Understanding MS Approaches to Peptide Characterization

Alessandra L. Ferzoco

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry

Chapel Hill
2011

Approved by:

Dr. Gary L. Glish

Dr. Tomas Baer

Dr. John Papanikolas

Dr. Maurice Brookhart

Dr. Thomas J. Meyer

Abstract

ALESSANDRA L. FERZOCO: Understanding MS Approaches to Peptide Characterization
(Under the direction of Gary L. Glish)

Proteomics studies using mass spectrometry have become routine. The overarching goal of proteomics is to understand how the proteome changes within individuals over time, in disease states, and between individuals. Currently the field is limited by the quality of the data that can be obtained for low concentration proteins in complex biological mixtures, and by the lack of chemical knowledge that needs to be incorporated into the automated data analysis protocols that are needed to handle the massive volumes of data generated by proteomics studies. The work in this dissertation addresses both limitations. IRMPD was found to increase in efficiency with the size of a peptide, unlike CID. Attempts to quantify the increase in internal energy responsible for the increase in IRMPD efficiency with size were precluded by non-Boltzmann internal energy distributions in the population of trapped ions. FMOC derivatized peptides were found to promote sodium binding, and thereby facilitate C-terminal dissociation patterns that are easily interpreted. Infrared spectroscopy was used to measure the structure of b_3 and a_4 peptide fragments to gain insight into sequence- and size-dependent dissociation patterns. Comparison of the structures found by spectroscopy with dissociation patterns seen in CID and IRMPD gave further information about the dissociation kinetics of b_3 and a_4 peptide fragments.

To keV

Acknowledgements

I would first and foremost like to thank Dr. Gary L. Glish for his conscious and patient (often *very patient*) guidance. His dedication to the comprehensive education of each of his students is a priority, which is exceptional among scientists of his caliber. His commitment to putting his students first consistently produces well-rounded scientists that are prepared for successful careers in the field. Often this work goes without recognition, so I hope that collectively the community of scientists that you have cultivated can grow into a demonstration of your legacy in the field of mass spectrometry. I wish you many happy years taming ions…and students.

Past and current Glish lab students form a unique community that both challenges and supports each other. Together with the Miller and Bowers lab students, I have had no shortage of examples of people who are simply happy to do "good science."

I am fortunate to have a wonderful family who are my closest friends, and a collection of friends that are so close they are more like family. Anything is possible with the level of unconditional support you all provide. Mom, Dad, Tim, and Robin, I would like to thank you specifically for teaching me that work ethic is all I need to achieve my goals, but that kindness is the most important achievement. (And thank you to Nonna, Nonno, Bob, and Elaine for conveying the same). Morgan and Richie, I look up to the both of you as examples of kindness and grace in how you move through the world.

Table of Contents

List of Tables

List of Figures

Figure

List of Terms

***a_n, b_n, c_n* and *x_n, y_n, z_n* peptide sequence ions:** A nomenclature used to describe the dissociation of peptides along the backbone. If the peptide fragment retains the N-terminus from the original peptide, the ions are termed a_n, b_n, c_n, depending on where the backbone is cleaved. If the peptide fragment retains the C-terminus from the original peptide, the ions are termed x_n, y_n, z_n, depending on where the backbone is cleaved.

Action spectroscopy: The use of a chemical signature, such as dissociation or induced reaction, to indicate resonant absorption of photons. The method is an alternative to direct absorption experiments for obtaining infrared spectra, since gas phase ion densities are too low for direct absorption.

Bottom-up proteomics: The practice of using a proteolytic enzyme to digest a mixture of proteins into peptides, and then analyzing the sequence of the peptides.

CID: Collision induced dissociation: The term can be more general, but in this dissertation it refers to the use of low energy collisions with a helium bath gas to increase the internal energy of a parent ion to the point of dissociation in MS^n experiments.

CID-type products: This term refers specifically to the type of product ions formed by collision induced dissociation of the b_3 ions from the analogs of YGGFL discussed in Chapters 6 and 8. This class of product ions is commonly observed.

***C_p*:** The constant pressure heat capacity. C_p was used to measure the change in enthalpy of an ion after a period of IR irradiation, as discussed in Chapter 4.

Critical energy, ε_0: The minimum energy required for dissociation given an infinite amount of time.

Dissociation efficiency: A measure of the amount of dissociation in an MS^n experiment. Dissociation efficiency is calculated as the sum of the product ion intensities divided by the sum of the product ion intensities plus the parent ion intensity.

Dissociation threshold: The minimum energy required to observe unimolecular dissociation given the time frame of the experiment.

FEL: Free electron laser. A type of laser that uses electrons at relativistic speeds wiggled through a series of magnets to generate light. High output powers are achieved, which makes

FELIX: Free electron laser for infrared experiments. A specific free electron laser located in Nieuwegein in the Netherlands. This term is also used to refer to the facility in general.

FMOC: Fluorenylmethyloxycarbonyl, which is a chemical moiety that can be used to derivatize peptides at the N-terminus to increase the binding affinity for sodium cations.

Fragment ion: An ion that is formed from larger ions breaking apart in between the ion source and the mass analyzer. It is distinct from a product ion because no specific parent ion was isolated, and therefore the origin of the fragment may not be known.

FT-ICR: Fourier transform ion cyclotron resonance. A type of high resolution mass spectrometer that operates at very low pressures (below 10^{-10} mbar) and over long time frames. This is the type of mass spectrometer that was used to obtain infrared spectra at FELIX.

***h*:** Planck's constant, equal to 6.626×10^{-34} J s^{-1}.

***hν*:** Used to represent the energy from a photon.

ΔH: The change in enthalpy. Specifically refers to the change in enthalpy of an ion after a period of IR irradiation discussed in Chapter 4.

IR: Infrared.

IRa-CID: Infrared activation collision induced dissociation. The practice of pre-heating ions by irradiation with an infrared laser before resonant excitation for collision induced dissociation.

IRMPD: Infrared multiphoton photodissociation. An MS^n method where the absorption of multiple infrared photons causes dissociation of a parent ion.

IRMPD-type products: This term refers specifically to the type of product ions formed by infrared multiphoton photodissociation of the b_3 ions from the analogs of YGAFL and YAAFL discussed in Chapters 6 and 8. This class of product ions is distinct from the CID-type products commonly observed, and were therefore unexpected.

k: Boltzmann's constant, which has a value of 1.38065×10^{-23} J K^{-1}.

***k_{cc}*:** The rate of collisional cooling.

***k_{diss}*:** The rate of unimolecular dissociation. Synonymous with *k(ε)*.

***k(ε)*:** The rate of a unimolecular dissociation reaction as a function of internal energy in the parent ion.

***k_{IR}*:** The rate of internal energy increase due to IR irradiation.

***k_{rad}*:** The rate of radiative cooling by both stimulated and spontaneous emission in an ion.

Kinetic shift: The amount of energy above the critical energy required to reach the dissociation threshold.

Leucine enkephalin: A peptide with the sequence tyrosine-glycine-glycine-phenylalanine-leucine (YGGFL). In the body it acts as a neurotransmitter. In mass spectrometry it often serves as a model peptide for fundamental studies, along with analogs of the form YXZFL where X and Z can be any amino acid.

Loose transition state: A transition state with a relatively high density of states. A simple cleavage reaction is an example of a reaction with a loose transition state.

M^+ or $[M+H]^+$: Used to denote a parent ion. M^+ is general, while $[M+H]^+$ refers spectifically to a protonated molecular ion.

Macrocycle: A peptide structure formed by nucleophilic attack on the C-terminal carbonyl carbon by the N-terminus. This structure is implicated in the sequence scrambling reactions seen in peptide dissociation chemistry.

MS: Mass spectrometry.

MS^n: Tandem mass spectrometry or multiple stages of mass spectrometry. An ion is isolated, dissociated, and then the product ions are analyzed by mass spectrometry. Further stages of isolation and dissociation can be carried out on the product ions, and "n" represents the number of stages of isolation and dissociation.

***m/z*:** The ratio of the mass of an ion to its charge. The property of an ion that governs it's characteristic frequency of motion in trapping mass analyzers.

***N*:** The number of atoms in an ion.

$N^{\ddagger}(\varepsilon-\varepsilon_0)$: The number of states between the critical energy and internal energy of the parent ion in a dissociation channel transition state.

Oxazolone: A five-member ring structure often found at the C-terminus of b_n ions. The structure is formed by nucleophilic attack of the C-terminal carbonyl by the carbonyl on the adjacent residue.

Parent ion: An ion selected for isolation and dissociation in an MS^n experiment.

Product ion: An ion generated from dissociation of a parent ion in an MS^n experiment.

Proteomics: The study of the network of proteins in an organism and how that network evolves over time and in disease states.

QET: Quasi-equilibrium theory. A theory used to describe unimolecular dissociation kinetics. For gas phase ions the theory is identical to RRKM theory.

QIT: Quadrupole ion trap. A type of mass spectrometer that operates with a relatively high pressure of helium (mTorr) and is particularly useful for MS^n experiments. Ions are confined in all three dimensions with a quadrupolar electric field.

R: The gas constant, which has a value of 8.314 J mol^{-1} K^{-1}.

Rearrangement: As the term is used in this dissertation, it refers to a type of unimolecular dissociation reaction where the product ions have a significantly different bonding structure than the parent ions, and which proceeds through a tight transition state that requires the participation of many atoms in the parent ion.

r-IRMPD: Resonant infrared multiphoton photodissociation. A method for obtaining infrared spectra of gas phase ions by action spectroscopy. Dissociation is the signature of resonance and multiple photons are required to reach the dissociation threshold.

RRKM: Rice, Ramsperger, Kessel, and Marcus. A theory used to describe unimolecular dissociation kinetics. For gas phase ions the theory is identical to quasi-equilibrium theory.

Simple cleavage: As the term is used in this dissertation, it refers to a type of unimolecular dissociation reaction where only a single bond is broken and so the bonding structure of the product ions remains the same as that of the parent ion except the broken bond. Reactions of this type proceed through a loose transition state.

T: Temperature.

t_{50}**:** The time required to achieve 50% dissociation of a parent ion.

TA-CID: Thermally activated collision induced dissociation. A CID experiment performed in the presence of a heated helium bath gas to pre-heat the ions before resonant excitation for collision induced dissociation.

Tight transition state: A transition state that is formed by the participation of many atoms, and therefore has a low density of states.

TS: Transition state. The point corresponding to the highest energy along a reaction coordinate.

ε**:** The amount of internal energy in a parent ion.

ν_i**:** The list of vibrational frequencies in an ion.

$\rho(\varepsilon)$**:** The density of states in the parent ion as a function of the energy of the parent ion.

σ**:** The degeneracy of unimolecular dissociation reaction. For example, in the case of dissociation of a hydrogen atom from methane, the reaction degeneracy is four.

Chapter 1

Motivation for the Study of Gas Phase Peptide Chemistry

1.1 The use of mass spectrometry for proteomics studies

Proteomics is the study of the structures and functions of the complement of proteins in an organism. Similarly to genomics, the field of proteomics seeks to quantitatively understand how the network of proteins transforms through aging and in disease states, both within an individual and between individuals[1]. Proteomics studies begin by identifying the proteins present in a sample by their sequence. Most frequently protein identification is done by a bottom-up approach where the proteins in a sample are first subjected to proteolytic digestion with an enzyme, and then the resulting peptides are sequenced[2,3]. Mass spectrometry (MS) is a good tool for peptide sequencing because of the sensitivity, speed, and dynamic range of the technique[4]. Moreover, and importantly for large-scale studies, sample preparation and data analysis can be automated[5,6]. Unfortunately the technology is not without its complications.

An MS experiment on a peptide or mixture of peptides yields knowledge of the mass-to-charge of the species present in the sample, which gives information about the amino acids present in the peptide(s). However there is no information about the sequence of amino acids contained in the mass-to-charge value. In the example shown in Figure 1, there are 840 permutations of the correct sequence that would result in measurement of the same

mass-to-charge in the mass spectrum (shown in red in Figure 1). There are also additional combinations of amino acids that could have a different elemental formula, but may have a mass-to-charge that is indistinguishable from the correct sequence given the resolution of the mass spectrometer, so in practice there is a large number of sequences that could correspond to a single mass-to-charge value. To obtain sequence information multiple stages of mass spectrometry (MS^n) are used[7-9]. In an MS^n experiment, ions of a selected mass-to-charge ratio are first isolated, and then through mechanisms described in Chapter 2, are dissociated. The products of the dissociation are analyzed or subjected to further stages of isolation and dissociation. The ions selected for isolation and dissociation are termed parent ions, and the ions formed from the dissociation are termed product ions. The "n" in MS^n is used to denote the number of stages of mass analysis, and is equal the number of cycles of isolation and dissociation plus one.

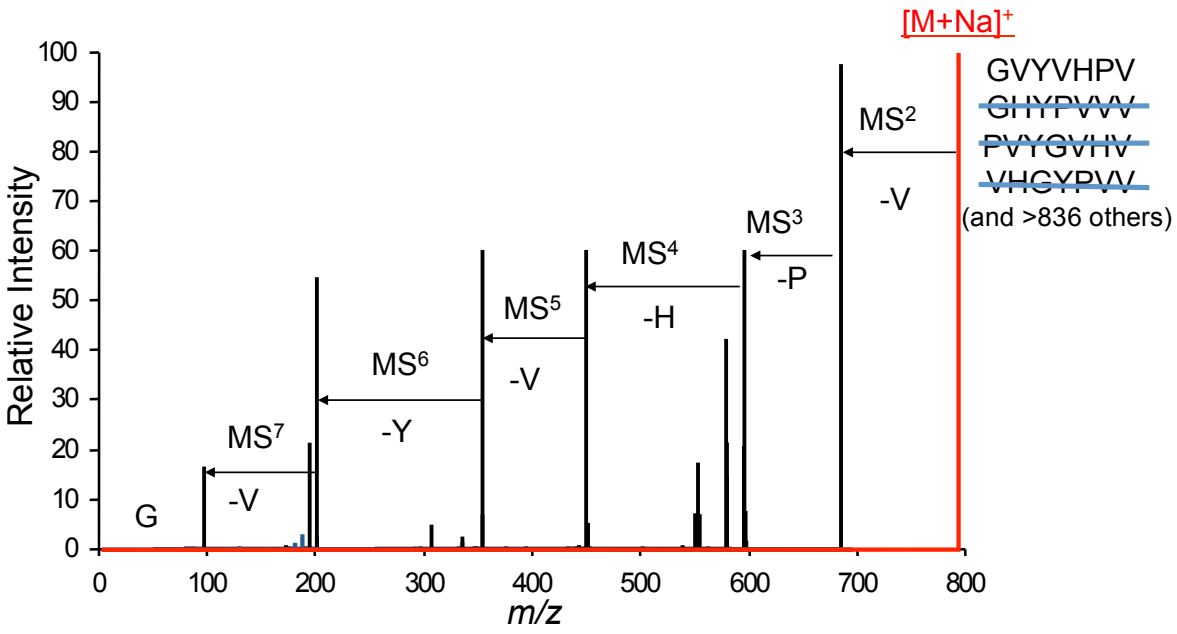


Figure 1. A composite spectrum of a series of MS^n spectra for C -terminal sequencing of the sodium cationized peptide $[GVYYVHPV+Na]^+$. The first stage of mass analysis is shown in red. From this information the amino acid composition is known, but not the sequence. The sequence of the peptide is identified from the series of MS^n spectra. Figure adapted from Lin et al[12].

1.1.1 Simple C-terminal sequencing of sodium cationized peptides

If the charge carrier on the peptide is a sodium cation, inferring the amino acid sequence from the product ions formed in MS^n experiments is straightforward[10-12], as illustrated in Figure 1. For sodium cationized peptides, dissociation in an MS^n experiment is by neutral loss of the C-terminal residue. In many cases it is possible to carry out sequential stages of MS^n experiments until the entire peptide is sequenced[10]. Using sodium cationized peptides for C-terminal sequencing is effective, but unfortunately sodium cationized peptides are scarce in MS experiments so the sensitivity of this method is insufficient for many applications.

1.1.2 The complicated dissociation chemistry of protonated peptides

In the peptide solutions used for MS sequencing experiments, protonated peptides are significantly more abundant than the sodium cationized species, and thus protonated peptides are most often used for sequencing. Unlike sodium cationized peptides, the dissociation of protonated peptides occurs through multiple competing dissociation pathways and so sequence interpretation is not as simple as it is for sodium cationized peptides[13]. Dissociation is often by small neutral losses or cleavage along the peptide backbone. There exists a nomenclature used to denote which bond on the backbone was broken to form a particular product ion (Figure 2)[14,15]. Product ions containing the N-terminus are termed a_n, b_n, or c_n ions depending respectively on whether the C_α-C, C-N, or N-C_α bond is cleaved, and where n denotes the number of residues contained in the product ion. Product ions containing the C-terminus are termed x_n, y_n, or z_n for bond cleavage at the C_α-C, C-N, or N-C_α, respectively.

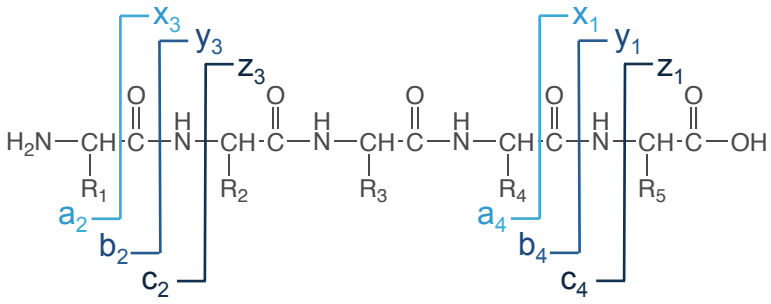


Figure 2. Peptide fragment nomenclature based on which bond along the backbone is cleaved and where the charge is retained. Product ions containing the N-terminus are termed a_n, b_n, or c_n ions depending respectively on whether the C_α-C, C-N, or N-C_α bond is cleaved, and where n denotes the number of residues contained in the product ion. Product ions containing the C-terminus are termed x_n, y_n, or z_n for bond cleavage at the C_α-C, C-N, or N-C_α , respectively.

An MS^2 spectrum of the protonated peptide $[YGAFL+H]^+$ is shown as an example in Figure 3. The identity of the peptide is known in this case, so the product ion peaks can be identified. If this were an unknown peptide, identification software would be used to match these peaks against the predicted dissociation patterns of peptides in a database generated from the genome of a particular species[16,17]. Identification by a database search works by applying rudimentary dissociation rules to generate a list of expected product ions from catalogs of peptides organized by species[18]. The experimental spectra are compared to the predicted spectra and confidence scores are assigned to potential matches. A difficulty with database searches is that the dissociation rules used by the software are not representative of the complicated chemistry that can occur in the experiment, so the predicted spectra do not correlate well with the experimental data[19]. A particularly troublesome example is when amino acids are occasionally lost from the interior of the peptide, resulting in what appears to be a scrambled sequence[20-23]. The peaks labeled with asterisks in Figure 3 are an example. These peaks at worst contribute to incorrect peptide matches and at best decrease the confidence score of the correct matches.

Alternatives to database searches are de novo sequencing programs that look for gaps between peaks that correspond to the mass of amino acids[24,25]. For example the b_n ions series shown in green in Figure 3 is separated by mass-to-charge values that correspond to sequential amino acids in the YGAFL sequence. Again, if this were an unknown peptide, the type of ion that each peak corresponds to would be unknown, and a de novo sequencing program must consider the mass-to-charge gaps between all of the peaks. In the example shown in Figure 3, the amino acids proline, asparagine, arginine, and methionine would be recognized even though these amino acids are not present in YGAFL. These algorithms do take into account how many amino acids can be linked in a sequence in assigning a score to the confidence of the identification. Nevertheless, the misidentified amino acids represent

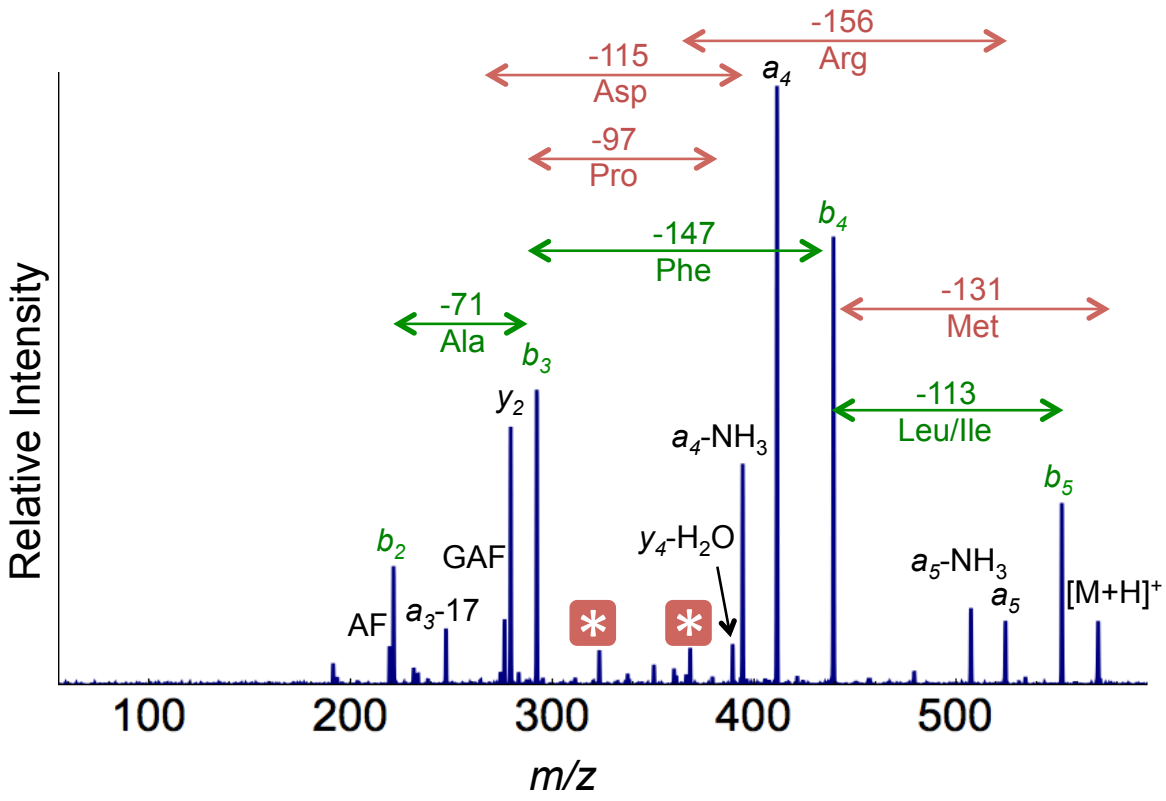


Figure 3. MS^2 spectrum of the protonated peptide [YGAFL+H]$^+$. If this were an unidentified peptide, the features labeled in black and green would contribute information helpful for identification while the features labeled in red would contribute misinformation.

conflicting information.

Two other features of the MS^2 spectrum in Figure 3 are the amount of dissociation, and the wide variety of product ions formed. In peptides with many stabilizing intramolecular interactions[26-28], and/or as the peptide size increases[29], dissociation can become difficult and limit the number and variety of product ions observed. Having a wide variety of product ions in MS^n spectra does complicate the interpretation of spectra, but the complexity also represents information content. Beyond the model peptide shown as an example here, peptides from a proteolytic digest often have longer sequences, multiple charges, and post-translational modifications. Even so, the problem is tractable computationally if the dissociation rules used in the sequencing algorithms were able to accurately predict experimental dissociation patterns[19].

The overarching goal of the work presented in this dissertation is to gain an understanding of how peptides react in mass spectrometry experiments. A better understanding should enable the improvement of sequencing algorithms. It will also facilitate the development of new methods and instrumentation that improve the quality of experimental data presented to the peptide sequencing algorithms. Additionally, many researchers use mass spectrometry as a tool to measure not just peptide sequences, but also the three-dimensional structure of peptides and proteins. Characterizing the three dimensional structure of a peptide is necessary to elucidate dissociation mechanisms, but is also important for fundamental studies on the structure/function relationships of proteins.

Discussed in Chapter 2 are methods for measuring structure in mass spectrometry. In Chapter 3, a method to address the size limitations of common MS^n experiments that utilize the increase in the infrared absorption cross section with the size of a peptide is suggested.

Presented in Chapter 4 is an attempt to quantify the change in infrared absorption cross section with size. As mentioned above, the results from MS^n experiments on sodium cationized peptides are much easier to interpret than MS^n experiments on protonated peptides, but the abundance of sodium cationized peptides is insufficient for many sequencing applications. A chemical modification strategy to increase the abundance of sodium cationized species and to give insight into the mechanism of dissociation in sodium cationized peptides is presented in Chapter 5. Infrared spectroscopy experiments to measure how the structure of peptide fragments depend on sequence are described in Chapters 6 and 7. Also in Chapters 6 and 7 are examples of how the use of multiple structural techniques can provide complementary information. A summary and suggestions for future experiments is given in Chapter 8.

References

(1) Godovac-Zimmermann, J.; Brown, L. R. *Mass Spectrometry Reviews* **2001**, *20*, 1.

(2) Ning, Z. B.; Zhou, H.; Wang, F. J.; Abu-Farha, M.; Figeys, D. *Analytical Chemistry* **2011**, *83*, 4407.

(3) Sardiu, M. E.; Washburn, M. P. *Journal of Biological Chemistry* **2011**, *286*, 23645.

(4) Yates, J. R., III *Journal of Mass Spectrometry* **1998**, *33*, 1.

(5) Khan, M. F.; Bennett, M. J.; Jumper, C. C.; Percy, A. J.; Silva, L. P.; Schriemer, D. C. *Journal of Pharmaceutical and Biomedical Analysis* **2011**, *55*, 832.

(6) Siu, S. O.; Lam, M. P. Y.; Lau, E.; Kong, R. P. W.; Lee, S. M. Y.; Chu, I. K. *Proteomics* **2011**, *11*, 2308.

(7) Hunt, D. F.; John R. Yates, I.; Shabanowitz, J.; Winston, S.; Hauer, C. R. *Proceedings of the National Academy of Science of the United States of America* **1986**, *83*, 6233.

(8) Johnson, R. S.; Taylor, J. A. *Molecular Biotechnology* **2002**, *22*, 301.

(9) Busch, K. L.; Glish, G. L.; McLuckey, S. A. *Mass Spectrometry / Mass Spectrometry: Techniques and Applications of Tandem Mass Spectrometry*; VCH: New York, 1988.

(10) Lin, T.; Payne, A. H.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 497.

(11) Lin, T.; Glish, G. L. *Analytical Chemistry* **1998**, *70*, 5162.

(12) Lin, T., The University of North Carolina at Chape Hill, 1997.

(13) Paizs, B.; Suhai, S. *Mass Spectrometry Reviews* **2005**, *24*, 508.

(14) Roepstorff, P. *Biomedical Mass Spectrometry* **1984**, *11*, 601.

(15) Biemann, K. *Biomedical and Environmental Mass Spectrometry* **1988**, *16*, 99.

(16) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. *Electrophoresis* **1999**, *20*, 3551.

(17) MacCoss, M. J.; Wu, C. C.; Yates, J. R. *Analytical Chemistry* **2002**, *74*, 5593.

(18) Sadygov, R. G.; Cociorva, D.; Yates, J. R. *Nature Methods* **2004**, *1*, 195.

(19) Zhang, Z. *Analytical Chemistry* **2004**, *76*, 3908.

(20) Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L. *Journal of the American Chemical Society* **1997**, *119*, 5481.

(21) Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128*, 10364.

(22) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *Journal of the American Chemical Society* **2008**, *130*, 17774.

(23) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2174.

(24) Seidler, J.; Zinn, N.; Boehm, M. E.; Lehmann, W. D. *Proteomics* **2010**, *10*, 634.

(25) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *PNAS* **2000**, *97*, 10313

(26) Vachet, R. W.; Asam, M. R.; Glish, G. L. *Journal of the American Chemical Society* **1996**, *118*, 6252.

(27) Tsaprailis, G.; Somogyi, A.; Nikolaev, E.; Wysocki, V. *Int. J. Mass Spectrom.* **2000**, *195/196*, 467.

(28) Tabb, D. L.; Huang, Y. Y.; Wysocki, V. H.; Yates, J. R. *Analytical Chemistry* **2004**, *76*, 1243.

(29) Dongre, A. R.; Somogyi, A.; Wysocki, V. H. *Journal of Mass Spectrometry* **1996**, *31*, 339.

Chapter 2

Structural Measurement by Tandem Mass Spectrometry

In trapping mass spectrometers, such as quadrupole ion trap (QIT) mass spectrometers and Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometers, ions oscillate within the trap at a frequency characteristic of their mass-to-charge ratio[1]. Consequently, oscillating electric fields can be used to manipulate ion motion in a mass-to-charge selective manner[2]. For example, in MS^n experiments the first task is to isolate the species selected to be the parent ion. The isolation is accomplished by destabilizing the trajectories of ions of all other mass-to-charge ratios such that they are ejected from the trapping volume, and only ions of the selected mass-to-charge ratio remain trapped. Isolated ions can be subjected to reactions, most often ion activation followed by the unimolecular dissociation reactions described in the remainder of this chapter[3]. The reaction products, which are related to the functional groups present in the parent ion, can then be mass analyzed to give information about the structure of the parent ion[4-6].

All MS^n experiments are cycles of isolation and reaction followed by mass-to-charge analysis of the reaction products. While the field of multiple stage mass spectrometry encompasses a wide variety of applications and techniques, this chapter is limited in scope to the specific techniques used in the experiments presented in this dissertation.

2.1 Unimolecular Dissociation

The most common type of reaction used in MS^n experiments is unimolecular dissociation, which is the second step of a two-step mechanism[7-9]. As applied to ions in an MS^n experiment, the first step is an increase in the internal energy of a selected parent ion, followed by dissociation of the activated parent ion into smaller product ions[10,11]. The rate of the dissociation step depends on the amount of internal energy in the parent ion and the product ions formed[9,10]. Rice, Ramsperger, Kessel, and Marcus (RRKM) and quasiequilibrium (QET) theories were independently derived to quantify the rate of dissociation as a function of internal energy in the parent ion, and are equivalent theories when applied to activated ions[12-16]. The dissociation rate, $k(\varepsilon)$ is given in equation 1 where σ is the reaction degeneracy, $N(\varepsilon-\varepsilon_0)$ is the number states in the transition state between the critical energy of a particular product ion channel (ε_0) up to the internal energy of the parent ion (ε), h is Planck's constant, and $\rho(\varepsilon)$ is the density of states at the internal energy of the parent ion.

$$k_{diss}(\varepsilon) = \frac{\sigma}{h}\frac{N^{\neq}(\varepsilon-\varepsilon_0)}{\rho(\varepsilon)} \qquad \text{equation 1}$$

2.1.1 Formation of activated parent ions

For the experiments presented in this dissertation the increase in internal energy of the parent ion is accomplished by collisions with an inert bath gas or irradiation with an infrared laser, and these methods will be discussed in subsequent sections of this chapter. The RRKM rate equation is arrived at by assuming conditions of thermal equilibrium. Whether or not the condition of thermal equilibrium is met in a mass spectrometer is debatable[11,17,18], and this subject is discussed further in Chapter 4. The RRKM dissociation rate is governed by the amount of internal energy in the parent ion. So as long as the energy

is statistically distributed throughout the ion, the mechanism by which the energy is imparted to the parent ion should have no consequences for the dissociation rate. Discussed in Chapters 6 and 7 are results from two different activation methods that initially seem to contradict the concept of a dissociation rate that is independent of the activation mechanism. In fact, combining the information from the two activation methods actually yields additional information about parent ion structure and competing product ion channels.

