

DEPENDENCE OF HEAD-TO-TAIL CYCLIZATION ON PRIMARY STRUCTURES OF
PEPTIDES IN COLLISION-INDUCED DISSOCIATION

By

XIAN CHEN

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2010

© 2010 Xian Chen

To my parents and Yilin

ACKNOWLEDGMENTS

Throughout my graduate study at the University of Florida, I have benefited a lot from the help from a number of people. First and foremost, I would like to express my deep gratitude to my advisor, Dr. Nicolas Polfer, for his guidance, support, and patience over the past three years. He encouraged and supported me to present our work at the national conferences. I also thank him for giving me the opportunity to travel to the FOM institute in The Netherlands to carry out the IRMPD experiments and to the University of Amsterdam for peptide synthesis. It was a great experience in my life. I want to acknowledge my committee members, Drs. David Powell, Richard Yost, Kirk Schanze, and David Barber for helpful suggestions. I also appreciate Dr. John Eyler's suggestions given for our research and my practice talks for conferences and seminars.

I would like to thank all the people with whom I collaborated because this dissertation will not be finished without their kind help. I thank Dr. David Powell for generously allowing us to use his FTICR mass spectrometer, with which we performed all the gas-phase H/D exchange experiments. He is always very patient for all my questions. Drs. Jos Oomens and Jeff Steill at FOM institute are greatly acknowledged for their help on the IRMPD experiments and helpful discussions. I want to thank Dr. Britta Redlich for arranging the guesthouse for me while I was working at the FOM institute. I also want to thank Dr. Jan van Maarseveen for welcoming me to carry some peptide synthesis experiments in his lab. Jochem Rutters is acknowledged for teaching me how to synthesize cyclic peptides in solution and answering my questions. My third project will not be accomplished without their help. I would like to acknowledge Mr. Alfred Chuang from the ICBR at University of Florida. He helped us synthesize and

characterize some peptides and generously shared his experience on peptide synthesis with me.

My colleagues in the Polfer group are acknowledged for the discussion and the suggestions on my research. They also helped me preparing my every presentation. I would like to thank Yu Long for doing the calculations to generate the theoretical spectra for the first project. Some former Eyler group members provided a lot of help to me. I thank Dr. Julia Rummel for teaching me how to use the FTICR instrument, generously sharing her experience, and being patient for my questions. I also want to thank Dr. Sarah Stefan and Dr. Michelle Sweeney for offering help and suggestions.

Finally, I want to thank my family. I thank my parents, Shanfeng Chen and Suyun Liao, for their love and support through these years. They are always very supportive for the decisions I made, including studying abroad for a PhD degree. I also want to acknowledge my husband, Yilin Meng, for being a wonderful friend and companion. It is his love, patience, support and encouragement makes this dissertation possible.

TABLE OF CONTENTS

	<u>page</u>
ACKNOWLEDGMENTS.....	4
LIST OF TABLES.....	9
LIST OF FIGURES.....	10
ABSTRACT	15
CHAPTER	
1 INTRODUCTION	17
Peptide Chemistry	17
Amino Acids	17
Peptide Synthesis.....	18
Fmoc solid-phase peptide synthesis	18
Synthesis of macrocyclic peptides	20
Protein Identification	21
Edman Degradation	22
Protein Identification with Mass Spectrometry	23
Peptide Fragmentation.....	24
Mass Spectrometry	24
Electrospray ionization.....	25
Tandem mass spectrometry.....	27
Fragmentation Methods.....	28
Mobile Proton Model	29
Structures of b Fragment Ions	30
Objective of This Research.....	32
2 INSTRUMENTATION AND METHODS.....	43
Fourier Transform Ion Cyclotron Resonance Mass Spectrometry	43
Principles.....	43
Ion cyclotron motion.....	43
Trapping motion	46
Magnetron motion	47
Apparatus	48
Magnet.....	48
Analyzer cell.....	49
Ultra-high vacuum system.....	49
Data system.....	50
Excitation and Detection.....	50
Mass Accuracy	51
Mass Resolution.....	52

	Space Charge Effects	54
	Infrared Multiple Photon Dissociation Spectroscopy	55
	Gas-Phase Hydrogen/Deuterium Exchange	58
3	INVESTIGATION OF THE INFLUENCE OF PEPTIDE B FRAGMENT IONS ON THE FORMATION OF MACROCYCLE STRUCTURES THROUGH HEAD-TO-TAIL CYCLIZATION IN THE CASE OF OLIGOGLYCINES	69
	Background.....	69
	Experimental.....	70
	Sample Preparation.....	70
	Mass Spectrometry and Hydrogen/Deuterium Exchange.....	71
	Mass Spectrometry and Infrared Photodissociation Spectroscopy.....	72
	Results and Discussion.....	73
	Scrambling in Isotope-Tagged Peptide	73
	Infrared Spectroscopy Results	74
	b ₂ ion.....	74
	b ₅ ion.....	76
	b ₈ ion.....	77
	Summary of IRMPD results.....	78
	Hydrogen-Deuterium Exchange (HDX) Experiments.....	78
	Kinetic Fitting of HDX Data.....	79
	Chemical Basis for Differences in HDX Rates.....	81
	Summary	82
4	STRUCTURAL ANALYSIS FOR LEU-ENKEPHALIN B2-B4 WITH INFRARED MULTIPLE-PHOTON DISSOCIATION SPECTROSCOPY AND GAS-PHASE HYDROGEN/DEUTERIUM EXCHANGE	96
	Background.....	96
	Experimental.....	97
	Sample Preparation.....	97
	Mass Spectrometry and Hydrogen/Deuterium Exchange (HDX).....	97
	Mass Spectrometry and Infrared Photodissociation Spectroscopy.....	98
	Computations	99
	Results and Discussion.....	100
	Infrared Spectroscopy Results	100
	Hydrogen-Deuterium Exchange (HDX) Experiments	103
	Kinetic Fitting of HDX Data.....	104
	Comparison with Oligoglycines Study	105
	Summary	107
5	DEPENDENCE OF HEAD-TO-TAIL CYCLIZATION ON PRIMARY STRUCTURE OF PEPTIDES IN COLLISION-INDUCED DISSOCIATION	115
	Background.....	115
	Experimental.....	116

Materials.....	116
Preparation of Peptides.....	117
Fmoc synthesis of linear peptides.....	117
Synthesis of head-to-tail cyclic peptides.....	117
Purification of synthesized head-to-tail cyclic peptides.....	118
Mass Spectrometry.....	118
Results and Discussion.....	119
Commercial Cyclo(QWFGLM) as Direct Reference.....	119
IRMPD of cyclo(QWFGLM) and analogs.....	119
HDX of cyclo(QWFGLM) and analogs.....	120
The Effect of Proline on Head-to-Tail Cyclization.....	122
Charaterization of synthetic peptides.....	122
IRMPD results.....	122
The Effect of Glutamine Side Chain on the Head-to-Tail Cyclization.....	123
Summary.....	125
6 CONCLUSIONS AND FUTURE DIRECTIONS.....	134
LIST OF REFERENCES.....	138
BIOGRAPHICAL SKETCH.....	148

LIST OF TABLES

<u>Table</u>	<u>page</u>
1-1	Effect of yield per coupling step on final product yields. ³ 39
1-2	Sources for MS-based protein identification tools. ⁷³ 39
2-1	Proton affinities of common amino acids. Numbers are adapted from Ref. 129. 67
3-1	Mass-to-charge ratios of all b ions of interested. 85
3-2	Photofragments of b ₂ , protonated cyclo(Gly-Gly) and b ₅ 85
3-3	Kinetic fitting results for the ln[d ₀ /Σd _n] plots vs. H/D exchange time for the oligoglycine fragments b ₂ -b ₈ 94
4-1	Energies for the lowest-energy conformers of each chemical structure of b ₂ . The electronic energy for each conformer at the MP2/6-31G+(d,p) level was corrected for the zero-point energy (ZPE) derived at the B3LYP/6-31G+(d,p) level. 109
4-2	Exchange rates and relative abundances of “fast” and “slow”-exchanging structures for the Leu-enkephalin fragments b ₂ -b ₄ 114
4-3	Kinetic fitting results for the ln[d ₀ /Σd _n] plots vs. H/D exchange time for the Leu-enkephalin fragments b ₂ -b ₄ 114
5-1	Mass-to-charge ratios (<i>m/z</i>) of the ions of interest in this project..... 126

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
1-1 The chemical structure of an alpha-amino acid. (A) An alpha-amino acid which has one amine group, one carboxylic group and one distinct functional group (R) attached to one alpha carbon atom; (B) Fischer projection formulas for an L-amino acid; (C) Fischer projection formulas for a D-amino acid; (D) Stereo-representations for an L-amino acid; and (E) Stereo-representations for a D-amino acid.	34
1-2 A dipeptide is formed through a condensation reaction. Peptide bond is bracketed in red, and the N- and C-termini are indicated in blue and purple, respectively.....	35
1-3 The twenty amino acids found in proteins. Their names, three-letter and one-letter abbreviations are given below each structure.....	36
1-4 The general procedure of SPPS.....	37
1-5 Mechanism of Edman degradation.	38
1-6 The general procedure of MS-based protein analysis. Figure is adapted from Ref 10 with permission.	38
1-7 Scheme of ESI source. Reproduced with permission from John Wiley and Sons, Inc. from Cech, N.B.; Enke, C.G. <i>Mass Spectrom. Rev.</i> 2001, 20, 364....	40
1-8 Comparison of tandem mass spectrometry in space and in time. Figure is modified from Ref 20.	40
1-9 Nomenclature of common peptide fragment ion types for a protonated pentapeptide.....	41
1-10 Structures of b ions that have been proposed.	41
1-11 Sequence scrambling of a b ₅ ion generated from a hexapeptide. Figure modified with Harrison et al. <i>J. Am. Chem. Soc.</i> 2006, 128, 10364.....	42
2-1 Ion cyclotron motion. A spatially uniform magnetic field has a direction that is perpendicular to and going into the plane of the paper and an ion is moving in the plane of the paper. The moving path of the ion is a circle resulting from the magnetic Lorentz force. Positive and negative ions travel in opposite directions.	61
2-2 Ion path resulting from the combination of cyclotron, tapping, and magnetron motions. Figure reproduced from Ref.80 with permission.	61

2-3	Schematic diagram of Bruker APEXII 4.7T FTICR instrument.	62
2-4	FTICR-MS performance as a function of magnetic field strength B_0 . Figure reproduced from Ref. 80 with permission.	62
2-5	Schematic presentation of a cubic (left) and a cylindrical (right) FTICR-MS analyzer cell. Both types of cell have six plates, and each pair functions for excitation, trapping, and detection.	63
2-6	Schematic showing cross section of cylindrical ICR cell. Ions are excited to a larger radius (shown in blue) by applying an rf potential to the excitation plates. The motion of the ion packet is then detected on the detection plates in the form of an image current.	63
2-7	Schematic showing how time-domain data is converted to a frequency-domain spectrum, followed by conversion to a mass spectrum. Figure taken from Ref 91 with permission of The Royal Society of Chemistry. http://dx.doi.org/10.1039/b403880k	64
2-8	Demonstration of the effect of dataset size on resolution. The three mass spectra correspond to different acquisition dataset sizes for the same sample. Figure reproduced from Ref 91 with permission of The Royal Society of Chemistry. http://dx.doi.org/10.1039/b403880k	65
2-9	The dependence of resolution on both measured m/z and the lowest m/z cutoff. Figure reproduced from Ref 91 with permission of The Royal Society of Chemistry. http://dx.doi.org/10.1039/b403880k	65
2-10	Schematic presentation of IRMPD mechanism.	66
2-11	Schematic representation of the laboratory-constructed FTICR instrument equipped with an ESI source coupled to FELIX. Figure taken from Ref 104 with permission.	66
2-12	Gas-phase H/D exchange mechanisms proposed by Beauchamp and co-workers. Figure is adapted from Ref 127 with permission.	68
3-1	HDX mass spectrum with HDX products labeled.	85
3-2	Inserts from the nozzle-skimmer CID spectrum of octaglycine labeled with a ^{13}C -Gly as the N-terminal glycine residue, showing the b isotope distributions for (A) b_4 , (B) b_5 , (C) b_6 and (D) b_7 . The ^{13}C -labeled b_n peaks denote b ions that incorporate the ^{13}C -Gly label, whereas ^{12}C -only b_n peaks are entirely composed of ^{12}C -Gly residues.	86
3-3	The cartoon mechanism rationalizes the appearance of the ^{12}C -only b_7 peak. ...	86

3-4	IRMPD spectrum of the b_2 fragment generated from Gly-Gly-Gly, compared to computed spectra for (A) diketopiperazine structure protonated on a carbonyl O, (B) <i>oxazolone</i> structure protonated on the <i>oxazolone</i> ring N, and (C) <i>oxazolone</i> structure protonated on the N-terminus.....	87
3-5	IRMPD spectrum of protonated cyclo(Gly-Gly), compared to calculated spectra for (A) <i>diketopiperazine</i> with proton pointing to CH_2 group and (B) <i>diketopiperazine</i> with proton pointing to the amide N-H group.	88
3-6	Overlaid IRMPD spectra of protonated cyclo(Gly-Gly) and b_2 from triglycine.....	89
3-7	Possible structures for b_5 -G ₈ considered in theoretical study. (A) <i>macrocycle</i> structure protonated on backbone carbonyl, (B) <i>oxazolone</i> structure protonated on N-terminus, and (C) <i>oxazolone</i> structure protonated on <i>oxazolone</i> ring N.....	89
3-8	IRMPD spectrum of b_5 generated from octaglycine, compared to the two lowest-energy conformers for the various chemical structures: (A) <i>macrocycle</i> structure protonated on backbone carbonyl, (B) <i>oxazolone</i> structure protonated on N-terminus, (C) <i>oxazolone</i> structure protonated on <i>oxazolone</i> ring N. The chemical diagnostic bands and the relative energies to the lowest conformer are indicated.....	90
3-9	Overlaid mid-IRMPD spectra of b_5 -G ₈ , b_5 -G ₅ and b_8 -G ₈ . The chemically diagnostic modes are indicated.	91
3-10	Representative H/D exchange (10^{-8} Torr CH_3OD) mass spectra for (A) b_2 generated from tri-glycine, (B) b_3 generated from octa-glycine, (C) b_4 generated from penta-glycine, (D) b_5 generated from octa-glycine, (E) b_6 generated from octa-glycine, (F) b_7 generated from octa-glycine, and (G) b_8 generated from octa-glycine.	92
3-11	Kinetic fitting of the HDX results for glycine-based b fragment ions. (A) b_2 generated from triglycine, (B) b_3 generated from octaglycine, (C) b_4 generated from pentaglycine, (D) b_5 generated from octaglycine, (E) b_6 generated from octaglycine, (F) b_7 generated from octaglycine, and (G) b_8 generated from octaglycine. (H) Relative abundance of “slow”-exchanging structure as a function of b_n fragment size.....	93
3-12	The relative abundance of slow-exchanging structure as a function of b fragment size.	95
4-1	Chemical structure of Leu-enkephalin.	109
4-2	Experimental IRMPD spectrum of b_2 compared with theoretical spectra by computational studies. (A) IRMPD spectrum of the b_2 fragment generated from protonated Leu-enkephalin, (B) <i>diketopiperazine</i> structure protonated on a carbonyl O, (C) <i>oxazolone</i> structure protonated on the on the N-	

	terminus, and (D) <i>oxazolone</i> structure protonated on the <i>oxazolone</i> ring N. The scaling factor is 0.965. Corresponding structures are presented on the right. The site of proton attachment (red arrow) and <i>oxazolone</i> rings (black arrow) are indicated.....	110
4-3	Overlaid mid-IRMPD spectra of b_2 , b_3 and b_4 . The spectrum of b_4 is adapted from previous publication of Polfer at el. ³⁹ . The chemically diagnostic modes are indicated.....	111
4-4	H/D exchange (10^{-8} Torr CH_3OD) mass spectra for (A) b_2 , (B) b_3 , and (C) b_4 for different exchange times.	112
4-5	Structure of b_2 of Try-Gly of an <i>oxazolone</i> structure protonated at the N-terminus. Exchangeable hydrogens are shown in red.	112
4-6	Kinetic fitting of the HDX results for (A) b_2 , (B) b_3 , and (C) b_4	113
5-1	Overlaid IRMPD spectra of protonated cyclo(QWFGLM) and b_6 from QWFGLMPG and Ac-QWFGLMPG.....	126
5-2	Natural logarithm of relative d_0 depletion, $\ln[d_0/\Sigma d_n]$, as a function of time for (A) protonated cyclo(Gln-Trp-Phe-Gly-Leu-Met), and (B) b_6 generated from protonated linear QWFGLMPG.	127
5-3	Mass spectra of Ac-QWFGLMPG b_6 ion after exchanging with CH_3OD for (A) 0s and (B) 90s. No exchange is seen after 90s of exchange, indicating that such reaction is extremely slow.	127
5-4	Schematic presentation of an <i>oxazolone</i> exchanges with CH_3OD	128
5-5	LC/MS spectrum of linear Q(Trt)PFGLM after purification.....	128
5-6	LC/MS spectrum of aliquot taken from cyclization reaction solution. The top figure is the HPLC spectrum of the aliquot. The middle figure is the MS spectrum of the synthetic cyclo(QPFGLM) with a retention time of 6.14 min. The bottom figure is the MS spectrum of the synthetic cyclo(QPFGLM) ₂ with a retention time of 6.95 min.	129
5-7	IRMPD spectra of (A) QWFGLMPG b_6 , (A) QPFGLMPG b_6 , and (A) QWPFGLMPG b_7 . In the inserts, in the intensities are magnified by 5x in the range of $1750\text{-}1930\text{ cm}^{-1}$	130
5-8	Overlaid IRMPD spectra of cyclo(QPFGLM) and QPFGLMPG b_6 . The <i>oxazolone</i> C=O stretch band is pointed by the red arrow.	131
5-9	The structures of the b_6 ions from four peptides of QWFGLG system. The possible attack from the N-terminus is labeled in blue, and the one from the glutamine side-chain is labeled in orange.....	132

5-10 IRMPD spectra of b_6 ions generated from (A) QWFGLGPG, (B) Ac-QWFGLGPG, (C) Q(N- γ -ethyl)WFGLGPG, and (D) Ac-Q(N- γ -ethyl)WFGLGPG. The range for *oxazo/one* bands is highlighted in pink..... 133

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

DEPENDENCE OF HEAD-TO-TAIL CYCLIZATION ON PRIMARY STRUCTURES OF
PEPTIDES IN COLLISION-INDUCED DISSOCIATION

By

Xian Chen

December 2010

Chair: Nicolas C. Polfer
Major: Chemistry

Collision-induced dissociation (CID) of peptides to deduce their sequences is the key technology in identifying proteins by mass spectrometry. However, during activation the linear peptide structures can undergo a head-to-tail cyclization reaction, where the N- and C-terminus of a peptide “b” fragment fuse into a *macrocycle* structure. When this *macrocycle* structure opens up at a different site than where it was originally formed, a scrambling of the sequence information will occur.

There are few techniques that yield direct structural information on the minute quantities of gas-phase ions inside mass spectrometers, such as infrared spectroscopy, ion mobility, and hydrogen/deuterium exchange (HDX). IR spectroscopy confirms chemical structures based on diagnostic vibrations, however it is difficult to obtain relative abundances. Conversely, ion mobility can yield relative abundances of structures, but the structural interpretation is often ambiguous. HDX had so far not produced either information. In our research, we combined infrared multiple photon dissociation (IRMPD) spectroscopy and gas-phase HDX to structurally characterize and quantify *macrocycle*, as well as *oxazolone*, structures for a series of “b” ions from selected peptide systems.

