ions. *J. Am. Chem. Soc.* **2009,** *131* (32), 11503-11508.

171. Riba-Garcia, I.; Giles, K.; Bateman, R.; Gaskell, S., Evidence for structural variants of a- and b-type peptide fragment ions using combined ion mobility/mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2008,** *19* (4), 609-613.

172. Polfer, N. C.; Bohrer, B. C.; Plasencia, M. D.; Paizs, B.; Clemmer, D. E., On the dynamics of fragment isomerization in collision-induced dissociation of peptides. *J. Phys. Chem. A* **2008,** *112* (6), 1286-1293.

173. Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B., Sequence-scrambling fragmentation pathways of protonated peptides. *J. Am. Chem. Soc.* **2008,** *130* (52), 17774-17789.

174. Pan, S.; Verhoeven, K.; Lee, J. K., Investigation of the initial fragmentation of oligodeoxynucleotides in a quadrupole ion trap: Charge level-related base loss. *J. Am. Soc. Mass Spectrom.* **2005,** *16* (11), 1853-1865.

175. Schulten, H. R.; Beckey, H. D.; Boerboom, A. J. H.; Meuzelaar, H. L. C., Pyrolysis field desorption mass spectrometry of deoxyribonucleic acid. *Anal. Chem.* **1973,** *45* (14), 2358-2362.

176. Green-Church, K. B.; Limbach, P. A., Mononucleotide gas-phase proton affinities as determined by the kinetic method. *Journal of the American Society for Mass Spectrometry* **2000,** *11* (1), 24-32.

177. Bythell, B. J.; Knapp-Mohammady, M.; Paizs, B.; Harrison, A. G., Effect of the His residue on the cyclization of b ions. *J. Am. Soc. Mass Spectrom.* **2010,** *21* (8), 1352-1363.

178. Molesworth, S. P.; Van Stipdonk, M. J., Apparent inhibition by arginine of macrocyclic b ion formation from singly charged protonated peptides. *J. Am. Soc. Mass Spectrom.* **2010,** *21* (8), 1322-1328.

179. Campbell, J. M.; Collings, B. A.; Douglas, D. J., A new linear ion trap time-of-flight system with tandem mass spectrometry capabilities. *Rapid Commun. Mass Spectrom.* **1998,** *12* (20), 1463-1474.

180. Stucki, S.; Désiron, C.; Nyakas, A.; Marti, S.; Leumann, C.; Schürch, S., Gas-phase dissociation of homo-DNA oligonucleotides. *J. Am. Soc. Mass Spectrom.* **2013,** *24* (12), 1997-2006.

181. Kee, J.-M.; Villani, B.; Carpenter, L. R.; Muir, T. W., Development of stable phosphohistidine analogues. *J. Am. Chem. Soc.* **2010,** *132* (41), 14327-14329.

182. Attwood, P. V., Pn bond protein phosphatases. *BBA - Proteins Proteom.* **2013,** *1834* (1), 470-478.

183. Attwood, P. V.; Piggott, M. J.; Zu, X. L.; Besant, P. G., Focus on phosphohistidine. *Amino Acids* **2007,** *32* (1), 145-156.

184. Ruotolo, B. T.; Giles, K.; Campuzano, I.; Sandercock, A. M.; Bateman, R. H.; Robinson, C. V., Evidence for macromolecular protein rings in the absence of bulk water. *Science* **2005,** *310* (5754), 1658-1661.

185. Scarff, C. A.; Thalassinos, K.; Hilton, G. R.; Scrivens, J. H., Travelling wave ion mobility mass spectrometry studies of protein structure: Biological significance and comparison with x-ray crystallography and nuclear magnetic resonance spectroscopy measurements. *Rapid Commun. Mass Spectrom.* **2008,** *22* (20), 3297-3304.

186. Wyttenbach, T.; Bowers, M. T., Structural stability from solution to the gas phase: Native solution structure of ubiquitin survives analysis in a solvent-free ion mobility–mass spectrometry environment. *J. Phys. Chem. B* **2011,** *115* (42), 12266-12275.

187. Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T., Ion mobility–mass spectrometry reveals a conformational conversion from random assembly to β-sheet in amyloid fibril formation. *Nat. Chem.* **2011,** *3* (2), 172-177.