2.1.2 Dissociation of activated ions

The dissociation rate depends on the states available to the parent ion and transition state, which in turn depend on the structure of those species[9,19]. The reaction coordinates for two types of dissociation reaction are depicted in the top half of the Figure 1: a simple cleavage reaction where a single bond in the parent ion is broken to produce products that closely resemble the bonding structure of the parent, and a more complicated rearrangement reaction where multiple bonds are broken and formed to produce products that may have a significantly different bonding structure than the parent. Because bonds are broken and formed in the rearrangement reaction, the energy of the transition state and products are lower than those of the simple cleavage reaction. For the simple cleavage reaction, the bonding structure remains unchanged except for the cleaved bond so the states available to the transition state are essentially those of the parent ion with one elongated bond. In the case of a simple cleavage reaction, the transition state is said to be "loose." The rearrangement reaction requires the participation of several atoms, and the number of structures that will facilitate such a reaction are limited. Hence the density of states in the rearrangement transition state is limited, and the transition state is said to be "tight." Hypothetical RRKM rates as a function of the internal energy in the parent for the two

reaction types are plotted in the bottom half of Figure 1[19]. Below the critical energy for the simple cleavage, only the rearrangement can occur. The rearrangement reaction will be faster than the simple cleavage until the cumulative number of states in the transition state of

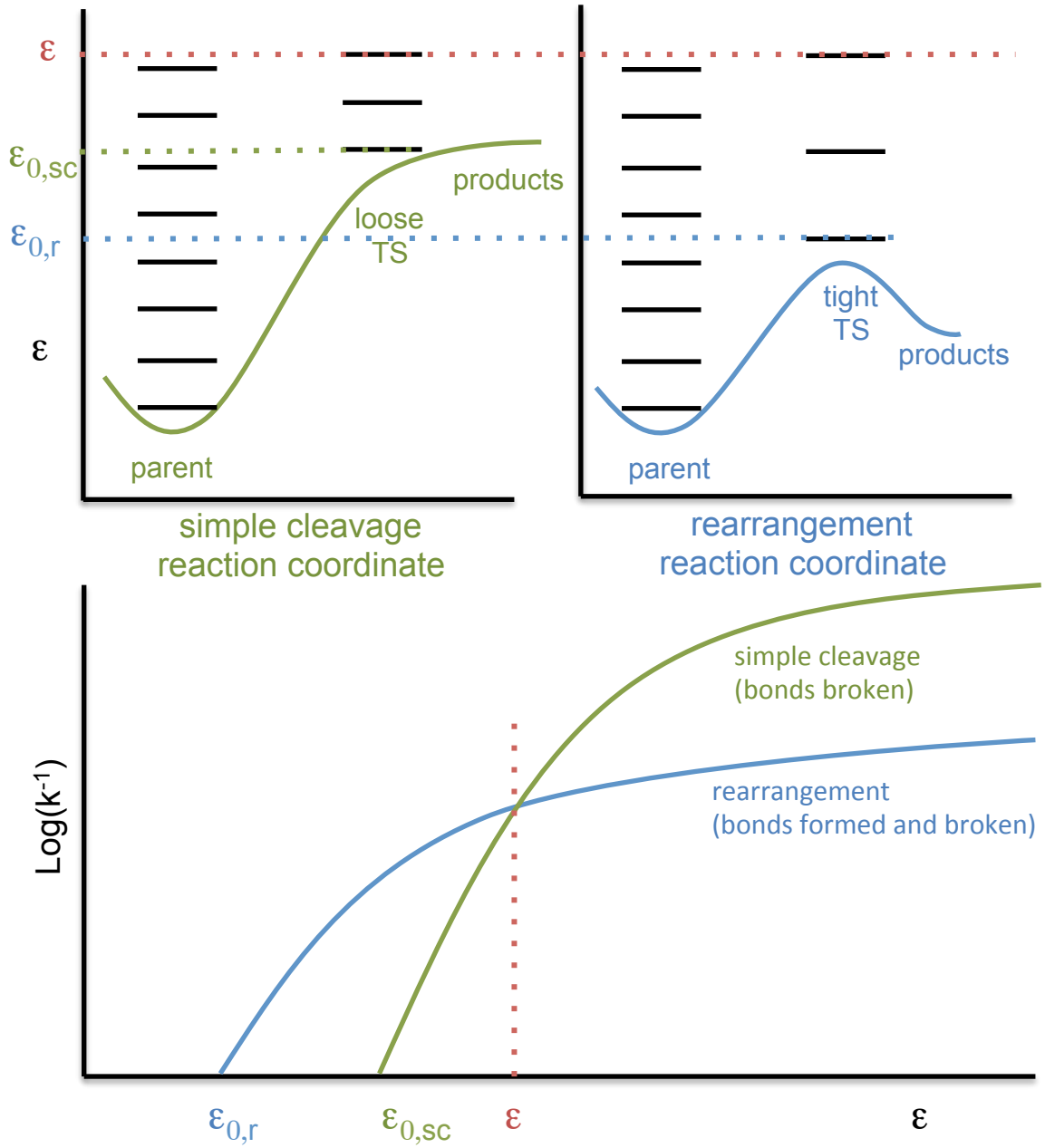


Figure 1. An illustration of how the rate of a unimolecular dissociation reaction depends on the structure of the parent ion and TS. The top plots are reaction coordiates for two competitive dissociation channels, and on the bottom the corresponding rate curves are plotted as a function of internal energy in the parent ion. Please see the text for a complete description of the figure.

the simple cleavage exceeds that of the rearrangement. The point at which the two rates are equal is denoted in red on the three plots. Because dissociation rate is so closely related to structure, we can use the relative dissociation rates to infer relative structural information about the transition states and product ions of competitive dissociation channels[20]. An example of such a study is presented in Chapter 7.

The critical energy is the minimum amount of energy needed for a reaction to occur. But even an energy above the critical energy does not guarantee the observation of product ions in an MS^n experiment because the rate of dissociation may be slower than the experimental time frame at that energy[21]. It is often more useful to speak of a dissociation threshold (Figure 2). The dissociation threshold is the minimum amount of energy required to observe product ions on the time frame of an MS^n experiment. The difference between the

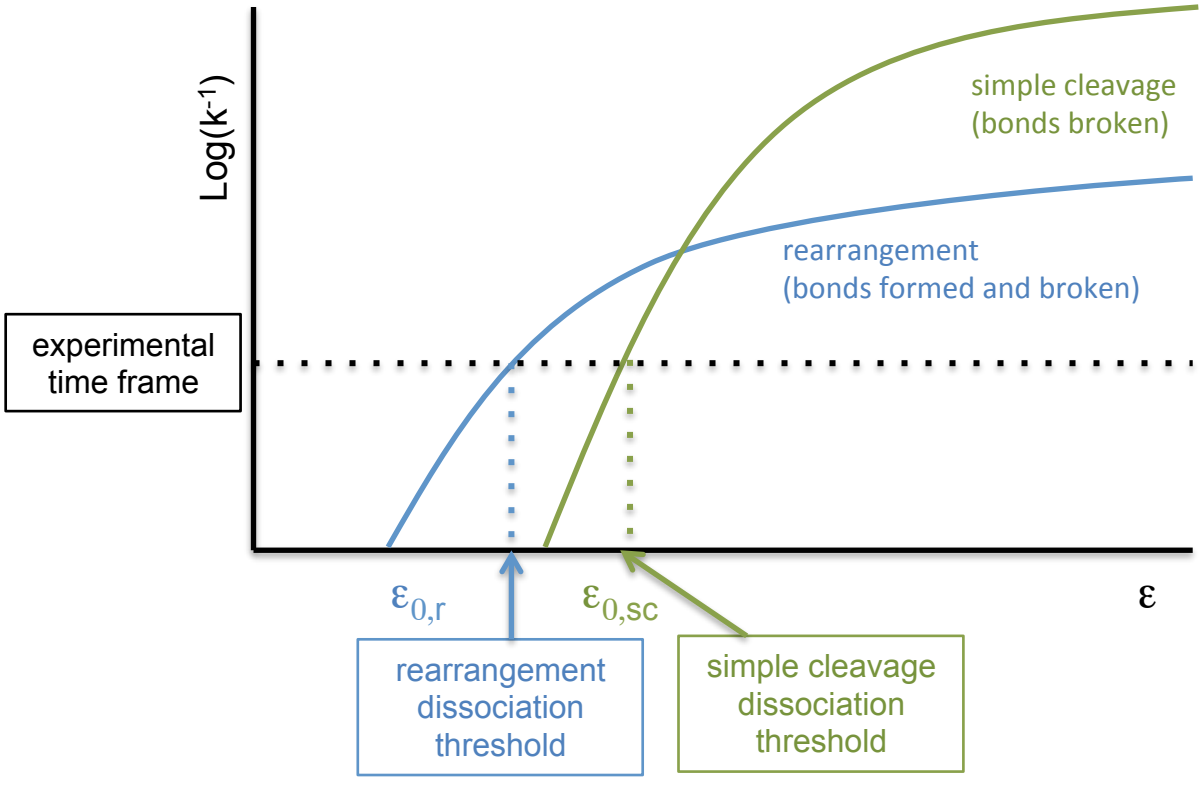


Figure 2. An depiction of the difference between the critical energy, and the dissociation threshold which takes the experimental time frame into account.

dissociation threshold and the critical energy is the kinetic shift.

2.2 Collision induced dissociation in a QIT

The most common type of reaction used in MS^n experiments is unimolecular dissociation, and the most common means of inducing dissociation in MS^n experiments is by collisions with an inert background gas[3]. QITs operate with a background pressure of ~ 1 mTorr of helium which serves to kinetically damp ion trajectories to the center of the trap, resulting in improved trapping efficiencies, resolution, and sensitivity[22,23]. While ions are trapped they frequently undergo collisions with the helium bath gas. When the collisions are inelastic some of the kinetic energy of the ion can be converted into internal energy of the ion, and vice versa[24]. Under the experimental conditions used to obtain a mass spectrum, as opposed to an MS^n experiment, the energy gained by the collisions is insufficient to cause dissociation[23].

In a collision induced dissociation (CID) experiment, the mass-to-charge specific frequency of a trapped ion is utilized. An oscillating electric field in resonance with the frequency of parent ion oscillation is used to increase the amplitude of the parent ion trajectory. The collisions with the helium bath gas now occur at a greater kinetic energy, and the amount of internal energy gained by the parent ion can be sufficient to cause dissociation[23]. The kinetic energy of an ion trapped in a QIT, even under CID conditions, is too low to result in conversion of enough internal energy for dissociation after a single collision. The internal energy gained per collision is on the order of 0.1 to 1 eV[24], so many collisions are needed to cumulatively increase the internal energy of the parent ion above the dissociation threshold.

The power used to drive the increase in parent ion kinetic energy is chosen as a balance between causing sufficient excitation to cause dissociation without causing ejection of the ion from the trapping volume[23]. The trapping parameters used to optimize the balance between excitation and ejection also happen to impose a lower limit on the mass-to-charge of the ions that can be trapped[23,25]. Ions below a mass-to-charge value of 0.28 that of the selected parent ion are not trapped and lost from the experiment. These lost product ions represent lost information.

2.3 Infrared multiphoton photodissociation

An alternative to heating ions by collisions with an inert gas is irradiation with an infrared laser. Similar to a single collision event, absorption of a single infrared photon is insufficient for a parent ion to reach the dissociation threshold. Since multiple infrared photons are required to effect dissociation, the technique is termed infrared multiphoton photodissociation (IRMPD).

IRMPD is energetically similar to CID (vide infra), but there are some differences in the results obtained from the two MS^n techniques. Unlike CID, activation in IRMPD is independent of trapping parameters. Therefore low mass-to-charge product ions that may be lost in a CID experiment are observed in an IRMPD experiment. In CID, the excitation voltage is only in resonance with the parent ion. Therefore product ions formed from CID begin to cool once they are formed. Conversely product ions formed from IRMPD may still be in the path of the laser. Thus when a product ion is formed its internal energy may continue to increase by photon absorption, leading to sequential dissociation. So while the types of product ion channels observed in CID and IRMPD are the same, the differences in trapping parameters and sequential dissociation may lead to different MS^n spectra.

2.4 Both CID and IRMPD are considered "slow heating" techniques

Neither a single collision event nor absorption of a single infrared photon is sufficient to cause dissociation, so both CID and IRMPD proceed through many excitation steps. In between excitation events, the energy imparted to the ion is rapidly redistributed throughout all vibrational modes. Because the time between excitation events is long compared to the time required for vibrational redistribution, both collisional activation and infrared irradiation are considered "slow heating" techniques[9]. The time for vibration energy redistribution means that the internal energy of the ion is statistically distributed across all vibrational modes and when the ion dissociates there is no "memory" of how the energy was imparted in the first place. A statistical energy distribution meets the criteria of an activated ion that will dissociate by RRKM kinetics. As RRKM theory predicts, dissociation of a parent ion through a particular dissociation channel will depend only on the internal energy of the parent ion. Hence the same types of product ions are formed by CID and IRMPD, despite the difference in the mechanism of conferring internal energy to the parent ion.

2.5 The difference between IRMPD in a QIT and an FT-ICR

IRMPD experiments described in this dissertation were performed in both a QIT and an FT-ICR, and there are some differences in performing IRMPD in these two analyzers. The IRMPD experiments in the QIT have been described in detail[26]. One particular feature of IRMPD in a QIT is a laser beam focused to approximately the size of the ion cloud. The optimal overlap between the ion cloud and the laser results in efficient dissociation. In the FT-ICR cell, a multipass configuration was used[27], which means overlap between the ion cloud and the laser was incomplete. Ions in the FT-ICR will pass into and out of the laser beam, and will therefore oscillate between heating and cooling periods. The cooling

mechanism for these ions, however, is radiative. For the size ions used in the experiments described herein, the rate of radiative cooling is slow compared to an oscillation period[28,29]. Thus sufficient heating for dissociation is still achieved.

A feature of QITs in general, as mentioned above, is the ~ 1 mTorr of a helium bath gas. In between photon absorption events, ions will undergo cooling collisions with the ambient helium. Unlike the radiative decay seen in FT-ICRs, collisional cooling is competitive with infrared heating, and can inhibit dissociation efficiencies by IRMPD[26,30]. In fact, several techniques have been developed to circumvent or compensate for collisional cooling during IRMPD in QITs[26,31-37].

2.6 Structural measurement by infrared spectroscopy at the Free Electron Laser for Infrared Experiments

Most IRMPD experiments are performed with a fixed-wavelength (10.6 mm) CO_2 laser because the output powers are high and the spectral profile is broad enough to ensure dissociation of the vast majority of analytes. If, however, the laser had a tunable wavelength of sufficiently narrow resolution, dissociation would only occur when the wavelength of the laser was in resonance with a vibrational transition in the trapped ions. The wavelength specific form of IRMPD is denoted as r-IRMPD. In practice a gas phase r-IRMPD spectrum is composed from a series of MS^n experiments (Figure 3). The amount of dissociation at each wavelength is indicative of the absorption intensity at that wavelength. The amount of dissociation in an MS^n spectrum can be quantified by the amount of parent ion depletion (equation 2), the amount of product in formed (equation 3), or the dissociation efficiency (equation 4). The practice of using a chemical signature, such as unimolecular dissociation, to indicate the resonant absorption of a photon is termed action spectroscopy.

The development of action spectroscopy to obtain the infrared spectra of gas phase ions trapped in a mass spectrometer was a significant development[38-41]. Trapped gas phase ions do not have a sufficient number density for the traditional direct absorption scheme where a decrease in the intensity of light passed through a sample is used to measure resonant absorption. There simply are not enough ions to produce a measurable difference in the light

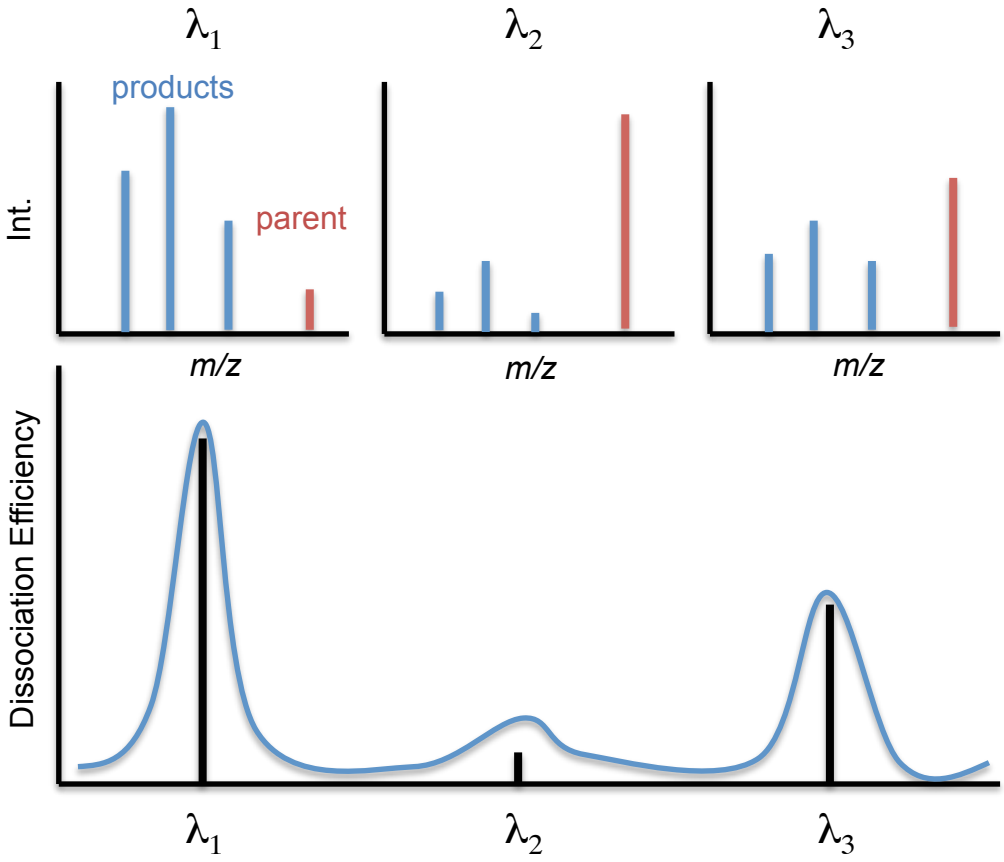


Figure 3. An illustration to describe the action spectroscopy method of using a series of MS^n spectra at different wavelengths to compile an infrared spectrum.

$$\text{parent depletion} = 1 - \frac{\text{parent}}{\text{parent}_0} \qquad \text{equation 2}$$

$$\text{product formation} = \sum \text{product ions} \qquad \text{equation 3}$$

$$\text{DE} = \frac{\sum \text{product ions}}{\sum \text{product ions} + \text{parent ion}} \qquad \text{equation 4}$$

intensity, even at the strongest absorption bands. To obtain sufficient sensitivity in infrared action spectroscopy by r-IRMPD, high laser powers are required. Free electron lasers (FEL) were the first to accomplish r-IRMPD of trapped ions in a mass spectrometer, and are well suited to these experiments because of their high output powers and tunability across the fingerprint region of the infrared spectrum (~800 to 2000 cm^{-1}). FELs work by accelerating electrons to relativistic speeds. The fast electrons are then wiggled through a series of magnets of alternating poles, and the oscillation of the electron beam generates light[42,43]. Unfortunately, FELs require a substantial infrastructure to support the underground accelerators used to generate the relativistic electrons that most universities cannot support. There are a few facilities that allow visiting scientists to apply for beam time and use of their mass spectrometers to perform spectroscopy experiments. One such facility is the Free Electron Laser of Infrared Experiments (FELIX) in the Netherlands, which was used for some of the experiments described in Chapters 6 and 7.

Prior to the development of ion spectroscopy in trapped analyzers, experiments were performed in molecular beam instruments[44-47]. Only a handful of these instruments were able to perform the spectroscopy on a single species of ion rather than a mixture, and none had the ability to form ion species other than what was produced from the ion source. Combining IR spectroscopy and trapping mass spectrometers enables the combination of the MS^n abilities of trapping mass spectrometers to prepare many types of ions with the detailed information provided by infrared spectra. Both of these attributes are necessary for the study of the structure of peptide fragments (Chapters 6 and 7).

References

(1) Allison, J.; Stepnowski, R. M. *Analytical Chemistry* **1987**, *59*, 1072A.

(2) Fulford, J. E.; Hoa, D.-N.; Hughes, R. J.; March, R. E.; Bonner, R. F.; Wong, G. J. *J. Vac. Sci. Technol.* **1980**, *17*, 829.

(3) McLuckey, S. A.; Mentinova, M. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 3.

(4) Glish, G. L.; Hemberger, P. H.; Cooks, R. G. *Analytica Chimica Acta* **1980**, *119*, 137.

(5) Yost, R. A.; Enke, C. G. *Journal of the American Chemical Society* **1978**, *100*, 2274.

(6) Johnson, J. V.; Yost, R. A.; Kelley, P. E.; Bradford, D. C. *Analytical Chemistry* **1990**, *62*, 2162.

(7) Lindemann, F. A.; Arrhenius, S.; Langmuir, I.; Dhar, N. R.; Perrin, J.; Lewis, W. C. M. *Transactions of the Faraday Society* **1922**, *17*, 598.

(8) Hinshelwood, N. *Proceedings of the Royal Society of London Series A* **1926**, *113*, 230.

(9) McLuckey, S. A.; Goeringer, D. E. *Journal of Mass Spectrometry* **1997**, *32*, 461.

(10) Baer, T.; Hase, W. L. *Unimolecular Reaction Dynamics*; Oxford University Press: New York, 1996.

(11) Baer, T.; Mayer, P. M. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 103.

(12) Kassel, L. S. *The Journal of Physical Chemistry* **1928**, *32*, 225.

(13) Marcus, R. A.; Rice, O. K. *Journal of Physical and Colloid Chemistry* **1951**, *55*, 894.

(14) Rice, O. K.; Ramsperger, H. C. *Journal of the American Chemical Society* **1927**, *49*, 1617.

(15) Rice, O. K.; Ramsperger, H. C. *Journal of the American Chemical Society* **1928**, *50*, 617.

(16) Rosenstock, H. M.; Wallenstein, M. B.; Wahrhaftig, A. L.; Eyring, H. *Proceedings of the National Academy of Sciences of the United States of America* **1952**, *38*, 667.

(17) Grant, E. R.; Schulz, P. A.; Sudbo, A. S.; Shen, Y. R.; Lee, Y. T. *Physical Review Letters* **1978**, *40*, 115.

(18) Asano, K. G.; Goeringer, D. E.; McLuckey, S. A. *Int. J. Mass Spectrom.* **1999**, *185/186/187*, 207.

(19) Sleno, L.; Volmer, D. A. *Journal of Mass Spectrometry* **2004**, *39*, 1091.

(20) Hart, K. J.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 250.

(21) Colorado, A.; Brodbelt, J. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 1116.

(22) Wells, J. M.; Gill, L. A.; Ouyang, Z.; Patterson, G. E.; Plass, W.; Badman, E. R.; Amy, J. W.; Cooks, R. G.; Schwartz, J. C.; Stafford, G. C.; Senko, M. W. In *The 46th ASMS Conference on Mass Spectrometry and Allied Topics*; ASMS: Orlando, FL, 1998.

(23) Louris, J. N.; Cooks, R. G.; Syka, J. E. P.; Kelley, P. E.; Stafford, G. C.; Todd, J. F. J. *Analytical Chemistry* **1987**, *59*, 1677.

(24) Goeringer, D. E.; McLuckey, S. A. *Journal of Chemical Physics* **1996**, *104*, 2214.

(25) Enyenihi, A. A.; Griffiths, J. R.; Glish, G. L. In *56th Conference on Mass Spectrometry and Allied Topics* Denver, CO, 2008.

(26) Newsome, G. A.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1127.

(27) Oomens, J.; Sartakov, B. G.; Meijer, G.; Von Helden, G. *Int. J. Mass Spectrom.* **2006**, *254*, 1.

(28) Price, W. D.; Williams, E. R. *Journal of Physical Chemistry A* **1997**, *101*, 8844.

(29) Price, W. D.; Schnier, P. D.; Williams, E. R. *Analytical Chemistry* **1996**, *68*, 859.

(30) Black, D. M.; Payne, A. H.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 932.

(31) Payne, A. H.; Glish, G. L. *Analytical Chemistry* **2001**, *73*, 3542.

(32) Boué, S. M.; Stephenson, J. L.; Yost, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1391.

(33) Colorado, A.; Shen, J. X.; Vartanian, V. H.; Brodbelt, J. *Analytical Chemistry* **1996**, *68*, 4033.

(34) Hashimoto, Y.; Hasegawa, H.; Yoshinari, K.; Waki, I. *Analytical Chemistry* **2003**, *75*, 420.

(35) Crowe, M. C.; Brodbelt, J. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1581.

(36) Pikulski, M.; Hargrove, A.; Shabbir, S. H.; Anslyn, E. V.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 2094.

(37) Pikulski, M.; Wilson, J. J.; Aguilar, A.; Brodbelt, J. S. *Analytical Chemistry* **2006**, *78*, 8512.

(38) Oomens, J.; Roij, A. J. A. v.; Meijer, G.; Helden, G. v. *The Astrophysical Journal* **2000**, *542*, 404.

(39) Stearns, J. A.; Mercier, S.; Seaiby, C.; Guidi, M.; Boyarkin, O. V.; Rizzo, T. R. *Journal of the American Chemical Society* **2007**, *129*, 11814.

(40) Stearns, J. A.; Guidi, M.; Boyarkin, O. V.; Rizzo, T. R. *Journal of Chemical Physics* **2007**, *127*.

(41) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *Journal of the American Chemical Society* **2007**, *129*, 13820.

(42) Putter, M.; von Helden, G.; Meijer, G. *Chemical Physics Letters* **1996**, *258*, 118.

(43) von Helden, G.; Holleman, I.; Putter, M.; Meijer, G. *Nuclear Instruments & Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms* **1998**, *144*, 211.

(44) Lisy, J. M. *Journal of Chemical Physics* **2006**, *125*.

(45) Huang, X. C.; McCoy, A. B.; Bowman, J. M.; Johnson, L. M.; Savage, C.; Dong, F.; Nesbitt, D. J. *Science* **2006**, *311*, 60.

(46) Duncan, M. A. *Int. J. Mass Spectrom.* **2000**, *200*, 545.

(47) Headrick, J. M.; Diken, E. G.; Walters, R. S.; Hammer, N. I.; Christie, R. A.; Cui, J.; Myshakin, E. M.; Duncan, M. A.; Johnson, M. A.; Jordan, K. D. *Science* **2005**, *308*, 1765.

Chapter 3

Efficient Dissociation of Large Peptides by IRMPD

3.1 Introduction

Once volatilization and ionization of large analytes became routine the breadth of samples studied with mass spectrometry has expanded to include large biopolymers. Analyzer technology has, for the most part, been able to keep up with this new field with sufficient resolution, accuracy, and speed. Dissociation techniques, however, remain too inefficient to obtain enough structural information for confident identification[1-4]. CID has endured as the dominant dissociation technique because it is effective for a wide variety of ions and comparatively easy to implement[5-8]. Nonetheless there is a fundamental limitation on CID efficiency for large analytes due to the large number of internal modes into which energy can be partitioned[2,9-11].

CID is considered a "slow-heating" technique because multiple collisions are required to cause dissociation. Similarly dissociation in IRMPD requires the absorption of multiple infrared photons. For both techniques, the energy absorption event is followed by statistical redistribution of the energy throughout the ion, and dissociation once there is enough energy in the dissociative reaction coordinate to account for the critical energy plus the kinetic shift[9,10]. For larger analytes, energy is distributed into so many modes that achieving a significant probability of accumulating enough energy in one mode requires a large amount

of total energy in the ion, which explains the aforementioned limitation on CID efficiency with size.

It is commonly believed that since IRMPD operates by a similar mechanism to CID it would suffer the same size limitation. Furthermore, in IRMPD, activation by photons competes with deactivating collisions with the bath gas, so an overall greater amount of energy must be absorbed as compared to a CID experiment[12]. The work presented in this chapter demonstrates an increase in IRMPD efficiency with the size of the analyte, which is the opposite of the trend observed for CID. Many techniques to improve IRMPD efficiencies have been published, including: decreasing pressure to limit collisional deactivation[13,14]; heating the bath gas for thermally assisted IRMPD[12]; combining IRMPD with collisional activation[15]; attaching chromogenic labels that increase the absorption cross-section and/or decrease dissociation thresholds[16-18]; supercharging to reduce the dissociation threshold and increase the number of mobile protons[19]; and focusing the laser to increase photon density in the region of overlap between the ion cloud and the laser beam[20]. The increase in efficiency with size observed in the work presented here, however, is due to the particular combination of time, pressure, and photon flux that enables the increase in IR absorption rate to exceed competing deactivation processes.

3.2 Experimental Methods

Samples were purchased from Sigma Chemical Co. (St. Louis, MO.) and used without further purification. FLEEV, YGGFLRR, bradykinin, granuliberin-R, and cardiodilatin were prepared as 100 μM solutions in 49:49:2 methanol:water:acetic acid by volume. Melittin and bovine ubiquitin were prepared as 50 mM solutions in 75:20:5 and 75:22:3 methanol:water:acetic acid, respectively. YGGFL was prepared as a 100 μM

solution in 75:20:5 methanol:water:acetic acid. A custom-built nano-electrospray source was used to generate ions. The parent ions used in for IRMPD experiments were as follows: $[YGGFL+H]^+$ (556 Da, *m/z* 556, 79 atoms); $[FLEEV+H]^+$ (636 Da, *m/z* 636, 99 atoms); $[YGGFLRR+2H]^+$ (869 Da, *m/z* 435, 128 atoms); $[bradykinin+2H]^{2+}$ (1062 Da, *m/z* 531, 156 atoms); $[granuliberin\text{-}R+2H]^{2+}$ (1425 Da, *m/z* 713, 210 atoms); $[cardiodilatin+2H]^{2+}$ (1831 Da, *m/z* 915, 249 atoms); $[melittin+4H]^{4+}$ (2850 Da, *m/z* 712, 433 atoms); $[ubiquitin+11]^{11+}$ (8.5 kDa, *m/z* 778, 1265 atoms).

Experiments were performed on a modified Finnigan ITMS controlled with ICMS software[21]. Base pressure in the quadrupole ion trap mass spectrometer (QITMS) is 2×10^{-5} Torr. Helium was added to the trapping volume to a constant pressure of 6.7×10^{-4} Torr. IRMPD experiments were performed with a 50 W Synrad CO_2 laser (10.6 μM wavelength) triggered by a TTL pulse from the ICMS software. The laser beam is focused to the center of the trap using a 38.1 cm focal length ZnSe lens, and passes into the vacuum housing through a ZnSe window. There are two holes in the ring electrode: one that allows the laser beam to enter the trapping volume, and another on the opposite side of the trap lets the beam pass into an Aerodag-coated Cajon fitting used as a beam dump. The laser is pre-aligned using reflected beams, then finely adjusted by optimizing dissociation efficiency.

The ICMS software allows for complete customization of the scan function, including rf trapping and timing parameters. Ions are injected into the trap for as long as required to obtain sufficient ion signal for isolation and dissociation, generally from 10 to 200 ms. The ions are then isolated using the amplitude of the rf trapping field and resonance ejection. Following isolation, ions are allowed to cool for 10 ms to ensure thermal equilibrium with the helium bath gas, and to allow ions to kinetically cool to the center of the trap after the

isolation ramps. After the cool time, ions are irradiated with the CO_2 laser for dissociation experiments, or left for an extra cool period to measure the initial parent ion intensity. Ten scans are taken and averaged with the laser off followed by ten scans with the laser on for a particular irradiation time. Percent dissociation is calculated as P/P_0, where P is the intensity of the parent ion in the laser-on experiment and P_0 is the intensity of the parent ion in the preceding laser-off experiment.