For a series of glycine-based “b” fragment ions, this approach shows a size-dependency, where smaller “b” fragments exclusively adopt *oxazolones*, while larger fragments display a mixture of *oxazolones* and *macrocycles*. The results are consistent with the finding from HDX, where smaller “b” fragments display a single HDX rate, whereas larger fragments show two distinct rates. Relative abundances of *oxazolones* and *macrocycles* are approximated from HDX kinetic fitting. Similar trends are found in the peptide Leucine-enkephalin (YGGFL).

The correlation between peptide sequence and propensity for *macrocycle* formation for the b₆ motif QWFGLM is investigated. The IRMPD spectrum for the b₆ fragment of QWFGLMPG is nearly identical to that for protonated cyclo(QWFGLM), which confirms the exclusive presence of *macrocycle* structures for b₆ (from QWFGLMPG). The incorporation of a proline in for instance QPFGLMPG is found to reduce the propensity of the formation of *macrocycle* structures in the corresponding b₆ fragment. A systematic chemical protection study was performed to the QWFGLG system. IRMPD spectra of b₆ ions generated from QWFGLGPG, Ac-QWFGLGPG, Ac-Q(N-ethyl)WFGLGPG, and Q(N-ethyl)WFGLGPG indicate that no cyclization from the glutamine side chain occurs.

CHAPTER 1 INTRODUCTION

Peptide Chemistry

The four main groups of molecules that play crucial role in cellular function are proteins, nucleic acids, lipids, and carbohydrates. A protein is a molecule that has more than 50 amino acid residues, while a peptide is composed of less than 50 residues.

Amino Acids

Amino acids are the building blocks of all peptides and proteins. The structure of a typical amino acid is illustrated in Figure 1-1A. Bonded to the alpha carbon are one amino group, one carboxylic group, one hydrogen, and one distinct functional group R on the side chain. It is the side chain group R that differs for each amino acid. An alpha-amino acid is formed when all these three functional group are attached to one alpha carbon atom. Except for the glycine, which has a hydrogen atom at the side chain, the four groups attached to the alpha carbon atom are different, the carbon atom then allows for stereoisomerism. As a result, amino acids are chiral. There are two enantiomers: L- and D- amino acids based on the position of the amino group and hydrogen. Amino acids having amino group locating at the left hand side of the chiral center atom have L-configuration, and D-amino acid are those with a hydrogen atom on the left hand side, as illustrated in Figure 1-1 B-E. All amino acids found in proteins have the L-configuration.

The amine group from one amino acid can react with the carboxylic group from another amino acid, eliminating a water molecule and forming a molecule where the two amino acid residues link via a covalent amide bond, which is commonly called the peptide bond in peptides and proteins. Peptides are formed through a series of such

condensation reactions. Figure 1-2 shows how a dipeptide is formed through the loss of a water molecule.

There are twenty amino acids commonly found in human proteins, and the vast majority of proteins in nature are composed of these twenty amino acids (with the exception of some fungi). The structures and abbreviations of these common amino acids are shown in Figure 1-3. Commonly, amino acids are classified into four categories based on the polarity of the side chains: nonpolar or hydrophobic amino acids, neutral (uncharged) but polar amino acids, acidic amino acids, and basic amino acids.¹

Peptide Synthesis

As mentioned above, peptides are formed through the polymerization of amino acids. However, during synthesis it is possible that internal reaction will occur and result in unintended products. For example, to synthesize a dipeptide of Ala-Gly, simply mixing the two amino acids Ala and Gly together will yield not just the target peptide Ala-Gly, but also by-products Ala-Ala and Gly-Gly. In order to direct synthesis, it is important to have protecting groups. Nowadays, solid-phase peptide synthesis is the most popular method.

Fmoc solid-phase peptide synthesis

Solid-phase peptide synthesis (SPPS) was developed by Robert Bruce Merrifield,² whose contribution to the solid-phase chemical synthesis was recognized by the 1984 Nobel Prize in Chemistry. Peptides are synthesized from the C-terminus, or carboxyl group side. The general procedures of SPPS are shown in Figure 1-4. The assembly of a peptide chain involves repeating cycles of “deprotection–washing–coupling–washing”. The C-terminal amino acid residue of the desired peptide is attached to an insoluble

porous solid bead (resin) via its carboxyl group. Functional groups on amino acid side chains are blocked by permanent protecting groups that cannot be removed under the conditions during assembly. The temporary protecting group on the alpha-amino group during the initial resin loading is first removed, and this is called the “deprotection”. The “coupling” step is achieved by introducing an excess of second amino acid, whose amino group is temporarily protected while its carboxyl group is activated, along with coupling reagents. The two amino acids are then linked via an amide bond, resulting in a dipeptide that is protected on the N-terminus. Notably, the N-terminal protected amino acids and coupling reagents are all soluble, whereas the peptide is immobilized on the resin, thus, after each coupling step, excess reagents and soluble by-products can be removed by washing. Following the washing step, the protecting group on the amino group of the dipeptide is then removed, prior to addition of a third amino acid residue. This process is repeated until the desired peptide sequence is attained. The deprotection of the final N-terminus protecting group is done prior to the final step, “cleavage”, where the peptide is released from the resin and the side-chain protecting groups are removed. Typically, the resins and side-chain protecting groups are chosen so that they can be removed under the same conditions.

The major limitation in SPPS is the final yield. To obtain pure and high-yield product requires extremely high yield in every step. The effect of chemical efficiency of each step on the yield of the final product is shown in Table 1-1. The by-products, resulting from incomplete coupling, often have very similar properties as the target peptide, hence making the purification difficult.³

The two major SPPS methods involve *Fmoc* and *Boc* chemistries, and differ from each other by the N-terminal protecting groups and the resins. The *Boc* method name is derived from the protecting group on the amine, *tert*-butoxycarbonyl (Boc), the removal of which necessitates trifluoroacetic acid (TFA). The resin used in the *Boc* method is hoxymethylphenylacetamidoethyl polystyrene (PAM) resin, and to release the peptide requires the use of hydrogen fluoride (HF). The *Boc* method can give very high yield and thus is powerful in synthesis of larger peptides and proteins.⁴ However, the need to use toxic HF requires special equipment and caution.

Fmoc chemistry is a more popular method because it is safe and easy to operate. Similarly to the *Boc* method, the *Fmoc* method is named after its amino protecting group 9-fluorenylmethoxycarbonyl (Fmoc). The deprotecting reagent used is 20–50% piperidine in N,N-dimethylformamide (DMF). Coupling is achieved by addition of Fmoc protected amino acids and coupling reagents diisopropylethylamine (DIPEA) and 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU). Cleavage of the peptide from the resin and removal of the protecting groups on side chains is done with 95% TFA.

Synthesis of macrocyclic peptides

Macrocyclic peptides are cyclic polypeptides whose amino and carboxyl termini are linked together via a peptide bond in a circular chain. Cyclic peptides often play important roles in biological processes, such as for instance gramicidin S, which is an antibiotic agent. This explains the interest for synthetic approaches in making cyclic peptides. Cyclization of linear peptides with side-chain protected in the solution phase is the widely-used method. Synthesis of intramolecular head-to-tail cyclic peptides is normally performed in dilute conditions (10^{-3} – 10^{-4} M)⁵ to minimize the competing

polymerization of linear peptides. Given the dilute solution conditions and competing reactions, cyclization reactions are typically slow and give low yield.

Protein Identification

Each protein is unique from others based on its distinctive sequence of amino acids in the polypeptide chain. Note that by convention, the amino acid sequence of a peptide or protein is read from the N-terminus of the polypeptide chain to the C-terminal end.¹ The primary structure of the protein also gives rise to higher-order structure, which account for its distinctive functions in biological system. The amino acid sequence is based on the genetic information encoded by DNA. Given the human genome project, and other genome sequencing projects, a wealth of information is now available on expected protein sequences, and this information is stored in DNA/protein databases.

Nonetheless, information on the genetic make-up of cells does not directly correlate with biological function. Instead, the primary actors in cells are proteins, the expression of which is steered by highly in complex cellular regulatory networks.⁶ The central theme in *proteomics* is to identify the role of each protein in these processes. For that purpose, it is essential that proteins can be identified from minute and highly heterogeneous biological samples. Since the identity can be established from its sequence, analytical techniques are required that can confirm the primary structures of proteins.

Twenty years ago, *Edman degradation*, an automated, stepwise chemical degradation, was used to digest and identify the amino acid sequence of proteins. The method has been largely replaced by mass spectrometry. Since the development of two soft ionization methods electrospray ionization (ESI)⁷ and matrix-assisted laser desorption/ionization (MALDI)^{8,9} in the late 1980s, mass spectrometry (MS) has become

a key approach in the field of proteomics. These methods solved the difficult problem of generating ions from large, nonvolatile analytes such as proteins and peptides without inducing analyte fragmentation. Compared to Edman degradation, MS is much more sensitive, can analyze peptides in seconds, does not require proteins or peptides to be purified, and has no problem in identifying blocked or modified proteins.¹⁰ Tandem MS (MS/MS) involves mass isolating a peptide of interest and subjecting it to fragmentation through collisions. The mass-to-charge ratios (m/z) of the fragment ions are then employed to derive information on the amino acid sequence of the peptide.

Edman Degradation

Edman degradation, developed by Peer Edman, was the key technique in protein and peptide sequencing until the 1990's. In this method, amino acids are chemically cleaved in a stepwise way from the amino terminus of the proteins, followed by identification. The mechanism of Edman degradation is illustrated in Figure 1-5. In Edman degradation, phenylisothiocyanate (PITC) reacts with the alpha amino group at the N-terminus of a peptide, forming a phenylthiocarbamyl (PTC) adduct. Under anhydrous acidic conditions the N-terminal amino acid residue is then cleaved from the peptide chain, resulting in a heterocyclic derivative through the attack of the sulfur of the PTC adduct on the carbonyl component of the first peptide bond. The cleaved amino acid derivative is separated from the residual peptide by extraction with an organic solvent, and then gets identified by ultraviolet (UV) absorbance spectroscopy. The remaining peptide chain is then subjected to further cycles of coupling of PITC, cleavage and identification.¹¹

Edman degradation is capable of reasonable sensitivity (i.e., 10-100 picomoles). However, its major limitation is the need of a free amino group to couple PITC. Hence,

peptides cannot be identified if they are acetylated. In addition to needing purified samples, one of the main drawbacks of Edman degradation is the time required for analysis, which can stretch from hours to days for a single peptide/protein.

Protein Identification with Mass Spectrometry

Mass spectrometry (MS) is currently the most popular method for peptide/protein identification due to its higher sensitivity (i.e., attomol, 10^{-15} mol) and that the identification of a single peptide can be done in the timeframe of seconds. Furthermore, compared to Edman degradation, MS does not require purification of peptides or proteins, and can identify modified proteins (e.g., N-terminal acetylated proteins).

The general procedure for MS-based proteomics is illustrated in Figure 1-6. A protein is first prepared from a biological sample, by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The gel bands correspond to individual proteins and can be cut out using a scalpel. The protein is digested into peptides using an enzyme (e.g. trypsin). The extracted peptides are then ionized by ESI or MALDI for mass analysis. In order to reduce sample complexity, the peptides may be further separated by high-performance liquid chromatography (HPLC) prior to ionization. Hyphenated HPLC separation and ESI can in fact be carried out on-line with mass spectrometric analysis. In the mass spectrometer, the mass of the ionized peptide of interest is first measured. The peptide is then isolated and subjected to fragmentation, where the ionized peptide is broken into smaller pieces (i.e., fragments). The mass spectra of these fragment ions contain sequence-specific information. Lastly, the peptide-sequencing data that are obtained from the mass spectra are searched against protein databases using one of a number of database-searching programs,¹⁰ which are listed in Table 1-2.

The method described above is also called the “bottom-up” approach, as the peptide sequence information from an enzymatic digestion is used to identify a whole protein. One difficulty with this method is that the complete protein sequence can only be obtained when the peptide coverage is also complete; often, this is not possible, as some peptides have low ionization efficiencies.^{12,13} An alternative approach for protein identification involves ionizing and fragmenting the entire protein in the mass spectrometer, and is hence referred to as the “top-down” method.¹⁴⁻¹⁸ Given the highly complex mass spectra in this approach, this requires the use of high resolution and high mass accuracy mass spectrometers, such as Fourier transformed ion cyclotron resonance (FTICR) mass spectrometers. Note that the theory of FTICR-MS will be discussed in Chapter 2. The “bottom-up” approach remains the most popular method in protein identification by mass spectrometry, as it is compatible with all mass analyzers. Moreover, the “top-down” approach is limited to proteins of moderate size.

Peptide Fragmentation

As described above, in the approach of “bottom-up”, to get the sequence information of a peptide, the peptide has to undergo fragmentation, where it is broken up into smaller pieces. Peptide fragmentation in the mass spectrometer is the crucial step in peptide sequencing. In this section, the basic concept and theory of mass spectrometry techniques will be introduced, and techniques for peptide fragmentation will be discussed.

Mass Spectrometry

Among all the analytical techniques, mass spectrometry is the most widely applicable. This technique can be used to get the elemental composition of samples of matter, to confirm the structures of inorganic, organic and biological molecules, to do

both qualitative and quantitative analysis of complex mixtures, and to get isotopic ratios of atoms in samples.¹⁹ A mass spectrometer has three major parts: an ion source, a mass analyzer and a detector. In a typical analysis with a mass spectrometer, analytes have to be first transformed into gaseous ions, which is done at the ion source. The generated analyte gas-phase ions are then transferred through a series of ion optics to the mass analyzer, where the mass-to-charge ratios (m/z) are measured. A detector is used to count the ions from the mass analyzer and to measure their abundance.

Electrospray ionization

A number of ionization techniques have been developed and they are adopted based on different research needs. These techniques can be classified as “hard” and “soft” ionization methods. With hard ionization method, samples are not just get ionized, but also dissociated during the ionization process. Soft methods produce little or no fragmentation during the ionization process.

One of the main examples as a hard ionization technique is electron ionization (EI), which is particularly suitable in the analysis of organic compounds. In this method, electrons are emitted from a heated filament by thermionic emission. The electron beam is intersected with gaseous molecules of the sample of interest. In positive ion mode, the incident electron ejects an electron from the molecule, generating an ion in the process. The number of ions I produced per unit time can be calculated using the equation²⁰:

$$I = NpiV \tag{1-1}$$

where N is a constant proportionality coefficient, p is the pressure, i is the electron current, and V is the volume.

In terms of peptide/protein analysis, EI is not a useful technique. EI causes extensive dissociation of amino acids. Further, the vapor pressure of peptides/proteins is much too low to bring sufficient densities into the gas phase. In fact, heating the sample to achieve higher vapor pressures results in dissociation of the protein, as opposed to their transfer into the vapor phase. The ionization techniques that have revolutionized biological mass spectrometry of macromolecules are electrospray ionization (ESI)⁷ and matrix-assisted laser desorption/ionization (MALDI)^{8,9}. Both of these are soft ionization techniques, as they are capable of transferring large molecules in the gas phase while leaving them intact. This dissertation will focus on ESI, since it is the only ionization technique employed throughout the research.

ESI is an atmospheric pressure ionization method. While the phenomenon of electrospray has been known for more than a century, Fenn and co-workers first coupled electrospray to a mass analyzer for the purpose of ionizing large molecules in late 1980's. Since then, the technique has been employed to routinely ionize a wide range of molecules, including lipids, carbohydrates, polymers, peptides, proteins, protein complexes and even entire viruses. The capability of ESI for transferring very large complexes into the gas phase is unparalleled, and no upper mass limit has been determined so far. ESI allows very high sensitivity to be reached and is easy to couple to separation techniques, such as high-performance liquid chromatography (HPLC) or capillary electrophoresis.²⁰

In ESI, gaseous ionized samples are produced by applying a high electric field to solution-phase molecules contained in droplets. Figure 1-7 illustrates the ESI process. A high voltage drop (2-5 kV) is applied between the capillary and the metal plate, which

is near the entrance to the mass spectrometer. The electric field induces a charge accumulation at the liquid surface located at the end of the capillary, which will break to form highly charged droplets.²⁰ An “onset voltage” has to be achieved in order to generate the droplets, and the onset voltage depends on the surface tension of the solvent. At the end of the capillary, when the onset voltage is reached, the pressure is higher than the surface tension, then the shape of the flow will change to a Taylor cone and small droplets are released. These droplets contain solvent molecules, sample molecules, and charges. As they drift toward the mass spectrometer entrance, solvent evaporation causes the charged droplets to undergo Rayleigh explosions, since the surface tension of the droplet can no longer withstand the Coulombic repulsion. This solvent evaporation is aided by a heated nebulizing gas (normally inert gas, e.g. N₂).^{21,22} The process of Rayleigh explosions is repeated until the analyte ion is brought into the gas phase. There is still contention on the exact mechanism for the final step, when the ion is brought into the vapor phase. It is generally accepted that smaller ions may be capable to escape the droplet from the surface, while larger ions require a boiling off of all solvent molecules.^{23,24}

Tandem mass spectrometry

The charged analyte produced with ESI enters the mass spectrometer through the inlet, and its mass-to-charge ratio is then measured in the mass analyzer. To solve the problem of coupling an atmospheric pressure ionization source to a mass analyzer, which must be operated in vacuum conditions (i.e., $<10^{-5}$ mbar), the inlet has to restrict air flow into the mass spectrometer, and multiple pumping stages are required. For proteomics measurements, mass spectrometer can be used either to measure simply the molecular mass of a polypeptide, or to determine additional structural features

including the amino acid sequence or the site of attachment and type of post-translational modifications.⁶ For the first purpose, a single stage mass spectrometer is used. In the second case, tandem mass spectrometry has to be utilized.

Tandem mass spectrometry (MS/MS) is a method that involves at least two stages of mass analysis. In such an experiment, after determining the mass of a precursor ion, the ion is isolated and then subjected to fragmentation. The analysis of the fragment ions is essential in structural elucidation. There are two types of instruments that can carry out MS/MS measurements: tandem mass spectrometers in *space* or in *time*. A tandem mass spectrometer in space has two distinct mass analyzers coupling together, while the one in time is done via performing a sequence of events inside one ion storage device. The comparison of a product ion scan performed by a space-based and a time-based instrument is shown in Figure 1-8.

In our research, we use Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS). Its mass analyzer is the ICR cell, which is an ion-trapping device, and hence, the tandem mass spectrometry experiment is performed in time.

Fragmentation Methods

To get the sequence of a peptide, one has to break the peptide into fragment ions. Generally, this is done for protonated peptide cations. The nomenclature²⁵ for peptide fragment products is shown in Figure 1-9. Cleavage of a chemical bond along the backbone will result in either **a**, **b**, or **c** ions if the charge is retained on the N-terminal side, or **x**, **y**, and **z** ions if the charge is on the C-terminal side. The subscript indicates the number of amino acid residues present in the ion, counting from either the N- or C-terminus. For the pentapeptide in Figure 1-10, when the peptide bond that links the third and the fourth residue dissociates, either a b_3 or y_2 ion will be produced.

There are several methods to fragment peptides, including electron capture dissociation (ECD),²⁶⁻²⁹ blackbody infrared radiative dissociation (BIRD),³⁰ surface-induced dissociation (SID),³¹ electron transfer dissociation (ETD),³² collision-induced dissociation (CID),^{33,34} and infrared multiple photon dissociation (IRMPD).³⁵

The most widely used ion activation and dissociation process is collision-induced dissociation (CID), where the mass-selected protonated peptide collides with neutral background gas molecules (typically He, N₂ and Ar) in a collision cell or ion trap. Generally, the masses of the CID product ions are expected to reflect the amino acid sequence of the peptide, and therefore be valuable for sequencing.^{36,37} In low energy CID, bond breakage mostly occurs at the peptide (i.e., amide) bonds, which are often the lowest-energy pathways, leading to **b** and **y** ions.