188. Zhong, Y.; Hyung, S.-J.; Ruotolo, B. T., Ion mobility–mass spectrometry for structural proteomics. *Expert Rev. Proteomics* **2012,** *9* (1), 47-58.

189. Jurneczko, E.; Barran, P. E., How useful is ion mobility mass spectrometry for structural biology? The relationship between protein crystal structures and their collision cross sections in the gas phase. *Analyst* **2011,** *136* (1), 20-28.

190. Chen, S.-H.; Russell, D., How closely related are conformations of protein ions sampled by im-ms to native solution structures? *J. Am. Soc. Mass Spectrom.* **2015,** *26* (9), 1433-1443.

191. Cumeras, R.; Figueras, E.; Davis, C.; Baumbach, J.; Gràcia, I., Review on ion mobility spectrometry. Part 2: Hyphenated methods and effects of experimental parameters. *Analyst* **2015,** *140*, 1391-1410.

192. Rokushika, S.; Hatano, H.; Baim, M. A.; Hill, H. H., Resolution measurement for ion mobility spectrometry. *Anal. Chem.* **1985,** *57* (9), 1902-1907.

193. Siems, W. F.; Wu, C.; Tarver, E. E.; Hill, H. H., Jr.; Larsen, P. R.; McMinn, D. G., Measuring the resolving power of ion mobility spectrometers. *Anal. Chem.* **1994,** *66* (23), 4195-4201.

194. Shvartsburg, A. A.; Smith, R. D., Fundamentals of traveling wave ion mobility spectrometry. *Anal. Chem.* **2008,** *80* (24), 9689-9699.

195. May, J. C.; Dodds, J. N.; Kurulugama, R. T.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A., Broadscale resolving power performance of a high precision uniform field ion mobility-mass spectrometer. *Analyst* **2015,** *140* (20), 6824-6833.

196. Dwivedi, P.; Wu, C.; Matz, L. M.; Clowers, B. H.; Siems, W. F.; Hill, H. H., Gas-phase chiral separations by ion mobility spectrometry. *Anal. Chem.* **2006,** *78* (24), 8200-8206.

197. Shvartsburg, A. A.; Smith, R. D., Accelerated high-resolution differential ion mobility separations using hydrogen. *Anal. Chem.* **2011,** *83* (23), 9159-9166.

198. Fasciotti, M.; Sanvido, G. B.; Santos, V. G.; Lalli, P. M.; McCullagh, M.; de Sá, G. F.; Daroda, R. J.; Peter, M. G.; Eberlin, M. N., Separation of isomeric disaccharides by traveling wave ion mobility mass spectrometry using CO_2 as drift gas. *J. Mass Spectrom.* **2012,** *47* (12), 1643-1647.

199. Liang, X.; Wang, X.; Wang, W.; Zhou, Q.; Chen, C.; Peng, L.; Wen, M.; Qu, T.; Wang, Z.; Zhao, K.; Li, J.; Li, H., Sensitive detection of black powder by stand-alone ion mobility spectrometer with chlorinated hydrocarbon modifiers in drift gas. *Talanta* **2014,** *121*, 215-219.

200. Purves, R. W.; Ozog, A. R.; Ambrose, S. J.; Prasad, S.; Belford, M.; Dunyach, J.-J., Using gas modifiers to significantly improve sensitivity and selectivity in a cylindrical faims device. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (7), 1274-1284.

201. May, J.; McLean, J., The influence of drift gas composition on the separation mechanism in traveling wave ion mobility spectrometry: Insight from electrodynamic simulations. *Int. J. Ion Mobil. Spec.* **2013,** *16* (2), 85-94.

202. Liu, Y.; Clemmer, D. E., Characterizing oligosaccharides using injected-ion mobility/mass spectrometry. *Anal. Chem.* **1997,** *69* (13), 2504-2509.

203. Clowers, B.; Dwivedi, P.; Steiner, W.; Hill, H.; Bendiak, B., Separation of sodiated isobaric disaccharides and trisaccharides using electrospray ionization-atmospheric pressure ion mobility-time of flight mass spectrometry. *J. Am. Soc. Mass Spectrom.* **2005,** *16* (5), 660-669.