The irradiation time is increased incrementally from 0.1 ms to the amount of time required to produce 100% dissociation of the parent ion. The irradiation time required to achieve 50% dissociation of the parent ion (t_{50}) is the metric used to compare dissociation efficiency between peptides and proteins of different sizes. It is rarely the case that an irradiation time that resulted in exactly 50% dissociation was selected during the experiment, so surrounding data points were used to interpolate t_{50}. The purpose was not to find a quantitative relationship between t_{50} and the number of atoms (N), and the conclusions drawn from the results are the same if we use this method, or a number of other methods tried.

3.3 Results and Discussion

Since the maximum value for dissociation efficiency varies with the charge state of an analyte, the time required to achieve 50% dissociation of the parent ion (t_{50}) was used to compare relative amounts of dissociation between analytes (Figure 1). As the size of the ion increases, t_{50} decreases, showing the IRMPD process gains efficiency with an increase in ion size. The particular charge state for each analyte was chosen to maintain as similar a m/z value as possible, so the observed effects are the result of N, and not a difference in charge density in the ions or distribution within the ion trap. The data was split into two classes of ions to show that arginine-containing peptides took longer to dissociate than comparably

sized peptides that did not contain an arginine residue. A slower rate for arginine-containing peptides was expected as arginine-containing peptides were previously observed to require higher energies to dissociate in CID experiments[22-26]. It is thought that a hydrogen bond is formed between a protonated arginine side chain and a carbonyl on the backbone, increasing the number of interactions that must be disrupted to form product ions[26-29].

In 1922 Lindemann[30], and later in 1926 Hinshelwood[31], used a two-step mechanism to describe unimolecular dissociation. The first step is an energy deposition step, followed by dissociation of the activated ion. The rate of the dissociation step can be described using Rice-Ramsberger-Kassel-Marcus (RRKM) kinetic theory, which predicts that the rate of dissociation will decrease given an increase in the density of states of the reactant relative to the number of states available in the transition state.

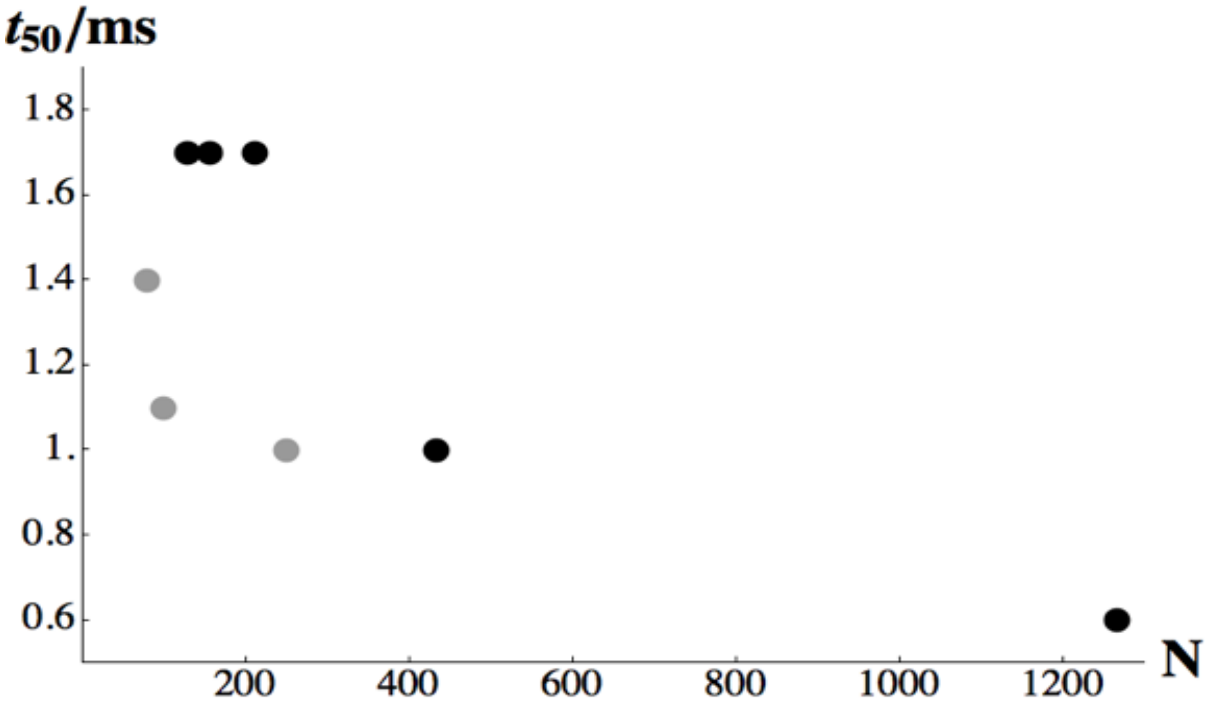


Figure 1. The time required to achieve 50% dissociation of a parent ion in milliseconds is plotted as a function of the number of atoms in the parent ion. The gray points correspond to peptides that do not contain arginine and the black points correspond to arginine-containing peptides, which are expected to require more energy for dissociation

$$k_{diss} = \frac{\sigma}{h}\frac{N^{\neq}(\varepsilon - \varepsilon_0)}{\rho(\varepsilon)}$$ equation 1

For a given energy, a larger reactant will have a larger density of states and a slower dissociation rate. The slower dissociation rate for larger analytes occurs because the number of states into which the available energy can be partition is increased, thereby decreasing the probability that a reactive state will have enough energy to dissociate. Consequently, to achieve an increase in dissociation efficiency with analyte size, the rate of energy absorption must be greater for larger analytes. Furthermore, the increase in the rate of dissociation due to energy absorption must outpace the decrease in the rate of dissociation due to a greater number of internal degrees of freedom.

While unimolecular reactions can be described as formation of an excited state followed by dissociation, there are multiple sub-reactions that govern the number of excited species present in the system. In an IRMPD experiment, the following are the elementary reactions involved in the excited state population:

$$M^+ + h\nu \xrightarrow{k_{IR}} M^{+*}$$
$$M^{+*} \xrightarrow{k_{rad}} M^+ + h\nu$$
$$M^{+*} + He \xrightarrow{k_{cc}} M^+ + He$$
$$M^{+*} \xrightarrow{k_{diss}} products$$

As discussed above, it is expected that the rate of dissociation of an activated species (k_{diss}) is smaller for larger ions. For product formation to be more efficient for larger ions, the population of activated species must be greater. The population of activated species is increased by IR irradiation (k_{IR}) and depleted by both spontaneous and induced radiative cooling (k_{rad}),and collisional (k_{cc}) cooling.

Most IRMPD experiments are performed using a CO_2 laser with a central wavelength of 10.6 μm. Peptides and proteins can be poor absorbers in the 10.6 μm region of the spectrum (see Chapter 6 and 8), but IRMPD experiments are successful because of the wavelength spread of the laser, and the high powers used. For polymeric species, the number of vibrations increases with size, which increases the probability of a vibrational transition that overlaps the CO_2 laser. The chormophores in a peptide that absorb around 10.6 μm are backbone carbonyl bending modes. Since a carbonyl functional group is present in all amino acids, the increase in absorption cross-section should change with the number of residues even if the amino acid side chains do not absorb in this region. Despite the small absorption cross-section in the wavelength region of a CO_2 laser, it is expected that energy gain due to photon absorption will increase as the size of the peptide or protein increases.

In these experiments a focused laser was used, enabling dissociation in approximately 1 ms. The relatively short time scale used here means that both the radiative and collisional cooling mechanisms are less significant here than in many other IRMPD experiments. The radiative cooling rate is expected to be slow compared to absorption and collisional cooling under these conditions. Therefore any changes in the radiative cooling rate as a function of size are likely unimportant to these experiments. The collisional cooling rate will depend on the number of collisions and the amount of energy removed per collision. Indeed, both of these values are expected to increase with the size of an ion so collisional deactivation should be more efficient for larger analytes. However, the short time scale of these experiments limits the amount of collisional deactivation.

To explain the observed increase in efficiency with size, the increase in the photon absorption rate must exceed the increase in the rate of competing cooling processes for this

particular set of experimental parameters. The relative importance of each of the activation and deactivation processes will change depending on time, pressure, and photon flux. IRMPD in a quadrupole ion trap using a focused 50 W CO_2 laser evidently works under experimental conditions that make dissociation of relatively large analytes efficient. Compared to other methods used to achieve dissociation of large analytes, the experimental configuration presented here is relatively inexpensive, easy to implement, and amenable to high-throughput applications.

3.4 Conclusions

IRMPD was shown to occur faster as the size of peptide and protein ions increased. It is hypothesized that this phenomenon is due to an increase in the energy absorption rate with size that exceeds competitive cooling processes and results in efficient dissociation despite a statistical energy distribution over a large number of modes. Ion activation methods that are able to impart large amounts of energy into an ion, such as surface induced dissociation (SID) and blackbody infrared radiative dissociation (BIRD), are able overcome the statistical limits imposed by large ions, but suffer time and instrumentation limitations that prevent their widespread use. Compared to SID and BIRD, IRMPD in a quadrupole ion trap is relatively easy to implement, fast, and inexpensive, which makes it a promising technique for high-throughput applications, including proteomics studies. In comparison to CID, IRMPD is more efficient for large analytes and the trapping parameters during IRMPD can be set to include smaller ions than that allowed by CID, allowing the acquisition of additional information in the low mass range of product ions that are normally lost in CID experiments.

While dissociation of the parent ion is highly efficient for large ions using IRMPD, it is not clear whether this technique will give rise to improved sequencing. Large peptides and proteins have access to many dissociation channels, such that MS^2 spectra are highly congested. Since large amounts of internal energy are required for dissociation it is likely that excess internal energy remains in the product ions after dissociation, making subsequent dissociation steps facile, especially since all ions remain in the path of the laser during IRMPD experiments. Selective broadband-IRMPD has been shown to prevent subsequent dissociation steps by shifting the laser position relative to the ion cloud such that only selected ions are irradiated. Efficient dissociation of the parent ion by IRMPD by the selective broadband method shows promise as a structural tool for large analytes.

References

(1) Kelleher, N. L. *Analytical Chemistry* **2004**, *76*, 196A.

(2) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Analytical Chemistry* **1994**, *66*, 2809.

(3) Jockusch, R.; Schnier, P.; Price, W.; Strittmatter, E.; Demirev, P.; Williams, E. *Analytical Chemistry* **1997**, *69*, 1119.

(4) Reid, G. E.; Wu, J.; Chrisman, P. A.; Wells, J. M.; McLuckey, S. A. *Analytical Chemistry* **2001**, *73*, 3274.

(5) Sleno, L.; Volmer, D. A. *Journal of Mass Spectrometry* **2004**, *39*, 1091.

(6) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599.

(7) Futrell, S. A. J. *Journal of Mass Spectrometry* **2000**, *35*, 1069.

(8) Jennings, K. R. *Int. J. Mass Spectrom.* **2000**, *200*, 479.

(9) Baer, T.; Mayer, P. M. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 103.

(10) Baer, T.; Hase, W. L. *Unimolecular Reaction Dynamics*; Oxford University Press: New York, 1996.

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

(12) Payne, A. H.; Glish, G. L. *Analytical Chemistry* **2001**, *73*, 3542.

(13) Boué, S. M.; Stephenson, J. L.; Yost, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1391.

(14) Colorado, A.; Shen, J. X.; Vartanian, V. H.; Brodbelt, J. *Analytical Chemistry* **1996**, *68*, 4033.

(15) Hashimoto, Y.; Hasegawa, H.; Yoshinari, K.; Waki, I. *Analytical Chemistry* **2003**, *75*, 420.

(16) Crowe, M. C.; Brodbelt, J. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1581.

(17) Pikulski, M.; Hargrove, A.; Shabbir, S. H.; Anslyn, E. V.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 2094.

(18) Pikulski, M.; Wilson, J. J.; Aguilar, A.; Brodbelt, J. S. *Analytical Chemistry* **2006**, *78*, 8512.

(19) Madsen, J. A.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 349.

(20) Newsome, G. A.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1127.

(21) Yates, N.; Yost, R.; 2.20 ed.; unpublished: University of Florida, Gainesville, FL, 1992.

(22) Huang, Y.; Triscari, J. M.; Pasa-Tolic, L.; Anderson, G. A.; Lipton, M. S.; Smith, R. D.; Wysocki, V. H. *Journal of the American Chemical Society* **2004**, *126*, 3034.

(23) Tabb, D. L.; Huang, Y. Y.; Wysocki, V. H.; Yates, J. R. *Analytical Chemistry* **2004**, *76*, 1243.

(24) Dongré, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *Journal of the American Chemical Society* **1996**, *118*, 8365.

(25) Tsaprailis, G.; Nair, H.; Somogyi, A.; Wysocki, V. H.; Zhong, W.; Futrell, J. H.; Summerfield, S. G.; Gaskell, S. J. *Journal of the American Chemical Society* **1999**, *121*, 5142.

(26) Vachet, R. W.; Asam, M. R.; Glish, G. L. *Journal of the American Chemical Society* **1996**, *118*, 6252.

(27) Jones, J. L.; Dongré, A. R.; Somogyi, A.; Wysocki, V. H. *Journal of the American Chemical Society* **1994**, *116*, 8368.

(28) Somogyi, A.; Wysocki, V. H.; Mayer, I. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 704.

(29) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *Journal of the American Chemical Society* **1996**, *118*, 8365.

(30) Lindemann, F. A.; Arrhenius, S.; Langmuir, I.; Dhar, N. R.; Perrin, J.; Lewis, W. C. M. *Transactions of the Faraday Society* **1922**, *17*, 598.

(31) Hinshelwood, N. *Proceedings of the Royal Society of London Series A* **1926**, *113*, 230.

Chapter 4

Measuring an Effective Temperature of Vibrationally Excited Ions

4.1 Introduction

As discussed in Chapter 3, it was observed that during IRMPD experiments proteins require less irradiation time to dissociate than peptides. This observation is contrary to what is observed in CID experiments, where proteins require longer activation times and/or higher excitation amplitudes. Moreover, a molecule with many degrees of freedom has more states to distribute energy into, so the probability that any single state contains enough energy to exceed the dissociation threshold is low. It is known that IR absorption cross sections will increase with the size of an ion[1]. The time/size relationship observed in IRMPD experiments led to the hypothesis that the increase in absorption cross section is more significant than the reduction in the rate of dissociation due to competing cooling factors and the increase in degrees of freedom. This chapter describes theoretical calculations and experimental measurements to determine the increase in enthalpy of an ion after a period of IR irradiation.

A change in enthalpy is given by equation 1, where ΔH is the change in enthalpy due

$$\Delta H = \int_{T_i}^{T_f} C_p \, dT \qquad \text{(equation 1)}$$

to irradiation, C_p is the constant pressure heat capacity, and T is the effective temperature of the parent ions before and after irradiation. The heat capacity can be found using equation 2, where the vibrational frequencies are found by ab initio calculations.

$$C_p(T) = 4R + R\sum_i \left(\frac{hc\tilde{\nu}_i}{kT}\right)^2 \left(\frac{e^{hc\tilde{\nu}_i/kT}}{\left(e^{hc\tilde{\nu}_i/kT}-1\right)^2}\right) \qquad \text{(equation 2)}$$

The effective temperatures, T_i and T_f in equation 1, are quantities derived experimentally by the method described below. Both equations 1 and 2 are only true for equilibrium conditions, but the term "effective temperature" is an acknowledgement of the fact that non-Boltzmann distributions of internal energies may exist in the experiment. Quadrupole ion traps are operated at relatively high pressures (mTorr), and ions are trapped for long periods of time (> ms) such that even in the presence of trapping rf fields it is expected that equilibrium is achieved[2-5]. It is thought that a statistical energy distribution is maintained under the conditions of a CID experiment, but with a truncated Boltzmann distribution because the population of ions above the dissociation threshold leave the parent ion population and form product ions[4]. It is unknown whether equilibrium is maintained during irradiation in an IRMPD experiment. In fact, non-Boltzmann populations in an IRMPD experiment have been reported[6].

Pre-heating ions to reduce the excitation amplitudes needed for dissociation in CID experiments forms the basis of the method used here for ascribing an effective temperature to a period of IR irradiation. Thermally-assisted CID (TA-CID) is one such method performed by heating the ion trap and the helium bath gas used for trapping and collisional excitation, which pre-heats trapped ions before collisional excitation[7]. Similarly, in infrared activation CID (IRa-CID), irradiation with an infrared laser is used to pre-heat ions. IRa-CID is performed using a series of irradiation times, one of which results in the same amount of pre-heating as the heated helium bath gas of a TA-CID experiment. By comparing the results from IRa-CID and TA-CID experiments, an irradiation time can be assigned to an effective temperature (Figure 1), which is then used to calculate the increase in enthalpy due to IR

irradiation. The method of ascribing a temperature to an IR activation period for enthalpy calculations yields reasonable results for the relatively small peptide $[YGGFL+Na]^{+}$. The pitfalls of the measurement become apparent in the measurement of the larger peptide $[Melittin+4H]^{4+}$. The assumption of thermal equilibrium fails for this analyte during multiphoton IR activation. Furthermore, sequential dissociation limits the CID parameters that can be used.

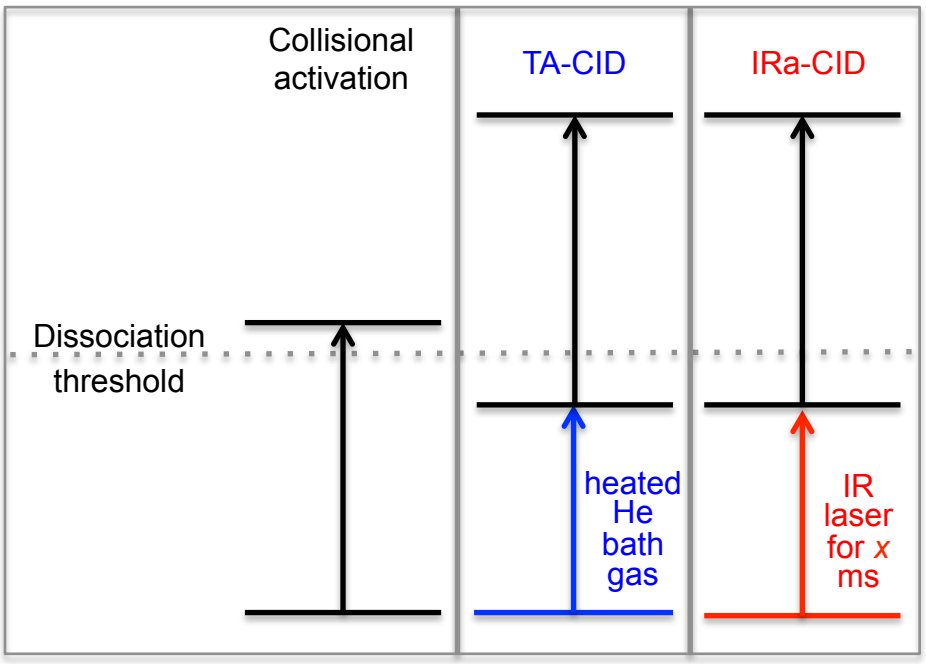


Figure 1. Depiction of how an IR irradiation time can be equated with a temperature in pre-activation CID experiments. One particular IR irradiation time should result in the same enthalpy gain as a heated bath gas, which will be evidenced by the same amount of dissociation.

4.2 Experimental Methods

4.2.1 Calculation of heat capacities

As stated in the introduction, heat capacity is calculated using equation 2. The assumptions implicit in using this equation are conditions of thermal equilibrium, the heat capacity is constant with temperature, and all ions exist in the ground electronic state. While it is doubtful that equilibrium conditions are met for all phases of this experiment, there is no

suitable non-equilibrium alternative for the calculation. Also, since the work is only semi-quantitative at this stage, the equilibrium assumption is acceptable to a first approximation. Electronic excitation is not expected in these experiments so only translational, rotational, and vibrational degrees of freedom are included in the heat capacity calculation, as reflected in equation 2, which greatly simplifies the required calculations.

Density functional theory calculations performed in Gaussian using the B3LYP functional were chosen as this method has consistently been shown to have a good balance of accuracy and efficiency[8,9]. To determine which basis set to use, calculations of heat capacity using basis sets of varying rigor were compared to experimental heat capacities published by

Table 1. Heat capacities calculated using different levels of theory compared to NIST values for alcohol test sytems.

Basis set		Calculated Cp (J/ mol-Kelvin)	RMS error
B3LYP/3-21G	methanol	44.7	
	ethanol	65.52	0.0127
	propanol	84.44	
B3LYP/6-31G	methanol	45.13	
	ethanol	64.51	0.0158
	propanol	84.49	
B3LYP/6-31+G(d,p)	methanol	45.2	
	ethanol	65.08	0.0094
	propanol	85.56	
B3LYP/6-31++G(d,p)	methanol	45.23	
	ethanol	65.08	0.01
	propanol	85.43	
B3LYP/6-311++G(d,p)	methanol	45.35	
	ethanol	65.09	0.0111
	propanol	85.75	
cc-aug-pVQZ	methanol	45.42	
	ethanol	65.2	0.0155
	propanol		
NIST Values	methanol	44.06	
	ethanol	65.21	
	propanol	85.56	

the National Institute of Standards and Technology[10]. Since no experimental heat capacity data for peptides was available methanol, ethanol, and propanol were used as a model system for the test calculations (Table 1). From this test it was determined that the 3-21g basis set would provide sufficient accuracy, and the B3LYP/3-21G method was used for geometry optimization and frequency calculations for $[YGGFL+Na]^+$.

The calculated heat capacity as a function of temperature is plotted in Figure 2. As the temperature increases more vibrational modes are accessible to the population thereby increasing heat capacity. Similar calculations were not done for larger analytes as the computation time would not have been practical. A larger analyte will have more vibrational degress of freedom. More importantly at these relativley low temperatures a larger analyte will have more low frequency vibrational modes. The heat capacity of a large molecule would therefore expected to increase more rapidly with temperature.

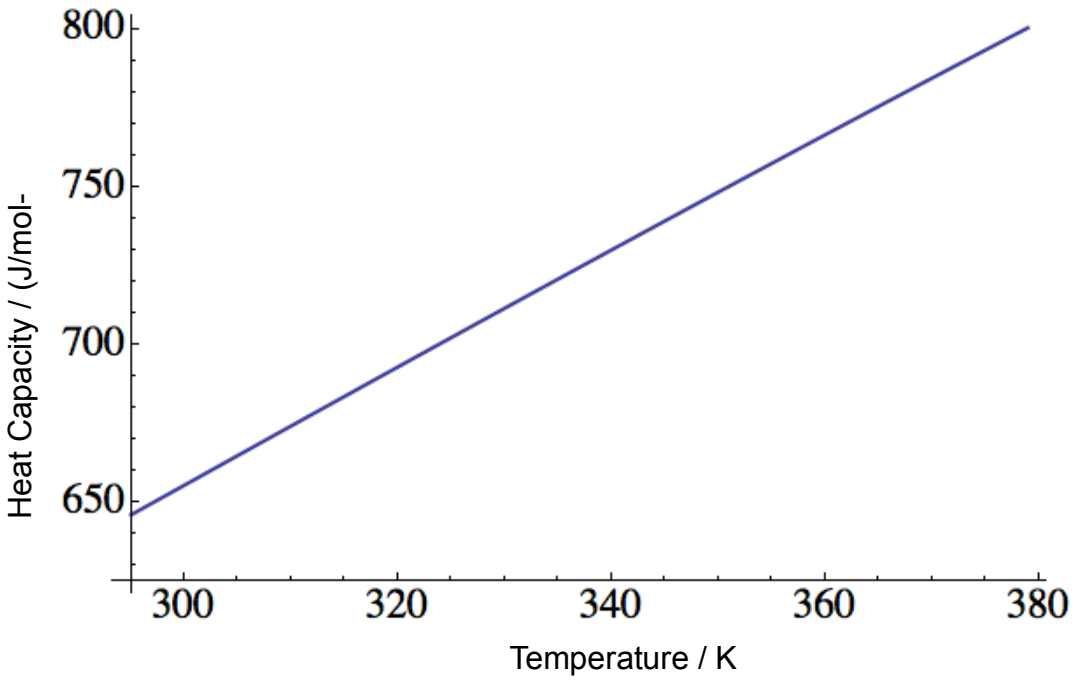


Figure 2. Heat capacity as a function of temperature calculated using equation 2.

4.2.2 TA-CID and IRa-CID experiments

The instrument used for these experiments has been described previously[11], as well as the method used for TA-CID experiments[7]. IR activation using a CO_2 laser is accomplished using the same equipment and software as the IRMPD experiments described in Chapter 3. IR activation times ranged from 0.1 ms up to just below an activation time that caused dissociation without the need for subsequent collisional activation. The series of spectra obtained at the various activation times were then compared to 373 K TA-CID spectra.

The root mean square difference between TA-CID and IRa-CID spectra normalized to the parent ion intensity was used to assign a score that indicated the degree of similarity, where the lowest score corresponds to the best match (equation 3). Both the intensities of the

$$Score = \sum_i \sqrt{\left(TA_i - IRa_i\right)^2} \qquad \text{(equation 3)}$$

entire spectrum and the parent ion plus selected product ions were used to calculate scores for each of the analytes. The score calculation is not weighted by the relative intensity of the parent or product ions, and especially for larger analytes the score based on the entire spectrum was dominated by many low intensity product ions. Hence only the few most intense product ions were selected for the calculation.

The CID parameters particular to each analyte were chosen to ensure observable dissociation, but not so much that the further increase in internal energy from pre-heating resulted in only a small observable change in the amount of dissociation. For $[YGGFL+Na]^+$, an amplitude of 400 mV over 4 ms was used. For $[Melittin+4H]^{4+}$, 200 mV over 10 ms was used, but 190 mV and 210 mV were also used to verify the technique, and yielded the same results. It was also important to avoid sequential dissociation. In a CID

experiment only the parent ion is in resonance with the applied excitation waveform. Product ions formed from CID will be cooled by a room temperature bath gas in the IRa-CID experiment, but a heated bath gas in the TA-CID experiment. The slower rate of product ion cooling in TA-CID may result in more sequential dissociation, which could be mistaken as more internal energy in the parent ion[12]. Additionally, the helium bath gas pressure was varied with temperature to maintain a constant number density between the two experiments. The number of collisions between an ion and the bath gas depends on number density, and it is important that the collisional activation aspect be consistent between TA-CID and IRa-CID. For TA-CID experiments at 373 K a helium pressure of 8.9×10^{-4} mbar was used, and 4.9×10^{-4} mbar was used for 300 K IRa-CID experiments.

4.3 Results and Discussion

4.3.1 $[YGGFL+Na]^+$

The results from matching IRa-CID to TA-CID spectra are shown in Figure 3. The match scores indicate that a 1.6 ms irradiation time in an IRa-CID experiment was most similar to pre-heating with a 373 K bath gas. The spectrum match is demonstrated visually with the plots on the right side of Figure 3, and mathematically with the plots on the left side of Figure 3. The spectra plotted in the graph on the top right of the figure are the IRa-CID spectra at various IR irradiation times, normalized to the parent ion intensity. Shown in the plot on the bottom right of the figure is the similarity between the 1.6 ms IRa-CID and 373 K TA-CID spectrum. The match scores as function of IR irradiation time in the IRa-CID experiment are shown on the left of the figure, where the bottom is a zoomed in version of the top.

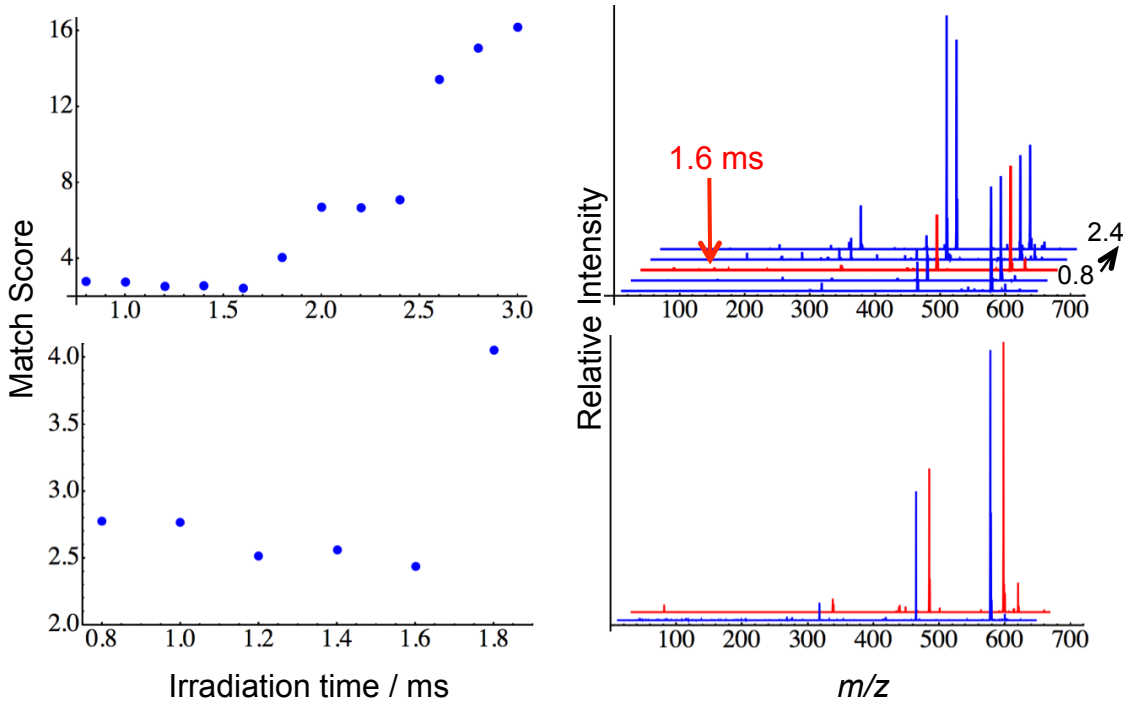


Figure 3. Results from TA-CID and IRa-CID spectra of $[YGGFL+Na]^+$. Shown on the left side of the figure are the scores for the match between 373 K TA-CID and IRa-CID at a series of irradiation times. In the top right IRa-CID spectra at various irradiation times are shown, with the matched spectrum highlighted in red. The bottom right shows the match between the 1.6 ms IRa-CID(red) and 373 K TA-CID spectra (blue).

Using the calculated vibrational frequencies the heat capacity was calculated to be 656 J mol^{-1} K^{-1}. Based on extrapolating the heat capacity values from the alcohol data as a function of the number of atoms, 656 J mol^{-1} K^{-1} is within expectations. Using a final temperature of 373 K, the increase in enthalpy after 1.6 ms irradiation was calculated to be 52.8 kJ mol^{-1}. This value seems within reason because it is less than the dissociation energy of a covalent bond, and neither 1.6 ms of irradiation nor 373 K bath gas cause dissociation.

4.3.2 [Melittin + 4H]$^{4+}$

The goal of the work presented in this chapter was to compare how the magnitude of the enthalpy increase due to IR irradiation changes with peptide size. The second analyte studied was the 2.8 kDa, 26 residue peptide Melittin. The results from the Melittin

experiments exemplify the difference between IR heating with a laser and collisional heating through a bath gas, which prevents comparison as originally intended.