In CID, the ion is vibrationally excited. Vibrational excitation of ions can also be achieved by absorption of infrared photons from an infrared laser. In infrared multiple-photon dissociation (IRMPD) similar fragmentation is observed as in CID. In Chapter 2, the technique of IRMPD spectroscopy will be introduced, which is a method to record the infrared spectra of ions. This technique is based on the premise that the ion absorbs a number of photon (i.e., tens to hundreds) when a vibrational mode of the molecule is resonant with the laser wavelength.^{38,39}

Mobile Proton Model

In the fragmentation of protonated peptides, the proton in fact plays an important role in the dissociation mechanism. The “mobile proton model”, first proposed by Wysocki and co-workers,⁴⁰ and Gaskell and collaborators,⁴¹ states that the proton migrates along the backbone to induce charge-site-initiated dissociation. In particular, the attachment of the proton to the amide NH is thought to weaken the amide bond, and

hence promote cleavage at that site. Basic amino acid side-chains (i.e., Arg, Lys, His) sequester the proton more tightly, and for such peptides higher activation energies are required. Theoretical research has made important contribution to the development of this theory. To evaluate the competing reaction pathways that take place, Paizs and Suhai suggested the “pathway in competition” model.⁴² Computations based on this model provide information about the competing fragmentation mechanisms after the liberation of the mobile protons on to the peptide backbone. In term of the different reaction mechanisms, experimental approaches are required to validate the chemical structures that are formed.⁴³

Structures of **b** Fragment Ions

The chemical structures of **b** ions had been the subject of a debate in the mass spectrometry community. b_1 ions are very instable in mass spectrometer, so it is normally not observed in MS spectra, whereas larger **b** ions are often abundant in the CID spectra. It was initially thought that **b** ions form *acylium* structures⁴⁴, as shown in Figure 1-10(A). However, acylium structures are not stable and are expected to spontaneously lose CO. In order to rationalize stable **b** ions, Harrison and co-workers suggested that *oxazolone* structures are formed, via nucleophilic attack from a backbone carbonyl (structure shown in Figure 1-10B).⁴⁵⁻⁵⁰ The first direct evidence for *oxazolone* structures was obtained by IRMPD spectroscopy by Polfer et al.³⁹ In their study, the IR spectrum of b_4 ions generated from Leu-enkephalin (YGGFL) showed that the *oxazolone* diagnostic C=O stretch bands were observed, whereas the diagnostic acylium C=O stretch mode was not seen.

A third possible **b** fragment structure was first proposed by Wesdemiotis and co-workers, involving a cyclization via the N-terminus, in a so-called *diketopiperazine* structure, as shown in Figure 1-10C. For smaller b_2 ions, the evidence so far has shown that diketopiperazine structures are rarely formed.⁵¹⁻⁵⁵ For larger **b** ions, however, there is increasing evidence that such head-to-tail *macrocycle* structures are possible. The presence of such *macrocycle* structures is particularly worrisome, since their presence can explain some of the sequence permutations that are observed in CID.⁵⁶⁻⁵⁸ With respect to peptide sequencing via MS/MS, it is possible to add substantial internal energy and the fragment ions of peptides can undergo rearrangement processes, leading to the formation of a *macrocycle* structure, as shown in Figure 1-10D.

Oxazolone structures can isomerize to the *macrocycle* structure. When the *macrocycle* structure is present, reopening of this *macrocycle* structure will lead to the scrambling of peptide sequence. The mechanism is proposed by Harrison and Paizs, as illustrated in Figure 1-11 for a b_5 ion made from protonated hexapeptide $R_1R_2R_3R_4R_5R_6$, where R_i is the name of an amino acid. The head-to-tail cyclization reaction results in a macrocyclic b_5 ion, and the reopening leads to *oxazolones* with original and permuted sequences. Further fragmentation of b_5 with original sequence will generate direct sequence ions, which contain information of the initial sequence. However, fragmentation of permuted *oxazolones* will lead to non-direct sequence ions, in which the original primary sequence information is lost. The phenomenon is problematic for MS-based peptide sequencing because the algorithms of current software for peptide sequencing do not take this into consideration, which might lead to false identification.

The phenomena of sequence scrambling have been observed in several peptide systems with MS/MS method. CID is performed to generate fragments from ions of interest and the masses of fragment ions are then collected and analyzed. Any internal neutral loss from an ion may suggest that a *macrocycle* structure is formed prior to the fragmentation, and may also give information about fragmentation pathways.⁵⁹⁻⁶²

Experimental methods for investigating **b** ion structures include IRMPD, isotope labeling, ion mobility spectrometry (IMS), and gas-phase hydrogen/deuterium exchange (HDX). Since it was shown that IRMPD spectroscopy can provide direct evidence for chemical moieties,⁶³ it has become a popular approach.^{39,51,53-55,64,65}

IMS is also a very useful technique that can distinguish isomers by their different collision cross sections in the drift tube. Polfer et al. reported a work done by ion mobility spectrometry (IMS) that showed further evidence of the presence of cyclic **b**₄ and **a**₄ ions of protonated YGGFL.⁶⁶ Gaskell and coworkers published a paper showing further evidence of cyclic **b**₅ ion of protonated YAGFL-NH₂ by IMS⁶⁷.

Gas-phase HDX has been employed to structurally characterize **b** ions.^{54,55,65,68-72} The premise of this approach rests on the fact that isomers of **b** ions are expected to exchange with the deuterating reagent (e.g. CH₃OD) at different rates. In other words, if multiple chemical structures are present, more than one HDX kinetic rate should be observed. The mechanism and theory of gas-phase HDX will be discussed in detail in Chapter 2.

Objective of This Research

This thesis describes a systematic study of the effect of primary structure on the formation of *oxazolone/macrocycle* **b** structures under CID conditions. All experiments were conducted on FTICR instrumentation, where the peptide of interest was generated

by ESI. IRMPD experiments were performed at the Free Electron Laser for Infrared eXperiments (FELIX) facility (Nieuwegein, The Netherlands) with the tunable free electron laser. HDX experiments were carried out at the University of Florida, by introducing CH₃OD vapor into ICR cell to react with isolated ions of interest.

Computations were conducted at the High-Performance Computing (HPC) center at the University of Florida.

In Chapter 2, the main experimental techniques employed here are introduced, namely FTICR mass spectrometry, IRMPD spectroscopy and HDX. In Chapter 3, a series of b ions generated by CID from oligoglycines were used to investigate the size effect on the formation of the *macrocycle* structure. Chapter 4 discusses a similar study for the peptide Leu-enkephalin (YGGFL). In Chapter 5, synthetic chemistry approaches are employed to create reference compounds for IRMPD spectroscopy and to investigate the effect of site mutations in the primary structure on the prevalence of *oxazolone/macrocycle* structures. Lastly, in Chapter 6, future directions are discussed that could lead to further insights in the dissociation chemistry of protonated peptides.

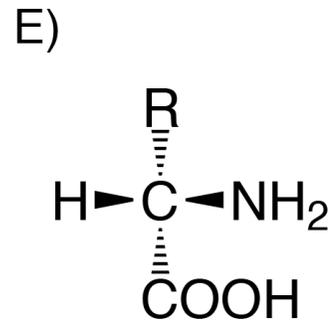
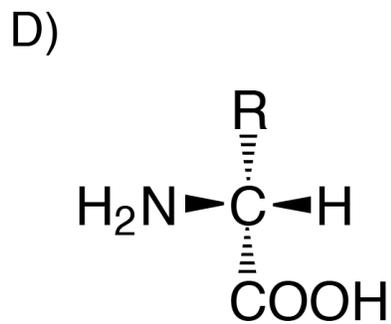
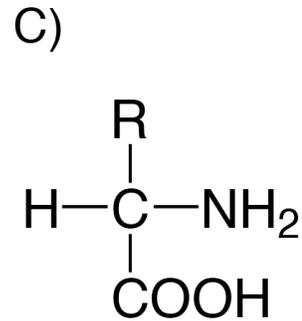
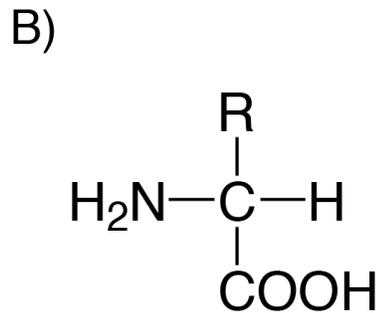
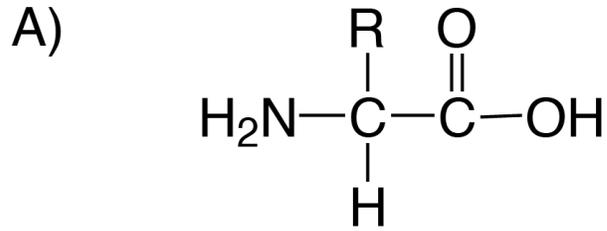


Figure 1-1. The chemical structure of an alpha-amino acid. (A) An alpha-amino acid which has one amine group, one carboxylic group and one distinct functional group (R) attached to one alpha carbon atom; (B) Fischer projection formulas for an L-amino acid; (C) Fischer projection formulas for a D-amino acid; (D) Stereo-representations for an L-amino acid; and (E) Stereo-representations for a D-amino acid.

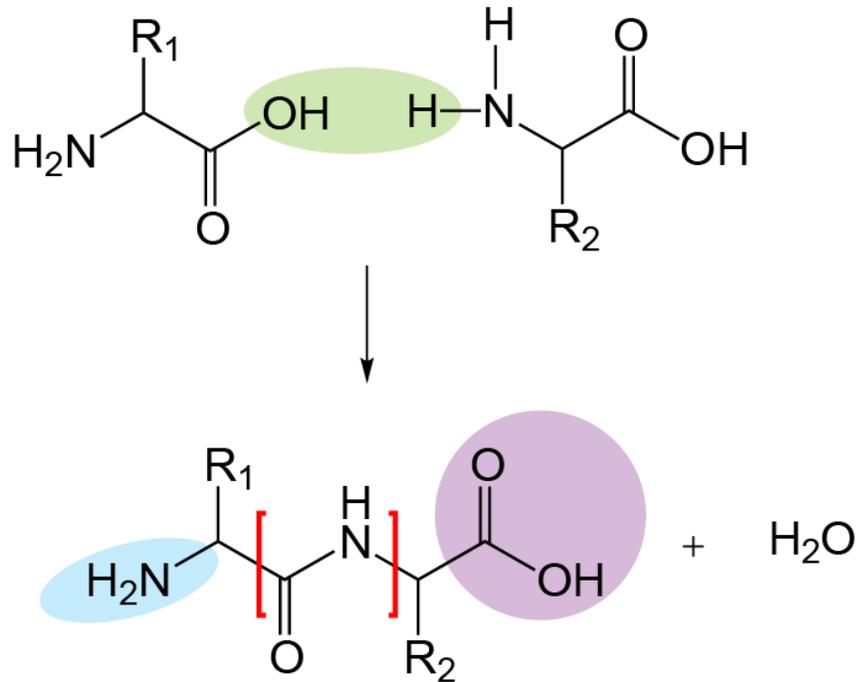
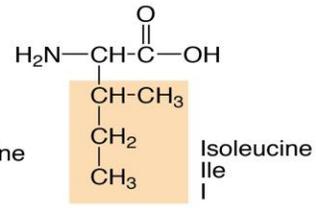
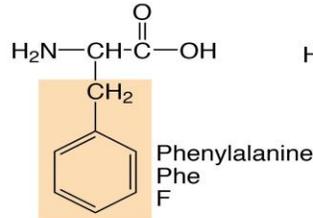
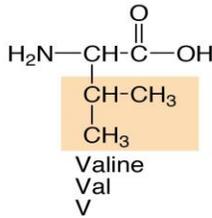
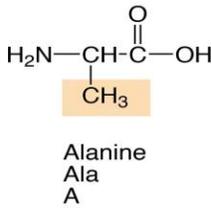
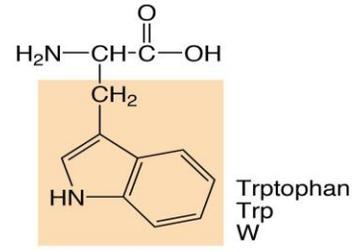
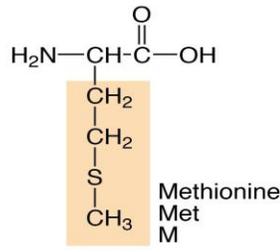
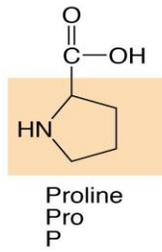
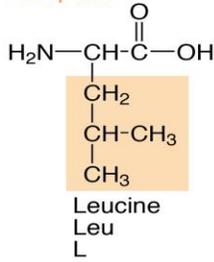
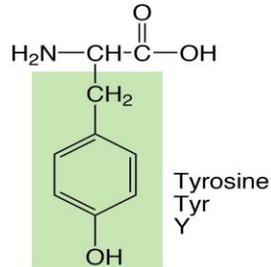
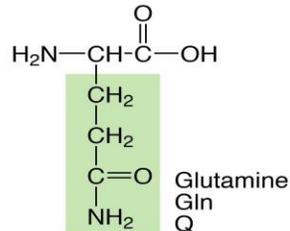
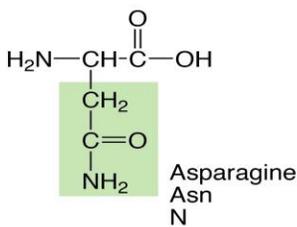
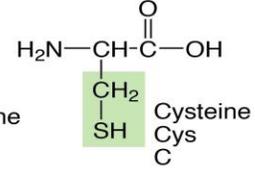
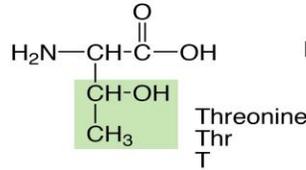
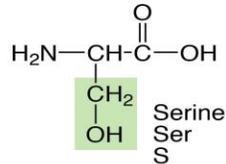
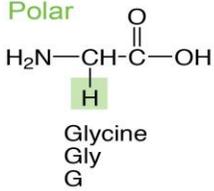


Figure 1-2. A dipeptide is formed through a condensation reaction. Peptide bond is bracketed in red, and the N- and C-termini are indicated in blue and purple, respectively.

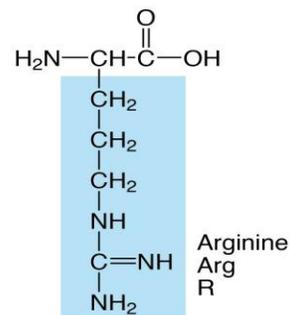
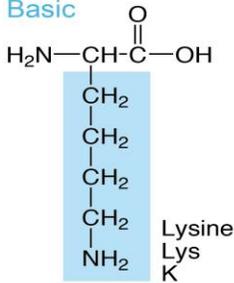
Nonpolar



Polar



Basic



Acidic

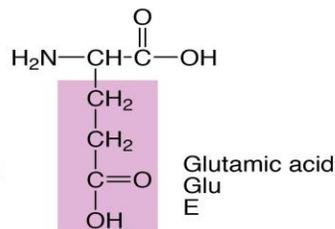
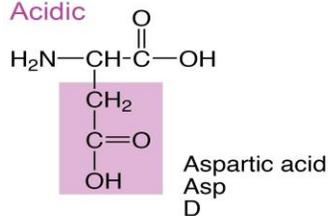


Figure 1-3. The twenty amino acids found in proteins. Their names, three-letter and one-letter abbreviations are given below each structure.

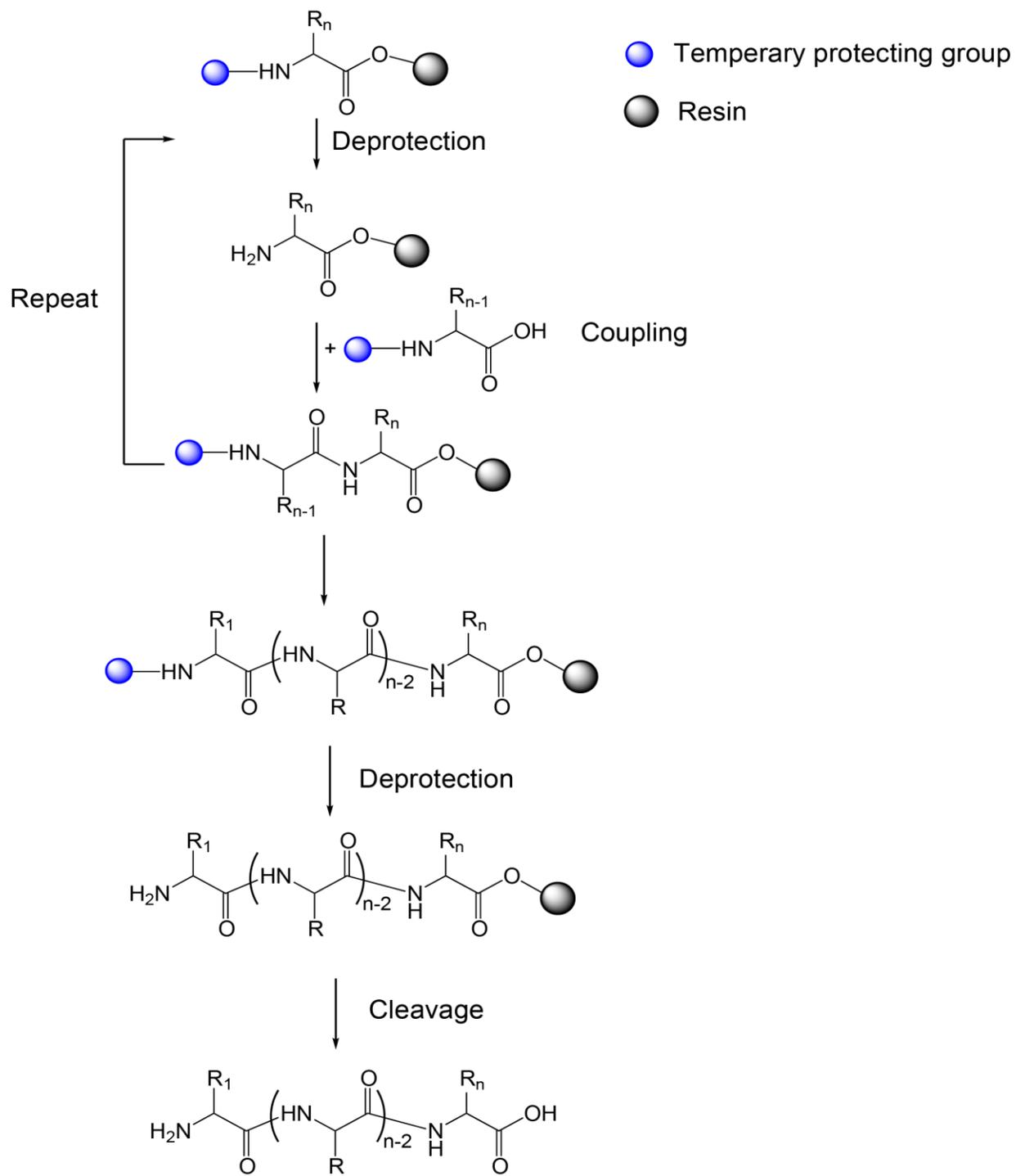


Figure 1-4. The general procedure of SPPS.