204. Clowers, B. H.; Ibrahim, Y. M.; Prior, D. C.; Danielson, W. F.; Belov, M. E.; Smith, R. D., Enhanced ion utilization efficiency using an electrodynamic ion funnel trap as an injection mechanism for ion mobility spectrometry. *Anal. Chem.* **2008,** *80* (3), 612-623.

205. Zucker, S.; Lee, S.; Webber, N.; Valentine, S.; Reilly, J.; Clemmer, D., An ion mobility/ion trap/photodissociation instrument for characterization of ion structure. *J. Am. Soc. Mass Spectrom.* **2011,** *22* (9), 1477-1485.

206. Hernandez, D. R.; DeBord, J. D.; Ridgeway, M. E.; Kaplan, D. A.; Park, M. A.; Fernandez-Lima, F., Ion dynamics in a trapped ion mobility spectrometer. *Analyst* **2014,** *139* (8), 1913-1921.

207. Lee, S.; Li, Z.; Valentine, S. J.; Zucker, S. M.; Webber, N.; Reilly, J. P.; Clemmer, D. E., Extracted fragment ion mobility distributions: A new method for complex mixture analysis. *Int. J. Mass Spectrom.* **2012,** *309*, 154-160.

208. Zhu, F.; Lee, S.; Valentine, S.; Reilly, J.; Clemmer, D., Mannose7 glycan isomer characterization by ims-ms/ms analysis. *J. Am. Soc. Mass Spectrom.* **2012,** *23* (12), 2158-2166.

209. Hoffmann, W.; Hofmann, J.; Pagel, K., Energy-resolved ion mobility-mass spectrometry—a concept to improve the separation of isomeric carbohydrates. *J. Am.Soc. Mass Spectrom.* **2014,** *25* (3), 471-479.

210. Khakinejad, M.; Kondalaji, S.; Maleki, H.; Arndt, J.; Donohoe, G.; Valentine, S., Combining ion mobility spectrometry with hydrogen-deuterium exchange and top-down MS for peptide ion structure analysis. *J. Am. Soc. Mass Spectrom.* **2014,** *25* (12), 2103-2115.

211. Harvey, D.; Crispin, M.; Bonomelli, C.; Scrivens, J., Ion mobility mass spectrometry for ion recovery and clean-up of MS and MS/MS spectra obtained from low abundance viral samples. *J. Am. Soc. Mass Spectrom.* **2015**, 1-14.

212. Meloun, M.; Čapek, J.; Mikšík, P.; Brereton, R. G., Critical comparison of methods predicting the number of components in spectroscopic data. *Anal. Chim. Acta* **2000,** *423* (1), 51-68.

213. Malinowski, E. R., Determination of the number of factors and the experimental error in a data matrix. *Anal. Chem.* **1977,** *49* (4), 612-617.

214. Bush, M. F.; Campuzano, I. D. G.; Robinson, C. V., Ion mobility mass spectrometry of peptide ions: Effects of drift gas and calibration strategies. *Anal. Chem.* **2012,** *84* (16), 7124-7130.

215. Pettit, M. E.; Harper, B.; Brantley, M. R.; Solouki, T., Collision-energy resolved ion mobility characterization of isomeric mixtures. *Analyst* **2015,** *140* (20), 6886-6896.

216. May, J. C.; Goodwin, C. R.; Lareau, N. M.; Leaptrot, K. L.; Morris, C. B.; Kurulugama, R. T.; Mordehai, A.; Klein, C.; Barry, W.; Darland, E.; Overney, G.; Imatani, K.; Stafford, G. C.; Fjeldsted, J. C.; McLean, J. A., Conformational ordering of biomolecules in the gas phase: Nitrogen collision cross sections measured on a prototype high resolution drift tube ion mobility-mass spectrometer. *Anal. Chem.* **2014,** *86* (4), 2107-2116.

217. May, J. C.; McLean, J. A., A uniform field ion mobility study of melittin and implications of low-field mobility for resolving fine cross-sectional detail in peptide and protein experiments. *Proteomics* **2015**.

218. Mason, E. A.; Schamp Jr, H. W., Mobility of gaseous lons in weak electric fields. *Ann. Phys.* **1958,** *4* (3), 233-270.