Shown in Figure 4 are the results from matching TA-CID and IRa-CID spectra. The scores show that 1.0 ms of irradiation was most similar to heating with a 373 K bath gas, and the matching spectra are on the right side of the figure. The conclusion that follows from the data in Figure 4 is that the internal energy distribution that results from equilibration at 373 K is the same internal energy distribution that results from 1.0 ms of irradiation. The conclusion cannot be correct, however, because 1.0 ms of irradiation causes dissociation without subsequent collisional activation (Figure 5), whereas no dissociation is observed at 373 K without collisional activation. A possible explanation for these apparently conflicting

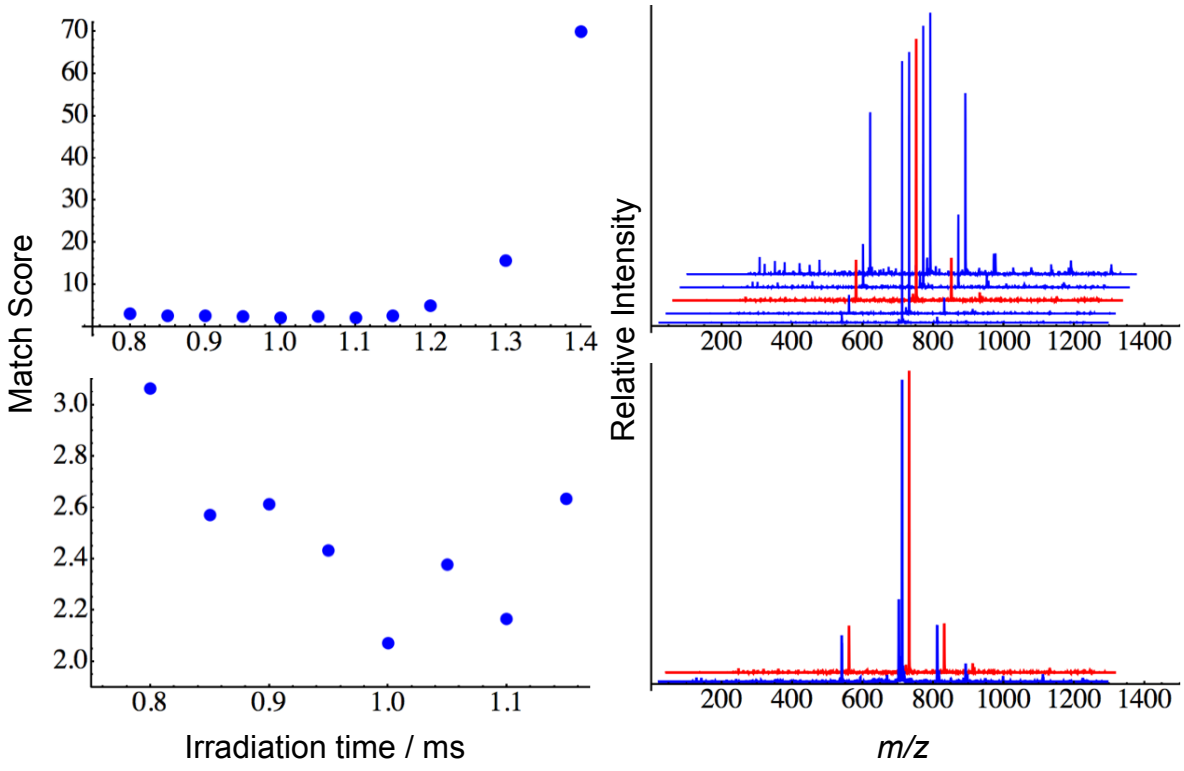


Figure 4. Results from TA-CID and IRa-CID spectra of $[\text{Melittin+4H}]^{4+}$. Shown on the left side of the figure are the scores for the match between 373 K TA-CID and IRa-CID at a series of irradiation times. In the top right IRa-CID spectra at various irradiation times are shown, with the matched spectrum highlighted in red. The bottom right shows the match between the 1.0s ms IRa-CID(red) and 373 K TA-CID spectra (blue).

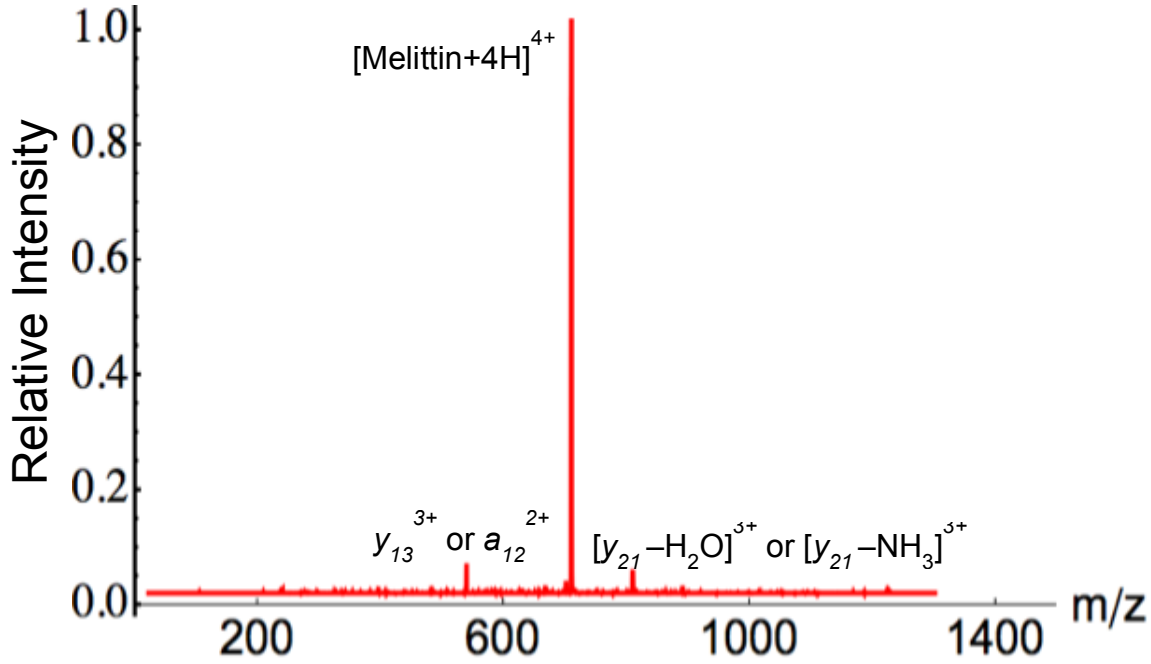


Figure 5. MS^2 spectrum of $[Melittin+4H]^{4+}$ after 1.0 ms IR irradiation with a CO_2 laser and no collisional activation.

results can be found by considering the shape of the internal energy distribution.

Figure 6 is an illustration of how the shape of the internal energy distribution can influence the amount of dissociation observed. Two parent ion internal energy distributions with the same average internal energy but different shapes are shown at the top of the figure. A hypothetical rate curve for a single dissociation channel is shown in the middle of the figure. Given an experimental time frame, the dissociation threshold can be found from the rate curve. The parent ion population with an internal energy greater than the dissociation threshold will form product ions. Given the two internal energy distributions depicted, a larger portion of the parent ions on the right will dissociate to product ions. Non-Boltzmann internal energy distributions have been previously observed in IRMPD processes[6]. They arise from absorption bands that lie higher in energy than the laser, but close enough to shift into resonance with increasing internal energy due to anharmonicity.

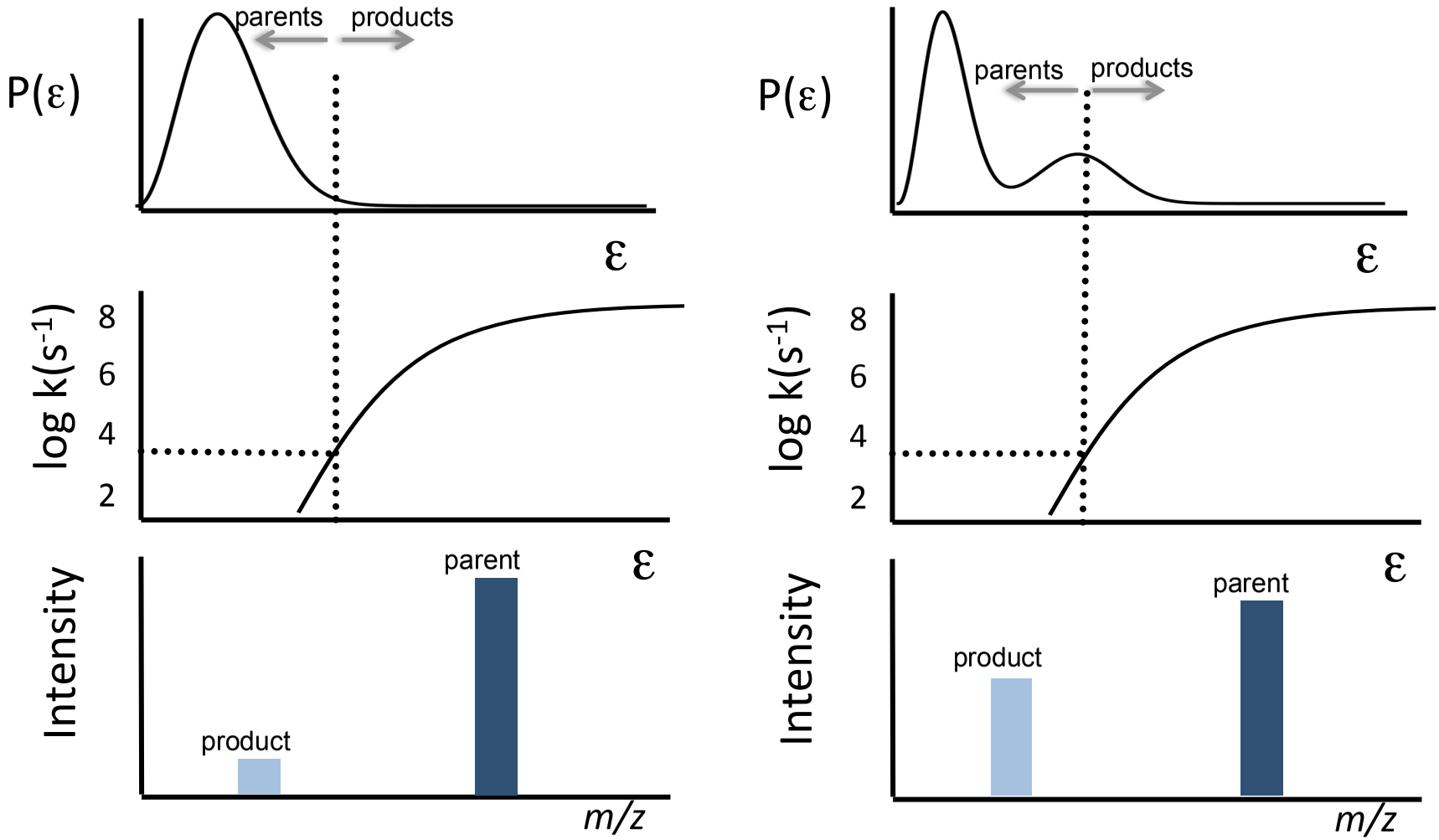


Figure 6. Illustrated in this figure is how the shape of internal energy distributions can influence the amount of dissociation observed. Two internal energy distributions with the same average internal energy, but different shapes are depicted at the top of the figure. The population of parent ions with an internal energy above the dissociation threshold at a given rate will dissociate to product ions, so a larger portion of the population of ions depicted on the right will dissociate to product ions.

IR activation of Melittin in these experiments may result in the same average internal energy as a 373 K bath gas. However the shape of the internal energy distribution after IR activation may be non-Boltzmann. In which case a high-energy subset of the population may lie above the dissociation threshold and form product ions, whereas a Boltzmann distribution of the same average internal energy may not have a sufficient population above the dissociation threshold to form products. The enthalpy calculations that were the initial goal are based on the assumption of thermal equilibrium, and it seems the non-Boltzmann distribution found for Melittin would preclude the use of this method.

The fact that non-Boltzmann energy distributions were observed for $[Melittin+4H]^{4+}$, which contains 477 atoms, and not $[YGGFL+Na]^{+}$, which contains 79 atoms, prompts similar size-dependent questions that motivated this work in the first place. Are non-Boltzmann internal energy distributions more likely to be observed in larger ions? A bimodal distribution, such as that previously calculated[6] and depicted in Figure 6, comes from an extra boost of IR absorption from the bands that shift into resonance with the laser for the subset of the population that has a high enough internal energy to reach the anharmonic part of the potential well. A larger analyte has a larger number of vibrational modes, and is therefore more likely to have modes that are close enough to shift into the laser frequency. It does make some sense, then, that the non-Boltzmann distribution was observed for $[Melittin+4H]^{4+}$ and not $[YGGFL+Na]^{+}$.

4.4 Conclusions

The method of comparing TA-CID to IRa-CID spectra to measure the increase in enthalpy of an ion after a period of IR irradiation does seem to work as long as certain pitfalls are avoided. First, sequential dissociation typical of IRMPD experiments was avoided due to

the different cooling rates of product ions in a room temperature bath gas, as in the IRa-CID experiments, versus a heated bath gas, as in the TA-CID experiments. Since the goal is to use dissociation to measure the energy of the parent ion, sequential dissociation is uninformative. The product ion lineage in protonated peptides is often complicated, and many product ions are formed by multiple pathways, so it would be difficult to account for sequential dissociation. Conversely the product ion lineage of sodiated peptides is well characterized, and product ions are formed by only one pathway. It would be possible to account for differences in sequential dissociation if sodiated peptides were used. Second, non-Boltzmann internal energy distributions violate the assumptions implicit in the calculation and can give nonsensical results. These types of distributions are more likely for larger ions, so it may be that there is an upper size limit to these measurements.

References

(1) Price, W. D.; Williams, E. R. *Journal of Physical Chemistry A* **1997**, *101*, 8844.

(2) Gronert, S. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 845.

(3) Hart, K. J.; McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 250.

(4) Goeringer, D. E.; McLuckey, S. A. *Journal of Chemical Physics* **1996**, *104*, 2214.

(5) Remes, P. M.; Glish, G. L. *Int. J. Mass Spectrom.* **2007**, *265*, 176.

(6) Oomens, J.; Sartakov, B. G.; Meijer, G.; Von Helden, G. *Int. J. Mass Spectrom.* **2006**, *254*, 1.

(7) Racine, A. H.; Payne, A. H.; Remes, P. M.; Glish, G. L. *Analytical Chemistry* **2006**, *78*, 4609.

(8) Moision, R. M.; Armentrout, P. B. *Journal of Physical Chemistry A* **2002**, *106*, 10350.

(9) Stepanian, S. G.; Reva, I. D.; Radchenko, E. D.; Rosado, M. T. S.; Duarte, M.; Fausto, R.; Adamowicz, L. *Journal of Physical Chemistry A* **1998**, *102*, 1041.

(10) http://webbook.nist.gov/chemistry/

(11) Berkel, G. J. V.; Glish, G. L.; McLuckey, S. A. *Analytical Chemistry* **1990**, *62*, 1284.

(12) Black, D. M.; Payne, A. H.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 932.

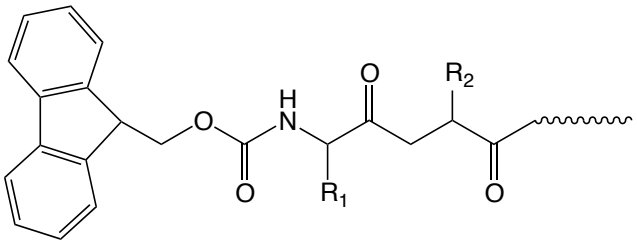
Chapter 5

C-terminal Sequencing of Sodium Cationized FMOC Derivatized Peptides

5.1 Introduction

Protonated peptides tend to dissociate by multiple pathways to form a range of product ions unpredicted by a simple backbone cleavage dogma. Conversely, sodium cationized peptides dissociate primarily by loss of the C-terminal residue[1,2]. The predictable dissociation products of sodium cationized peptides have been shown to be informative of the peptide sequence[2,3]. Unfortunately peptides exhibit a much lower binding affinity for a sodium cation than a proton, which limits the usefulness of this species for de novo identification. The purpose of this work is to begin to explore the usefulness of FMOC-derivatized peptides (structure I) in binding sodium cations for de novo sequencing and to gain insight into the mechanism of dissociation of sodium cationized peptides.

Derivatized peptides have been used to facilitate sequencing by improving ionization efficiencies and manipulating dissociation patterns[4,5]. In the work shown here, the FMOC derivatization is only used to change the charge carrier from a proton to a sodium cation. As



I

long as the FMOC does not interfere with the predictable C-terminal dissociation of sodium cationized peptides, the strategy is to increase the abundance of the sodium cationized species so that de novo sequencing may be carried out with sufficient sensitivity.

Multiple mechanisms have been proposed for the dissociation of sodium cationized peptides[6-13]. Commonly they involve sodium binding at the C-terminus and intramolecular nucleophilic attack of the carbonyl carbon on the residue adjacent to the C-terminus. The mechanisms differ in the proposed attacking group; some suggest a hydroxide anion[8] while others propose the C-terminal carbonyl oxygen[6]. Spectroscopic as well as theoretical evidence exists for various structures, all of which involve solvation by C-terminal and backbone carbonyls[14]. Occasionally, if the peptide is small enough, there is evidence for solvation by the N-terminus in addition to backbone carbonyls[14]. It should be noted, though, that these calculations were done for peptides that are shorter than those commonly used in de novo sequencing. Furthermore, the C-terminal dissociation pattern does not always hold for peptides of three or fewer residues. The calculations do bring up questions about the role of the FMOC-sodium interaction in dissociation of $[FMOC\text{-}peptide+Na]^+$. The FMOC rings and carbamate linkage seem to be the most plausible site for sodium binding, but in the previous studies mentioned the sodium participates in the dissociation mechanism at the C-terminus.

There is of course fundamental interest in how the dissociation mechanism occurs, but irrespective of the mechanism are the practical implications of $[FMOC\text{-}peptide+Na]^+$ species. To be practical for most sequencing applications, the derivatization chemistry must be able to be coupled in-line with LC. Additionally, since a new stage of MS^n is required for each amino acid sequenced in the peptide, the analytes need to be abundant enough in the

mass spectrometer for sufficient sensitivity for many MS^n experiments, or must be identifiable with just a few of the C-terminal residues.

This chapter will describe methods used to derivatize peptides with an FMOC group at the N-terminus. It will go on to show that FMOC-peptides do have a higher binding affinity for a sodium cation than a proton, and undergo predictable dissociation by loss of the C-terminal residue. Finally, calculations will be used to determine where the sodium cation is bound in an attempt to understand the dissociation mechanism.

5.2 Experimental Methods

FMOC groups can be attached to the N-terminus of a peptide in two ways. FMOC is commonly used as the N-terminal protecting group in solid phase peptide synthesis procedures[15]. The penultimate step in the synthesis is to remove the final N-terminal FMOC before the finished peptide is cleaved off the resin (2-chlorotrityl chloride in this case). The FMOC group is stable under the acidic conditions used to cleave the peptide from the resin, hence the final deprotection step can simply be skipped to yield an FMOC-derivatized peptide (Structure I). This procedure was used to synthesize FMOC-GVFIPI. FMOC-derivatized amino acids and the 2-chlorotrityl chloride resin were obtained from C S Bio Co. (Menlo Park, CA) and used without further purification. The synthesis was carried out in a C S Bio Co. CS036 peptide synthesizer.

Chloroformates will undergo rapid condensation reactions with amines under mildly basic conditions, and this was used as a second method for synthesizing FMOC-peptides[16]. A 0.2 M boric acid solution was titrated with 0.1 M NaOH until a pH of 7.7 was read on a pH meter. 2 mL of 100 μM YGGFL was added to 0.5 mL of the borate buffer. To this solution

2.5 mL of a 15 mM FMOC chloride solution, allowed to react for 45 seconds, and then washed with two 10 mL washes of ethyl ether.

FMOC derivatized peptides were diluted to less than 6 μM for FMOC-GVFIPI and 1 μM for FMOC-YGGFL in methanol with 10 μM and 5 μM NaCl, respectively. Peptide concentrations are approximate because they were calculated assuming a pure substance. All samples were analyzed by direct infusion using a Bruker Esquire 3000 QIT.

Theoretical calculations were performed using density functional theory in Gaussian. First, neutral peptides and FMOC-peptides were optimized at the B3LYP/3-21g level of theory. FMOC-peptides can be built with the peptide chain either "cis" or "trans" to the FMOC plane, and both minimized structures were used in further calculations. Several starting structures for the sodium cationized species were generated by positioning the sodium atom at reasonable binding sites. These were then optimized at successively higher levels of theory up to B3LYP/6-311++g(2d,2p). No frequency calculations were performed so the relative energies do not include a correction for the zero-point energy.

5.3 Results and Discussion

5.3.1 Synthesis of FMOC-peptides and sodium cation binding affinity

The reverse binding preferences for a proton versus a sodium cation in peptides and FMOC-peptides can be seen in Figure 1. The products from the FMOC-GVFIPI synthesis were diluted in methanol with 10 μM of NaCl. There is some loss of the FMOC group during cleavage from the resin and storage, leaving some bare GVFIPI in the sample. Despite no additional acid in the solution, the ratio of sodium cationized to protonated GVFIPI is 0.44. Conversely, the peak corresponding to sodium cationized FMOC-GVFIPI is 7.8 times more intense than the protonated peak. This could be due to a combination of a

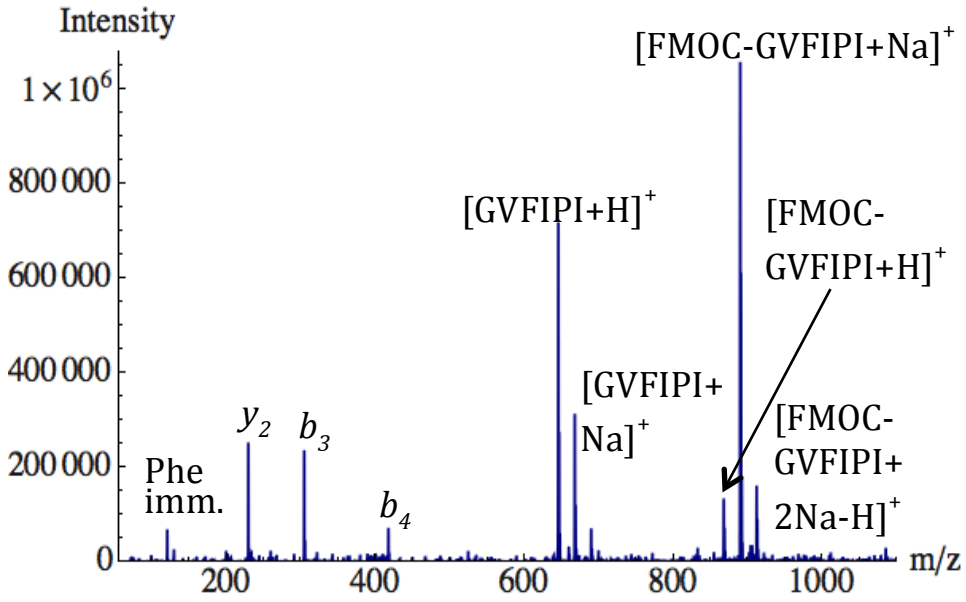


Figure 1. Mass spectrum of the reaction products from the synthesis of FMOC-GVFIPI. The solution is less than 6 mM FMOC-GVFIPI and 10 mM NaCl dissolved in methanol where the uncertainty in the peptide concentration comes from assuming a pure sample in the calculation. Some of the FMOC groups are lost leaving a population of underivatized GVFIPI. The difference binding can be seen in the intensities of the peaks corresponding to protonated and sodium cationized species.

reduction in the basicity of a carbamate group as opposed to a free amine and enhanced sodium binding at the FMOC group. For hydrophobic amino acids derivatized with an FMOC group the sodium cationized species is exclusively observed, even in solutions with acetic acid and no added sodium (Figure 2). It is thought that the FMOC group eliminates the N-terminus as a protonation site, but basic side chains and the rest of the peptide backbone are available for protonation depending on the acidity of the solution.

To be practical for sequencing, the FMOC derivatization procedure would ideally proceed to completion rapidly, under mild conditions, and with few side reactions. Forty-five seconds in a pH 7.7 borate buffer is sufficient for complete derivatization of leucine enkephalin, and there are few side reactions, as shown in Figure 3. The FMOC group can potentially react with any free amines and it is not yet known how the reaction would proceed for sequences containing lysine, arginine, asparagine, or glutamine. Further, if the

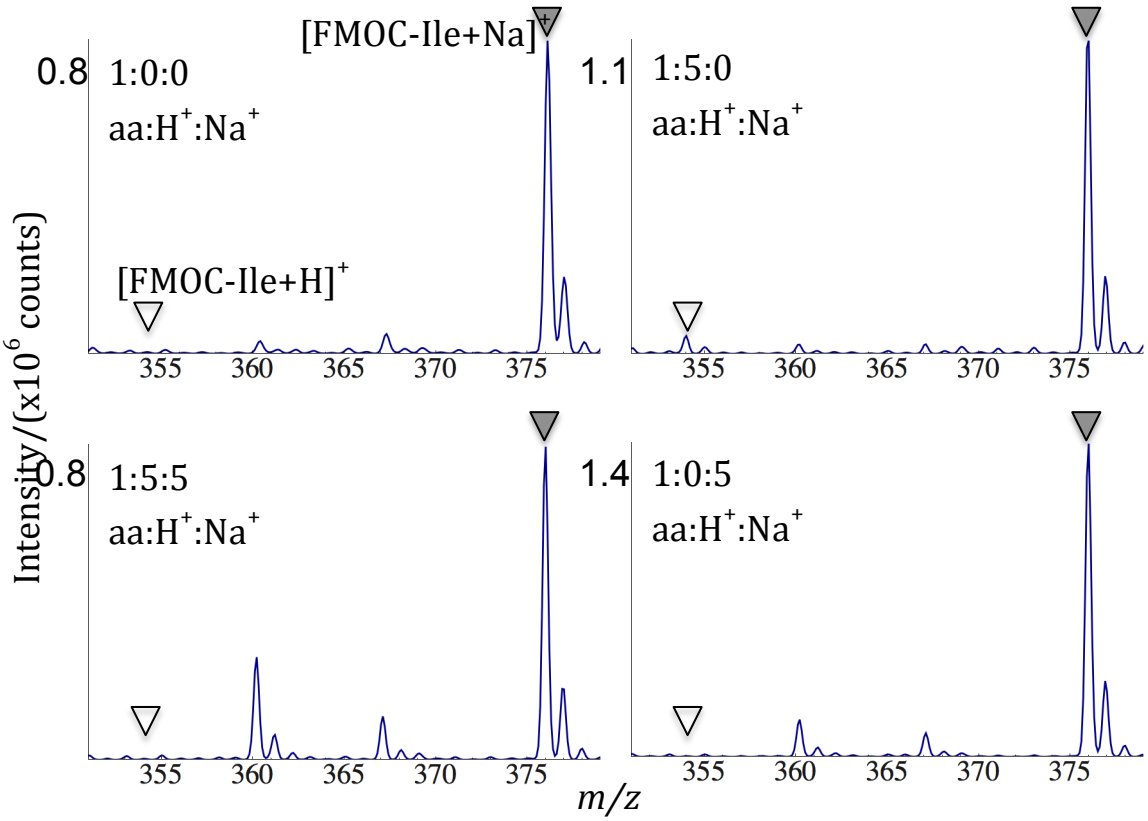


Figure 2. Mass spectra of 10 mM FMOC-Isoleucine in methanol solutions of varying acetic acid and sodium chloride concentrations. The ratios listed are the amounts of FMOC amino acid:acetic acid:sodium chloride. The mass spectrum in the top right shows that even in a solution acid present and no added sodium, the sodium cationized species is still preferred.

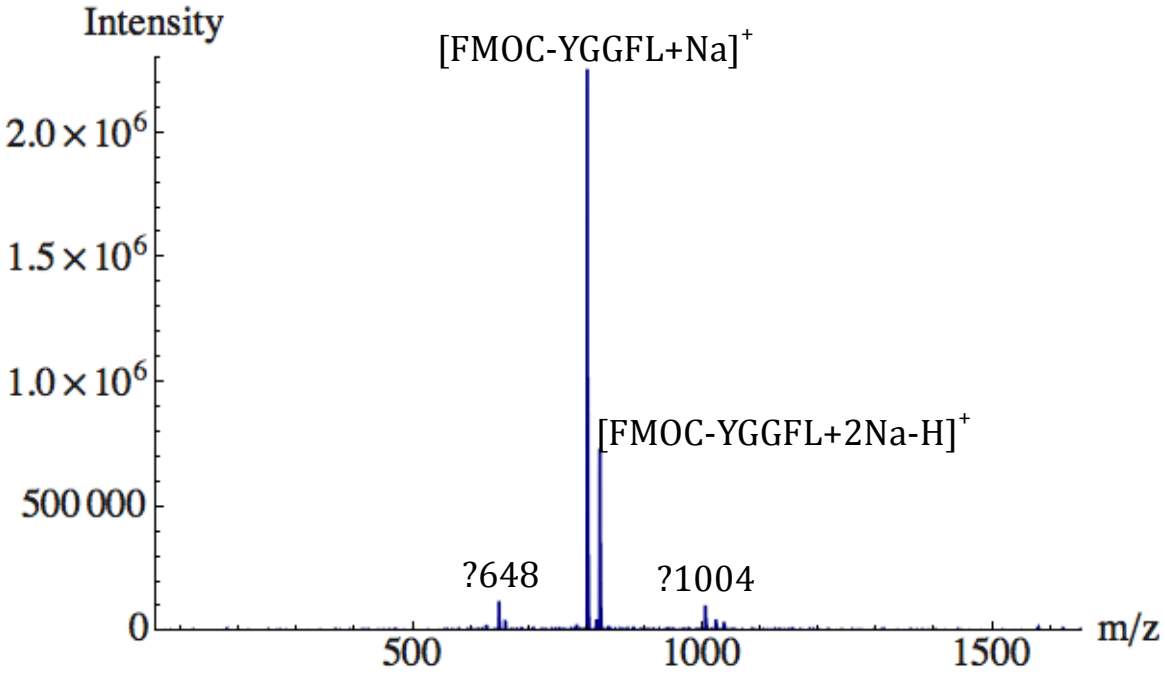


Figure 3. Products from the reaction of FMOC chloride with leucine enkephalin.

FMOC does react with the free amines on these side chains it is not known how that would affect dissociation chemistry. The FMOC group on the N-terminus remains stable through the entire series of MS^n experiments performed here, so it may be that and FMOC attached to the side chain would simply be lost with the residue.