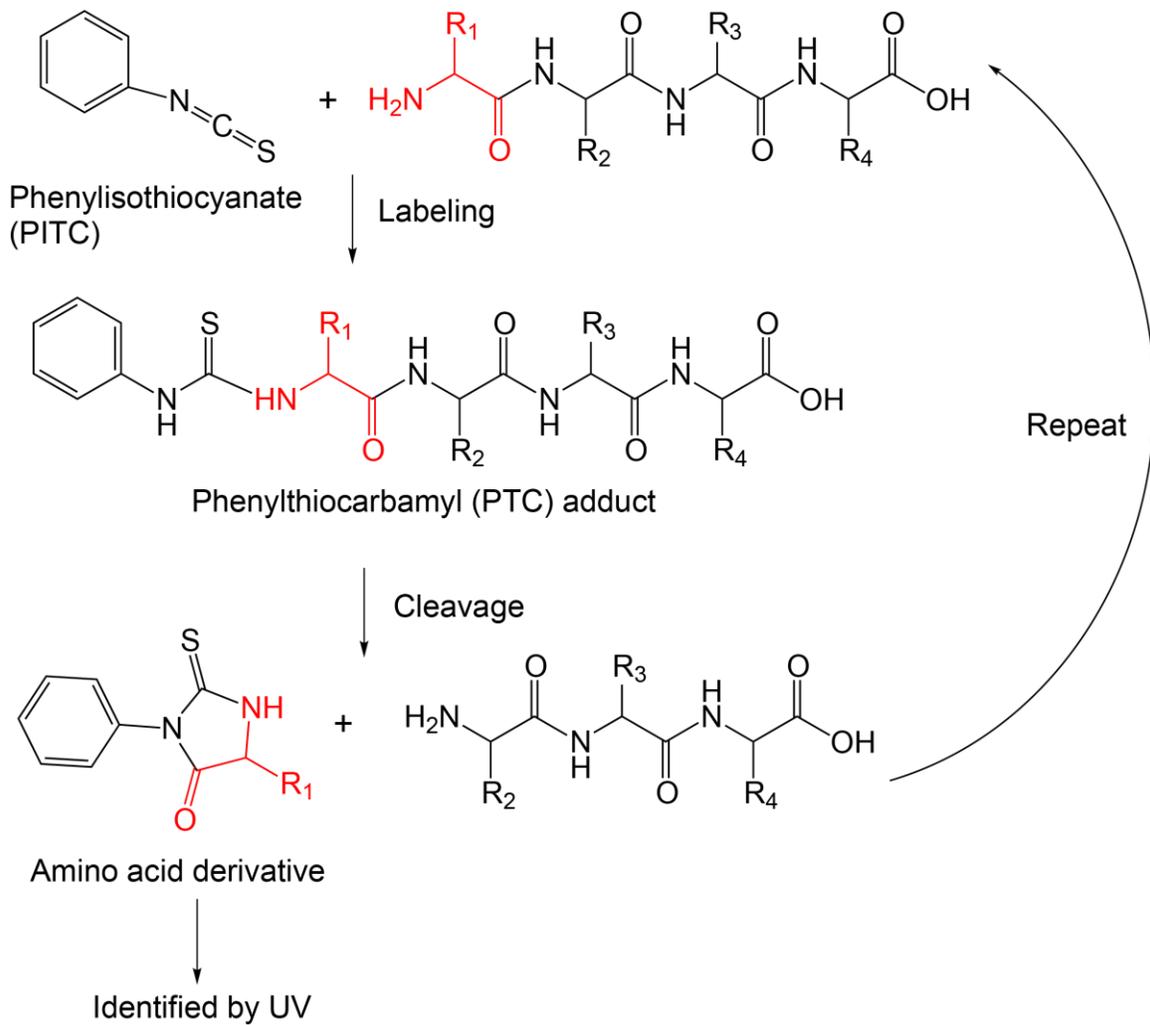


Figure 1-5. Mechanism of Edman degradation.

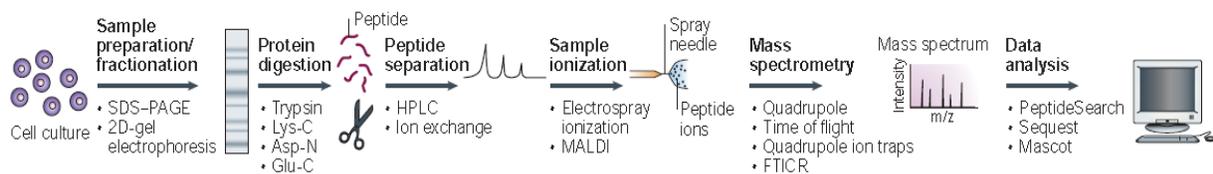


Figure 1-6. The general procedure of MS-based protein analysis. Figure is adapted from Ref 10 with permission.

Table 1-1. Effect of yield per coupling step on final product yields.³

	No. of coupling steps	Yield per coupling step (%)			
		100	99	95	90
	Overall yield	100	90	60	35
10		100	81	36	12
20		100	74	21	4
30		100	67	13	1
40		100	61	8	<1
50		100			

Table 1-2. Sources for MS-based protein identification tools.⁷³

Application	Sponsor	Uniform resource locator (URL)
MassSearch	Eidgenossische Technische Hochschule	http://cbrg.inf.ethz.ch
PeptideSearch	European Molecular Biology Laboratory	http://www.mann.emblheidelberg.de
ExpASY	Swiss Institute of Bioinformatics	http://www.expasy.ch/tools
Mascot	Matrix Science	http://www.matrixscience.com
PepFrag, ProFound	Rockefeller University	http://prowl.rockefeller.edu
MOWSE	Human Genome Research Center	http://www.seqnet.dl.ac.uk
MS-Tag, MS-Fit, MS-Seq	University of California	http://prospector.ucsf.edu
COMET	Institute for Systems Biology (COMET)	http://www.systembiology.org
SEQUEST	University of Washington	http://thompson.mbt.washington.edu/sequest

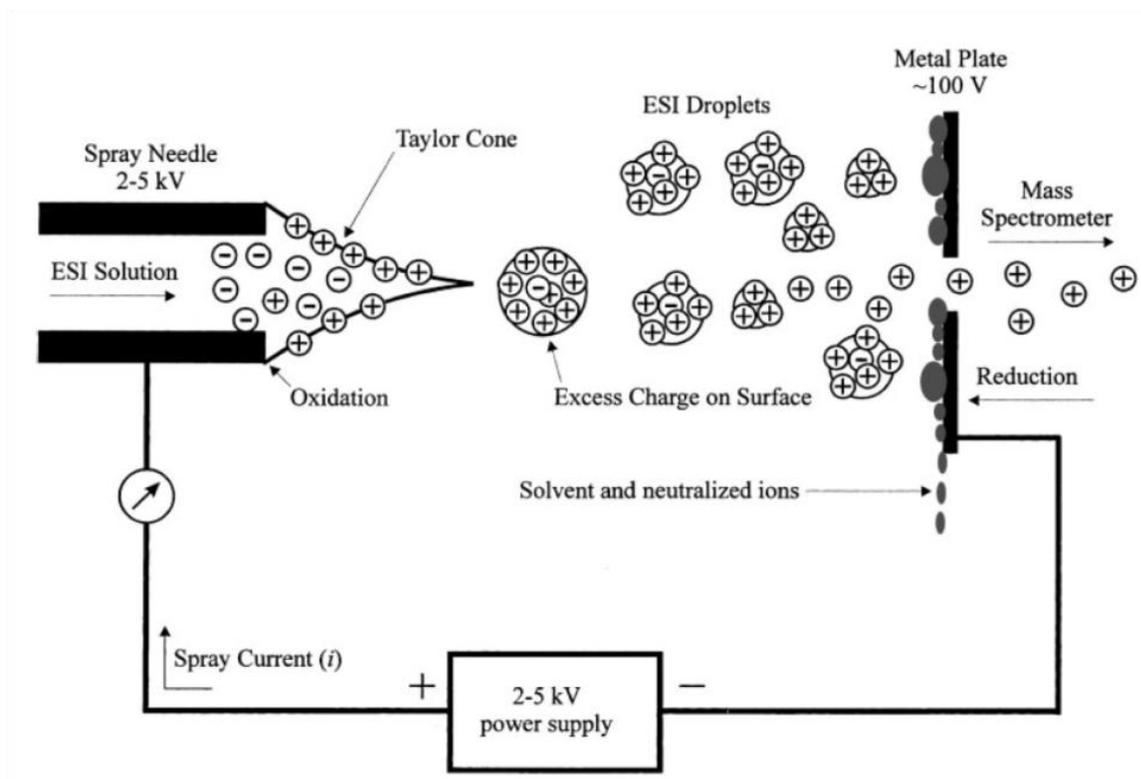


Figure 1-7. Scheme of ESI source. Reproduced with permission from John Wiley and Sons, Inc. from Cech, N.B.; Enke, C.G. *Mass Spectrom. Rev.* 2001, 20, 364.

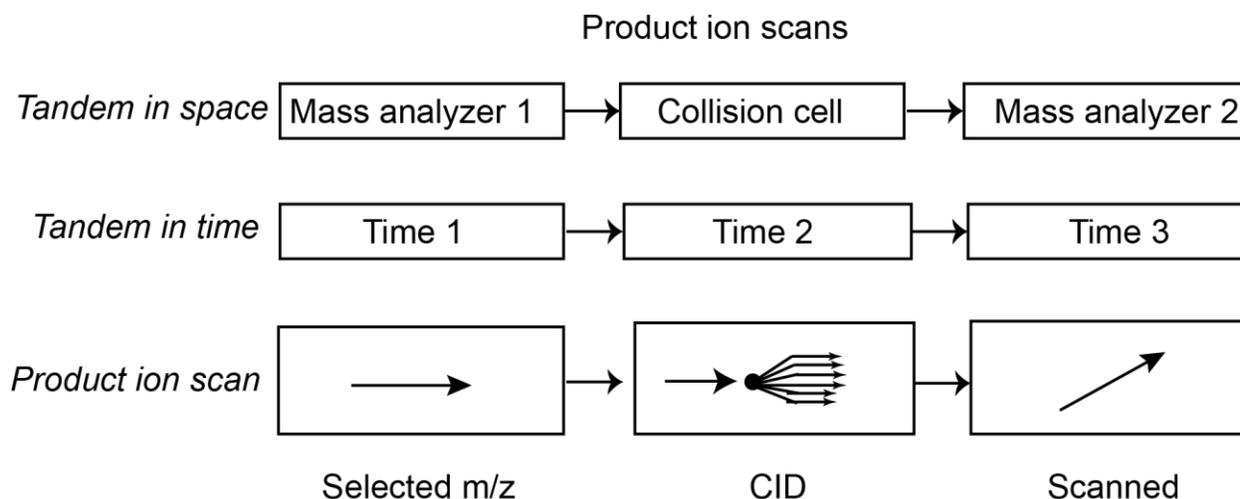


Figure 1-8. Comparison of tandem mass spectrometry in space and in time. Figure is modified from Ref 20.

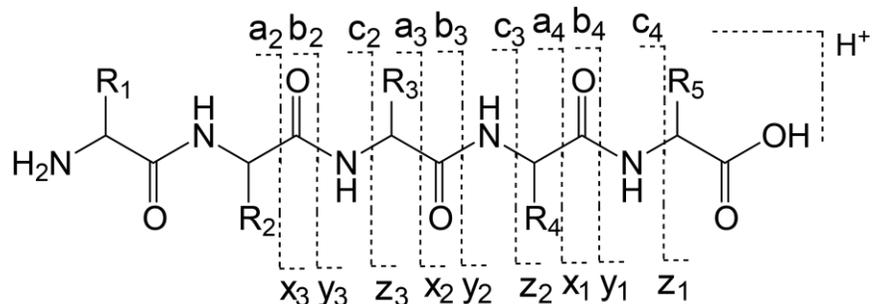
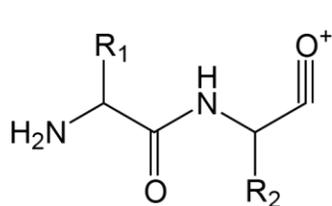
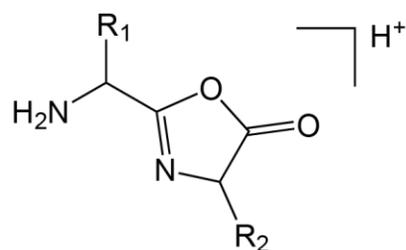


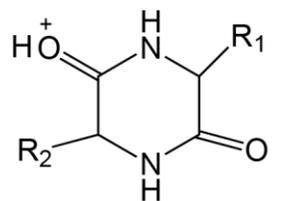
Figure 1-9. Nomenclature of common peptide fragment ion types for a protonated pentapeptide.



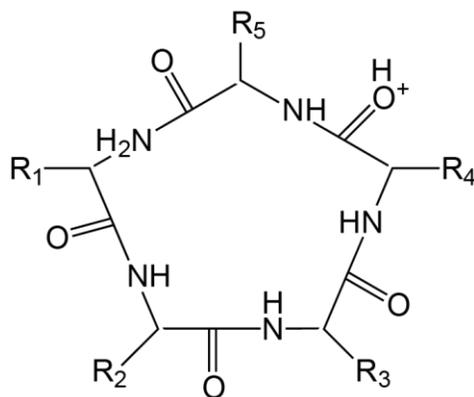
A Acylium



B Oxazolone



C Diketopiperazine



D Macrocyclic

Figure 1-10. Structures of b ions that have been proposed.

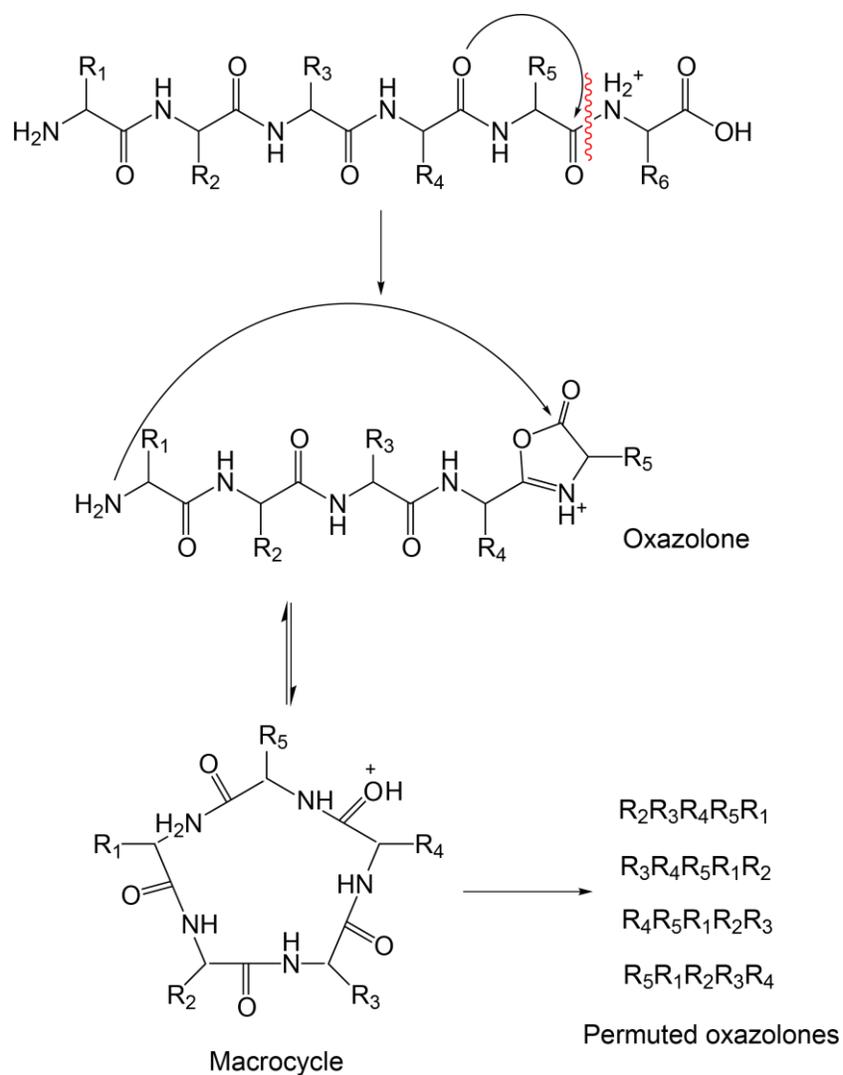


Figure 1-11. Sequence scrambling of a b₅ ion generated from a hexapeptide. Figure modified with Harrison et al. *J. Am. Chem. Soc.* 2006, 128, 10364.

CHAPTER 2 INSTRUMENTATION AND METHODS

Fourier Transform Ion Cyclotron Resonance Mass Spectrometry

Fourier transform ion cyclotron resonance mass spectrometry (FTICR-MS) is a “high-end” mass spectrometry technique. FTICR-MS has the highest resolving power and mass accuracy, and is thus very suitable for analysis of complex systems, e.g. petroleum⁷⁴ and proteins⁷⁵⁻⁷⁹. There are a few excellent reviews on FTICR-MS available.⁸⁰⁻⁸²

Modern FTICR-MS derives from the invention of the cyclotron by Lawrence in the 1930s, when Lawrence and co-workers built the first cyclotron accelerator and used it to investigate atomic fundamental properties.⁸³ Two decades later, Simmer and co-workers incorporated the theory of ICR to a mass spectrometer called the *omegatron*.⁸⁴ In 1974, Comisarow and Marshall first described the concept of combining the Fourier transform with ICR mass spectrometry, and built the first FTICR-MS instrument.^{85,86} Fourier transformation allows one to measure several ions in the cell at the same time, instead of only measuring one ion at a time. Since that time, interests in FTICR-MS have greatly increased, and so has the number of FTICR-MS instruments. Nowadays, commercial FTICR-MS instruments are offered by several companies on the market.

Principles

Ion cyclotron motion

The mass analyzer in FTICR-MS is the ICR cell, on which a magnetic field and an electric field is applied. The magnetic field is uniform, unidirectional, and homogeneous. The motion of ions in the ICR cell is determined by the combination of the magnetic and electric fields. When an ion moves in a magnetic field (B) and an electric field (E), it is

subject to the Lorentz force (F), as described in Equation 2-1, where m is the mass of the ion, a is the acceleration, q is the charge that the ion carries, v is the ion velocity, and F_E and F_M refer to the electric and magnetic components of the Lorentz force.

$$\vec{F} = m\vec{a} = \vec{F}_E + \vec{F}_M = q\vec{E} + q\vec{v} \otimes \vec{B} \quad (2-1)$$

The ion cyclotron motion arises from the interaction of an ion with the magnetic field. When a magnetic field has a direction that is perpendicular to the direction of the ion velocity, the direction of magnetic component, F_M , is perpendicular to the plane determined by v and B . If there are no collisions, the velocity of the ion will remain constant, and hence the ion will be forced by the magnetic field to move in a circular path with a radius of r , as shown in Figure 2-1.

If we set the z-axis along the direction of the magnetic field B , then an x-y plane can be defined perpendicular to z. The ion velocity on the x-y plane is denoted as v_{xy} , and hence the angular acceleration, dv_{xy}/dt , can be expressed as v_{xy}^2/r . When the electric field is missing, and we only consider the magnitude, Equation 2-1 then becomes:

$$F = F_M = qv_{xy}B = m\frac{v_{xy}^2}{r} \quad (2-2)$$

Equation 2-2 can then be rearranged to yield the cyclotron motion radius:

$$r = \frac{mv_{xy}}{qB} \quad (2-3)$$

Since the kinetic energy $KE = mv^2/2$, v can be written as $\sqrt{2KE/m}$, then Equation 2-3 can also be expressed as:

$$r = \frac{mv_{xy}}{qB} = \frac{\sqrt{2mKE}}{qB} \quad (2-4)$$

This equation shows that the radius of an ion in the ICR cell depends on the kinetic energy. Thus, in order to be detected, the ion will need to be excited so that the cyclotron radius is increased to a significant fraction of the dimensions of the analyzer cell.⁸²

The angular velocity (in rad/s), ω , is defined as:

$$\omega = \frac{V_{xy}}{r} \quad (2-5)$$

Substituting Equation 2-5 in Equation 2-3 then becomes:

$$m\omega^2 r = q\omega r B \quad (2-6)$$

Rearranging this equation gives the cyclotron frequency ω_c :

$$\omega_c = \frac{qB}{m} \quad (2-7)$$

The cyclotron frequency can also be defined in Hz as:

$$f_c = \frac{\omega_c}{2\pi} = \frac{qB}{2\pi m} = \frac{B}{2\pi(m/z)} \quad (2-8)$$

In Equation 2-8, m/z is the mass-to-charge ratio of an ion. For example, a singly charged ion with a mass of 151 at 4.7T magnetic field has a frequency of 484.9 kHz:

$$\frac{4.7T}{2\pi \times (151u \times 1.673 \times 10^{-27} kg u^{-1} / 1.602 \times 10^{-19} C)} = 484.9 kHz$$

Equation 2-8 shows that the ICR frequency depends on the m/z and the magnetic field strength, while it is independent of the velocity and hence the kinetic energy. In essence, all ions with the same mass-to-charge ratio have the same cyclotron frequency.