219. Case, D. A.; Babin, V.; Berryman, J.; Betz, R. M.; Cai, Q.; Cerutti, D. S.; Cheatham Iii, T. E.; Darden, T. A.; Duke, R. E.; Gohlke, H., *Amber 14*. **2014**.

220. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A., Gaussian 09, revision A.02, Gaussian Inc., **2009,** Wallingford, CT.

221. Forsythe, J. G.; Stow, S. M.; Nefzger, H.; Kwiecien, N. W.; May, J. C.; McLean, J. A.; Hercules, D. M., Structural characterization of methylenedianiline regioisomers by ion mobility-mass spectrometry, tandem mass spectrometry, and computational strategies: I. Electrospray spectra of 2-ring isomers. *Anal. Chem.* **2014,** *86* (9), 4362-4370.

222. Stow, S. M.; Onifer, T. M.; Forsythe, J. G.; Nefzger, H.; Kwiecien, N. W.; May, J. C.; McLean, J. A.; Hercules, D. M., Structural characterization of methylenedianiline regioisomers by ion mobility-mass spectrometry, tandem mass spectrometry, and computational strategies. 2. Electrospray spectra of 3-ring and 4-ring isomers. *Anal. Chem.* **2015,** *87* (12), 6288-6296.

223. von Helden, G.; Hsu, M. T.; Gotts, N.; Bowers, M. T., Carbon cluster cations with up to 84 atoms: Structures, formation mechanism, and reactivity. *J. Phys. Chem.* **1993,** *97* (31), 8182-8192.

224. Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F., Structural information from ion mobility measurements: Effects of the long-range potential. *J. Phys. Chem.* **1996,** *100* (40), 16082-16086.

225. Shvartsburg, A. A.; Jarrold, M. F., An exact hard-spheres scattering model for the mobilities of polyatomic ions. *Chem.Pphys. Lett.* **1996,** *261* (1), 86-91.

226. Campuzano, I.; Bush, M. F.; Robinson, C. V.; Beaumont, C.; Richardson, K.; Kim, H.; Kim, H. I., Structural characterization of drug-like compounds by ion mobility mass spectrometry: Comparison of theoretical and experimentally derived nitrogen collision cross sections. *Anal. Chemi.* **2012,** *84* (2), 1026-1033.

227. Bleiholder, C.; Wyttenbach, T.; Bowers, M. T., A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (I). Method. *Int. J.Mass Spectrom.* **2011,** *308* (1), 1-10.

228. Bleiholder, C.; Contreras, S.; Do, T. D.; Bowers, M. T., A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (II). Model parameterization and definition of empirical shape factors for proteins. *Int. J. Mass Spectrom.* **2013,** *345*, 89-96.

229. Anderson, S. E.; Bleiholder, C.; Brocker, E. R.; Stang, P. J.; Bowers, M. T., A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (III): Application to supramolecular coordination-driven assemblies with complex shapes. *Int. J. Mass Spectrom.* **2012,** *330-332*, 78-84.

230. Bleiholder, C.; Contreras, S.; Bowers, M. T., A novel projection approximation algorithm for the fast and accurate computation of molecular collision cross sections (IV). Application to polypeptides. *Int. J. Mass Spectrom.* **2013,** *354*, 275-280.

231. http://luschka.bic.ucsb.edu:8080/WebPSA/index.jsp.

232. Gidden, J.; Wyttenbach, T.; Jackson, A. T.; Scrivens, J. H.; Bowers, M. T., Gas-phase conformations of synthetic polymers: Poly(ethylene glycol), poly(propylene glycol), and poly(tetramethylene glycol). *J. Am. Chem. Soc.* **2000,** *122* (19), 4692-4699.

233. Forsythe, J. G.; Petrov, A. S.; Walker, C. A.; Allen, S. J.; Pellissier, J. S.; Bush, M. F.; Hud, N. V.; Fernandez, F. M., Collision cross section calibrants for negative ion mode traveling wave ion mobility-mass spectrometry. *Analyst* **2015,** *140* (20), 6853-6861.