5.3.2 C-terminal sequencing of [FMOC-peptide+Na]$^+$

Similarly to sodium cationized peptides, $[FMOC\text{-}peptide+Na]^+$ species dissociate almost exclusively by loss of the C-terminal residue, making the amino acid sequence of the peptide easy to determine from MS^n spectra (Figures 4 and 5). There are two peaks in Figure 4 labeled with asterisks that originate from an additional neutral loss of 46 Da, which could correspond to *a*-type ions[2,13]. A de novo sequencing algorithm may confuse glycine (57) for cysteine (103=57+46), but this would be easily accounted for because the smallest neutral

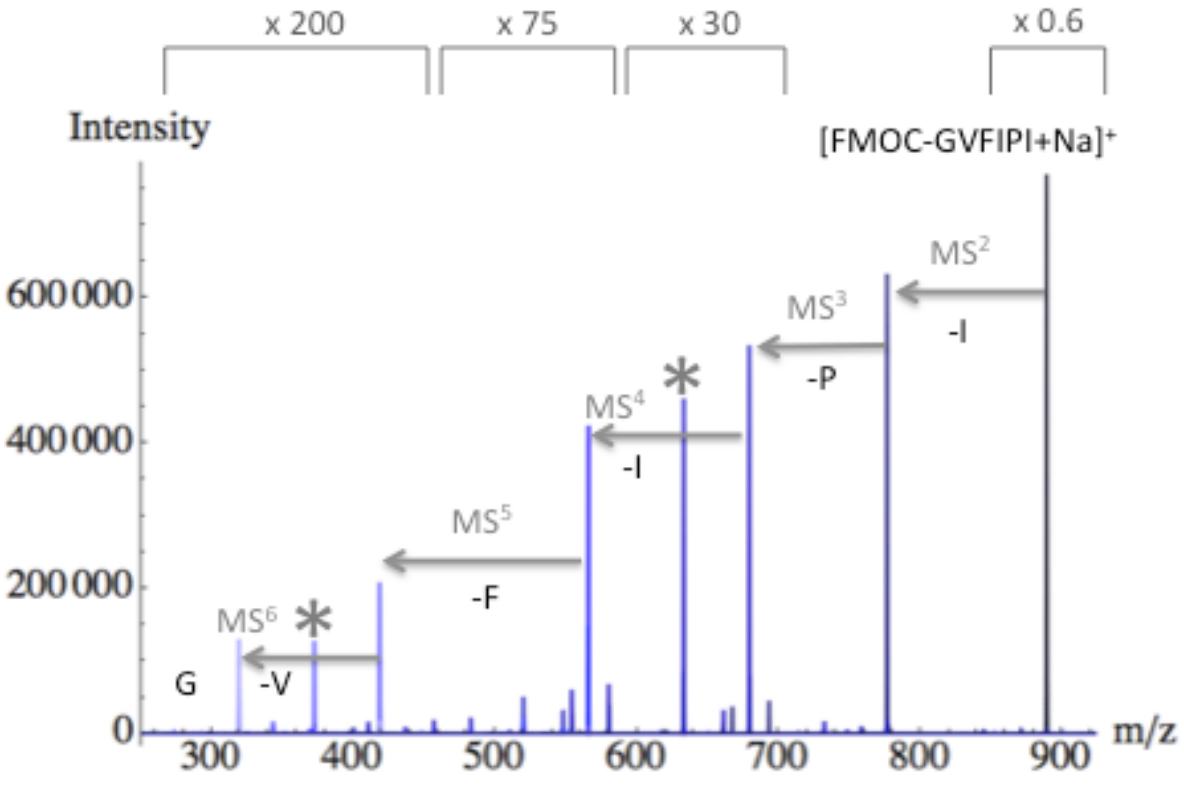


Figure 4. Sequential C-terminal residue loss in MS^n experiments on $[FMOC\text{-}GVFIPI+Na]^+$.

loss always corresponds to the correct residue for the peptides tested and accurate mass can distinguish these two species.

Extensive experiments have been carried out to determine the dissociation rules for many sodium cationized peptides. The dissociation chemistry always involves losses related to the C-terminal residue, but two categories of behavior were found, based on the identity of the C-terminal amino acid[2]. According to the nomenclature used by Lin et al, category I amino acids dissociate by loss of the C-terminal residue to form $[b_{n-1}+Na+OH]^+$, which has the structure of the original peptide only one residue shorter. Category II amino acids form a wider variety of product ions, including some small neutral losses, and the base peak is not always the same. Experiments here were performed exclusively with category I amino acids.

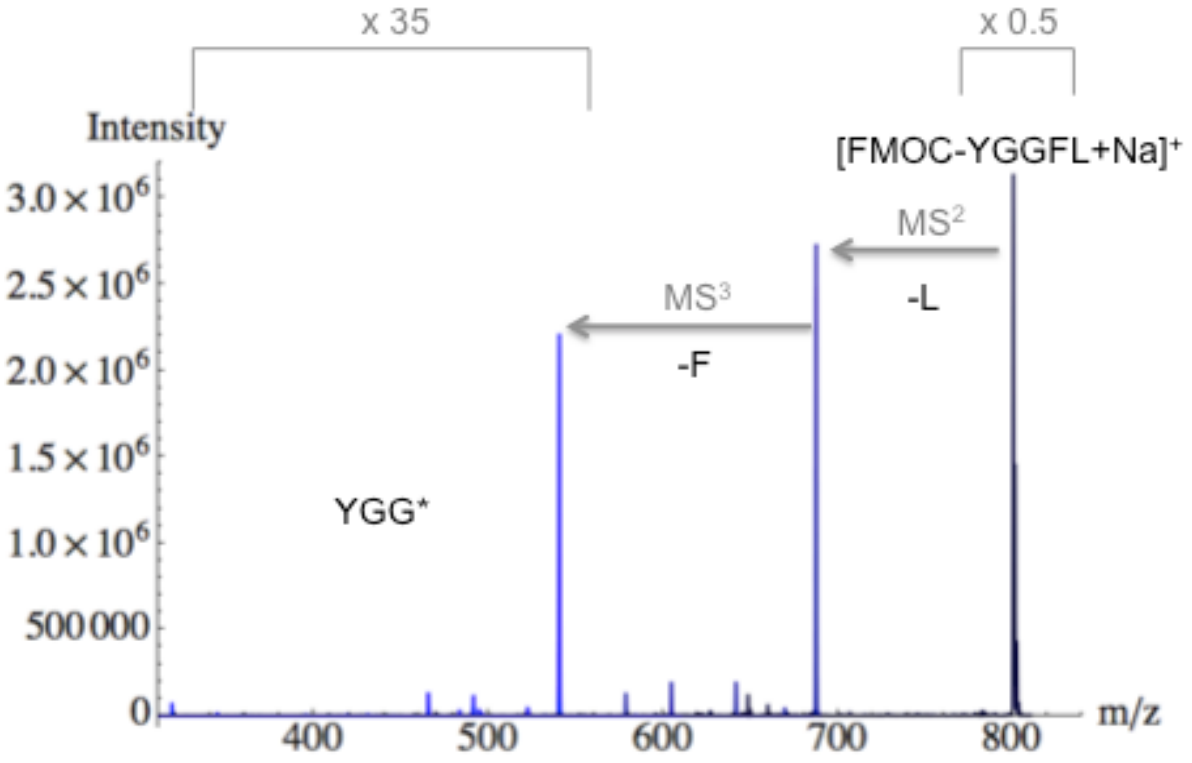


Figure 5. Sequential C-terminal residue loss in MS^n experiments on $[FMOC-YGGFL+Na]^+$. Stages beyond MS^3 may still yield predominantly C-terminal residue loss, but at the time sensitivity in the instrument was insufficient to proceed.

Future experiments will determine how FMOC derivatized peptides involving category II amino acids dissociate.

5.3.3 Where is the sodium bound?

For the case of $[peptide+Na]^+$, it has been proposed that the sodium is located at the C-terminus where it participates in the C-terminal residue loss mechanism. However, for $[FMOC\text{-}peptide+Na]^+$ it seems that the aromatic rings and ester group would provide a stable binding site for the sodium, which prompts the question of how the same C-terminal residue loss chemistry occurs. It could also be that the FMOC group only serves to eliminate the N-terminus as a protonation site, and the sodium is bound at the C-terminus as previously proposed for $[peptide+Na]^+$. To determine the location of sodium binding, calculations were used to determine the relative energies for various sodium locations in $[FMOC\text{-}Glycine+Na]^+$ and Gly-Val.

In the most stable structure of $[FMOC\text{-}Glycine+Na]^+$ the sodium is complexed between one of the FMOC rings, the ester carbonyl, and the C-terminal carbonyl (Figure 6a). This structure was started in the cis configuration, which enabled both carbonyls and an FMOC ring to be in a position to solvate the sodium, conferring additional stability. The second and third lowest energy structures (Figures 6b and 6c) also have the sodium solvated by the FMOC ring and either the ester carbonyl (+ 25 kJ mol^{-1}) or the whole ester group (+ 36 kJ mol^{-1}), respectively. There is also a structure (Figure 6d) with a relative energy 37 kJ mol^{-1} above the lowest energy structure where the sodium is complexed between the ester and C-terminal carbonyls.

The structure shown in figure 6d is similar to that which has been previously proposed for $[peptide+Na]^+$, where the C-terminal loss mechanism is initiated by

nucleophilic attack from either a hydroxide ion[8] or a carbonyl oxygen[6]. The high energy of the C-terminal bound structure calculated here and the fact that $[FMOC\text{-}peptide+Na]^+$ undergoes the same C-terminal dissociation chemistry as $[peptide+Na]^+$ indicate that these mechanisms do not provide a complete picture.

It is possible the sodium remains complexed at the N-terminus and dissociation is driven by neutral chemistry at the C-terminus or a non-reactive N-terminal bound structure first isomerizes to the reactive C-terminal bound structure. The former explanation is supported by mechanisms and structures proposed by Feng et al[13]. They use theoretical calculations to show that a neutral diglycine may dissociate by the same mechanism as lithium cationized diglycine. Furthermore, if a lithium cation is constrained to the

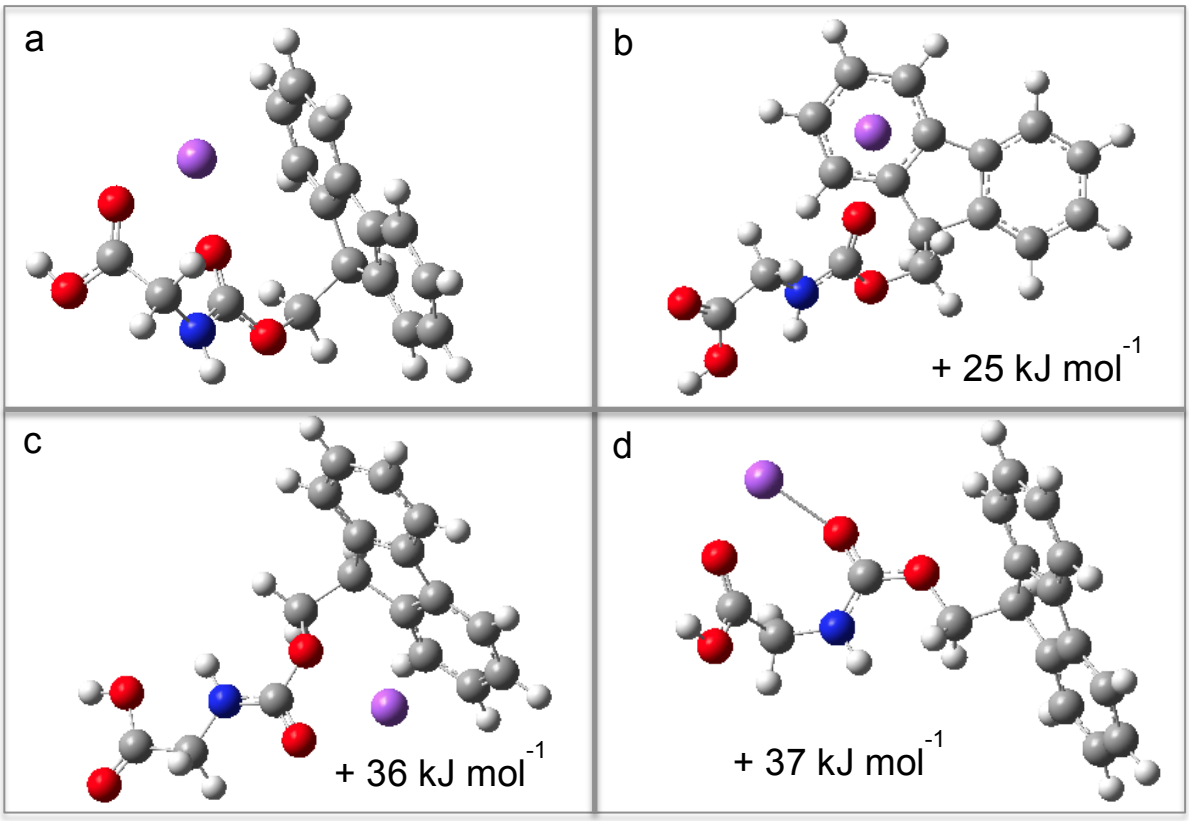


Figure 6. The four lowest energy structures found by varying the sodium binding configuration in B3LYP/6-311g++(2d,2p) calculations on $[FMOC\text{-}Gly+Na]^+$.

N-terminus, the transition state barriers are reduced below that of the neutral by approximately 40 kJ mol^{-1}. The calculation possibly demonstrates that a charge center in a peptide can catalyze charge-remote dissociation mechanisms though space by the electrostatic dipole it creates.

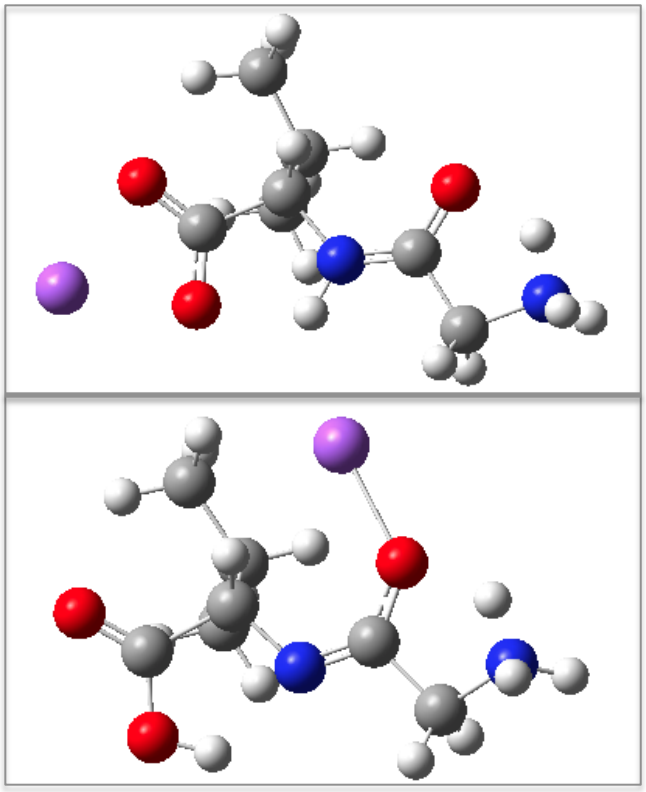


Figure 7. The two lowest energy structures found by varying the sodium binding configuration in B3LYP/6-311g++(2d,2p) calculations on Gly-Val.

Unlike $[FMOC\text{-}peptide+Na]^+$, calculations show that sodium cationized Gly-Val is almost certainly sodiated at the C-terminus, as the relative energy of the next lowest energy structure is +115 kJ mol^{-1} (Figure 7). $[Gly\text{-}Val+Na]^+$ species was shown by Lin et al to dissociate by loss of the C-terminal residue in the MS^7 stage of dissociation from the original GVYYVHPV sodium cationized peptide. Due to the size of a dipeptide, the C-terminal location is in accord with any of the previously proposed mechanisms, excepting the neutral driven chemistry discussed earlier.

5.3.4 Dissociation of $[FMOC\text{-}peptide\text{-}H+2Na]^+$

The singly charged, doubly sodiated FMOC derivatized species, $[FMOC\text{-}peptide\text{-}H+2Na]^+$, are also observed in the mass spectrum. Calculations have not been completed for $[FMOC\text{-}peptide\text{-}H+2Na]^+$, but it is reasonable to expect that one sodium is bound at the FMOC ring similarly to the singly sodiated species, and the other is

complexed at the deprotonated C-terminus as a zwitterion. According to proposed mechanisms for $[peptide+Na]^+$, the N-terminus does not participate in the dissociation and the C-terminus exists as a zwitterion, so it would be expected that $[FMOC\text{-}peptide\text{-}H+2Na]^+$ would dissociate by the same C-terminal residue loss channel as $[FMOC\text{-}peptide+Na]^+$ and $[peptide+Na]^+$. What is observed from dissociation of $[FMOC\text{-}YGGFL\text{-}H+2Na]^+$, however, is a completely different dissociation pattern than the singly sodiated species (Figure 8). MS^2 is still through a single loss channel to form an ion at 626 *m/z*, but the neutral mass lost is 196 Da which does not correspond to the 113 Da expected considering the C-terminal leucine. MS^3 of the species at 626 *m/z* forms many product ions including a neutral loss of 113 Da, which indicates the C-terminal leucine is retained in the MS^2 stage.

Dissociation of the doubly sodiated species is clearly unlike the singly sodiated species, but the identity of the species corresponding to the peak at 626 *m/z* is unknown as well as the origin of the neutral loss of 196 Da. The neutral losses labeled in Figure 6, with the exception of -196 Da, are speculative and only for guidance. Further MS^n experiments are needed to explicate the product ion branching relationships. It is possible that the 196 Da neutral has the elemental formula $C_{14}H_{12}O$, which could correspond to loss of 9-fluorenymethanol. Again, until further MS^n experiments are carried out the assignment is purely speculative, and the difference of 222 *m/z* in the MS^3 spectrum could also correspond to later loss of the full FMOC group.

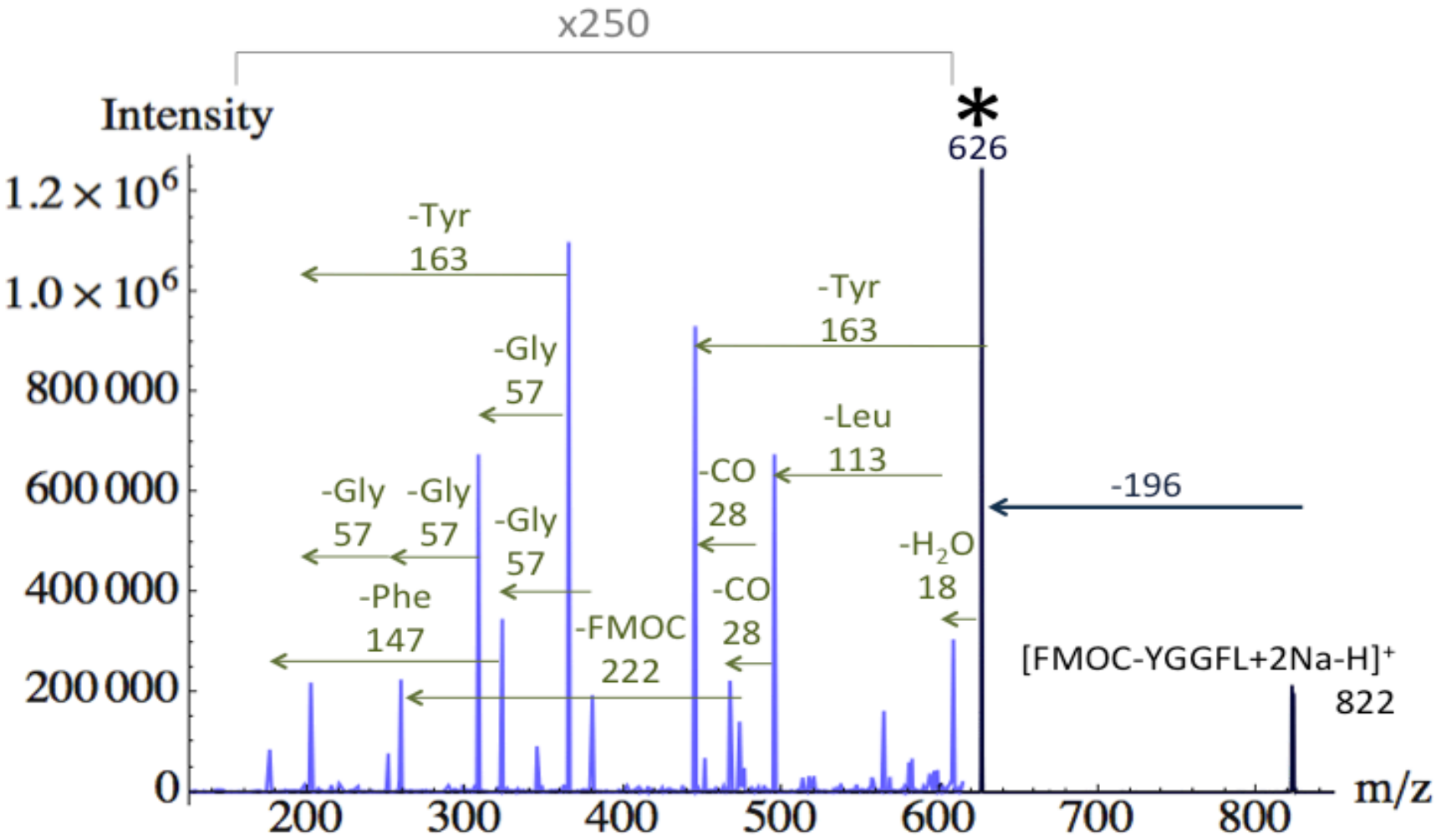


Figure 8. Superimposed MS^2 (dark blue) and MS^3 (light blue) spectra from dissociation of $[FMOC\text{-}YGGFL+2Na\text{-}H]^+$. Dissociation in the MS^2 stage is still by a single channel, although not the same C-terminal loss channel observed for the singly sodiated species. The neutral loss labels in the MS^3 spectra are speculative and for guidance only. Further stages of dissociation are needed to solidify product ion branching relationships.

5.4 Conclusions

Sodium cationized peptides have previously been shown to be useful for peptide sequencing due to their predictable C-terminal loss chemistry. The sodium cationized species is generally much less abundant than the protonated species, however, and therefore not useful in practice. An FMOC group attached to the N-terminus of a peptide facilitates sodium binding and these derivatized peptides undergo the same C-terminal loss chemistry as their underivatized counterparts. The derivatization procedure proceeds rapidly, under mild conditions, and with few side reactions, which makes it practical for many applications. Experiments will be done to determine whether the derivatization reaction can be done in-line with an LC separation for LC-MS^n sequencing experiments of protein digests, and to determine how many stages of MS^n are required for accurate identification.

It was previously thought that the sodium was located at the C-terminus where it facilitates the C-terminal loss mechanism. The fact that $[FMOC\text{-}peptide+Na]^+$ undergoes the same dissociation chemistry, but seems to have the sodium stably bound at the N-terminus supports the neutral dissociation chemistry already suggested by Feng et. Alternatively the C-terminal bound structure may be energetically attainable during a CID experiment, but calculations of the transition state energy would be needed to confirm that statement. In which case it is possible that the FMOC derivatized species isomerizes from the more stable N-terminal bound structure to a reactive C-terminal bound structure. Additionally, binding of a second sodium disrupts the dissociation chemistry. Based on the current understanding it is unclear why that should be the case.

References

(1) Lin, T.; Glish, G. L. *Analytical Chemistry* **1998**, *70*, 5162.

(2) Lin, T.; Payne, A. H.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 497.

(3) Ocana, M. F.; Jarvis, J.; Parker, R.; Bramley, P. M.; Halket, J. M.; Patel, R. K. P.; Neubert, H. *Proteomics* **2005**, *5*, 1209.

(4) Anderegg, R. J. *Mass Spectrometry Reviews* **1988**, *7*, 395.

(5) Karnezis, A.; Barlow, C. K.; O'Hair, R. A. J.; McFadyen, W. D. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 2865.

(6) Grese, R. P.; Cerny, R. L.; Gross, M. L. *Journal of the American Chemical Society* **1989**, *111*, 2835.

(7) Lee, S.-W.; Kim, H. S.; Beauchamp, J. L. *Journal of the American Chemical Society* **1998**, *120*, 3188.

(8) Tang, X.; Ens, W.; Standing, K. G.; Westmore, J. B. *Analytical Chemistry* **1988**, *60*, 1791.

(9) Anbalagan, V.; Patel, J. N.; Niyakorn, G.; Van Stipdonk, M. J. *Rapid Commun. Mass Spectrom.* **2003**, *17*, 291.

(10) Barr, J. M.; Van Stipdonk, M. J. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 566.

(11) Osburn, S. M.; Ochola, S. O.; Talaty, E. R.; Van Stipdonk, M. J. *Journal of Mass Spectrometry* **2008**, *43*, 1458.

(12) Cooper, T. J.; Talaty, E. R.; Van Stipdonk, M. J. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1305.

(13) Feng, W. Y.; Gronert, C.; Fletcher, K. A.; Warres, A.; Lebrilla, C. B. *Int. J. Mass Spectrom.* **2003**, *222*, 117.

(14) Balaj, O. P.; Kapota, C.; Lemaire, J.; Ohanessian, G. *Int. J. Mass Spectrom.* **2008**, *269*, 196.

(15) *Fmoc solid phase peptide synthesis : a practical approach*; Oxford University Press: Oxford ;New York, 2000.

(16) Einarsson, S.; Josefsson, B.; Lagerkvist, S. *Journal of Chromatography A* **1983**, *282*, 609.

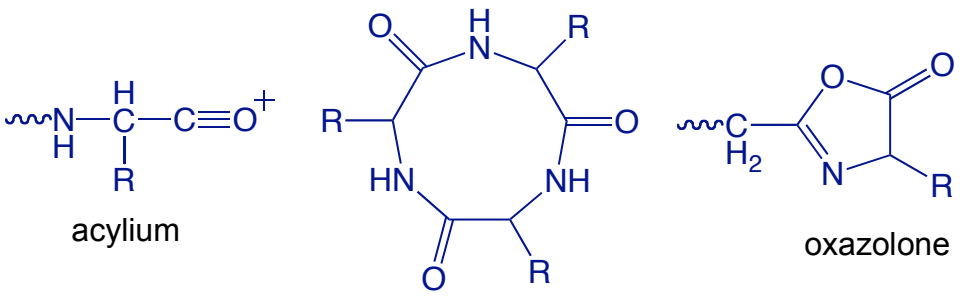
Chapter 6

Two Types of b_3 Structures Observed by Infrared Photodissociation Spectroscopy

6.1 Introduction

The inability to a priori predict collision induced dissociation (CID) tandem mass spectra given a peptide sequence has led to studies of the structures and dissociation mechanisms of peptide product ions[1,2]. The study of b_n type product ions is of interest because they are commonly observed stable product ions[3] and important intermediates in CID experiments on peptides. Additionally, b_n ions may undergo sequence-scrambling rearrangements[4,5] that may hinder peptide and protein identification[6-8]. The structure assigned to various b_n ions differs from experiment to experiment. To obtain a more general understanding of b_n ion chemistry various types of MS-based experiments have performed, including multiple stages of mass spectrometry (MS^n)[3,9-12], hydrogen-deuterium exchange[13,14], ion-molecule reactions[9], ion mobility spectrometry (IMS)[15-17], and infrared (IR) spectroscopy[12,13,18-21]. The b_n ion structures assigned in these experiments fall into different classes often depending on the number of amino acid residues[19].

Initially b_n ions were suspected to be acylium cations because of the high propensity to lose CO and form a_n ions[3,22,23]. However, kinetic energy release studies distinguished dissociation of protonated dipeptides from protonated amino acids where the acylium ion intermediate is known to be unstable[3]. These experiments seemed to indicate the presence of



a stable b_n structure that can isomerize to the more reactive acylium ion structure before dissociation. C-terminal oxazolone structures are more stable than acylium cations because electron density is donated to the electron-poor acylium carbon atom from an adjacent carbonyl. Alternatively electron donation may come from the N-terminal amine forming a diketopiperazine[24] in the case of dipeptides, or a macrocyclic structure for longer b_n ions.

Strong evidence for the oxazolone structure can be found in comparisons of CID spectra from b_n ions to synthesized oxazolones and spectroscopy experiments, with additional support from theoretical calculations[9,12,13,21,25-27]. Evidence for the presence of non-oxazolone structures in some IR spectroscopy experiments on b_n ions is clear, but the assignment of the additional structure as a macrocycle is less definite[5,13,18,19]. The assignment of a macrocycle is often based on the presence of the backbone CO-H^+ bending mode near 1440 cm^{-1}. However this band is less intense and in a more congested region of the IR spectrum than the distinctive carbonyl stretching mode of the oxazolone found above 1800 cm^{-1}. The presence of a macrocyclic structure is corroborated by CID, IMS and H/D exchange experiments: CID of a series of b_n ions with permutated sequences and the corresponding cyclic peptide yields the same dissociation pattern[5,11], the arrival time of b_n ions matches that of the cyclic peptide in mobility (IMS) experiments[15,16], and H/D exchange rates of oxazolone structures are nearly an order of magnitude faster than those of macrocyclic structures[11,16]. The conclusion that has been drawn from these experiments is

that in general b_n ions form both oxazolones and macrocycles, and it is energetically feasible for both structures to be present in a single experiment.

Despite the presence of multiple structures most b_n ions behave similarly in CID experiments and dissociate primarily by loss of CO to form the corresponding a_n[3,10,28-31]. However, a_3 ions are rarely observed in CID spectra, making b_3 ions a potential exception to the usual b_n-to-a_n principle[9,28,32]. Either the structure of b_3 ions is different than other b_n ions such that CO loss is no longer a preferred dissociation channel, or a_3 ions are formed as a transient species and rapidly dissociate further. Theoretical calculations support the idea that a_3 ions are transient[9]. In addition, collective information from a database of CID spectra show that a_3 ions are not observed as often as expected considering their size, while b_2 ions are observed more frequently than expected, which could also be explained if a_3 ions were a transient species[32]. However ion-molecule reactions with dimethylamine distinguish b_3 ions from other b_n ions lending credence to idea that b_3 ions may exist as a different type of structure[9]. Confirmation of either explanation requires structural information about the ion structures in the b_3 to a_3 dissociation pathway.

IR spectroscopy was performed to obtain structural information on the b_3 ions from the peptide leucine enkephalin (YGGFL) and analogs of leucine enkephalin formed by exchanging the glycines for alanines (YAGFL, YGAFL, and YAAFL). The sequence-dependent spectra and behavior of these ions during the experiment is presented herein.

6.2 Experimental Methods

IR spectra of the b_3 ions from YGGFL, YAGFL, YGAFL, and YAAFL were obtained via resonant IRMPD (r-IRMPD) spectroscopy, over an energy range of 800 to 1800 cm^{-1},

using the Free Electron Laser for Infrared Experiments (FELIX). Ions were generated by electrospray ionization from 10 μM solutions in 49.5:49.5:1 methanol:water:acetic acid. The b_3 ions from YAGFL and YGGFL were generated by in-source fragmentation using a high potential difference between the source block and skimmer cone of the Z-spray ESI source. Ions were accumulated in a linear hexapole trap and then transmitted to a Fourier transform ion cyclotron resonance mass spectrometer (FT-ICR) for isolation and analysis. Fragmentation using a high axial trapping amplitude in the hexapole generated b_3 fragments from YGAFL and YAAFL. IR spectra are produced by plotting the dissociation efficiency of the b_3 ion as a function of the FELIX wavelength, and spectra are baseline corrected to account for the power variation with the wavelength of FELIX. Non-resonant IRMPD experiments were performed on a custom quadrupole ion trap mass spectrometer (QITMS) using a CO_2 laser, as described previously[33]. CID experiments were performed on a Bruker Esquire 3000. For these experiments samples were diluted to 1 μM in 49.5:49.5:1 methanol:water:acetic acid.

Theoretical calculations for this work were performed in collaboration with Bela Paizs of the German Cancer Research Center in Heidelberg, Germany. The calculations began with molecular dynamics simulations using the Insight II program (Biosym Technologies, San Diego, CA) in conjunction with the AMBER force field[34], modified in-house[35,36] to enable the study of oxygen protonated amide bonds and oxazolone groups. During the dynamics calculations we used simulated annealing techniques to produce candidate structures for further refinement, applying full geometry optimization using the AMBER force field. The optimized structures were analyzed by a conformer family search program developed in Heidelberg. The program groups optimized structures into families for

which the most important characteristic torsion angles of the ion are similar. The most stable species in the families were then fully optimized at the PM3, HF/3-21G, B3LYP/6-31G(d), and finally at the B3LYP/6-31G(d,p) levels, and the conformer families were regenerated at each level. Such series of calculations were carried out for the oxazolone isomers with the YGG, YAG, YGA, and YAA sequences considering protonation at the N-terminal amino and C-terminal oxazolone nitrogens, respectively. Additionally we scanned the potential energy surface of cyclic YAA protonated at the amide oxygens. The Gaussian set of programs was used for all ab initio and DFT calculations[37,38]. Theoretical IR frequencies were scaled by 0.98 to account for anharmonicity, and convoluted with a 20 cm^{-1} Lorentzian. The scaling factor was chosen based on what gave the best experimental match, rather than 0.96 which is commonly used for density functional calculations using the B3LYP functional and the 6-31g+(d,p) basis set[39,40].