There are two more motions observed in ICR, the *trapping* and *magnetron* motions, which will be introduced in the next two sections. Of these three types of

motion, the ion cyclotron frequency has the highest magnitude and is the only one to be exploited in terms of mass analysis.

Trapping motion

When an ion travels along the z-axis, parallel to the magnetic field, there is no force from the magnetic field, and hence the ion motion along the z-axis is unconstrained. To prevent the ions from escaping from the cell along the z-axis, a trapping voltage is applied to the end caps or plates of the cell, and this electric potential causes another type of ion motion, the trapping motion. Typically, a three-dimensional axial quadrupolar electrostatic trapping potential is used, and it has a form as:

$$\Phi_{(r,z)} = V_{trap} \left(\gamma + \frac{\alpha}{2a} (2z^2 - r^2) \right), \quad (2-9)$$

where V_{trap} is the applied trapping voltage, r is the radial position of the ion in the x-y plane and equals to $\sqrt{x^2 + y^2}$, a is the trap size, and α and γ are constants that are determined by the trap shape.^{80 87}

Ion motion at the z-axis can then be solved using Equation 2-9:

$$F_{z-axis} = m \frac{d^2 z}{dt^2} = -q \nabla \Phi(x, y, z) \quad (2-10)$$

For an ion that oscillates along the z-axis sinusoidally with time, its position can be expressed as

$$z(t) = z(0) \cos(2\pi v_z t) \quad (2-11)$$

The trapping frequency of the ion can then be described as:

$$v_z = \frac{1}{2\pi} \sqrt{\frac{2qV_{trap}\alpha}{ma^2}} \quad (2-12)$$

Magnetron motion

A combination of electric and magnetic fields produces a three-dimensional trap, which allows ions to be stored in the ICR cell for extensive intervals of time (i.e., seconds, minutes, and even hours). The magnetic and electric fields generate cyclotron and trapping motions independently, and the two motions do not couple. However, the combination of magnetic and electric fields together create another ion motion, called the magnetron motion.

First, the trapping potential in Equation 2-9 creates a radial force, F_{radial} , expressed by:

$$F_{radial} = qE(r) = \frac{qV_{trap} \alpha}{a^2} r. \quad (2-13)$$

The radial electric field that operates on the ion generates an outward electric force, opposite to the inward Lorentz magnetic force from the applied magnetic field.

Combining Equation 2-13 and 2-5 gives

$$F = m\omega^2 r = qB\omega r - \frac{qV_{trap} \alpha}{a^2} r \quad (2-14)$$

Solving Equation 2-14, we can then get two results for ω :

$$\omega_+ = \frac{\omega_c}{2} + \sqrt{\left(\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}\right)} \quad (2-15)$$

$$\omega_- = \frac{\omega_c}{2} - \sqrt{\left(\left(\frac{\omega_c}{2}\right)^2 - \frac{\omega_z^2}{2}\right)} \quad (2-16)$$

in which

$$\omega_z = \sqrt{\frac{2qV_{trap} \alpha}{ma^2}} \quad (2-17)$$

ω_+ is called the *reduced* or *observed* cyclotron frequency, which is reduced by the magnetron frequency, ω_- . A calibration can relate the *observed* cyclotron frequency to the *unperturbed* cyclotron frequency, $\omega_c = Bq/m$.

Combination of the three ion motions in the cell – the cyclotron, trapping, and magnetron motions - makes ions behave in a way illustrated in Figure 2-2. As mentioned before, trapping and magnetron frequencies are much lower than the cyclotron frequency, particularly for low DC trapping voltages. The cyclotron frequency is most useful in terms of its simple relationship to the m/z of the ions⁸⁰

Apparatus

There are several types of FTICR instruments available, and they all have four specific components: a magnet, an analyzer cell, an ultra-high vacuum, and a data system, which will be introduced in this section. A schematic diagram of the commercial FTICR mass spectrometer (Bruker Daltonics, Billerica, MA) equipped with a 4.7 T superconducting magnet that is employed in this research is shown in Figure 2-3.

Magnet

A magnet is required to create a strong magnetic field. The performance of an FTICR mass spectrometer improves as the magnetic field strength increases, as shown in Figure 2-4. Some of these characteristics scale linearly, whereas others increase as a function of the square of the magnetic field strength. The magnet can either be a permanent magnet, an electromagnet, or a superconducting magnet. Permanent magnets have low field strengths (< 1 T), which limit the performance of the instrument and are hence not widely used. Electromagnets have higher field strengths (~ 1 -3T), which is good enough for ions of low mass-to-charge ratio. Superconducting magnets

have been used since they can provide much stronger fields, generally 3~12T, with the largest commercial magnets now reaching 14.5 T.⁸⁸ However, there are two main drawbacks about superconducting magnets. First, the cost of a superconducting magnet is high, as well as the maintenance for cryogenics. Second, the sizes of these magnets are large, thus imposing space restrictions.

Analyzer cell

For biomolecular mass spectrometry, ions are made externally, usually by ESI or MALDI, and then injected into the cell. The analyzer cell in FTICR-MS, also called the ICR cell, is surrounded by a magnet as shown in Figure 2-3. Ions are stored, mass-selected, activated, excited, and detected in the ICR cell. The advantage of an ion trap is that mass isolation and activation events can be carried out *in time*, as opposed to *in space*, and are hence only limited by ion signal. Historically, two common types of cell designs were employed, having cubic or cylindrical shape, as shown in Figure 2-5. The cubic cell was the first design, but has been largely phased out in favor of the cylindrical design. In the most basic design, ICR cells consist of three pairs of electrodes. The purpose of the trapping plates is to store ions inside the cell via a static DC potential. A dipolar excitation signal, of the same frequency as the cyclotron frequency, is applied to the excitation plates to excite ions to larger radii. The detection plates detect the image current from ions excited to larger radii. These concepts are explained in more detail in 2.1.3.

Ultra-high vacuum system

An ultra-high vacuum system is critical for all FTICR instruments. While all mass spectrometers require vacuum to analyze and detect ion, FTICR-MS is more sensitive

to background pressure, due to the long distances that the ions travel. To prevent collisions of ions with other molecules background pressures of 10^{-9} - 10^{-10} Torr are required for optimal performance. To obtain such low pressures in combination with higher-pressure ionization sources, a series of pumping stages are employed. In an ESI source, mechanical pumps are used for the first (rough) pumping stage and then turbo-molecular pumps are utilized for lower pressures. FTICR instruments are often equipped with pulsed valves or leak valves to allow introduction of a volatile neutral compound into the high-vacuum region. This neutral molecule can then be employed in ion-molecule reactions with the trapped ions, or can serve as a collision partner for dissociation. In this research, a Varian leak valve, equipped with a sapphire crystal, is used to introduce CH_3OD vapor into the ICR cell to do exchange with the ions of interest that are trapped in the cell. Dissociation experiments were performed by momentarily introducing a collision gas (e.g. Xe) into the cell via a pulsed valve, and applying an excitation pulse to the excitation plates.

Data system

Another main component of all FTICR instruments is a complex data system for acquisition and processing the data. With the data system, the signals produced by excited ions on the detection plates are collected and transformed into more straightforward information. The data system includes a frequency synthesizer, delay pulse generator, broadband radio frequency (r.f.) amplifier and preamplifier, a fast transient digitizer and a computer system, similar to those used for FT-NMR systems.

Excitation and Detection

When ions are transferred into the ICR cell, they are subject to motions discussed in Section 2.1.1. Initially, the radii of the cyclotron orbits are too small to be detected due

to the fact that their kinetic energy in the xy plane is low ($<1\text{eV}$).⁸² In order to be detected, the ions must be excited to a detectable radius. This is achieved by using a radio frequency (rf) pulse to the two excitation plates at the resonant cyclotron frequency of the ions, as shown in Figure 2-6. Following excitation, ions of the same m/z move as a coherent packet undergoing the cyclotron motion.

As the ion packet moves in between the detection plates, it induces a current in the detection circuitry, the image current.^{82,89} Since the cyclotron motion of the ions is periodic, the generated image current is sinusoidal and of the same frequency than the cyclotron frequency of the ions.

Multiple m/z 's of ions are excited to detectable radii by applying an rf frequency sweep. The image current on the detection plates then corresponds to a superposition of cyclotron frequencies. This time-domain signal is known as a transient. In order to extract the cyclotron frequencies that make up this signal, a “Fourier transform” (FT) is applied,⁹⁰ which then yields the frequency-domain spectrum, as shown in Figure 2-7. Since the cyclotron frequency is related to m/z , the frequency-domain spectrum can then be converted to a mass spectrum.

Mass Accuracy

FTICR-MS has the highest mass accuracy and resolution of any mass spectrometric technique. Mass accuracy is often expressed in ppm (parts-per-million), and is an indicator of how well the measured m/z correlates to the “true value”. A definition of mass accuracy is shown in Equation 2-18, by comparing the observed m/z with the “true value”:

(2-18)

$$\text{Mass accuracy} = \frac{m_{\text{observed}} - m_{\text{true}}}{m_{\text{true}}} \times 10^6$$

Other types of commercial mass spectrometers typically have mass accuracies of the orders of tens of ppm, whereas FTICR instruments usually give mass accuracies of 1 ppm or lower.⁹¹

Mass Resolution

Mass resolution is a measurement of how well adjacent peaks in a mass spectrum can be separated. It is very important for complex mixtures or multiply-charged ions, as in these cases the m/z 's of ions are closely spaced. For example, a fullerene derivative that contains nitrogen can be identified with a 9.4T FTICR mass spectrometer, as $^{12}\text{C}_{59}^{14}\text{N}^+$ (m/z 722.002525) and $^{12}\text{C}_{58}^{13}\text{N}_2^+$ (m/z 722.006161), but cannot be distinguished with a lower-resolution mass analyzer.^{91,92} Equation 2-19 defines the mass resolution:

$$\text{Mass resolution} = \frac{m}{\Delta m} \quad (2-19)$$

In Equation 2-19, m is the m/z of the peak of interest and Δm is either the difference in m/z between two adjacent peaks or the width of the peak. The resolution of many commercial time-of-flight mass spectrometers may achieve resolutions of 10^4 , while FTICR mass spectrometers can reach resolutions $>10^5$. The only mass analyzer that can in some cases approach the mass resolution of FTICR mass spectrometers are *orbitraps*.⁹³ Orbitraps measure frequencies of ions orbiting a central electrode by non-destructive image currents; the technique is hence very similar to FTICR-MS in terms of data acquisition and processing, even if the orbitrap is not suitable for actual trapping of the ions for extended periods of time.

Prior to the data acquisition, mass spectra are recorded as a finite number of data points in the case of FTICR-MS, which determines the number of data points used in

time-domain spectrum. The time for data acquisition depends on the sampling rate (or sampling frequency), and the number of data points:

$$T_{acq} = \frac{N}{S} \quad (2-20)$$

where T_{acq} is the acquisition time in seconds, N is the dataset size, and S is the sampling frequency (determined by the low m/z cutoff) in Hz. According to the Nyquist theorem, the sampling frequency must equal to at least twice the highest frequency (i.e., the low m/z limit) of interest.⁹⁰ Thus, the maximum resolution for a given dataset size is determined by:

$$R = \frac{f_c T_{acq}}{2} = \frac{BN}{4\pi(m/z)S} \quad (2-21)$$

in which R is the mass resolution, $f_c = \frac{B}{2\pi(m/z)}$ is the cyclotron frequency, and T_{acq} is the duration of acquisition. From Equation 2-21, it can be found that the resolution of FTICR-MS is proportional to the magnetic field strength, the dataset size, but is inversely proportional to m/z and the sampling rate. For a given FTICR instrument, to achieve better resolution, one can increase the dataset size or use a higher value for the lowest m/z cutoff. The effects of dataset size and lowest m/z cutoff on the resolution are illustrated in Figures 2-8 and 2-9, respectively. It can be seen from Figure 2-8 that the mass spectrum with larger dataset sizes (and hence longer acquisition times) are necessary for higher resolution. Figure 2-9 demonstrates that resolution increases when the value of the lowest m/z cutoff is increased. The mass resolution also increases inversely with m/z in the mass spectrum.

One trade-off with longer transients is that the signal-to-noise ratio is decreased. In Figure 2-8, the longer transients 128k and 512k show a clear decay of the signal,

whereas this is not apparent yet in the 32k transient. This decay is due to a de-phasing of ions inside of the ion packet as a result of Coulombic repulsion and collisions with background molecules. The latter effect demonstrates the importance of an ultra-high vacuum in FTICR-MS.

Space Charge Effects

Ions are stored in the analyzer cell for detection in the FTICR mass spectrometer, generating space charge. The ion space charge in the cell will affect the observed frequencies, and hence result in inaccurate measurements of masses. The frequency shifts were shown to correlate with the changes in space charge.⁹⁴ Jeffries and co-workers developed the theory of space charge induced frequency shifts in Penning cells with different geometries, and mass calibration can be approximated by the following equations.⁹⁵ They derived the following expression for the frequencies of the natural modes of single ion motion in a cubic cell:

$$\omega_{\pm} = \frac{\omega_c}{2} \left\{ 1 \pm \left[1 - 4 \left(\frac{2qV_T G_T}{m} + \frac{\rho q^2 G_i}{\epsilon_0 m} \right) / \omega_c^2 \right]^{1/2} \right\} \quad (2-22)$$

where ω_{\pm} is the angular frequency, B is the magnetic field, V_T is the trapping voltage, ρ is the charge density, m is the mass of the ion, and ω_c is the cyclotron frequency defined in Equation 2-7. An approximated mass calibration can then be derived at the point $m=0$, $V_T=0$, and $\rho=0$, and Equation 2-22 then becomes

$$\omega_{\text{eff}} = \omega_c - \frac{2\alpha\alpha_T}{a^2 B} - \frac{\rho q G_i}{\epsilon_0 B} \quad (2-23)$$

Thus, the frequency difference between two ions are

$$\omega'_{\text{eff}} - \omega''_{\text{eff}} = qB(1/m' - 1/m'') \quad (2-24)$$

When B is known, by measuring the value of $\omega'_{eff} - \omega''_{eff}$, the m/z of an unknown ion can be determined relative to a reference ion.

Infrared Multiple Photon Dissociation Spectroscopy

In Chapter 1, it has been mentioned that several techniques are used for fragmenting ions in mass spectrometers, and collision-induced dissociation (CID) is the most widely employed method. In ion trap mass spectrometry, such as FTICR-MS, infrared multiple photon dissociation (IRMPD) has been used as an alternative to CID, with the advantage that it does not require the introduction of a collision gas into the ion trap. Beauchamp and co-workers first demonstrated that trapped ions in an ICR cell could be dissociated with CO_2 lasers in the late 1970s,^{35,96-98} The irradiation by the IR light results in fragmentation by stepwise vibrational excitation of the molecules. Infrared photons are less energetic, therefore absorption of multiple IR photons are needed to induce dissociation in a molecule.

The mechanism of IRMPD is represented in Figure 2-10. The absorption of photons by trapped ions in IRMPD is a sequential process. An infrared photon with a frequency that is in resonance with the frequency of a vibrational mode (ν_i), is absorbed by an ion. This energy is rapidly dissipated to all other vibrational modes by intramolecular vibrational redistribution (IVR) ($< \text{ps}$).⁹⁹⁻¹⁰² A recycling of the ground state offers the advantage that multiple photon absorption can occur on the fundamental transition over and over again. This eliminates the problem of anharmonicity that is encountered in climbing the vibrational ladder, and hence a large number of photons can be absorbed. Once the dissociation threshold is reached, fragmentation is usually observed via the lowest-barrier channel. In the mass spectrum, this dissociation is

observed as a change in mass. The IRMPD yield reflects the efficiency of IR absorption at a particular wavelength. By employing a tunable laser and measuring the IRMPD yield as a function of wavelength, the IRMPD spectrum of an ion can be recorded.^{103,104}

Direct infrared spectroscopy measurements on trapped ions in mass spectrometers are not possible, due to the ultra-low ion densities. IRMPD spectroscopy hence offers an indirect approach to measure infrared spectra of gaseous ions. Such an approach is often referred to as “action” spectroscopy, as the absorption of photons causes a change in the species that can be observed, which in this case is a change in the m/z of the ion.

Since multiple photons are required for dissociation, moderately powerful lasers are necessary, and several types of lasers that have been successfully implemented in IRMPD spectroscopy studies. Historically, line-tunable CO₂ lasers were first employed. Such lasers cover the region from 9.2 to 10.8 μm . While there are several reports on IRMPD spectroscopy with CO₂ lasers,¹⁰⁵⁻¹⁰⁷ their main limitation is the limited wavelength range.

Free electron lasers (FELs) are another type of laser used for IRMPD spectroscopy studies. Several facilities have coupled FELs with FTICR-MS for IRMPD research, including the Free Electron Laser for Infrared experiments (FELIX) at the FOM Institute for Plasma Physics Rijnhuizen in The Netherlands, the Centre Laser Infrarouge Orsay (CLIO) in Orsay, France, and the FEL at the Science University of Tokyo (SUT) in Tokyo, Japan. FELs provide high power over a wide wavelength range (5-250 μm for FELIX), and are hence ideally suited for IRMPD experiments. Nonetheless, their cost is substantial and access to these lasers is limited. The IRMPD

experiments in this dissertation were done at the FELIX facility. Figure 2-11 shows the schematic representation of FEIXL coupled to laboratory-constructed FTICR mass spectrometer.

There are two beamlines at FELIX, FEL-1 and FEL-2. The free electrons are accelerated to either 15 to 25 or 25 to 45 MeV by one or two radio-frequency linear accelerators. Using the first accelerator allows FEL-1 to access wavelengths from 25 to 250 μm . When the two accelerators are used in conjunction with each other, FEL-2 can access wavelengths from 5 to 30 μm . The operation of a free electron laser rests on the principle that when a relativistic electron enters a magnetic structure, called an undulator, it can spontaneously emit radiation. This radiation can be tuned by adjusting the kinetic energy, since the laser wavelength in the labframe is substantially Doppler shifted from the mm to the μm range. For a particular kinetic energy setting, the wavelength can be tuned by a factor of 2-3 by adjusting the magnetic field in the undulator via positioning of samariumcobalt permanent magnets. FELIX is a pulsed laser composed of micro and macropulses that reflect the injection of electron bunches into the laser cavity. The micropulses are spaced 1 ns apart and have a duration of 1ps to multiple ps. Macropulses generally have a duration of 5-7 μs at a repetition rate of 5 Hz or 10 Hz.