234. Garmón-Lobato, S.; Abad-García, B.; Sánchez-Ilárduya, M. B.; Romera-Fernández, M.; Berrueta, L. A.; Gallo, B.; Vicente, F., Improvement using chemometrics in ion mobility coupled to mass spectrometry as a tool for mass spectrometry fragmentation studies: Flavonoid aglycone cases. *Anal. Chim. Acta* **2013,** *771* (0), 56-64.

235. Giles, K.; Pringle, S. D.; Worthington, K. R.; Little, D.; Wildgoose, J. L.; Bateman, R. H., Applications of a travelling wave-based radio-frequency-only stacked ring ion guide. *Rapid Commun. Mass Spectrom.* **2004,** *18* (20), 2401-2414.

236. Radionova, A.; Filippov, I.; Derrick, P. J., In pursuit of resolution in time-of-flight mass spectrometry: A historical perspective. *Mass Spectrom. Rev.* **2015**, E-published ahead of print.

237. Bush, M. F.; Hall, Z.; Giles, K.; Hoyes, J.; Robinson, C. V.; Ruotolo, B. T., Collision cross sections of proteins and their complexes: A calibration framework and database for gas-phase structural biology. *Anal. Chem.* **2010,** *82* (22), 9557-9565.

238. Mäkinen, M.; Sillanpää, M.; Viitanen, A. K.; Knap, A.; Mäkelä, J. M.; Puton, J., The effect of humidity on sensitivity of amine detection in ion mobility spectrometry. *Talanta* **2011,** *84* (1), 116-121.

239. Mayer, T.; Borsdorf, H., Accuracy of ion mobility measurements dependent on the influence of humidity. *Anal. Chem.* **2014,** *86* (10), 5069-5076.

240. Karpas, Z.; Berant, Z., Effect of drift gas on mobility of ions. *J. Phys. Chem.* **1989,** *93* (8), 3021-3025.

241. Asbury, G. R.; Hill, H. H., Using different drift gases to change separation factors (α) in ion mobility spectrometry. *Anal. Chem.* **2000,** *72* (3), 580-584.

242. Beegle, L. W.; Kanik, I.; Matz, L.; Hill Jr, H. H., Effects of drift-gas polarizability on glycine peptides in ion mobility spectrometry. *Int. J. Mass Spectrom.* **2002,** *216* (3), 257-268.

243. Srebalus Barnes, C. A.; Hilderbrand, A. E.; Valentine, S. J.; Clemmer, D. E., Resolving isomeric peptide mixtures: A combined HPLC/ion mobility-tofms analysis of a 4000-component combinatorial library. *Anal. Chem.* **2002,** *74* (1), 26-36.

244. Shaffer, S. A.; Tang, K.; Anderson, G. A.; Prior, D. C.; Udseth, H. R.; Smith, R. D., A novel ion funnel for focusing ions at elevated pressure using electrospray ionization mass spectrometry. *Rapid Commun. Mass Spectrom.* **1997,** *11* (16), 1813-1817.

245. Shaffer, S. A.; Prior, D. C.; Anderson, G. A.; Udseth, H. R.; Smith, R. D., An ion funnel interface for improved ion focusing and sensitivity using electrospray ionization mass spectrometry. *Anal. Chem.* **1998,** *70* (19), 4111-4119.

246. Kelly, R. T.; Tolmachev, A. V.; Page, J. S.; Tang, K.; Smith, R. D., The ion funnel: Theory, implementations, and applications. *Mass Spectrom. Rev.* **2010,** *29* (2), 294-312.

247. Szulejko, J. E.; McCulloch, M.; Jackson, J.; McKee, D. L.; Walker, J. C.; Solouki, T., Evidence for cancer biomarkers in exhaled breath. *IEEE Sens. J.* **2010,** *10* (1), 185-210.

248. Zou, W.; She, J.; Tolstikov, V. V., A comprehensive workflow of mass spectrometry-based untargeted metabolomics in cancer metabolic biomarker discovery using human plasma and urine. *Metabolites* **2013,** *3* (3), 787-819.

249. Asbury, G. R.; Hill, H. H., Evaluation of ultrahigh resolution ion mobility spectrometry as an analytical separation device in chromatographic terms. *J. Microcolumn Sep.* **2000,** *12* (3), 172-178.