6.3 Results and Discussion

6.3.1 Results from IR spectroscopy experiments at FELIX

The fact that a_3 ions are not observed as product ions from dissociation of b_3 was shown for 27 peptides out of the 29 studied by MS^3 experiments, including the four peptides studied here[9], and for a large number of peptides in a database of CID data[32]. Because b_3 ions exhibit different behavior than other b_n ions in CID and ion-molecule reactions, their structure has been suggested to be different from the oxazolone structure observed in similarly sized *b*-type fragments. Nonetheless, the IR spectrum of the b_3 ion from YGGFL recorded here matches well with the calculated spectrum for an oxazolone structure protonated on the N-terminus (Figure 1). The position of the oxazolone CO stretch, which typically appears as an intense band at the high-frequency end of the spectral range

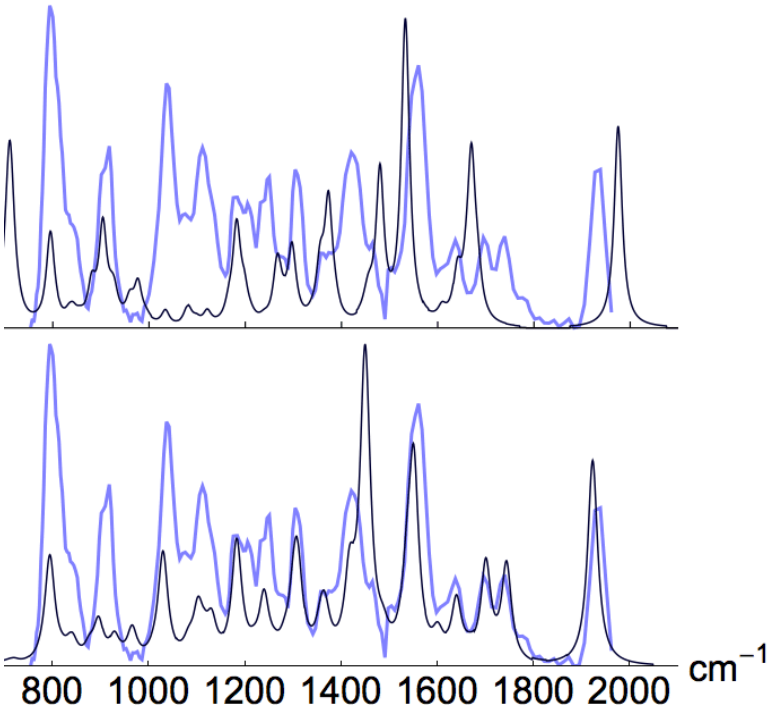


Figure 1. Experimental IR spectra of the b_3 ion from YGGFL (gray) compared to the theoretical spectrum calculated from an oxazolone structure protonated at the N-terminus (black, bottom) or protonated on the oxazolone ring (black, top). All spectra are normalized to the most intense peak and both theoretical spectra were frequency scaled by 0.98

investigated, facilitates reliable structural assignment because it is sensitive to the location of the proton and three-dimensional structure of the ion in general. Here the significance of the particular scaling factor is not well understood, but based on the overall match the position of the CO stretching band is not needed to conclusively assign the N-protonated oxazolone structure. Hence the band can be used to determine the 0.98 scaling factor.

MS-based studies of b_3 ions show no indication of structural differences within the b_3 ion class[9]. Additionally, the amount of energy required for dissociation of b_3 ions from YGGFL, YAGFL, YGAFL, and YAAFL in the Bruker Esquire 3000 changes only with molecular mass in this peptide series (Figure 2), also suggesting that all b_3 ions belong to the same structural class. It is important to note that although the spectroscopy and MS^n experiments were performed on different platforms, an FT-ICR and QITMS respectively,

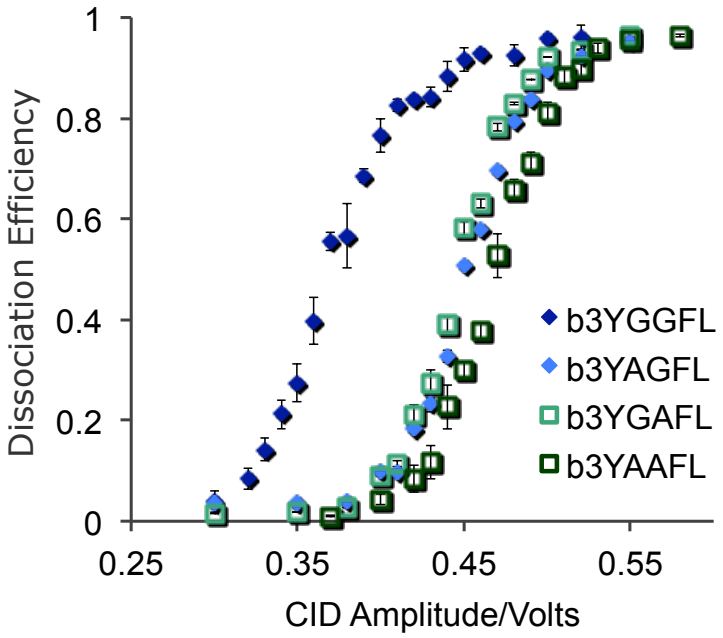


Figure 2. Dissociation curves for the b_3 ions from YGGFL, YAGFL, YGAFL, and YAAFL

structures of the b_3 ions are likely the same between the experiments since the product ion masses formed in both experiments were the same. In addition, the structure of the protonated YAGFL and YGAFL parent ions were also observed to be similar by infrared spectroscopy using the FT-ICR and FELIX (Figure 3), and by their dissociation kinetics in CID experiments on the Bruker Esq uire 3000 QITMS (Figure 4).

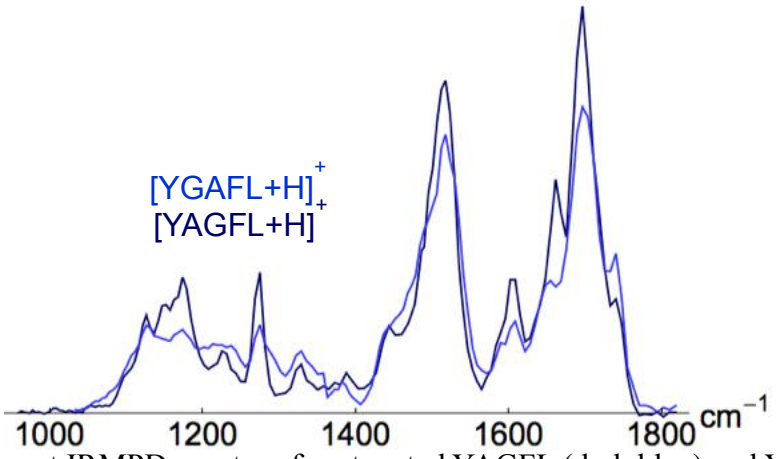


Figure 3. Resonant IRMPD spectra of protonated YAGFL (dark blue) and YGAFL (light blue), showing similar structures

To verify the oxazolone structure for b_3 ions, IRMPD spectra were recorded for the b_3 ions from YAGFL, YGAFL, and YAAFL. The structure of the b_3 ion from YAGFL is also assigned as an oxazolone protonated on the N-terminus (Figure 5), similar to that for the b_3

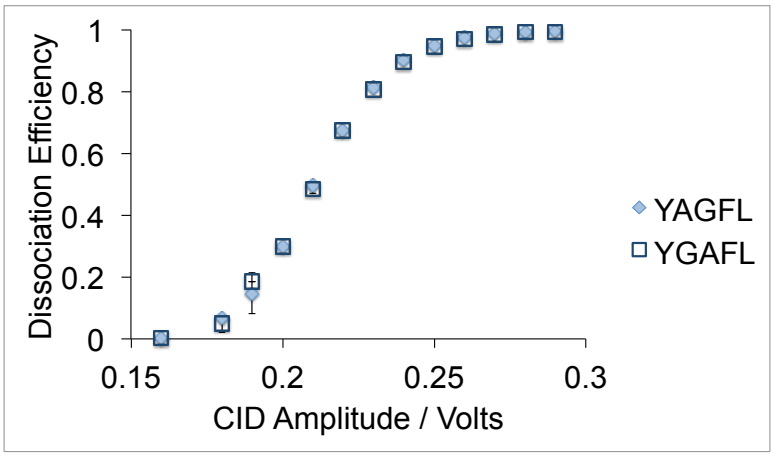


Figure 4. Dissociation efficiency as a function of the amplitude of the CID waveform in volts. Similar dissociation rates imply similar structures

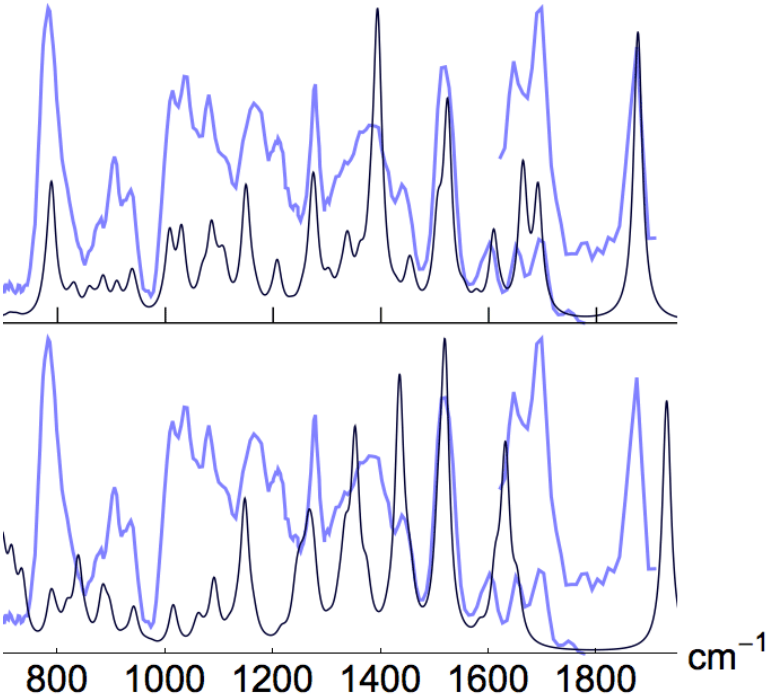


Figure 5. Experimental IR spectra of the b_3 ion from YAGFL (gray) compared to the theoretical spectrum calculated from an oxazolone structure protonated at the N-terminus (black, top) or protonated on the oxazolone ring (black, bottom). All spectra are normalized to the most intense peak and both theoretical spectra were frequency scaled by 0.97. Based on these results b_3-YAGFL is assigned as an oxazolone protonated at the N-terminus

ion from YGGFL. However, as can be seen in the IR spectra shown in Figure 6, the b_3 ions from YGAFL and YAAFL clearly fall into a distinctly different structural class than the oxazolones observed for b_3-YGGFL and b_3-YAGFL. Considering the apparently modest structural difference between a glycine and an alanine residue, such drastically different structures were not expected. Theoretical calculations for both the oxazolone and the macrocyclic structures give poor matches to the experimental spectra, although the macrocyclic structure does give a convincing match for the high-frequency region of the spectrum (Figure 7). This region of the spectrum contains the structurally diagnostic carbonyl stretches of the backbone amide linkages as well as the C-terminus.

In addition to the different spectral features for b_3-YGGFL/b_3-YAGFL versus b_3-YGAFL/b_3-YAAFL, the behavior during the spectroscopy experiments at FELIX also fell

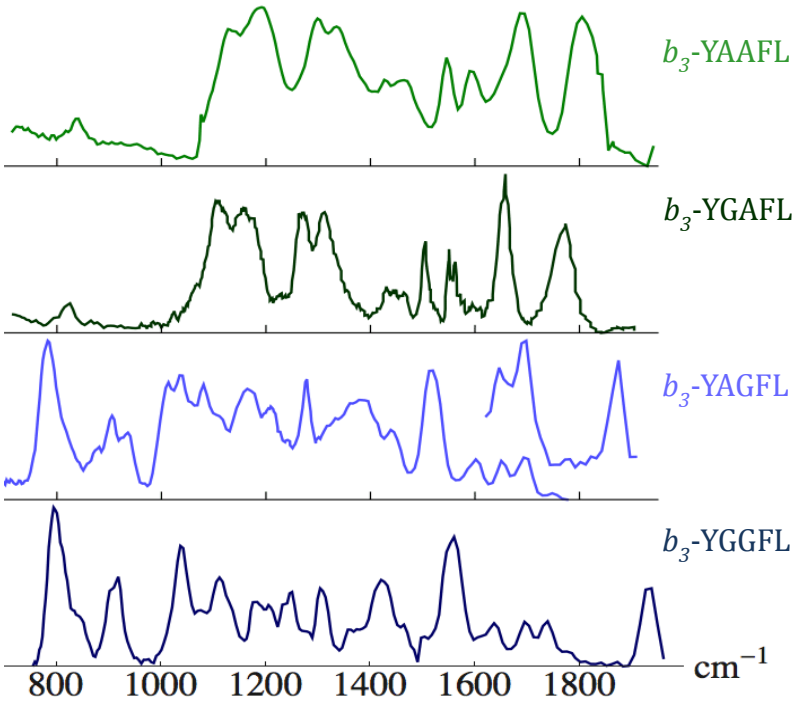


Figure 6. Experimental IR spectra of b_3 ions peptide fragments derived from YAAFL (top), YGAFL (second), YAGFL (third), and YGGFL (bottom). b_3-YGGFL and b_3-YAGFL match well to an N-protonated oxazolone structure while b_3-YGAFL and b_3-YAAFL appear to belong to a different class of structure that has yet to be assigned

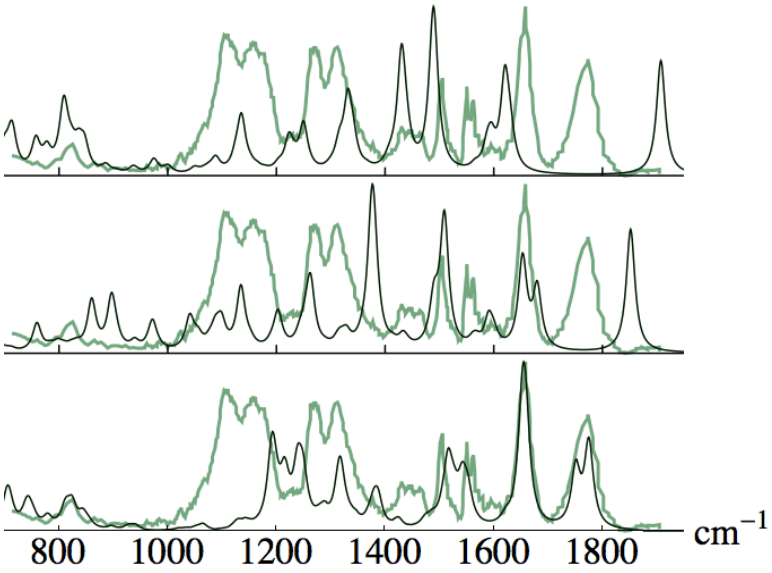


Figure 7. Experimental IR spectra of the b_3-YGAFL (gray) compared to theoretical spectra from an oxazolone protonated on the oxazolone ring (top, black), an oxazolone protonated on the N-terminus (middle, black), and a macrocyclic structure (bottom, black). Here the theoretical spectra were scaled by 0.96 as recommended by Scott, et al. 1996

into two classes. The b_3 ions from YGGFL and YAGFL were formed using a high potential difference between the source block and skimmer cone of the Z-spray source. The b_3-YGAFL/b_3-YAAFL ions could not be formed by nozzle/skimmer dissociation and instead required the use of a high axial trapping amplitude during accumulation in the hexapole trap. FELIX irradiation times up to 11 seconds were required to obtain spectra of b_3-YGAFL/b_3-YAAFL, as compared to the 2.5 seconds used for b_3-YGGFL and b_3-YAGFL. Furthermore, different types of product ions were formed from dissociation of b_3-YGAFL/b_3-YAAFL versus b_3-YGGFL/b_3-YAGFL during irradiation with FELIX (Table 1).

Table 1 Dominant product ions formed by IRMPD with an FT-ICR and FELIX or with a quadrupole ion trap and CO_2 laser at UNC

Parents	Products
b_3-YGGFL	a_3-17 (233 *m/z*), b_2 (221 *m/z*)
b_3-YAGFL	a_3-17 (247 *m/z*), b_2 (235 *m/z*
b_3-YGAFL	b_3-46 (246 *m/z*), a_3-17-28 (219 *m/z*), a_3-17-28-28 (191 *m/z*), Tyr immonium (136 *m/z*)
b_3-YAAFL	b_3-46 (260 *m/z*), a_3-17-28 (233 *m/z*), a_3-17-28-28 (205 *m/z*), Tyr immonium (136 *m/z*)

Conventional product ions that are also observed in CID were formed during r-IRMPD of b_3-YGGFL/b_3-YAGFL, whereas an entirely different subset of product ions are formed from b_3-YGAFL/b_3-YAAFL in the FT-ICR at FELIX as well as by IRMPD in a QITMS using a CO_2 laser.

6.3.2 What are other possible structures for b_3-*YGAFL*/b_3-*YAAFL?*

Further calculations are needed to find a match to the experimental b_3-YGAFL/b_3-YAAFL spectra, and perhaps there are hints for where to start conformational searches in the data already collected. The structure with a calculated spectrum that best fit the experimental data (Figure 7, bottom) was a macrocyclic structure protonated on the carbonyl between the alanine and tyrosine. The calculation suggests that the vibrational transitions in the high-frequency region of the spectrum at ~1770 and ~1650 cm^{-1}, which do match well with the experiment, are due to non-protonated backbone carbonyl stretches and modes involving C-N stretching of the protonated amide and C-C stretches in the tyrosine ring, respectively. Discrepancies between experimental and theoretical spectra include the band near 1440 cm^{-1} often used to identify a macrocyclic structure, which here is predicted to be located at at 1390 cm^{-1}. The intense transitions predicted between 1180 and 1310 cm^{-1} involve various vibrations of the tyrosine side chain. The observation that this region of the theoretical spectrum provides such a poor prediction of the experimental spectrum raises the question of whether rearrangement of the how tyrosine residue occurs in the b_3-YGAFL/b_3-YAAFL sequence ions.

CID data obtained in a QITMS further supports the hypothesis that the tyrosine side chain may play a significant role in the b_3-YGAFL/b_3-YAAFL structures as observed in these experiments. One of the product ions produced during IRMPD of b_3-YGAFL was

nominally the a_3-17-28 ion (*m/z* 219). This product ion can also be produced using CID, and further dissociation leads to a product ion at *m/z* 120 (Figure 8). There two neutral losses of 28 in the dissociation pathway between a_3-17 and the product ion at *m/z* 120. There are only two carbonyl groups in an a_3 ion so it is not obvious how consecutive COs could be lost from any structure where the peptide backbone is intact. The presence of a product ion at *m/z* 120 in peptide CID spectra is often assigned as the phenylalanine immonium ion, but there is no phenylalanine in b_3-YGAFL. Whether the product ion at *m/z* 120 is directly from a_3-17-28 or from sequential dissociation of a_3-17-28 is unknown. It is possible that the hydroxyl functional group is lost from the tyrosine side chain in a rearrangement reaction that yields a phenylalanine immonium ion, or some other isobaric/isomeric species is formed. There is precedence for loss of CO from protonated phenolic compounds in CID experiments performed in a triple quadrupole using 20 eV collisions[41]. Conversely CID of protonated tyrosine using 5 and 15 eV collisions in a triple quadrupole showed loss of CO and H_2O from the carboxylic acid group, but no evidence of losses from the side chain[42]. The experimental time frame in a QITMS or FT-ICR is much longer than that of a triple quadrupole, so even though the collision energies are lower it is possible that the longer time allows for more complicated rearrangement reactions to take

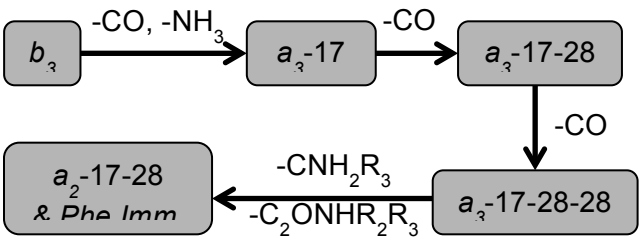


Figure 8. One of the dissociation pathways of the b_3 ions from YGGFL, YAGFL, YGAFL, and YAAFL. The boxes indicate the species isolated in the MS^n experiment. R_n denotes the side chain of the n^{th} residue. It is not known whether the product ion at 120 *m/z* is directly from a_3-17-28-28 or from sequential dissociation of a_2-17-28

place[43]. The particular structure of these ions may also facilitate neutral loss from the side chain.

Both the poor match between theoretical and experimental IR spectra as well as the CID data suggests that perhaps and as yet unidentified rearrangement of the Tyr side chain should be considered in searching for structures to match the experimental b_3-YGAFL/b_3-YAAFL spectra. In addition, one may consider different protonation sites and different types of backbone cyclization including some previously suggested 5- and 6-membered heterocyclic ring structures[44].

6.3.3 r-IRMPD is a more specific measure of structure than CID for b_3 *ions*

Due to the resonant nature of the r-RMPD experiment it is a more specific measure of structure than CID, so the fact that r-IRMPD was able to resolve structural differences that were not observed in a CID experiment is not surprising. The purpose of Figure 9 is to give an illustration of when r-IRMPD, but not CID, would identify the presence of multiple structures in an experiment. Excitation in CID is only specific to a mass-to-charge value, so

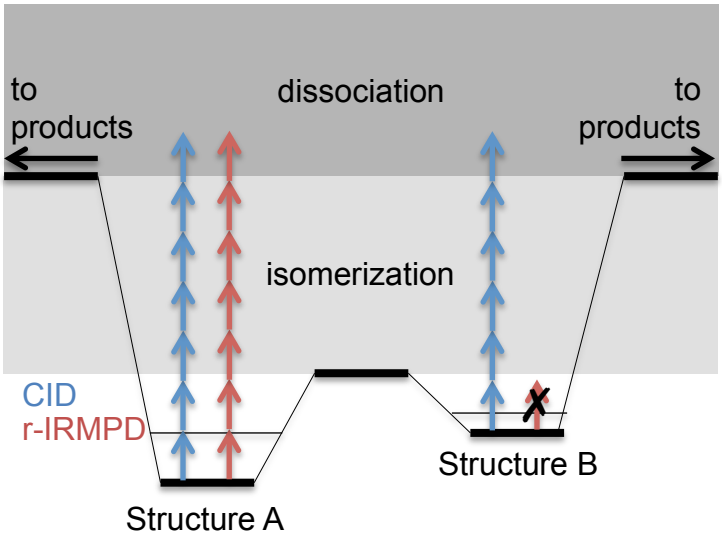


Figure 9. A potential energy diagram that depicts how r-IRMPD can identify multiple structures present in an experiment that are not observed by CID.

either structure will be collisionally activated. If the barrier to isomerization between two structures with the same *m/z* is below the dissociation threshold, it doesn't matter which structure the parent ions start in because they will dissociate from a common manifold of parent ion states and thus form the same product ion distribution. Conversely, dissociation in an r-IRMPD experiment will only occur for structures with vibrational states in resonance with the laser.

If the example is applied to the b_3 ion from YGGFL, YAGFL, YGAFL, and YAAFL, it can be supposed that "Structure A" is an oxazolone, and "Structure B" is the unassigned structure. CID experiments on the b_3 ion from all four peptides will result in the same product ion distribution, as is observed and will be discussed further in the next section. In r-IRMPD the b_3 ions from YGGFL and YAGFL, which are oxazolones, will only dissociate at wavelengths resonant with oxazolone transitions, and the b_3 ions from YGAFL and YAAFL will only dissociate at wavelengths resonant with transitions present in their structure. The end result of the scheme shown in Figure 9 is that b_3-YGGFL/b_3-YAGFL will be distinguished from b_3-YGAFL/b_3-YAAFL in r-IRMPD experiments but not in CID experiments.

If it wasn't for other observations made during the process of obtaining the infrared spectra of b_3 ions at FELIX, the potential energy diagram in Figure 9 would be a satisfactory description of how the two structures of b_3 ions relate to each other, the similar product ion distributions for all four b_3 ions observed in CID experiments, and why r-IRMPD distinguished the two structural classes. The next section will discuss the additional observations made during the FELIX experiments, results from non-resonant IRMPD experiments as compared to CID experiments, and go on to explain how the potential energy

diagram proposed in Figure 9 will need to be modified to account for all experimental observations[45].

6.3.4 The first account of different classes of product ions formed from CID and IRMPD

Recall that r-IRMPD is a form of action spectroscopy where the amount of product ions formed from a population of parent ions being irradiated at a particular wavelength indicates the absorption cross section of the ions at that wavelength. Even though Structure A and B will have different wavelength dependencies, and therefore generate different infrared spectra, when dissociation does occur it should be to the same product ions. Furthermore, the product ions observed should be the same for CID, IRMPD, and r-IRMPD, apart from some sequential dissociation that may occur in either of the IRMPD techniques.

Indeed, the product ions formed during CID, IRMPD, and r-IRMPD of the b_3 ions of YGGFL and YAGFL, those that have an oxazolone structure, are predominantly the a_3-17 and b_2 ions for all three techniques. These are commonly observed product ions from dissociation of b_3 ions (Figure 10). CID experiments on the b_3 ions from YGAFL and YAAFL, those with an unassigned structure, also show the expected a_3-17 and b_2 product ions. During r-IRMPD experiments at FELIX, however, an entirely different set of product ions was observed for b_3-YGAFL/b_3-YAAFL. It was mentioned previously that the product ions observed in CID and IRMPD/r-IRMPD may differ in the amount of sequential dissociation. To be clear, the product ions observed in r-IRMPD of the b_3 ions from YGAFL and YAAFL were in fact from a separate dissociation channel and not due to sequential dissociation of commonly observed CID product ions. Not only were the product ions observed in r-IRMPD different than CID, the identity of the product ions is unknown and not

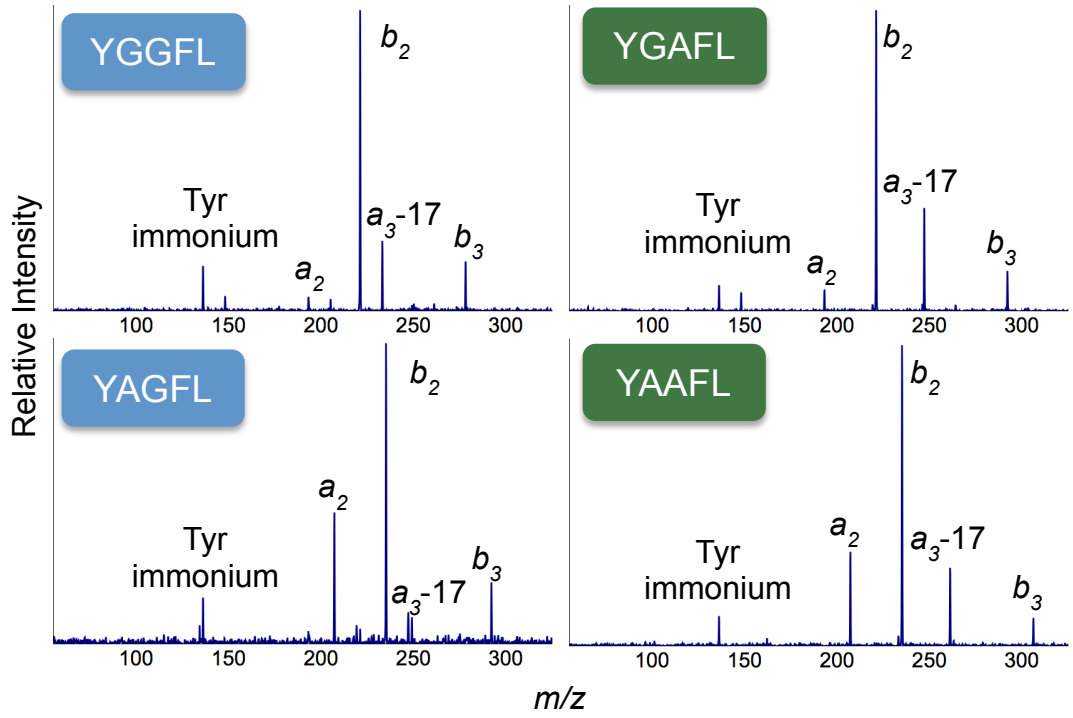


Figure 10. CID spectra of the the b_3 ions from all four leucine enkephalin analogs. Dissociation is to commonly observed sequence ions.

formed from common dissociation pathways known for protonated peptides and peptide fragments.

6.3.4.1 The differing experimental conditions of different types of mass spectrometers does not account for the CID/IRMPD product ion discrepancy

The experiments at FELIX are performed on an FT-ICR mass spectrometer, which operate at lower pressures ($<10^{-10}$ torr) and on a longer time frame (> 1s) than more commonly used QITs. Since CID experiments are often performed in QITs, what is considered a "commonly observed" product ion from CID of b_3 ions is established under different experimental conditions than those of the FT-ICR at FELIX.

To determine whether the differences in product ions from b_3-YGAFL and b_3-YAAFL were due to differences in the experimental conditions of a quadrupole-based mass spectrometer as opposed to an FT-ICR, IRMPD experiments on a quadruopole ion trap were performed. The product ions from IRMPD of b_3-YGAFL were again different from the

CID-type product ions and were the same as those observed during r-IRMPD in the FT-ICR at FELIX (Figure 11). Therefore the observed discrepancy between CID and r-IRMPD for b_3-YGAFL and b_3-YAAFL were not due to the differences in the experimental conditions of a quadrupole ion trap versus an FT-ICR. The spectra in Figure 11 are the first record of a

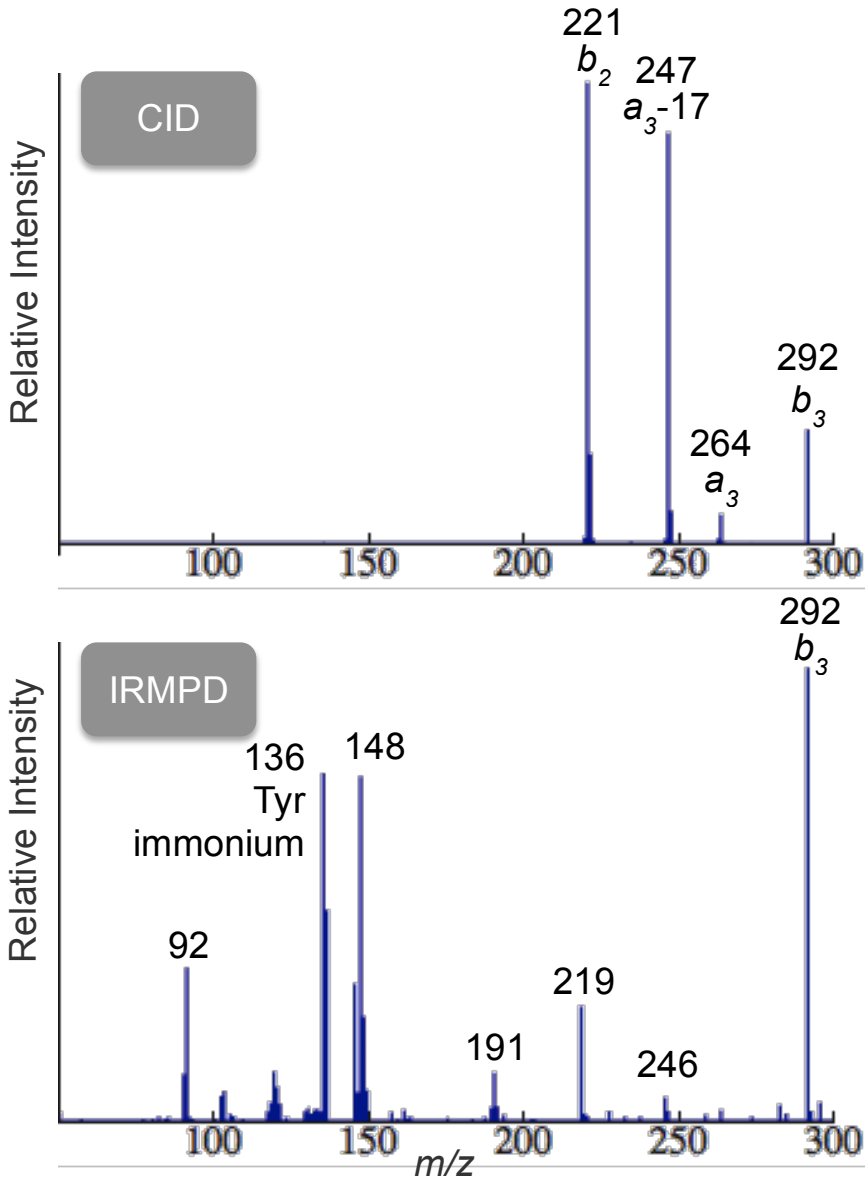


Figure 11. MS^n spectra from CID and IRMPD of the b_3 ion from YGAFL showing different types of product ions. Experiments were performed in a quadrupole ion trap mass spectrometer. The IRMPD results presented here were from dissociation with a CO_2 laser, but the same product ions were observed from r-IRMPD in the FT-ICR at FELIX.

different product ion distribution from CID and IRMPD that is not due to sequential dissociation.