More recently, our group has implemented a benchtop optical parametric oscillator (OPO) laser to carry out IRMPD experiments on mass-selected ions in an ICR cell. This laser was shown to be able to photodissociate metal-bound carbohydrate¹⁰⁸ and amino acid complexes.¹⁰⁹ In terms of IRMPD, the collision-less environment of an ICR cell is certainly advantageous; as the absorbed energy is not lost due to collisions with

background molecules. Moreover, the irradiation time can be extended to multiple seconds, thus increasing the IRMPD yield. All of the IRMPD results described in this thesis employ the free electron laser FELIX in conjunction with the laboratory-constructed FTICR instrument in Figure 2-11.

Gas-Phase Hydrogen/Deuterium Exchange

Hydrogen/deuterium exchange (HDX) is a powerful technique to investigate structures and conformations of peptides and proteins. Hydrogen atoms bonded to oxygen, nitrogen, and sulfur can exchange with deuterium when molecules are exposed to an environment of deuterated molecules. Typically, H/D exchange reactions are performed on proteins in solution, where a protein is incubated in a D₂O buffer, and the degree of subsequent deuterium incorporation determined by nuclear magnetic resonance (NMR) or mass spectrometry¹¹⁰⁻¹¹³. The isotopic exchange of these labile hydrogens is an important tool to determine protein structures, and to study dynamic processes such as protein folding.¹¹⁴⁻¹¹⁷ Exchanges at the side chain are typically too fast to be kinetically resolved, but hydrogens at the backbone amide positions exchange at rates that can be measured. The lability of amide hydrogens also offer most insights into the secondary and tertiary structure of proteins.

In gas-phase HDX, reaction rate constants are much lower than those in solutions due to the very diluted conditions¹¹⁸⁻¹²², so it has been proposed as an alternative method for structural analysis of amino acids.^{121,123-125} Mass spectrometers are very suitable for performing gas-phase HDX experiments because the deuterium incorporation can be monitored directly from the mass increase of the peptide. In addition, with ion trapping techniques, such as ion cyclotron resonance (ICR) and

quadrupole ion traps, ions can be trapped for extended time periods and thus allowing for observation of the exchange processes, so information about gas-phase H/D exchange kinetics can be obtained.^{121,126}

Systematic studies on model peptides had shown that the exchange rate in gas-phase HDX depends on the proton affinity (PA) and/or gas-phase basicity (GB) of the peptide molecule and the deuterating reagent.^{127,128} The PA is defined as the enthalpy of reaction 2-25, whereas the GB stands for the free energy.¹²⁹



Several groups have contributed to the evaluation of the gas-phase proton affinities for amino acids and peptides with either experimental or theoretical methods.¹³⁰⁻¹³⁴ The gas-phase proton affinities of typical deuterating reagents, including D₂O, CD₃OD, CD₃COOD, and ND₃, are 166.5, 181.9, 190.2, and 204.0 kcal/mol, respectively. Deuterating reagents with higher proton affinities display faster exchange. The evaluated proton affinities of amino acids are listed in Table 2-1. Proton affinities of peptides depend on the proton affinity of every single amino acid residue, as well as its chain length. For example, the proton affinities of Gly₂, Gly₃, Gly₄ and Gly₅ are 210.0, 213.0, 218.1, and 218.4 kcal/mol, respectively.^{129,133} There are several aspects that will affect the HDX: the difference in PAs between the analyte and the deuterating reagent, the size of the analyte molecule, and the number of deuterons on the deuterating reagent. When the PA is the main factor, if the deuterating reagent has a proton affinity that is similar to the sample, the sample's reactivity will be promoted. Of all the common deuterating chemicals listed before, ND₃ gives the fastest exchange rates when exchanging with peptides or amino acids.

The understanding to the mechanism of gas-phase HDX remains inadequate. In fact, Beauchamp and co-workers proposed a number of mechanisms that are active, depending on the deuterating molecule. For low-basicity reagents, such as CD_3OD or D_2O , a 'relay' mechanism was postulated. In this mechanism, the deuterating molecule forms two hydrogen bonds: one with the protonated site on the N-terminus, and one with an amide oxygen. The 'relay' mechanism then assumes a concerted movement of the N-terminal proton to the deuterating molecule, and movement of a deuteron to a carbonyl oxygen. In effect, the proton of the peptide ends up on the deuterating molecule, while a deuteron ends up on the peptide. For exchanges to occur, it was proposed that the energy gained by forming hydrogen bonds must be larger than the difference in the proton affinities of the exchange sites and the energy lost by opening the internally solvated structure.¹²⁸ More basic deuterating agents, such as ND_3 , are thought to operate via an 'onium' mechanism, where ND_3 is capable of solvating the proton as an intermediate. In this thesis, deuterated methanol (CH_3OD) is employed as deuterating reagent. For the purpose of the discussion, it will be assumed that the 'relay' mechanism is active.

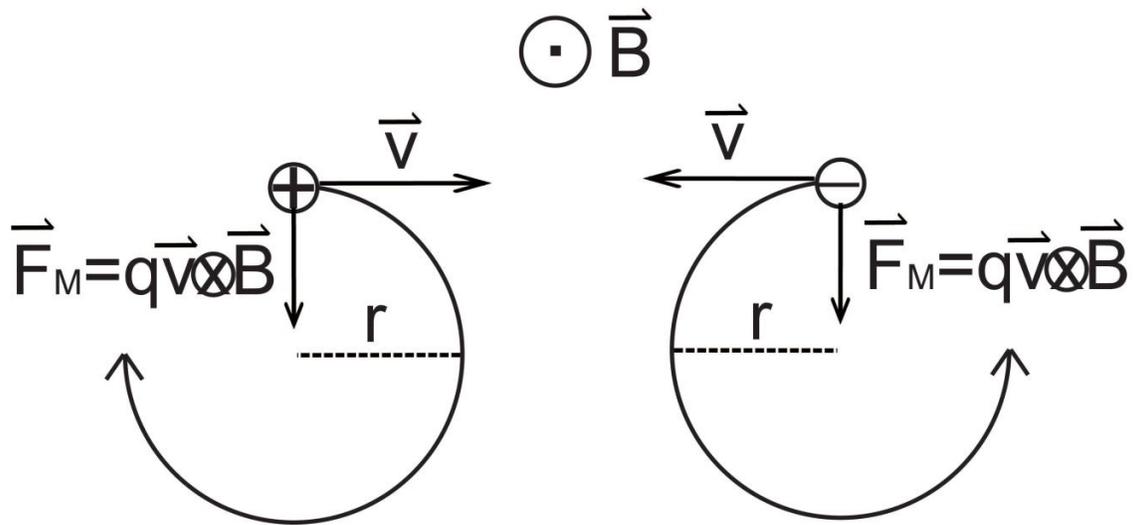


Figure 2-1. Ion cyclotron motion. A spatially uniform magnetic field has a direction that is perpendicular to and going into the plane of the paper and an ion is moving in the plane of the paper. The moving path of the ion is a circle resulting from the magnetic Lorentz force. Positive and negative ions travel in opposite directions.

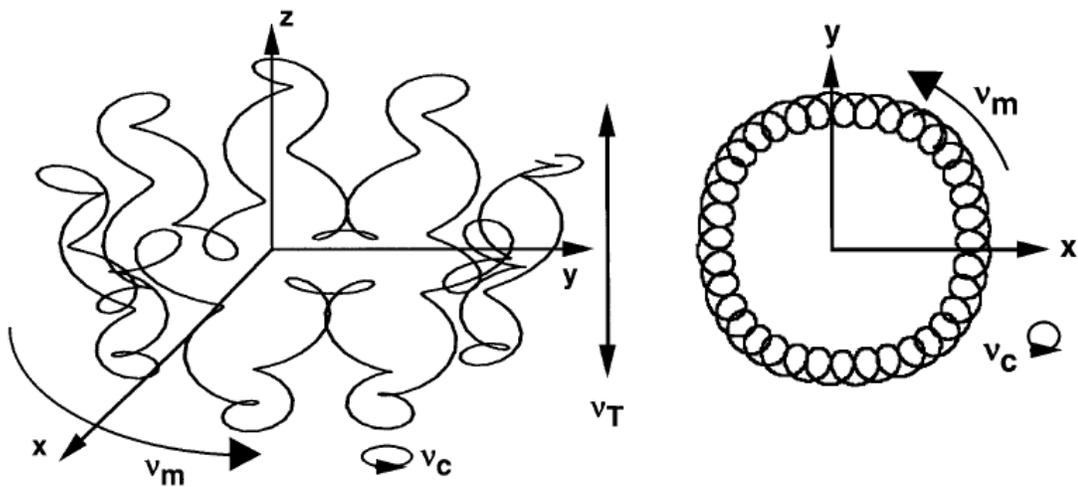


Figure 2-2. Ion path resulting from the combination of cyclotron, tapping, and magnetron motions. Figure reproduced from Ref.80 with permission.

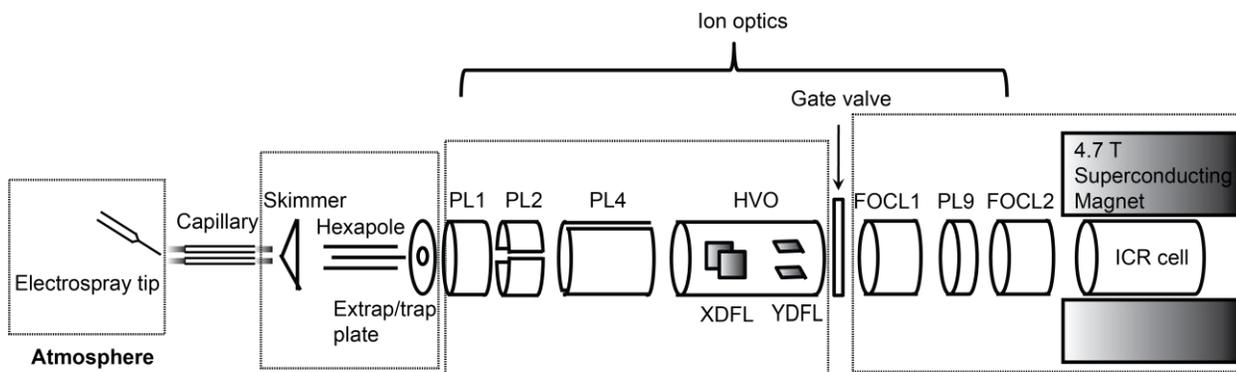


Figure 2-3. Schematic diagram of Bruker APEXII 4.7T FTICR instrument.

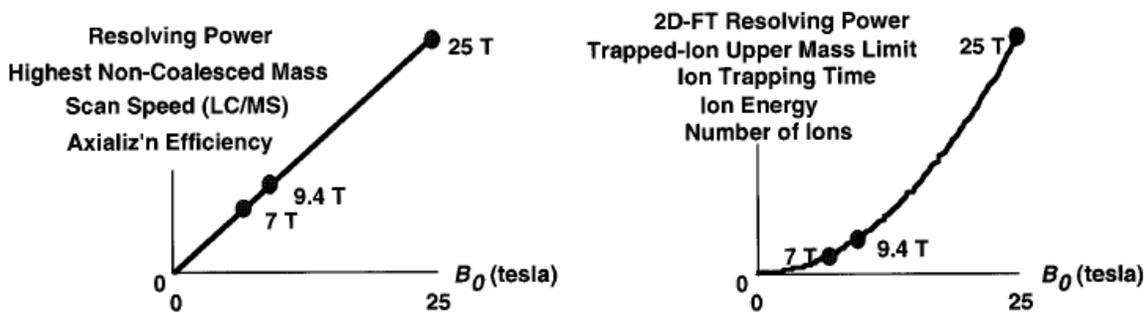


Figure 2-4. FTICR-MS performance as a function of magnetic field strength B_0 . Figure reproduced from Ref. 80 with permission.

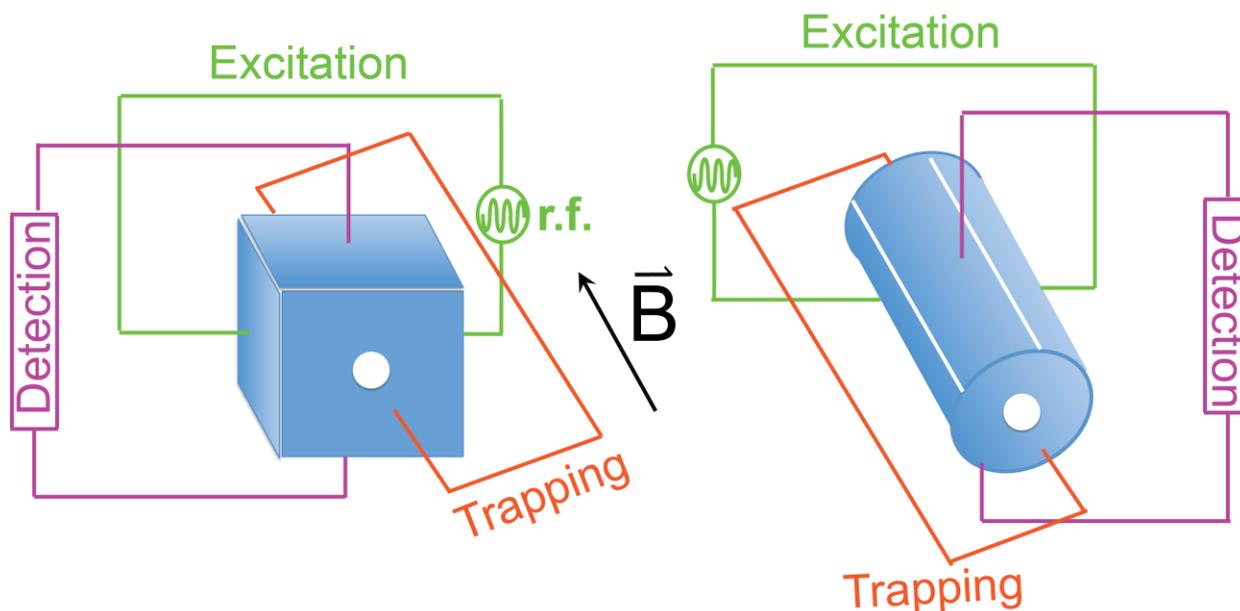


Figure 2-5. Schematic presentation of a cubic (left) and a cylindrical (right) FTICR-MS analyzer cell. Both types of cell have six plates, and each pair functions for excitation, trapping, and detection.

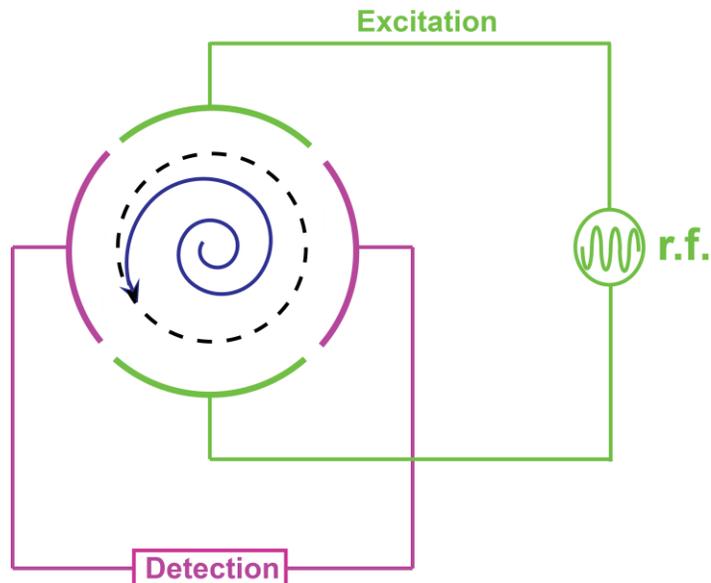


Figure 2-6. Schematic showing cross section of cylindrical ICR cell. Ions are excited to a larger radius (shown in blue) by applying an rf potential to the excitation plates. The motion of the ion packet is then detected on the detection plates in the form of an image current.

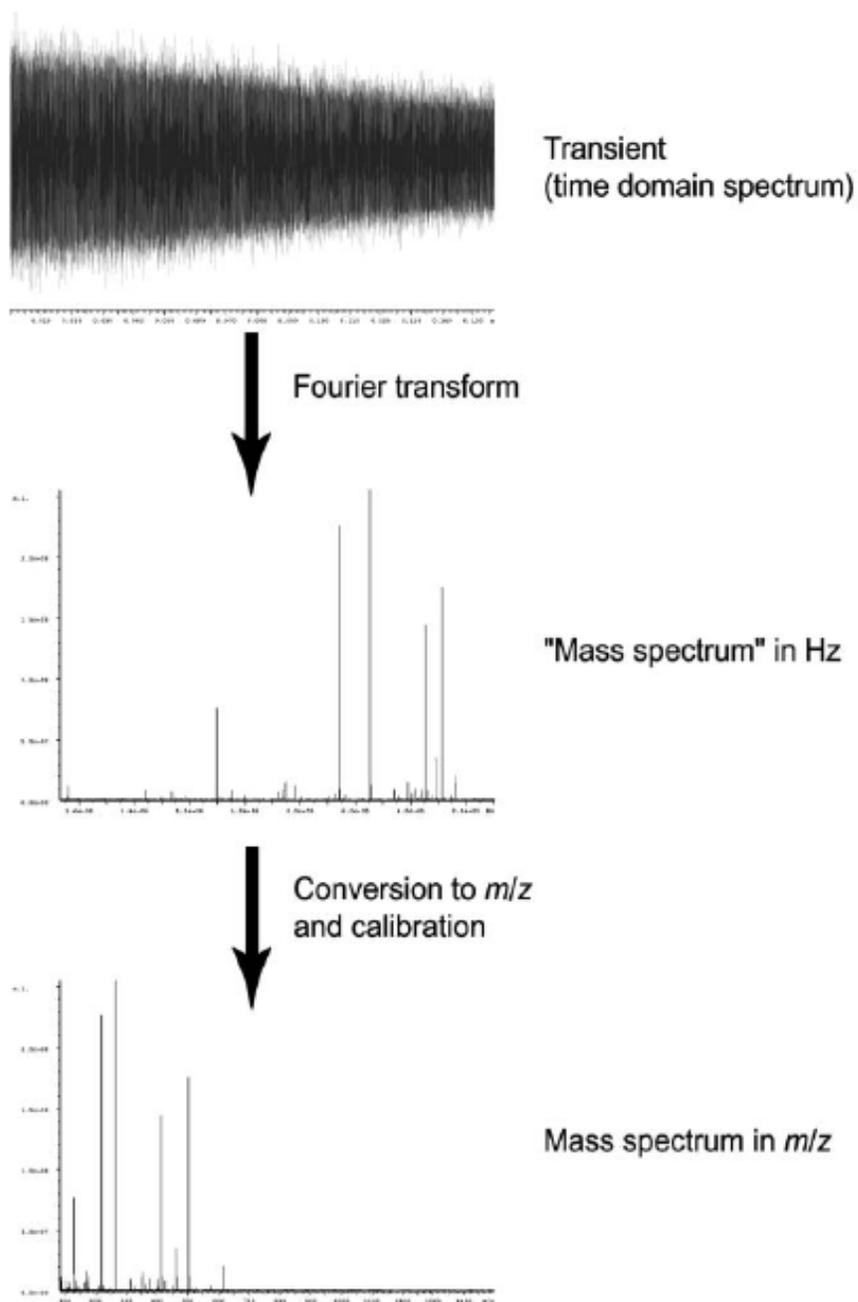


Figure 2-7. Schematic showing how time-domain data is converted to a frequency-domain spectrum, followed by conversion to a mass spectrum. Figure taken from Ref 91 with permission of The Royal Society of Chemistry.
<http://dx.doi.org/10.1039/b403880k>