6.3.4.2 The weak absorption cross section of b_3*-YGAFL and* b_3*-YAAFL results in a long IRMPD/r-IRMPD experiment time and consequent observation of product ion channels that are too kinetically constrained to be observed in CID experiments*

Dissociation kinetics and a modification of the potential energy diagram presented in Figure 9 will be used to explain the conflicting CID and IRMPD results for b_3 ions. The key to the explanation can be found in the infrared spectra of the four b_3 ions shown in Figure 6. Again the infrared spectra are shown, this time with the wavelength of the CO_2 laser used for IRMPD shown on the spectra (Figure 12). For the b_3-YGGFL and b_3-YAGFL oxazolones there is a significant transition within range of the CO_2 laser. Conversely, the absorption profile of b_3-YGAFL and b_3-YAAFL shows that there is little dissociation at that wavelength. In fact, dissociation at all wavelengths was more difficult for the non-oxazolone species. The spectra in Figure 12 are all normalized to the maximum intensity, which was between 80% and 100% dissociation efficiency for all four ions. However the b_3 oxazolones only required 2.5 seconds of irradiation to achieve those dissociation efficiencies, while b_3-YGAFL and b_3-YAAFL required 11 seconds.

6.3.5 A potential energy diagram proposed to satisfy the experimental observations

There are four observations that need to be accounted for in the proposed explanation:

(1) The b_3 ions from YGGFL and YAGFL are oxazolones while the b_3 ions from YGAFL and YAAFL have a different and currently unassigned structure.

(2) CID and IRMPD/r-IRMPD result in different product ions for b_3-YGAFL and b_3-YAAFL.

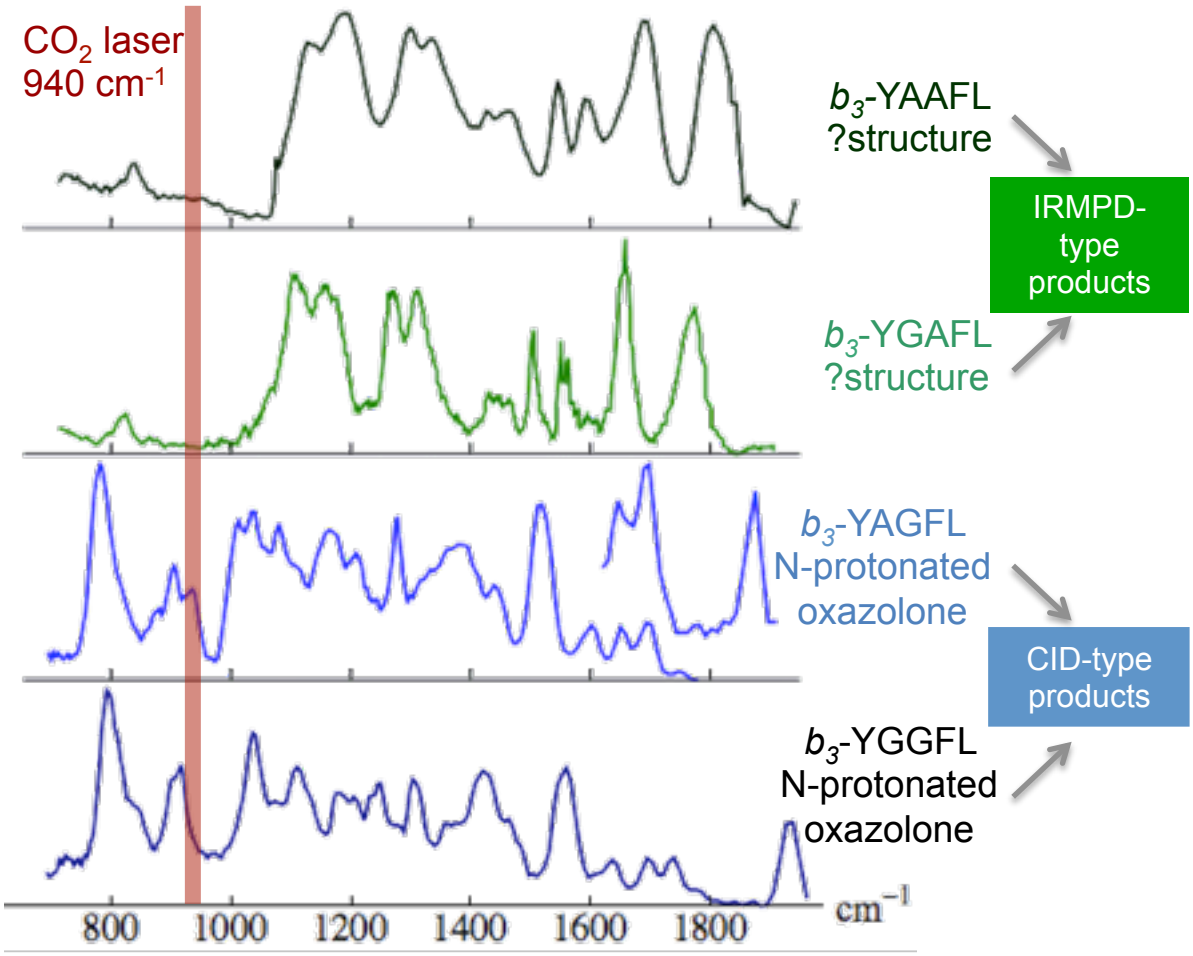


Figure 12. r-IRMPD spectra of the b_3 ions from YGGFL, YAGFL, YGAFL, and YAAFL taken at FELIX. The oxazolone strutures, b_3-YGGFL and b_3-YAGFL, dissociate to CID-type products during the experiment whereas b_3-YGAFL and b_3-YAAFL dissociate to IRMPD-type products. The output of a CO_2 laser is centered around 940 cm^{-1}, where there are no transitions in the b_3-YGAFL and b_3-YAAFL spectra.

(3) CID and IRMPD/r-IRMPD result in the same product ions for b_3-YGGFL and b_3-YAGFL.

(4) CID results in the same type of product ions for all four b_3 ions.

There are two important elements that emerge from these observations that can be used to redraw the proposed potential energy diagram. The first is that dissociation by IRMPD/r-IRMPD for b_3-YGAFL and b_3-YAAFL is slow compared to CID and compared to dissociation of the b_3-YGGFL and b_3-YAGFL oxazolones. The second is that there are two

distinct classes of product ions: The CID-type products formed from CID of all four b_3 ions and IRMPD/r-IRMPD of the b_3 oxazolones and the IRMPD-type products formed from IRMPD/r-IRMPD of b_3-YGAFL and b_3-YAAFL.

If dissociation to the IRMPD-type product ions is by a slow rearrangement channel, and the barrier to isomerization between the oxazolone and unassigned structure is comparable to or greater than the dissociation threshold, all of the stated experimental observations can be explained (Figure 13).

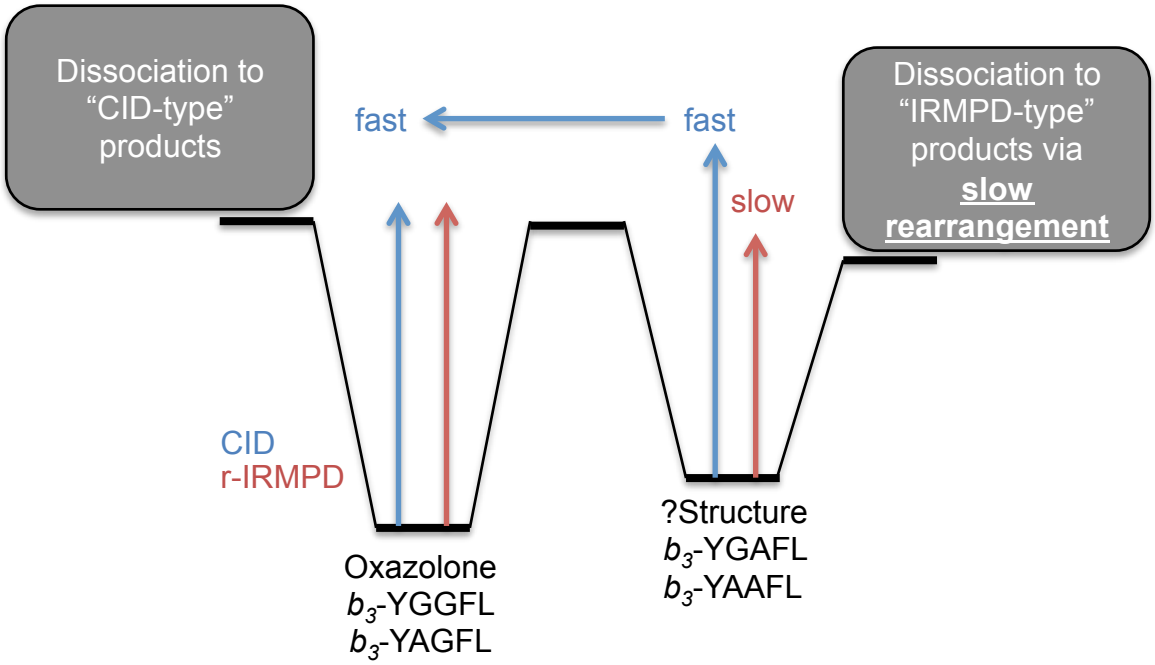


Figure 13. A potential energy diagram proposed to explain how CID and IRMPD could yield different product ions for the b_3-YGAFL and b_3-YAAFL, but not for the b_3-YGGFL and b_3-YAGFL.

(1) The oxazolone is the more stable structure for b_3-YGGFL and b_3-YAGFL, while the unassigned structure is more stable for b_3-YGAFL and b_3-YAAFL, and these populations can be distinguished by r-IRMPD. Most importantly, there is no interconversion between the structural classes.

(2) Dissociation of b_3-YGAFL and b_3-YAAFL by IRMPD/r-IRMPD is slow so the experimental time frame is sufficient for formation of IRMPD-type product ions. Dissociation of b_3-YGAFL and b_3-YAAFL by CID is fast enough that the barrier to CID-type products is surmounted before formation of the kinetically constrained IRMPD-type products.

(3) Dissociation of b_3-YGGFL and b_3-YAGFL by CID, IRMPD, and r-IRMPD is fast enough that CID-type products are always observed.

(4) The explanations in points 2 and 3 explain why CID experiments on all four species yields the same type of product ions.

So while CID and IRMPD operate by the same slow heating mechanism, and hence dissociate through the same product ion channels in the vast majority of cases, for b_3-YGAFL and b_3-YAAFL the absorption cross section is so weak that the experimental time frame of the IRMPD/r-IRMPD experiments is extended allowing dissociation through kinetically slow product channels that remain unobserved in CID experiments.

6.3.6 The relationship between CID- and IRMPD-type product ions

The explanation given in section 6.3.4 is predicated on there being two separate dissociation channels of differing rates instead of the usual sequential dissociation relationship between the product ions observed in CID and IRMPD. Often product ion relationships are complicated, however, and there can be many pathways to forming the same product ions. It turns out that for b_3-YGAFL, the IRMPD-type product ions can be formed from sequential dissociation of the CID-type product ions, but for the reasons discussed later in this section, sequential dissociation is not thought to be the reason for the observation of the IRMPD-type product ions in the FELIX experiments.

Sequential dissociation is limited in CID because the excitation waveform is in resonance with only the parent ion, so as soon as product ions are formed they begin to equilibrate with the room temperature helium bath gas. Application of a waveform in resonance with the secular frequency of the product ions should mimic the continuous excitation of an IRMPD experiment. The Bruker Esquire 3000 QITMS used for these experiments does not allow for the type of customized waveform that would simultaneously activate a parent and multiple product ions. However it is possible to use multiple CID steps without any isolation in between to excite selected ions in series. The results from sequential collisional activation of the b_3, a_3-17, and b_2 ions are shown in Figure 14. IRMPD-type product ions (219, 191, 148, and 136 *m/z*) are observed in this experiment, with the exception of the b_3-46 ions at 246 *m/z,* suggesting that the difference between the product ions observed

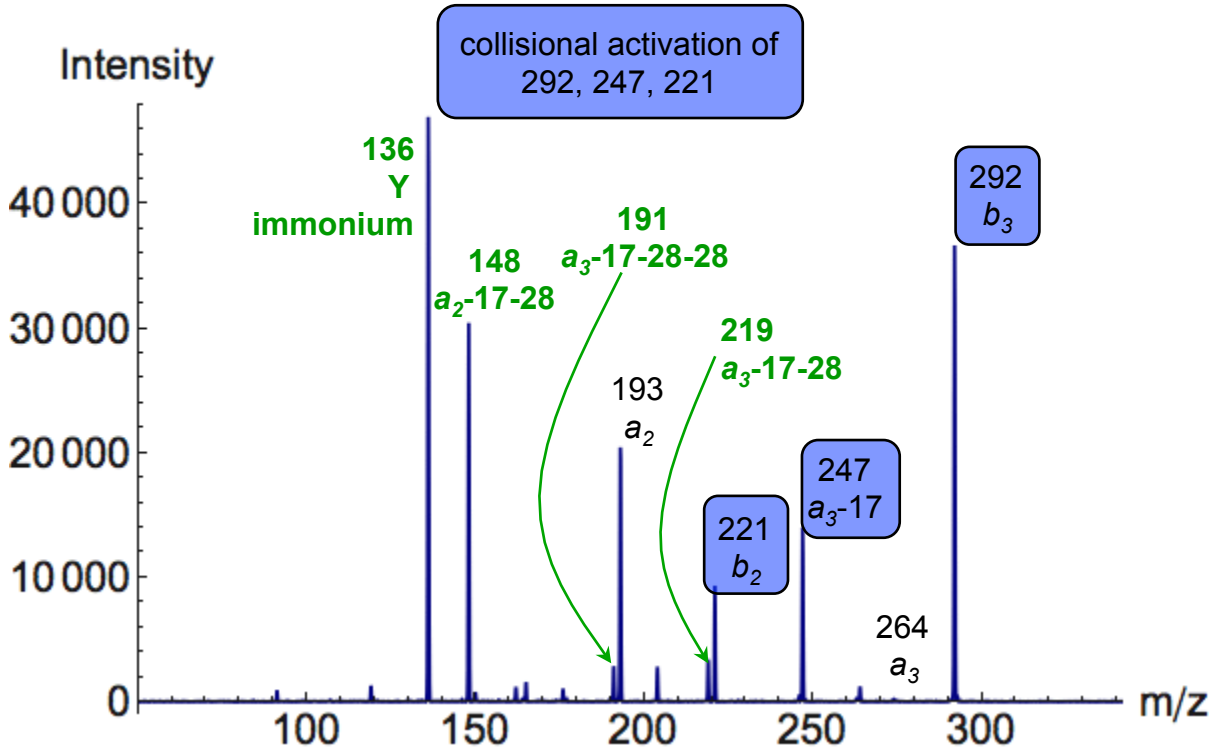


Figure 14. A CID experiment designed to mimic the sequential dissociation observed in IRMPD. Excitation waveforms in resonance with the b_3, a_3-17, and b_2 ions were applied sequentially resulting in the IRMPD-type product ions labeled in green

in CID and IRMPD/r-IRMPD could just be due to sequential dissociation. Regardless, this is not thought to be the reason for the observation of exclusively IRMPD-type product ions in the r-IRMPD experiments at FELIX or the IRMPD experiments performed in a QITMS with a CO_2 laser. There were no CID parameters that resulted in observation of solely IRMPD-type product ions. Some amount of b_3 and a_3-17 remained in the spectrum, along with the a_2 ion. In view of the fact that these ions are formed first and shown in these experiments not to be exceptionally labile they should have been observed before the IRMPD-type product ions during the FELIX experiments. Although the IRMPD-type product ions can be formed by sequential dissociation of CID-type product ions the fact that they were not observed during the FELIX experiments and an IRMPD-spectrum cannot be reproduced suggests that this is not the cause for the differences between CID and IRMPD/r-IRMPD of b_3-YGAFL.

Another experiment to test the potential energy diagram described in Figure 13 is to try to replicate the long experimental time frames of the FELIX experiment with a CID experiment by using low excitation amplitudes. The strategy is to create conditions that would allow access to the slow dissociation channel in a CID experiment. The maximum amount of excitation time allowed by the Bruker Esquire 3000 software is 9 seconds. Unfortunately isolated b_3-YGAFL ions that were trapped for this long formed adducts with an unknown 390 Da compound, likely a background impurity in the mass spectrometer (Figure 15). Again, there were no combinations of time and excitation waveform parameters that reproduced the IRMPD/r-IRMPD spectra.

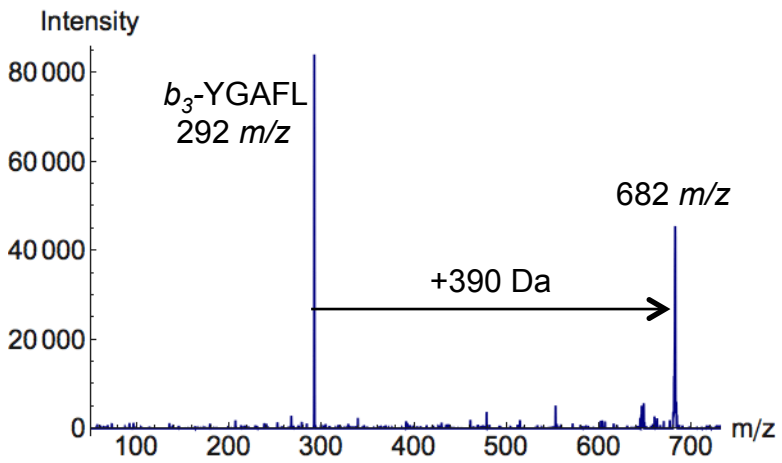


Figure 15. An adduct is formed, rather than IRMPD-type products, during a 9 second CID experiment in a Bruker Esquire 3000 QITMS

6.4 Conclusions

The structure of b_3 fragment ions from protonated leucine-enkephalin (YGGFL) and A/G substituted derivative has been investigate using CID and IRMPD spectroscopy. Both the IRMPD fragmentation as well as the spectral properties depend strongly on whether the third residue is glycine (YGGFL and YAGFL) or alanine (YGAFL and YAAFL). b_3 ions formed from YGGFL and YAGFL form an oxazolone structure protonated on the N-terminus, as concluded from their IR spectral signatures. In contrast, while no definite structure could be assigned to b_3-YGAFL and b_3-YAAFL ions, it is clear they are not oxazolone or conventional macrocyclic ions. CID data and the most closely matching theoretical IR spectra suggest that perhaps interaction with the tyrosine side chain is to be considered in searching for alternative structures or conformations of b_3-YGAFL/b_3-YAAFL.

The first instance of a different set of product ions formed from CID and IRMPD was also discussed. It is proposed that the IRMPD-type products are the result of a slow dissociation channel accessed by the unknown structure of b_3-YGAFL/b_3-YAAFL due to the inefficient IR absorption of this structure. The IRMPD-type product ions may also be

formed from sequential dissociation of CID-type product ions, which is a common result of continued irradiation of the product ions in an IRMPD experiment. Nevertheless the lack of an ability to reproduce and IRMPD-type MS^3 spectrum, and the lack of observation of any CID-type products in IRMPD experiments suggest that sequential dissociation does not fully account for the differences in the CID and IRMPD MS^3 spectra from b_3-YGAFL.

References

(1) Harrison, A. G. *Mass Spectrometry Reviews* **2009**, *28*, 640.

(2) Paizs, B.; Suhai, S. *Mass Spectrometry Reviews* **2005**, *24*, 508.

(3) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1165.

(4) Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128*, 10364.

(5) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *Journal of the American Chemical Society* **2008**, *130*, 17774.

(6) Goloborodko, A.; Gorshkov, M.; Good, D.; Zubarev, R. *J. Am. Soc. Mass Spectrom.* **2011**, *22*, 1121.

(7) Saminathan, I. S.; Wang, X. S.; Guo, Y. Z.; Krakovska, O.; Voisin, S.; Hopkinson, A. C.; Siu, K. W. M. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 2085.

(8) Yu, L.; Tan, Y. L.; Tsai, Y.; Goodlett, D. R.; Polfer, N. C. *Journal of Proteome Research* **2011**, *10*, 2409.

(9) Allen, J. M.; Racine, A. H.; Berman, A. M.; Johnson, J. S.; Bythell, B. J.; Paizs, B.; Glish, G. L. *J Am Soc Mass Spectrom* **2008**, *19*, 1764.

(10) Yalcin, T.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 233.

(11) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2174.

(12) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2005**, *127*, 17154.

(13) Chen, X. A.; Steill, J. D.; Oomens, J.; Polfer, N. C. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1313.

(14) Fattahi, A.; Zekavat, B.; Solouki, T. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 358.

(15) Garcia, I. R.; Giles, K.; Bateman, R. H.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1781.

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

(17) Polfer, N. C.; Bohrer, B. C.; Plasencia, M. D.; Paizs, B.; Clemmer, D. E. *Journal of Physical Chemistry A* **2008**, *112*, 1286.

(18) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Journal of the American Chemical Society* **2007**, *129*, 5887.

(19) Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. C. *Journal of the American Chemical Society* **2009**, *131*, 18272.

(20) Erlekam, U.; Bythell, B. J.; Scuderi, D.; Van Stipdonk, M.; Paizs, B.; Maitre, P. *Journal of the American Chemical Society* **2009**, *131*, 11503.

(21) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *Journal of the American Chemical Society* **2008**, *130*, 17644.

(22) Williams, D. H.; Bradley, C. V.; Santikarn, S.; Bojesen, G. *Biochemical Journal* **1982**, *201*, 105.

(23) Biemann, K. *Biomedical and Environmental Mass Spectrometry* **1988**, *16*, 99.

(24) Cordero, M. M.; Houser, J. J.; Wesdemiotis, C. *Analytical Chemistry* **1993**, *65*, 1594.

(25) Farrugia, J. M.; O'Hair, R. A. J.; Reid, G. E. *Int. J. Mass Spectrom.* **2001**, *210*, 71.

(26) Chen, X. H.; Turecek, F. *J. Am. Soc. Mass Spectrom.* **2005**, *16*, 1941.

(27) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 334.

(28) Harrison, A. G.; Young, A. B. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 1810.

(29) Cooper, T.; Talaty, E.; Grove, J.; Van Stipdonk, M.; Suhai, S.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1654.

(30) Bythell, B. J.; Molesworth, S.; Osburn, S.; Cooper, T.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1788.

(31) Bythell, B. J.; Maitre, P.; Paizs, B. *Journal of the American Chemical Society* **2010**, *132*, 14766.

(32) Savitski, M. M.; Hith, M.; Fung, Y. M. E.; Adams, C. M.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1755.

(33) Newsome, G. A.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1127.

(34) Pearlman, D. A.; Case, D. A.; Caldwell, J.; Seibel, G. L.; W=Singh, U. C.; Weiner, P.; Kollman, P. A.; University of California: San Francisco, CA, 1991.

(35) Wyttenbach, T.; Paizs, B.; Barran, P.; Breci, L.; Liu, D. F.; Suhai, S.; Wysocki, V. H.; Bowers, M. T. *Journal of the American Chemical Society* **2003**, *125*, 13768.

(36) Paizs, B.; Suhai, S.; Hargittai, B.; Hruby, V.; Somogyi, A. *Int. J. Mass Spectrom.* **2002**, *219*, 203.

(37) Ochterski, J. W. *Technical Support Information: Gaussian Inc.* **1999**.

(38) Frisch, M. J. *Gaussian 03, Revision C.02 Gaussian, Inc., Wallingford CT* **2004**.

(39) Scott, A. P.; Radom, L. *Journal of Physical Chemistry* **1996**, *100*, 16502.

(40) Irikura, K. K.; Johnson, R. D.; Kacker, R. N.; Kessel, R. *Journal of Chemical Physics* **2009**, *130*.

(41) Wood, K. V.; Burinsky, D. J.; Cameron, D.; Cooks, R. G. *Journal of Organic Chemistry* **1983**, *48*, 5236.

(42) Bourcier, S.; Hoppilliard, Y. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 93.

(43) Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L. *Journal of the American Chemical Society* **1997**, *119*, 5481.

(44) Verkerk, U. H.; Zhao, J. F.; Van Stipdonk, M. J.; Bythell, B. J.; Oomens, J.; Hopkinson, A. C.; Siu, K. M. *Journal of Physical Chemistry A* **2011**, *115*, 6683.

(45) Ferzoco, A. L.; Steill, J. D.; Oomens, J.; Bythell, B.; Paizs, B.; Glish, G. L. In *The 58th ASMS Conference on Mass Spectrometry and Allied Topics* Salt Lake City, Ut, 2010.

Chapter 7

The Structure of a_4 and a_4-17 Ions as Measured by CID and r-IRMPD

7.1 Introduction

In addition to b_n and y_n ions, a_n sequence ions are also commonly observed in CID spectra of peptides[1]. They are observed as stable product ions and important intermediates in the dissociation chemistry of both the full peptide[2] and b_n ions[3]. Similarly to b_n ions, a_n ions have been hypothesized to dissociate by cyclization reactions in which the original peptide sequence is scrambled in the product ions[4], producing data that is difficult to interpret. Whether an a_n sequence ion is observed, and which product ions are formed from sequential dissociation, are dependent on the structure of the a_n ion. Hence many structural studies have been performed to gain an understanding of the dissociation chemistry of a_n ions[4-11].

The structure of b_3 ions was shown to be sensitive to the amino acid sequence even though no sequence-dependent behavior was seen in the product ion distribution in CID experiments[3,12,13]. Conversely, sequence dependent product ion distributions have been observed in CID experiments on a_4 ions. Results from CID experiments in a Bruker Esquire quadrupole ion trap indicate that the dissociation pattern of a_4 ions from leucine enkepahlin derivatives depends on glycine position[14] (Table 1, reproduced from reference 14 with permission). If a glycine is the second residue of the peptide, the a_4 ion dissociates to form mostly a_4-17 as the base peak, whereas if glycine is the third residue, a *b*-type rearrangement

ion is the most abundant product ion. Further, a_4-17 ions form a *b*-type rearrangement ion as the most abundant product ion only if glycine is not the third residue of the original peptide (Table 2). Later experiments performed on a Bruker Esquire 3000 quadrupole ion trap did not have the same sequence dependent dissociation pattern as the data shown in Tables 1 and 2, but an entirely new set of trends that depends on glycine position was observed[15]. Despite the inconsistencies in the CID data, a sequence dependence repeatedly emerges. Especially considering the sequence sensitive structures found for b_3 ions, it was hypothesized that IR spectroscopy experiments on a_4 ions would show the same sequence sensitive ion structure.

Table 1. Product ion distributions from CID of the a_4 ion from a series of leucine enkephalin analogs of the form YXZFL differing in glycine position.

Peptide	***a*$_4$-17**	**FYX**	**ZFY**	***b*$_3$**
YGGFL	63	**100**		15
YAAFL	**100**	26		12
YGVFL	**100**	16	12	10
YVGFL	57	**100**		17
YGLFL	**100**	19	19	13
YLGFL	49	**100**		22
YGAFL	**100**	26	12	12
YAGFL	78	**100**		27

Table 2. Product ion distributions from CID of the a_4-17 ion from a series of leucine enkephalin analogs of the form YXZFL differing in glycine position.

Peptide	***a*$_4$-17-28**	**FYX**	**FYX-28**	***a*$_3$-28**
YGGFL	100	83	33	33
YAAFL	76	**100**		36
YGVFL	73	**100**		29
YVGFL	17	21	27	100
YGLFL	80	**100**		35
YLGFL	72	39	64	100
YGAFL	76	**100**		36

7.2 Experimental Methods

IR spectra were obtained by resonant IRMPD (r-IRMPD) over an energy range of 800 to 1800 cm^{-1} using the Free Electron Laser for Infrared Experiments (FELIX). Ions were generated by electrospray ionization from 10 μM solutions in 49.5:49.5:1 methanol:water:acetic acid. Protonated peptides were formed directly from the electrospray process, and the a_4 and a_4-17 peptide fragment ions were generated by in-source fragmentation. After formation, ions were transmitted to an FT-ICR for isolation and analysis. IR spectra were generated by plotting the dissociation efficiency of the selected parent ion as a function of the FELIX wavelength.

Theoretical calculations of the a_4 fragment were performed using density functional theory in Gaussian using the B3LYP functional. Two trans imine structures, one with the proton located on the N-terminus and the other protonated at the C-terminal imine, were first optimized using the 3-12g basis set. Structures were optimized at successively higher levels of theory, up to 6-311++g(d,p), and the frequency calculation was performed at this final level of theory. Spectra were scaled by a factor of 0.98 to account for anharmonicity. This value was chosen based on published work[16] and because it gave the best fit of the experimental data for the structure protonated at the N-terminus while there were no scaling factors that fit the structure protonated at the C-terminal imine.

7.3 Results and Discussion

7.3.1 No sequence dependence is observed in infrared spectra

The r-IRMPD spectra of the a_4 ion of four leucine enkephalin analogs are shown in Figure 1. The spectra are congested and signal to noise is poor, but all four ions appear to belong to the same structural class. The similarities are particularly clear when compared

with the b_3 ion spectra that show distinctly different classes. Since appearance rates in MS^n spectra depend on both the stability, and therefore the structure, of each successive generation of parent and product ions, r-IRMPD spectra were also obtained for intact protonated peptides and a_4-17 ions (Figures 2 and 3). As in the a_4 ion spectra, no sequence dependent differences in structure were observed.

In previous studies of a_4 ions, their structure has most often been assigned as a linear imine protonated at either the N- or C-terminus[7] or the macrocyclic structure implicated in

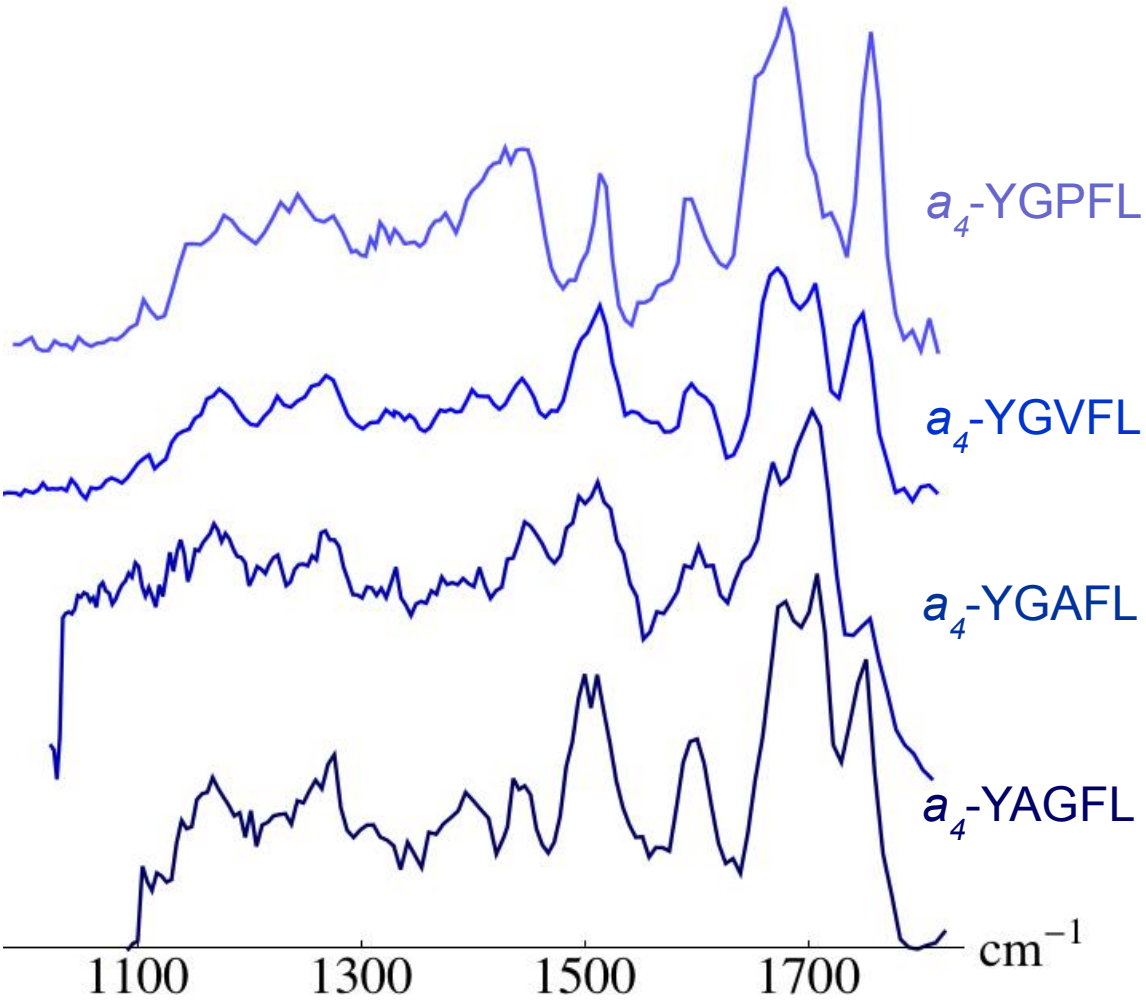


Figure 1. Resonant IRMPD spectra of a_4 ions from YAGFL, YGAFL, YGVFL, and YGPFL. The spectra show no sequence dependence in the structure of these peptide fragments.

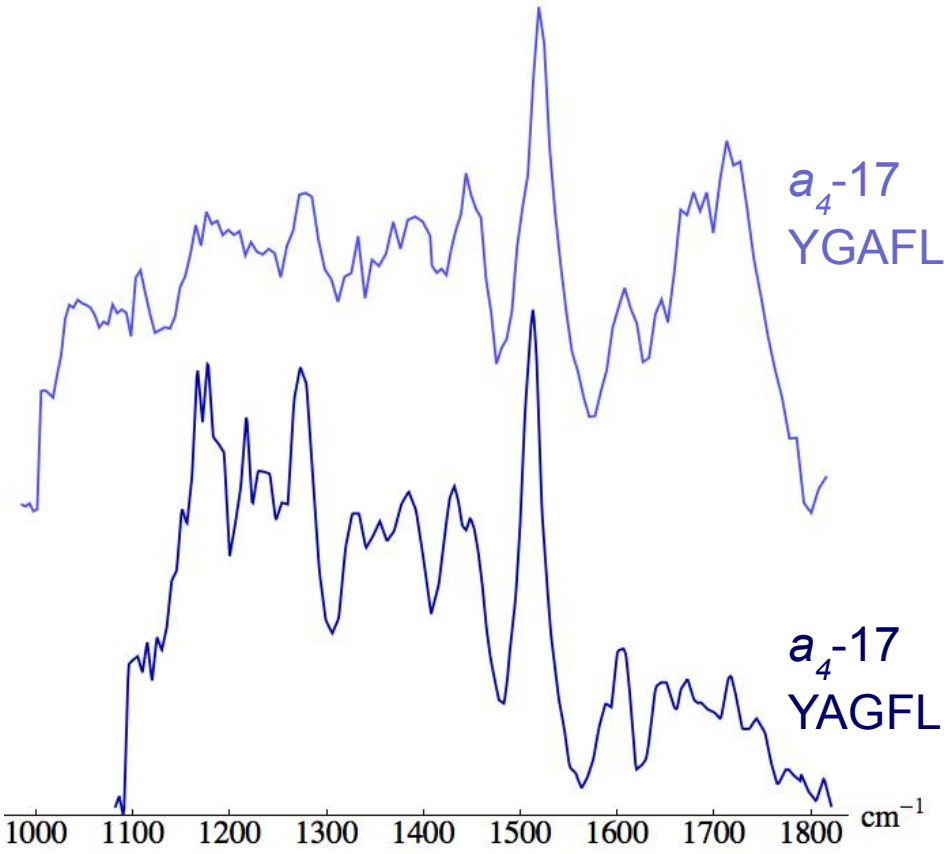


Figure 2. Resonant IRMPD spectra of a_4-17 peptide fragments from YAGFL and YGAFL. Again, the similarity of the spectra indicate no sequence dependence in the structure of these ions.

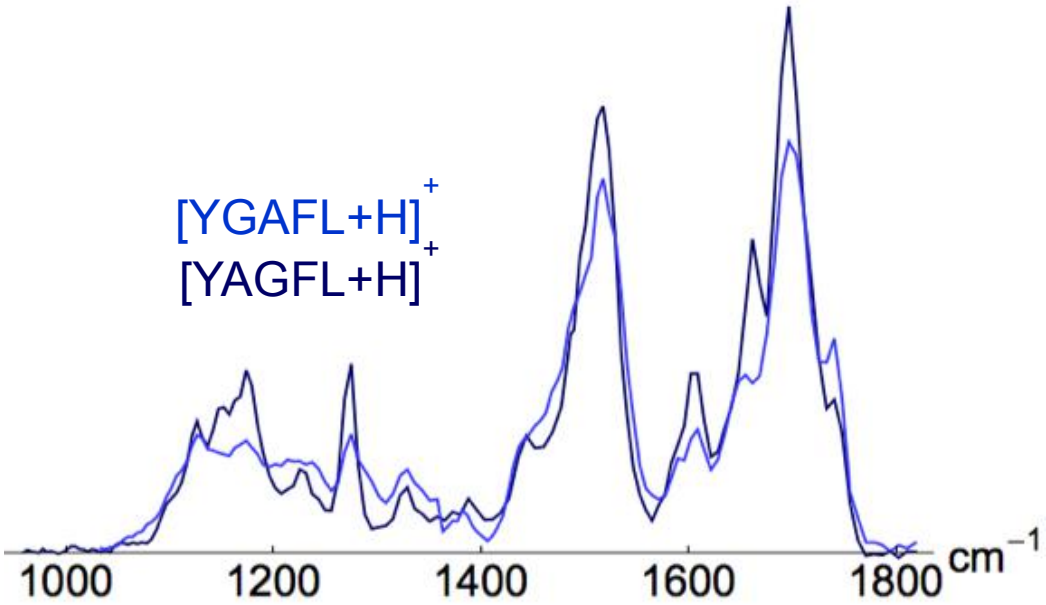


Figure 3. Resonant IRMPD spectra of protonated YAGFL (dark blue) and YGAFL (light blue), showing similar structures.

sequence scrambling rearrangements[4]. Shown in Figure 4 are calculated spectra of linear imine structures protonated at the N- (bottom) and C-terminus (top) compared to the experimental spectrum of the a_4 ion from YGAFL. The N-protonated imine gives the better match to the experimental spectrum and can account for most of the observed transitions. Some of the transitions, specifically those at approximately 1220, 1450, 1600, and 1680 cm^{-1}, do not appear in the theoretical spectrum. Previous results for the a_4 ion from YGGFL also indicate the presence of predominantly N-protonated imine with contributions from a macrocycle[7]. The band positions and intensities of the calculated macrocycle do not appear to account sufficiently for the unassigned transitions in the experimental spectrum. It is also possible that a smaller cyclic structure involving only part of the backbone is present in the experiment, but thus far such a structure has not been calculated or assigned to experimental data.

Multiple mechanisms for the formation of a_n-17 ions from a_n ions have been proposed[4,10,17]. The Vachet-Glish cyclization mechanism begins with an N-protonated imine, but proceeds through a high energy transition state (214 kJ mol^{-1})[4,17]. The mechanism proposed by Blieholder et al proceeds through a lower energy transition state (76.1 kJ mol^{-1}), but begins with a proton bound dimer structure that has not been observed as a stable species in spectroscopy experiments[17]. Neither mechanism is particularly convincing, the former because of the high energies involved and the latter because of the unusual starting structure. No mechanism will be proposed here, but there may be mechanistic information in the spectra of a_4-17 ions as compared to a_4 ions (Figure 5). The main difference between these spectra is in the region above 1600 cm^{-1}. For the a_4 ion, the transitions in that region are largely due to backbone carbonyl and amide stretches, but there is also a transition involving

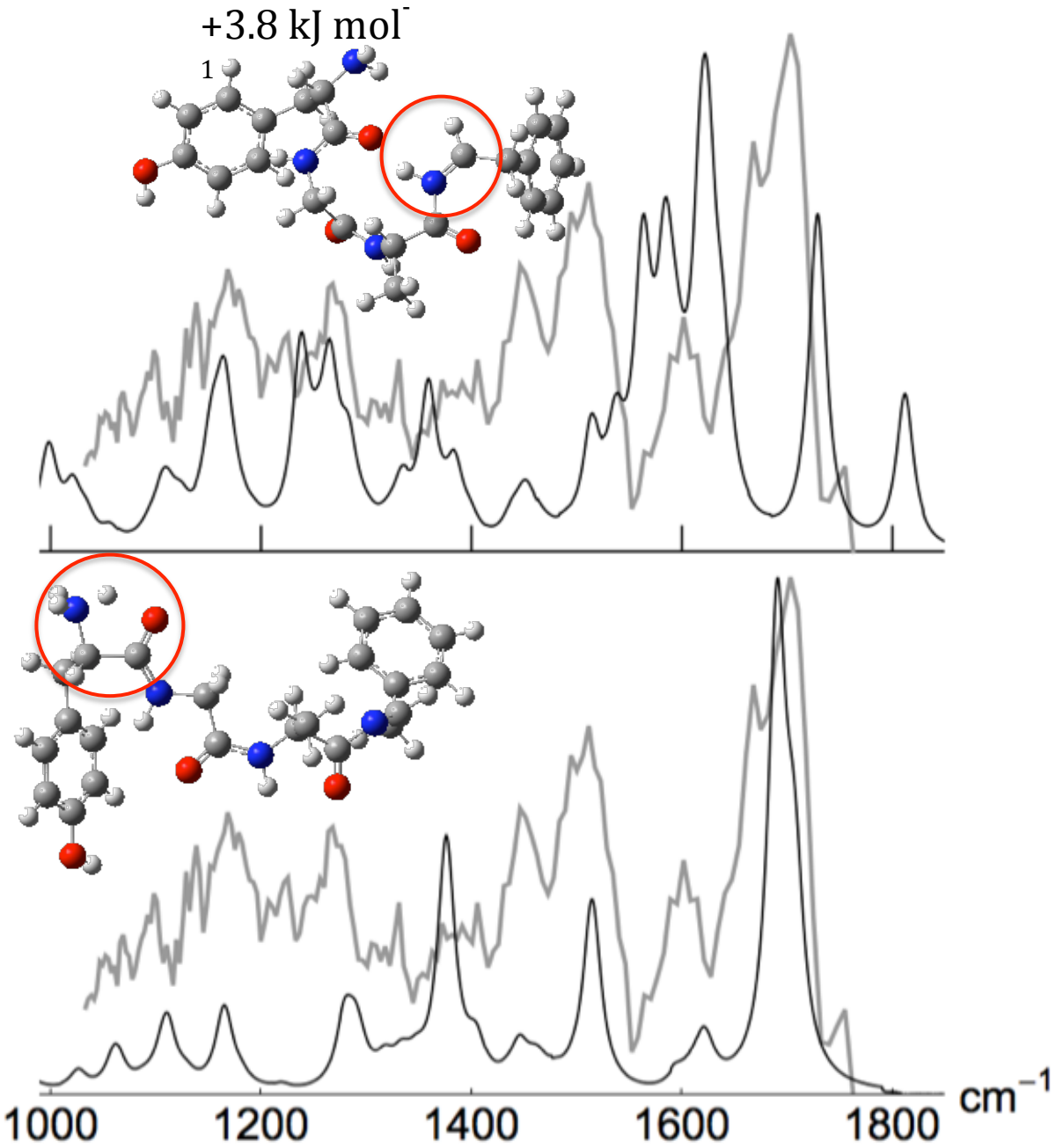


Figure 4. Experimental rIRMPD spectra of the a_4 ion from YGAFL (gray) compared to calculated spectra for a C-terminal imine structure protonated at the C-terminus (top) and the N-terminus (bottom). The calculations were performed at the B3LYP/6-311g++(d,p) level of theory, and the optimized structures are shown with their corresponding spectra, with the protonation site highlighted.

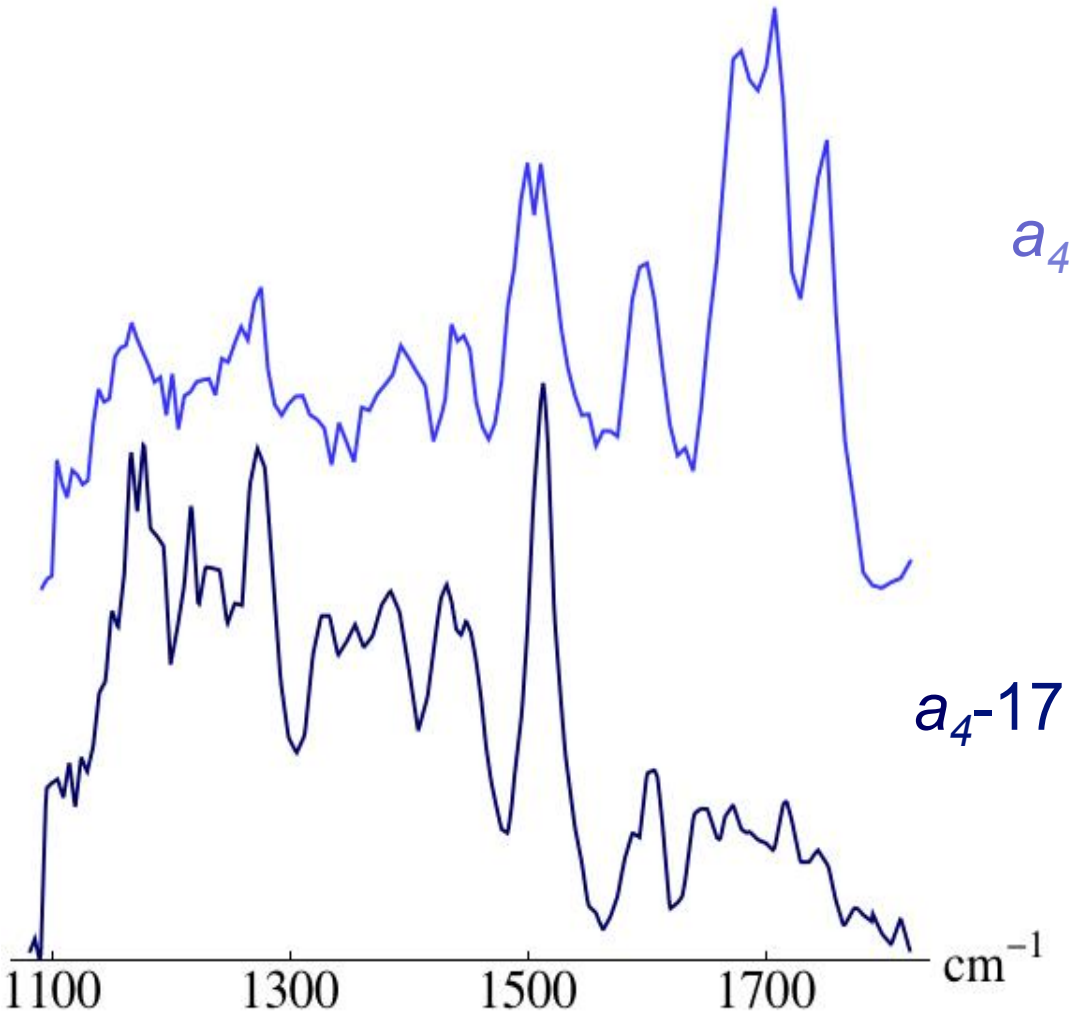


Figure 5. Experimental rIRMPD spectra of the a_4 ion from YAGFL (top) compared to the a_4-17 ion (bottom). The major difference is in the region above 1600 cm^{-1}, which are bands due to vibrational modes involving the peptide backbone and the tyrosine ring.

Table 3. Band assignments for the high energy modes of the a_4 ion from YAGFL.

Wavenumber (cm^{-1})	Intensity (km mol^{-1})	Assignment
1653.17	96	C=C str. Tyrosine ring
1724.09	835	CO and CN stretches on the backbone
1727.00	229	
1738.26	160	
1742.61	286	

symmetric C-C stretching modes of the tyrosine ring (Table 3). The lack of these bands in the a_4-17 spectra could be evidence of backbone rearrangement to form the a_4-17 product ion, or a different protonation site that shifts the backbone stretches above the range of the laser.

7.3.2 CID and r-IRMPD provide complementary information for a_4 *ions*

CID experiments on the a_4 ions of leucine enkephalin analogs showed sequence dependent product ion distributions[15,18], but r-IRMPD spectra show no sequence dependent structures. Both CID and r-IRMPD are experiments used to obtain information about ion structure and so initially the results appear to contradict one another. In fact the discrepancy in this case is due to dissociation kinetics, which CID is sensitive to, but r-IRMPD is not.

The r-IRMPD experiments establish that dissociation of a_4 ions is from the same starting structure, and that they dissociate to the same a_4-17 structure for the peptide sequences studied (although infrared spectra were not obtained for the other product in channels). Since the parent and product ion structures don't change with the peptide sequence, the dissociation kinetics established by the transition states must be the factor determining the sequence dependent differences in the relative abundances of product ions in CID experiments. The rates of product ion formation, and therefore the abundance of product ions observed in MS^n spectra, are sensitive to the geometry of the transition state (recall section 2.1.1)[19]. Dissociation reactions that require the participation of many atoms in the transition state are geometrically constrained and therefore have a low density of states. The low transition state density results in a slow dissociation rate and low product ion abundance as compared to an energetically similar channel with a less constrained transition state. Hence the product ion abundances for competitive dissociation channels in CID

experiments give information about the relative energetics and geometry of the reaction coordinate.

At a first the CID and r-IRMPD results for a_4 ions seemed to contradict one another, but they are in fact complementary. Because the r-IRMPD experiments establish the structural similarity between the different sequences the sequence dependent product ion distributions in CID experiments can be interpreted as resulting from differences in the dissociation kinetics due to the transition state geometries. If a glycine is the third residue from the N-terminus, dissociation is predominantly to product ions that must be formed from complicated rearrangements, as opposed to the simpler ammonia loss channel to form a_4-17. Therefore glycine either stabilizes the rearrangement transition state, or destabilized the transition state to formation of a_4-17.

7.4 Conclusions

The structure of a_4 ions from leucine enkephalin analogs was found to predominantly have an imine structure at the C-terminus with the protonation site at the N-terminus, although there does appear to be some contributions in the spectra from a second structure. A macrocyclic structure or a cyclic structure involving only part of the backbone may account for the unassigned transitions. Unlike b_3 ions, which show distinct sequence dependence in their structures, no sequence dependence was found for a_4 or a_4-17 sequence ions, or the full peptide.

The lack of a sequence dependence of ion structure observed in IR spectroscopy experiments appears to contradict the sequence dependence seen in CID data, and raises the question of what is the cause of the sequence-dependent product ion distributions seen in CID. Recall from Chapter 2 that the rate of unimolecular dissociation of a parent ion, which

governs the product ion distributions observed in CID experiments, depend on the energetics and structure of the transition state (tight versus loose). Neither CID nor r-IRMPD experiments have the ability to directly measure the barriers to formation of various products. While the parent and product ion structures related to a_4 ions may show no sequence dependence, differences in transition state structures may explain sequence dependent behavior in CID experiments.

References

(1) Savitski, M. M.; Hith, M.; Fung, Y. M. E.; Adams, C. M.; Zubarev, R. A. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1755.

(2) Vachet, R. W.; Ray, K. L.; Glish, G. L. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 341.

(3) Allen, J. M.; Racine, A. H.; Berman, A. M.; Johnson, J. S.; Bythell, B. J.; Paizs, B.; Glish, G. L. *J Am Soc Mass Spectrom* **2008**, *19*, 1764.

(4) Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L. *Journal of the American Chemical Society* **1997**, *119*, 5481.

(5) Bythell, B. J.; Molesworth, S.; Osburn, S.; Cooper, T.; Paizs, B.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1788.

(6) Vachet, R. W.; Winders, A. D.; Glish, G. L. *Analytical Chemistry* **1996**, *68*, 522.

(7) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *Journal of the American Chemical Society* **2007**, *129*, 5887.

(8) Garcia, I. R.; Giles, K.; Bateman, R. H.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1781.

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

(10) Bythell, B. J.; Maitre, P.; Paizs, B. *Journal of the American Chemical Society* **2010**, *132*, 14766.

(11) Cooper, T.; Talaty, E.; Grove, J.; Van Stipdonk, M.; Suhai, S.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1654.

(12) Ferzoco, A. L.; Steill, J. D.; Oomens, J.; Bythell, B.; Paizs, B.; Glish, G. L. In *The 58th ASMS Conference on Mass Spectrometry and Allied Topics* Salt Lake City, Ut, 2010.

(13) Allen, J. M.; Black, D. M.; Glish, G. L. In *The 52nd ASMS Conference on Mass Spectrometry and Allied Topics* Nashville, TN, 2004.

(14) Black, D.; Glish, G. In *The 50th ASMS Conference on Mass Spectrometry and Allied Topics* Orlando, Florida, 2002.

(15) Spencer, S. E.; Ferzoco, A. L.; Arnold, M. A.; Steill, J. D.; Oomens, J.; Glish, G. L. In *The 59th ASMS Conference on Mass Spectrometry and Allied Topics* 2011.

(16) Irikura, K. K.; Johnson, R. D.; Kacker, R. N.; Kessel, R. *Journal of Chemical Physics* **2009**, *130*.

(17) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *Journal of the American Chemical Society* **2008**, *130*, 17774.

(18) Glish, L. L.; Payne, A. H.; Glish, G. L. In *The 50th ASMS Conference on Mass Spectrometry and Allied Topics* Orlando, Florida, 2002.

(19) Baer, T.; Mayer, P. M. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 103.

Chapter 8

Summary and Future Work

8.1 Summary

The goal of the work presented in Chapters 3 and 5 was to improve the quality of data obtained from MS^n experiments. The experiments presented in Chapter 3 revealed that IRMPD efficiency increases with peptide size. This is useful because the efficiency of CID, which is currently the most commonly used MS^n method[1-4], decreases with size and limits the amount of sequence information that can be obtained for larger peptides/proteins[5-8]. The reason for the increase in IRMPD efficiency with size was explored in Chapter 4. A combined experimental/theoretical method to measure the increase in enthalpy of ions was attempted, but non-Boltzmann internal energy distributions that can arise in IRMPD experiments impose a limit on the size of analyte for which the measurement can be accurately performed. Work presented in Chapter 5 demonstrated how derivatizing the N-terminus of a peptide with an FMOC group increases the abundance of the sodium-cationized species relative to the protonated species. Further, $[FMOC\text{-}peptides+Na]^+$ were shown to dissociate by the same predictable C-terminal loss channel as $[peptide+Na]^+$, which means the peptide sequence can be easily read from a successive MS^n experiments.

The goal of the work presented in Chapters 6 and 7 was to improve the understanding of the dissociation chemistry of protonated peptide fragments through measurements of

structure. Spectroscopy experiments revealed a sequence-sensitive structure for b_3, but not a_4 ions. The position of glycine and alanine in leucine enkephalin analogs determined whether the structure of b_3 ions was a C-terminal oxazolone or a hitherto unassigned structure, but had no influence on the N-protonated imine structure found for a_4 ions. The sequence sensitivity of b_3 ions was particularly surprising because CID experiments do not reveal a structural dependence, and the difference between the glycine and alanine side chains is minor, especially considering the chemically interesting functional groups present in the side chains of other amino acids. The complementary information acquired by combining the results from CID, IRMPD, and r-IRMPD experiments leads to a more complete understanding of peptide fragment dissociation chemistry because relative rate information can begin to be incorporated.

For both b_3 and a_4 ions, there is spectral and CID evidence that the tyrosine side chain may be involved in the structure of these ions in a way that was not anticipated nor understood, and may be the cause of complicated rearrangements upon dissociation. Except in the case of the basic or acidic amino acids, which form salt bridge structures[9-11], amino acid side chains are generally disregarded as bystanders in peptide dissociation chemistry. The involvement of tyrosine and the sensitivity of b_3 structure to alanine and glycine, all of which are hydrophobic residues, indicates that side chain participation should be considered more closely.

Lastly, presented in Chapter 6 was the first report of the formation a different set of product ions from CID versus IRMPD/r-IRMPD. The two techniques are energetically and mechanistically similar and are therefore expected to access similar dissociation channels. In fact experiments on many peptides and peptide product ions have been performed using both

dissociation techniques, and the same sets of product ions are repeatedly observed. Any differences that arise are generally attributed to the sequential dissociation that occurs in IRMPD because of the continued activation of product ions by the laser. Sequential dissociation does not, however, does not account for all experimentally observed differences between CID and IRMPD for b_3-YGAFL.

8.2 Suggested future work

8.2.1 IRMPD efficiency as a function of peptide size (Chapter 3)

The efficiency of IRMPD was shown to increase with peptide ion size in a QIT for the series of peptides studied. Many proteomics experiments are performed in FT-ICRs, however, where important experimental parameters differ from those in a QIT. In FT-ICRs, as compared to QITs, the laser overlaps with a smaller portion of the ion cloud, there are far fewer collisions with the neutral background gas, and the experiments are carried out on a timescale where radiative processes become efficient[12]. To determine the applicability of IRMPD for obtaining sequence information on large peptides/proteins, these experiments should be repeated in an FT-ICR. Further, rather than just considering the amount of parent ion dissociation, whether or not the additional product ions result in a higher probability of identification by sequencing software should be explored.

8.2.2 Measuring the effective temperature of activated parent ions (Chapter 4)

Sequential dissociation will differ between TA-CID and IRa-CID. Using sodiated analytes, for which the product ion lineages are well understood, could make it so differences in sequential dissociation can be extracted from dissociation of the parent ion. The root mean square difference between TA-CID and IRa-CID was used to assign match scores. Programs

are available for comparing spectra that weight the calculation by the peak intensity, and can compute confidence intervals[13], which would improve quantitation of the method.

8.2.3 C-terminal sequencing of [FMOC-peptide+Na]$^+$ (Chapter 5)

To further the practical application of this technique, experiments should be performed that more closely resemble the types of analytes and experimental procedures encountered in many proteomics studies. For example, it is unknown how the FMOC group would react with free amines and other functional groups found in non-hydrophobic amino acids, and further it is unknown how the dissociation chemistry would be influenced if a reaction does occur. The derivatization procedure would need to be incorporated into LC-MS^n protocols, eventually in an automated manner. Studies on how many stages of MS^n are required for identification with a specified amount of confidence should be performed to see if this method can be implemented on the time scale of an LC-MS^n experiment.

To further the understanding of the dissociation mechanism, calculations of longer peptides where the C-terminus is farther from the FMOC group should be performed to find the sodium binding site in an ion more likely to be encountered in a proteomics experiment.

8.2.4 The sequence-sensitivity of b_3 *ions (Chapter 6)*

Identification of the structure of b_3-YGAFL and b_3-YAAFL and the IRMPD-type products is necessary to confirm whether the proposed potential energy diagram is accurate. Further MS^n experiments may help determine the structure of the IRMPD-type product ions. The mass accuracy of an FT-ICR could determine the elemental formula of the IRMPD-type products, which is a starting point for finding the structure. Searches for structures other than the oxazolone and macrocycle structures that are usually suggested should be performed to try to determine the structure of b_3-YGAFL and b_3-YAAFL. N-terminal oxazolones, and

ring structures that are formed by incorporating only part of the backbone are good candidates to begin the search. It would also be interesting to see if multiple structures are observed in hydrogen-deuterium exchange, ion mobility, and ion-molecule reactions. The strategy behind these experiments is to decouple structural measurement from ion activation. These experiments do not involve activation of the parent ion, so they may determine whether populations of both structures exist for each b_3 species at room temperature and observation of only one or the other is an artifact of the activation methods used. MS^n experiments on peptides containing tyrosine-related amino acids that have modified hydroxyl groups on their side chains would also lend insight into how the side chain participates in the structure of b_3 ions.

8.2.5 The structure of a_4 *ions (Chapter 7)*

Calculations of transition states, dissociation rates, and how they are influenced by sequence depend on knowing the structure of the parent and product ions. Theoretical calculations of additional a_4 structures are needed to determine what structures contribute to the unassigned bands in the experimental infrared spectrum. Additionally, calculations to match the experimental spectra of a_4-17 ions from YGAFL and YAGFL are needed to identify the structure of these product ions.

References

(1) Sleno, L.; Volmer, D. A. *Journal of Mass Spectrometry* **2004**, *39*, 1091.

(2) Jennings, K. R. *Int. J. Mass Spectrom.* **2000**, *200*, 479.

(3) McLuckey, S. A. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 599.

(4) Futrell, S. A. J. *Journal of Mass Spectrometry* **2000**, *35*, 1069.

(5) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Analytical Chemistry* **1994**, *66*, 2809.

(6) Baer, T.; Mayer, P. M. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 103.

(7) Baer, T.; Hase, W. L. *Unimolecular Reaction Dynamics*; Oxford University Press: New York, 1996.

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

(9) Drayss, M. K.; Blunk, D.; Oomens, J.; Polfer, N.; Schmuck, C.; Gao, B.; Wyttenbach, T.; Bowers, M. T.; Schafer, M. *Int. J. Mass Spectrom.* **2009**, *281*, 97.

(10) Wyttenbach, T.; Bushnell, J. E.; Bowers, M. T. *Journal of the American Chemical Society* **1998**, *120*, 5098.

(11) Lee, S.-W.; Kim, H. S.; Beauchamp, J. L. *Journal of the American Chemical Society* **1998**, *120*, 3188.

(12) Jockusch, R. A.; Paech, K.; Williams, E. R. *Journal of Physical Chemistry A* **2000**, *104*, 3188.

(13) Olson, M. T.; Blank, P. S.; Sackett, D. L.; Yergey, A. L. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 367.