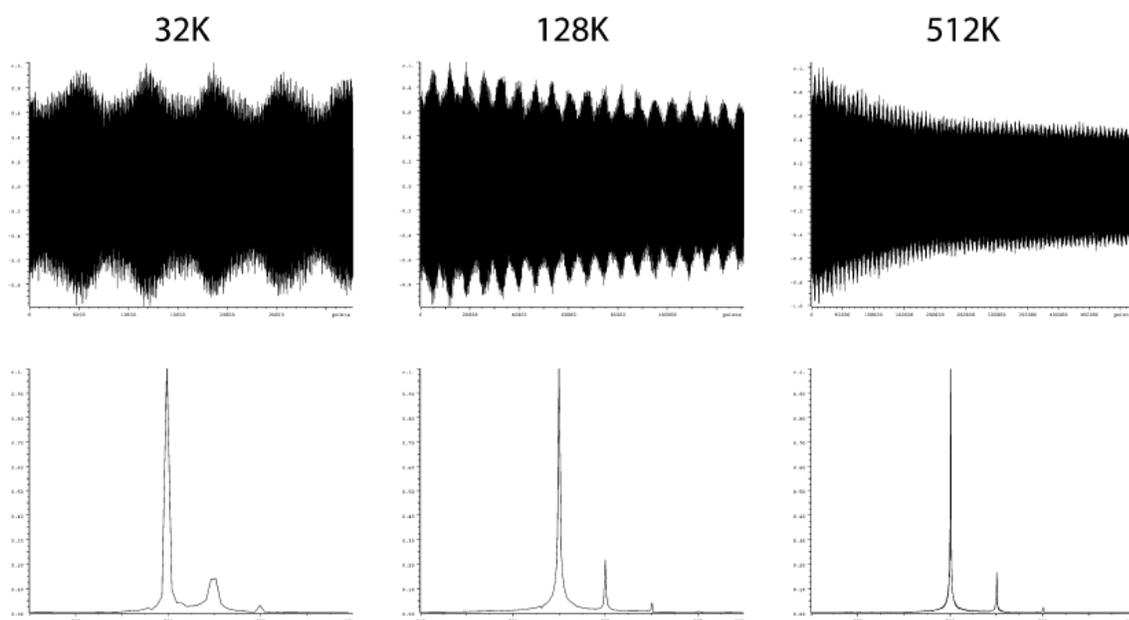


Figure 2-8. Demonstration of the effect of dataset size on resolution. The three mass spectra correspond to different acquisition dataset sizes for the same sample. Figure reproduced from Ref 91 with permission of The Royal Society of Chemistry. <http://dx.doi.org/10.1039/b403880k>

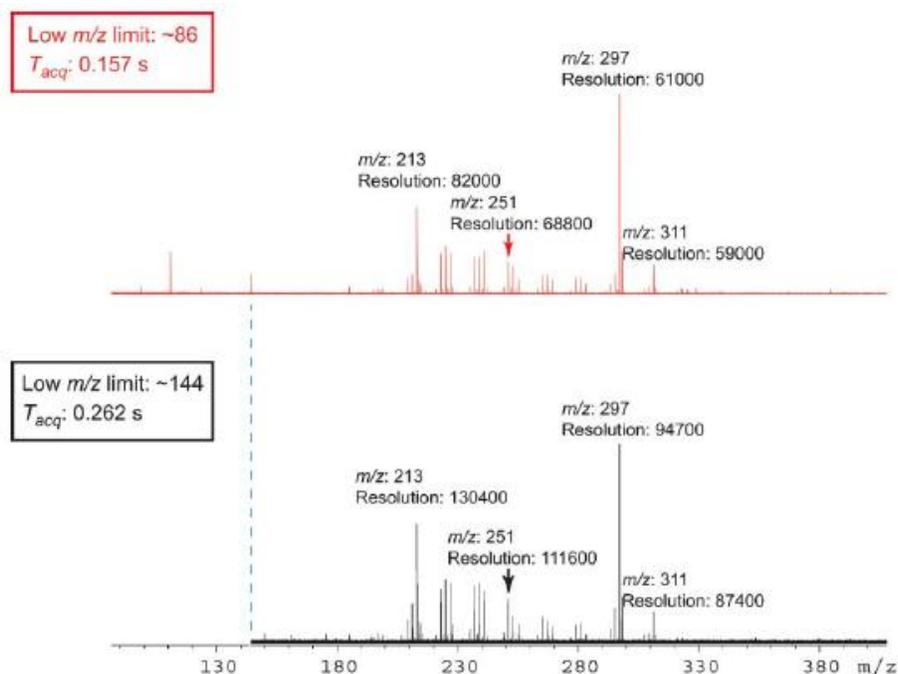


Figure 2-9. The dependence of resolution on both measured m/z and the lowest m/z cutoff. Figure reproduced from Ref 91 with permission of The Royal Society of Chemistry. <http://dx.doi.org/10.1039/b403880k>

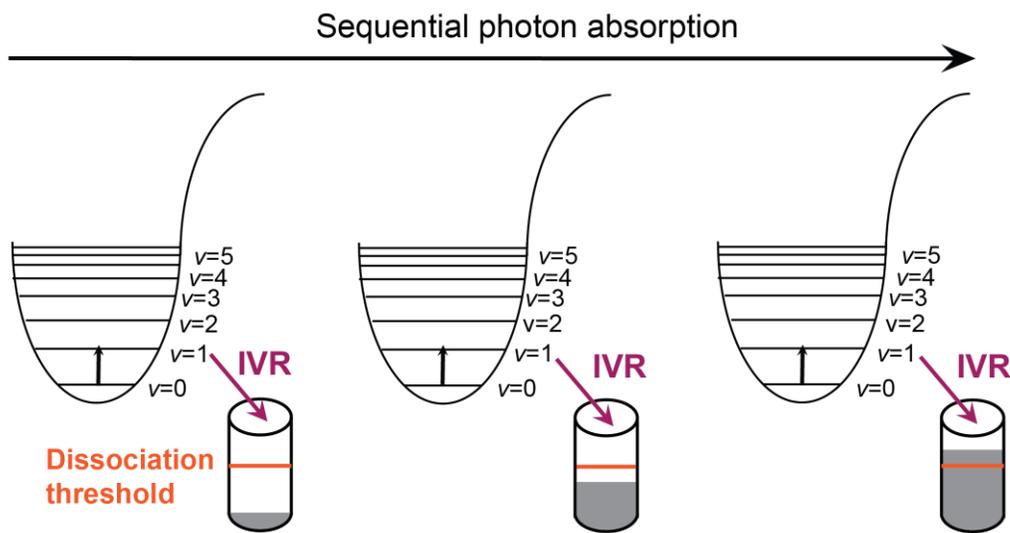


Figure 2-10. Schematic presentation of IRMPD mechanism.

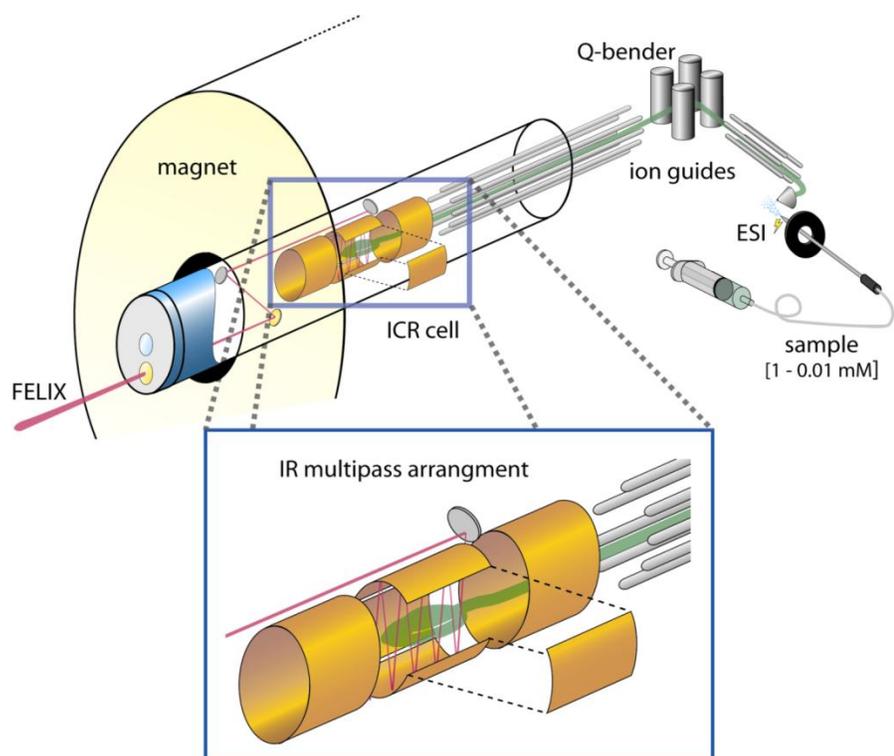


Figure 2-11. Schematic representation of the laboratory-constructed FTICR instrument equipped with an ESI source coupled to FELIX. Figure taken from Ref 104 with permission.

Table 2-1. Proton affinities of common amino acids. Numbers are adapted from Ref.129.

Amino Acid	Proton Affinity (kcal/mol)	Amino Acid	Proton Affinity (kcal/mol)
Gly	210.5	Asn	220.6
Cys	214.0	Tyr	220.9
Ala	214.2	Met	221.1
Ser	215.2	Gln	222.0
Asp	216.4	Pro	222.1
Val	216.5	Glu	223.4
Leu	217.4	Trp	223.9
Ile	218.6	Lys	235.6
Thr	219.5	His	231.5
Phe	219.9	Arg	244.8

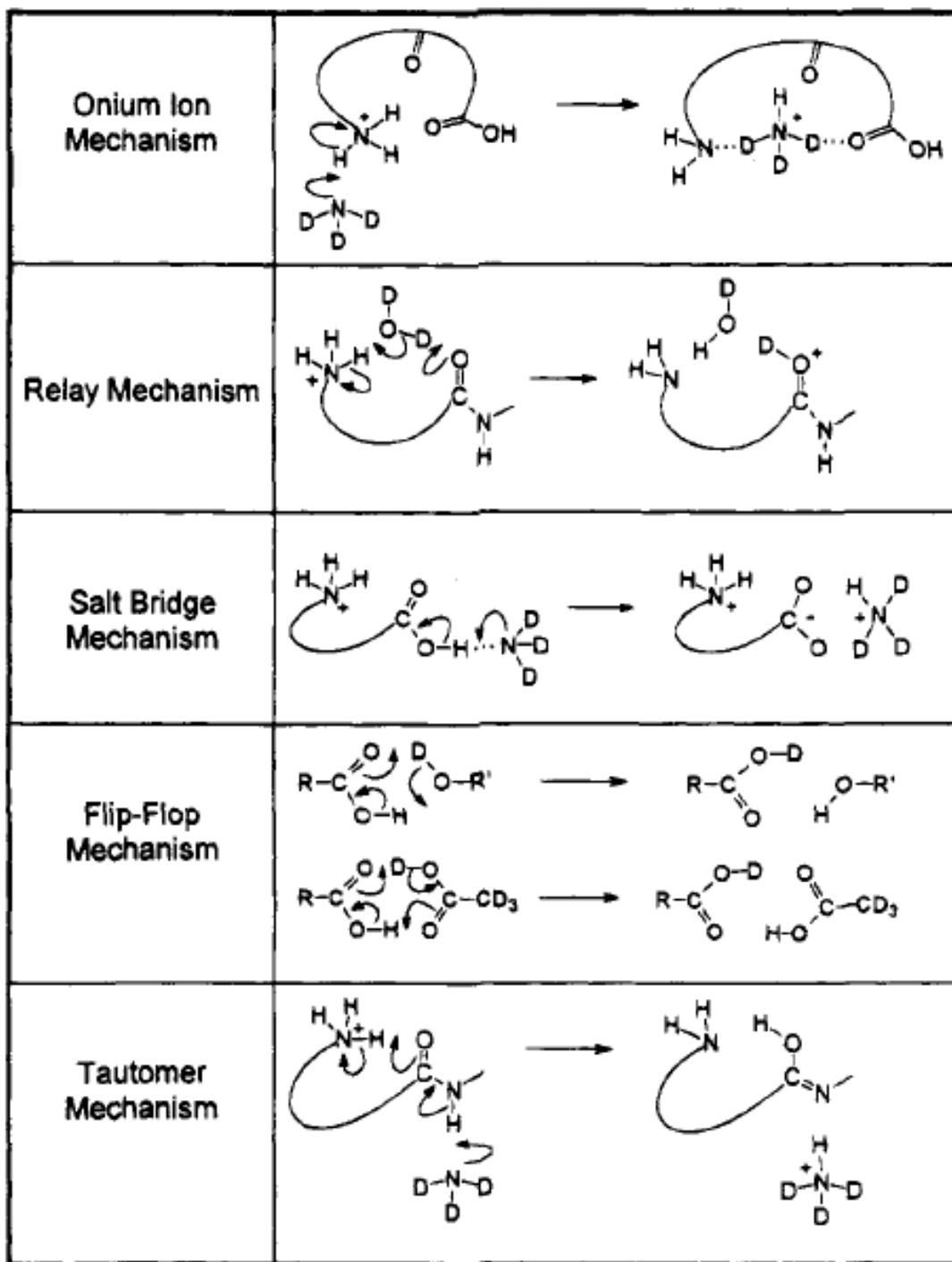


Figure 2-12. Gas-phase H/D exchange mechanisms proposed by Beauchamp and co-workers. Figure is adapted from Ref 127 with permission.

CHAPTER 3
INVESTIGATION OF THE INFLUENCE OF PEPTIDE B FRAGMENT IONS ON THE
FORMATION OF MACROCYCLE STRUCTURES THROUGH HEAD-TO-TAIL
CYCLIZATION IN THE CASE OF OLIGOGLYCINES*

Background

It has been discussed in Chapter 1 that b ions generated by collision-induced dissociation will undergo head-to-tail cyclization and form macrocyclic **b** ion isomers, which will then lead to sequence scrambling due the reopening of the ring structures at different sites. So far, to fully investigate the driving forces of the formation of *macrocycle* structures, systematic studies are required. In this chapter, the size effect of b ions will be studied.

As mentioned before, infrared photodissociation spectroscopy can provide direct evidence for structures of b ions, and one example is that the IRMPD spectrum of **b₄** fragment of the pentapeptide Leu-enkephalin (Tyr-Gly-Gly-Phe-Leu) shows prominent bands for both *oxazolone* and *macrocycle* structures.³⁹ However, IRMPD spectroscopy cannot give information about relative abundances because the intensity of the band does not directly relate to the concentration of the structures.

Gas-phase hydrogen/deuterium exchange (HDX) has been employed to study the ion structures in the gas phase. O'Hair and co-workers employed HDX to investigate a series of **b** fragments made from oligoglycines, where they considered several water loss reaction channels.¹³⁵ Wysocki and Somogyi have used HDX to determine that **b** ions often exhibit bimodal distributions, which can be rationalized by the presence of

* Reproduced in part with permission from Chen, X.; Yu, L.; Steill, J.D.; Oomens, J.; Polfer, N.C. Effect of Peptide Fragment Size on the Propensity of Cyclization in Collision-Induced Dissociation: Oligoglycine **b₂**–**b₈**, *J. Am. Chem. Soc.* 2009, 131, 18272-18282. Copyright 2009 American Chemical Society.

two different chemical structures.^{71,136} Paizs and Somogyi have used H/D exchange and DFT calculations for tryptic digest peptide **b₂** fragments to show that exclusively *oxazolone* structures are formed.⁷⁰ Since the inadequate understanding of the HDX mechanism, the interpretation of HDX data is complicated, and hence making direct structural assignment difficult.

Stable isotope labeling techniques have been used by many groups for the mechanistic elucidation of gas-phase reactions and fragmentations over the years.¹³⁷⁻¹⁴²

Influence of size on scrambling has been studied by Van Stipdonk and co-workers using tandem mass spectrometry on a series of peptides, ranging from tetrapeptides to decapeptides. Minimal scrambling was observed in **b₃** ion, while significant nondirect ions were found in all other ions.¹⁴² However, since the peptides they used had different residues, there might be influence from the size chain of some amino acid. In addition, with just MSⁿ method, they were not able to characterize the structures. Finally, the quantification of the isomeric structures was not obtained.

In this project, we combine isotope labeling, infrared spectroscopy, gas-phase HDX, and computational approaches to structurally analyze a series of **b** product ions generated by CID of oligoglycine, from **b₂** to **b₈**. Glycine is the simplest amino acid, and no side-chain nucleophiles are present, and hence the competition between *oxazolone* and *macrocycle* formation can be studied as a function of size of the fragment generated.

Experimental

Sample Preparation

The triglycine (Sigma Aldrich, St. Louis, MO), pentaglycine (Sigma Aldrich, St. Louis, MO), cyclo(Gly-Gly) (Bachem, Torrance, CA) and deuterating reagent CH₃OD

(Sigma Aldrich, St. Louis, MO) were employed without further purification. Octaglycine was prepared by solid-phase synthesis and was isotopically labeled by incorporating a ^{13}C -Gly as the N-terminal residue (ICBR, University of Florida), and the peptide was purified by high-performance liquid chromatography (HPLC). All peptides were used at 20 μM solution 49:49:2 (vol:vol:vol) water/methanol/acetic acid solutions. The acetic acid was added to aid the protonation.

Mass Spectrometry and Hydrogen/Deuterium Exchange

The hydrogen/deuterium exchange (HDX) experiments were carried out at the University of Florida using a commercially available Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer (4.7 T actively shielded APEX II, Bruker Daltonics, Billerica, MA) in Dr. David Powell's lab. The protonated peptide precursor ions were activated by "nozzle-skimmer" dissociation in the ESI source region by adjusting the voltage drop between the metal-plated glass capillary and the first skimmer. The CID product ions were then accumulated in the hexapole (3–4 s), prior to transfer to the ICR cell. The exception to this was the b_2 fragment, which was generated in the ICR cell by sustained off-resonance irradiation collision-induced dissociation (SORI-CID) from protonated triglycine by pulsing in Xe gas, as low- m/z (<130) ions could not be transferred from the hexapole to the ICR cell successfully. The mass-to-charge ratios of all the ions of interest and the peptides used are listed in Table 3-1.

The most abundant isotope peak of the b_n product ion of interest was mass isolated and subjected to gas-phase hydrogen/deuterium exchange (HDX) with CH_3OD , which was leaked into the vacuum chamber at a constant pressure of 10^{-8} Torr through a Varian leak valve. Mass spectra with different exchange times were recorded, and the abundances for the undeuterated (d_0), singly deuterated (d_1), doubly deuterated (d_2),

etc. peaks in the resulting distributions were determined, as Campbell showed in their HDX kinetic study^{119,128} (Figure 3-1). The normalized abundance of d_0 is defined as

$$[d_0] = \frac{I_{d_0}}{\sum_i I_{d_i}} \quad (3-1)$$

in which I refers to the relative abundance of each product, and i to the number of deuterium exchanges. For each fragment, the natural logarithm of $[d_0]$ was plotted against the exchange time,^{65,68} and then linear fits were performed to determine the reaction kinetics and abundances of the two structures. Since the exchange reagent is considered to be in great excess of the analyte, the H/D exchange reaction can be assumed as a pseudo first order reaction. The slope obtained represents the exchange rate constant. The larger the slope, the faster the exchange is. For fragments with two chemical structures, two different slopes should be obtained.

Mass Spectrometry and Infrared Photodissociation Spectroscopy

The infrared photodissociation experiments were conducted at the FOM institute for Plasma Physics “Rijnhuizen” using the free electron laser FELIX¹⁴³ coupled to a laboratory-built FT-ICR mass spectrometer, described in detail in previous publications.^{144,145} The b_2 , b_5 , and b_8 CID products were generated by “nozzle-skimmer” dissociation, in a similar scheme as described in section 3.2.2, using protonated triglycine, pentaglycine, and octaglycine, respectively. The notation b_2 -G3 for instance relates to the b_2 ion generated from triglycine. The fragment ion was accumulated in a hexapole, prior to transfer to the ICR cell. As a control experiment for b_2 , proton-attached cyclo(Gly-Gly) was produced by ESI in a separate experiment to measure its IR spectrum. Following mass isolation, the ion of interest was irradiated with the tunable output from the free electron laser. FELIX produces macropulses (5 μ s) that consist of a

train of micropulses at a GHz repetition rate. The pulse energy per macropulse is dependent on wavelength, reaching maximum values of 60 mJ at 12 μm . Typically, 20–30 macropulses were employed to induce efficient photodissociation.

Infrared multiple-photon dissociation (IR-MPD)¹⁴⁶ spectroscopy works on the premise that when the laser frequency is in resonance with a fundamental vibration in the molecule, many photons (i.e., tens to hundreds) are absorbed, leading to photodissociation. This manifests itself in the mass spectrum by a depletion of the precursor ion and appearance of photofragments. The photofragments of b_2 , protonated cyclo(Gly-Gly) and b_5 observed in the photodissociation are listed in Table 3-2.

The IR photodissociation spectrum was obtained by plotting the IRMPD yield as a function of the wavelength, using the relation shown below. Given the many CID product ions generated from octaglycine, the photodissociation products for $b_5\text{-G}_8$ and $b_8\text{-G}_8$ were not detectable, and hence the IRMPD depletion spectra are shown here.

$$\text{yield} = -\ln\left(1 - \frac{\sum \text{photofragments}}{\sum \text{all_ions}}\right) \quad (3-2)$$

In the photodissociation experiments presented here, the wavelength was scanned from 1200–2000 and 2500–3400 cm^{-1} . Note that in the latter experiments, mirrors with a dichroic coating were employed in the laser cavity, to specifically allow reflection (and hence lasing) at the third harmonic, but not at the fundamental wavelength.

Results and Discussion

Scrambling in Isotope-Tagged Peptide

In order to test for scrambling in oligoglycine peptides, the octaglycine peptide was isotopically labeled at the N-terminal position with a ^{13}C glycine residue. The mass

spectra of CID products b_4 - b_7 of the protonated precursor ion (m/z 476) are shown in Figure 3-2. For all these b ions, in addition to the b fragment ions that have this residue incorporated, ^{13}C -labeled b_n , b ions with masses that relates to the loss of the ^{13}C -Gly residue, ^{12}C -only b_n , are also observed in the mass spectra. The possible explanation for the case of b_7 are illustrated in Figure 3-3. The loss of N-terminal ^{13}C -Gly from b_7 can be rationalized by a rearrangement process, where a macrocyclic b_8 is first formed and then the reopening leading to an *oxazolone* with ^{13}C -Gly on the C-terminal side of the molecule. Upon amide bond cleavage between the seventh and eighth residues, ^{13}C -Gly is then lost as a neutral fragment, yielding ^{12}C -only b_7 . The relative abundance of these ^{12}C -only b ions increases, as one goes down the series from b_7 to b_5 . This is not surprising, as the loss of ^{13}C -Gly becomes statistically more likely for smaller fragments. All of these results are consistent with the hypothesis that cyclization in these peptides occurs, followed by sequence scrambling and the appearance of nondirect b ion.

Infrared Spectroscopy Results

b_2 ion

The IRMPD spectrum of b_2 ion (m/z 115) made from protonated triglycine with CID was recorded. To identify the structure of the ion, this spectrum is compared to theoretical spectra generated through a computational study. All the theoretical spectra in this chapter were calculated by Long Yu from the Polfer group. For this b_2 - G_3 fragment, a number of chemical structures have to be considered, including an *oxazolone* structure protonated on the N-terminus (*oxazolone N-prot*), an *oxazolone* structure protonated on the *oxazolone* ring N (*oxazolone ox-prot*), and a cyclic *diketopiperazine* structure protonated on a carbonyl O (*diketopiperazine O-prot*). Figure 3-4 presents a comparison of the experimental spectrum recorded in the mid-IR range

(1200–2000 cm^{-1}) with the calculated spectra of the lowest-energy conformer for each chemical structure, along with their structures and relative energies. Based on the prominent band centered at 1960 cm^{-1} , *oxazolone ox-prot* (Figure 3-3 B) can be identified unambiguously. Conversely, the *oxazolone N-prot* structure (Figure 3-3 C) gives a poorer match, suggesting that this structure either is not present at all or is at a much reduced abundance relative to *oxazolone ox-prot*. In fact, the predicted energy gap ($>23 \text{ kJ mol}^{-1}$) between both of these structures is considerable, supporting the claim that merely *oxazolone ox-prot* is populated at room temperature. For the *diketopiperazine O-prot* structure, there is no match between the experimental spectrum and the calculated spectrum, despite the fact this structure is marginally lower in energy than *oxazolone ox-prot*. Similar results have also been observed by others for related b_2 fragments, Ala-Ala (from Ala-Ala-Ala) by Oomens and Van Stipdonk and co-workers,⁵¹ Ala-Gly (from Ala-Gly-Gly) by Wysocki and co-workers,⁶⁴ and results by Paizs and Maitre and co-workers.⁵³ The only exception so far to this rule has been the very recent study of b_2 of His-Ala by Wysocki and co-workers, where the evidence from IR spectroscopy and HDX suggested a mixture of *oxazolone* and *diketopiperazine* structures.¹⁴⁷

To further confirm that b_2 from Gly-Gly-Gly does not adopt a diketopiperazine structure, the IRMPD spectrum of protonated cyclo(Gly-Gly), cyclo(Gly-Gly) H^+ , was recorded with FELIX. As shown in Figure 3-5, theory predicts the experimental spectrum qualitatively well, even if the scaling factor (0.98) is not optimal for all of the vibrations considered. Subtle differences in the interaction of the proton with the side-chain CH (Figure 3-5A), as opposed to the amide N-H (Figure 3-5B), can be

distinguished. The overlaid spectra of protonated cyclo(Gly-Gly) and $b_2\text{-G}_3$ are shown in Figure 3-6. It is clear that they adopt different structures. This result indicates that no diketopiperazine structure is present in $b_2\text{-G}_3$.

b_5 ion

As for the study of b_2 , the experimental mid-IRMPD ($1200\text{--}1940\text{ cm}^{-1}$) spectra for b_5 (m/z 286) generated from octaglycine, $b_5\text{-G}_8$ was recorded and then compared to the theoretical spectra in order to identify the structures. The candidate structures considered in this case are shown in Figure 3-7. The experimental spectra are compared with the two lowest-energy conformers for the various chemical structures, as shown in Figure 3-8.

The band at 1430cm^{-1} , which is the chemical diagnostic band for *macrocycle* structures, suggests that there is a *macrocycle* structure present in the b_5 ion population. This is also confirmed by the computational results (Figure 3-8), which predict the $\text{CO}\text{---}\text{H}^+$ bending mode to occur in this region. In the *macrocycle* structure, the proton is partially shared between two carbonyl sites. This results in a flat anharmonic potential, which is consistent with the broad nature of the $2500\text{--}2700\text{ cm}^{-1}$ band (not shown). In fact, shared-proton modes are often observed as broad features in IR spectra. For instance, in IR measurements by Johnson and co-workers on proton-bound dimers, the shared proton stretching band was found to be reasonably broad, even for these cold argon-tagged complexes formed in a supersonic expansion.^{148,149} For higher temperature proton-bound complexes, the IRMPD spectral features typically become very broad.¹⁵⁰⁻¹⁵³ The broadening of OH stretches as a consequence of strong H-bonding has even been observed in several neutral amino acid and peptide systems.^{154,155}

Oxazolone N-prot can be identified, based on the band at 1825 cm^{-1} , assigned to C=O stretch mode associated with the *oxazolone* ring moiety, which is in agreement with calculated spectra for this chemical structure. The corresponding *oxazolone* C=O stretch of *oxazolone ox-prot* is not observed, and this may be due to our limited scan to 1940 cm^{-1} . Proton attachment on a backbone carbonyl oxygen would also be possible in principle for an *oxazolone* structure; however, such structures are not likely to be substantially populated at room temperature, given their lower proton affinity and, hence, relatively high energetic penalty ($>30\text{ kJ mol}^{-1}$) determined from previous studies.³⁹

b₈ ion

IRMPD spectra of b₅-G₈ and b₈-G₈, are overlaid in Figure 3-9. Clear differences between the spectra can be seen, and these serve as useful guidance in the interpretation of the results. The spectrum for b₈ is clearly the simplest, lacking some of the spectral features present for b₅. For instance, the prominent band at 1775 cm^{-1} for b₅ is not observed for b₈. The weaker feature at 1825 cm^{-1} is observed for b₅, but not for b₈. Recently, Maitre and co-workers showed IR-MPD evidence that b₅ from G₅R exclusively gives rise to a *macrocycle* structure, based on the absence of *oxazolone* bands.⁵⁸ The positive identification of the *macrocycle* rests upon unambiguous assignment of modes associated to the proton attachment site on a backbone carbonyl, since this structure is chemically analogous to a peptide backbone and, therefore, lacks other diagnostic chemical moieties. In our comparison of b₅ and b₈, only b₈ is compatible with the exclusive presence of a *macrocycle* structure, as no *oxazolone* band is observed. Moreover, the presence of the band at 1430 cm^{-1} indicates the presence of the *macrocycle* structure, as this mode has previously been suggested to be due to the

CO—H⁺ (i.e., (C=)O—H⁺) bending mode of a cyclic peptide structure.³⁹

Summary of IRMPD results

From the IRMPD results for b₂, b₅ and b₈, it can be seen that there are a number of differences between these b ions. b₂ exclusively adopts *oxazolone ox-prot* structures. Conversely, b₅ appears to give a mixture of *oxazolone* and *macrocycle* structures. The absence of *oxazolone* band in IRMPD spectrum of b₈ indicates the exclusive presence of *macrocycle* structures.

Hydrogen–Deuterium Exchange (HDX) Experiments

HDX is employed as a method to quantify the relative amounts of *oxazolone* and *macrocycle* structures formed in CID, as a function of the b fragment size. Note that the fragment ions were made by nozzle-skimmer dissociation and that these CID products were accumulated in a hexapole prior to transfer to the ICR cell, as opposed to generating the fragments in the ICR cell by SORI CID (with the exception of b₂-G₃, as explained in the Experimental Section and Calculations). This approach presents a number of advantages over in-cell CID. Relatively large and constant number densities of CID product ions can be generated in this manner. Moreover, the pressure of the deuterating agent can be held at a constant pressure, without the need to pulse in a collision gas for SORI CID. Finally, this approach is expected to yield thermalized fragment ions more readily, given the higher-pressure environment in the storage hexapole. As the ion temperature is likely to affect the HDX kinetics, it is important to control the ion temperature in the interest of reproducibility.

The mass spectral distributions for different HDX times for the series b₂–b₈ are shown in Figure 3-10. In previous HDX studies, bimodal mass distribution in the mass spectra were observed for analytes with two or more isomeric structures, and this can

be explained by different exchange rates of the isomers.^{71,136} In all these b ions, no obvious bimodal distributions are seen here, however, a kinetic fitting of the HDX data may reveal the presence of more than one kinetic rate.

Kinetic Fitting of HDX Data

As discussed in the Experimental section, HDX can be considered as a pseudo-first-order reaction due to the fact that the amount of deuterating reagent is in great excess than the analyte inside the ICR cell. To determine pseudo-first-order HDX kinetics, the natural logarithm of the relative depletion of the undeuterated peak, $\ln[d_0/\sum d_n]$, is plotted as a function of the HDX time. The complete series of HDX measurements from b_2 to b_8 is summarized in Figure 3-11. A single kinetic rate ($k = 0.51 \text{ s}^{-1}$) is observed in the case of b_2 , whereas two distinct kinetic rates are required to fit the data for the larger b fragments (b_4 - b_7), with the notable exception of b_8 .

The reason for the appearance of two exchange rates may be the presence of two isomeric structures, which can exchange with CH_3OD at different rates that can be distinguished with FTICR mass spectrometer. These two structures can then be called “fast”- and “slow”-exchanging structures. By the assumption that the “fast”-exchanging population is fully depleted at longer HDX times, the slower exchange rate can be assigned to the “slow”-exchanging structure. Hence the faster rate is the combination of both “fast”- and “slow”- exchanging structures, and the rate of “fast”-exchanging structure can then be determined. A least-squares linear regression fit is employed to determine both the pseudo-first-order rate constant and the intercept. More conveniently, the depletion of d_0 can also be represented as remaining $d_0(\%)$ on a natural logarithm scale. It is then found that the intercept of the “slow”-exchanging linear regression fit represents the relative abundance of the “slow”-exchanging structure at

time zero. In the case of b_5 , the intercept equates to a relative abundance of 23% for the “slow”-exchanging structure at the beginning of the experiment. By default, the “fast”-exchanging structure accounts for the remainder (77%). Both “fast” and “slow” rates can be distinguished, as the difference in rate constant is almost an order of magnitude: $k_{\text{slow}} = 0.019 \text{ s}^{-1}$ vs $k_{\text{fast}} = 0.17 \text{ s}^{-1}$ ($= 0.19 - k_{\text{slow}}$). Table 3-4 summarizes the kinetic fitting results for b_2 - b_8 .

The exclusive presence of *oxazolone ox-prot* structures for b_2 , as confirmed by the IRMPD results, correlates well with the presence of one rate of HDX exchange, which happens to be “fast”. On the other hand, b_5 is shown to have a mixture of *oxazolone* and *macrocycle* structures, based on the IRMPD results, and this is confirmed by the presence of two distinct HDX rates, where one is “fast”- and the other one is “slow”-exchanging. Finally for b_8 , the IRMPD results are consistent with the exclusive presence of a *macrocycle* structure, while the HDX results for b_8 exhibit merely one rate of exchange, which happens to be “slow”. All of these results are consistent with the view that the “slow”-exchanging ions correspond to the *macrocycle* structure, whereas the “fast”-exchanging structure corresponds to the *oxazolone* structure.

The relative abundances of “slow”-exchanging structures from b_2 to b_7 are summarized in Figure 3-12, based on the kinetic fitting procedure, described above. The relative abundance of the “slow”-exchanging structure reaches 30% for b_7 , which indicates that the “fast”-exchanging structure still accounts for the majority of the ion population. Note that the error bars in are determined from the standard deviation of the intercept for the linear regression fit for the “slow”-exchanging structure.

The gradual increase of the relative amount of the *macrocycle* structure from b_4 to

b_7 shows that *macrocycle* formation becomes more favorable for larger structures. A sudden disappearance of the “fast” (= *oxazolone*) structure for b_8 is observed. A recent study by Harrison showed that b_9 ions of Tyr(Ala)₉, (Ala)₄Tyr(Ala)₅, and (Ala)₈TyrAla are exclusively *macrocycle*.⁵⁹ As the size gets larger, the *macrocycle* structures become more favorable, and this probably due to the fact that more flexible backbone makes it easier for the N-terminus to make the attack. It would be interesting to see what would be observed for larger b_n fragments. Preliminary experiments to generate larger b_n product ions ($n > 8$) from deca and dodecaglycine were not successful, instead resulting in dehydrated b ions, such as b_9 -H₂O and b_{10} -H₂O (not shown). Further experiments are required to determine under what conditions such fragment ions can be produced.

Chemical Basis for Differences in HDX Rates

Due to the obvious structural difference between *oxazolone* and *macrocycle* b fragment structures, it is expected that their proton affinities (PAs) should be different, and hence their HDX rates would differ. Previous studies have shown that *oxazolones* have higher PAs than *macrocycle* structures,^{39,42,156} due to the presence of more basic proton attachment sites (i.e., N-terminus and *oxazolone* ring N, vs backbone carbonyl). In detailed studies by Beauchamp and co-workers^{127,128} and Lebrilla and co-workers,^{157,158} it has been shown that a smaller difference in proton affinity between the deuterating reagent and the peptide results in faster HDX rates. The results for *oxazolone* and *macrocycle* b_5 seem to show the opposite effect, as the less basic *macrocycle* structure exchanges more slowly than the more basic *oxazolone* structure. These trends have also been observed by Wysocki and co-workers in their combined IRMPD and HDX study on HA b_2 fragments.¹⁴⁷ Clearly, differences in PAs between reagent and analyte ions are not the only parameter affecting HDX rates. Another

important parameter is the geometry of the analyte molecule, which enables (or inhibits) the gas-phase HDX mechanism to occur. In the case of low-basicity CH₃OD, HDX is thought to operate via a “relay” mechanism.^{122,128} This mechanism works on the premise that the deuterating molecule simultaneously hydrogen bonds to the proton and a basic site on the peptide. The proton/deuteron transfer then takes place in a concerted mechanism, where the proton is transferred to the deuterating agent, while the deuteron is transferred to the basic hydrogen-bonded site on the peptide. In light of the “relay” mechanism and the results above, it is plausible that a more flexible *oxazolone* structure more readily assumes the correct geometry for such a concerted mechanism, whereas a *macrocycle* structure is more prone to ring torsion strain, thus raising the transition state energy for this process.

Summary

In this chapter, the effect of chain length on peptide fragment structure formation is investigated using a range of gas-phase techniques. Using stable isotopic labeling, sequence scrambling is confirmed in the oligoglycine system. The data from IRMPD and HDX indicate that smaller b fragments (b₂, b₃) exclusively form *oxazolone* structures, whereas mid-sized fragments (b₄–b₇) display a mixture of *oxazolone* and *macrocycle* structures. Mixtures of both structures are in fact expected, given that a low-energy isomerization pathway is available to interconvert *oxazolone* and *cyclic b* fragment structures.^{56,159} This pathway is, however, apparently not available to smaller b fragments. In the case of b₈, the IRMPD and HDX results indicate an exclusive presence of *macrocycle* structures.

The occurrence of larger *macrocycle* structures in CID is particularly unsettling in

terms of sequence “scrambling”, since from a statistical point of view larger cyclic structures are more likely to open up at a different site than where they were originally put together. In fact, recent studies by Harrison⁵⁹ and Van Stipdonk¹⁴² seem to confirm that larger b fragments give rise to considerable sequence scrambling.

Complementary information from a range of gas-phase structural techniques was required to establish the qualitative and quantitative trends in the dissociation chemistry presented here. In particular, IR photodissociation spectroscopy allows identification of the chemical species that are generated, whereas H/D exchange enables quantification of the relative amounts that are made. Previous studies had shown that *oxazolone* structures could be readily identified based on the diagnostic *oxazolone* ring C=O stretch.^{39,51,52,160} Identification of the *macrocycle* can be established based on the chemically diagnostic CO-H⁺ bending mode (1430 cm⁻¹), even if this mode appears in a potentially more congested region of the IR spectrum. While IRMPD spectroscopy gives valuable information on the chemical structures of reaction products, such as the chemical moieties formed^{39,160,161} and the site of proton attachment,¹⁶²⁻¹⁶⁴ the IR-MPD yield cannot typically be related to relative abundances of structures in a mixture.

It is demonstrated here that the relative abundances of *oxazolone* and *macrocycle* can be inferred from the kinetic fitting of HDX data. The complementary information from these IR spectroscopy and HDX results allows identification of the *oxazolone* and *macrocycle* for b₅ fragment structures. The results from HDX kinetic analysis show that b₂ and b₈ display single exchange rate, characterized as “fast” and “slow”, respectively. Combining the results from IRMPD that b₂ is exclusively *oxazolone* and that b₈ is exclusively *macrocycle*, it is then clear that *oxazolone* structures display HDX rate

constants that are an order of magnitude higher than those of *macrocycle* structures. In addition, the relative abundance of “slow”-exchanging structure, or *macrocycle*, can be deduced by the intercept from the kinetic fitting of “slow” exchanging. And hence the trend of the formation of *macrocycle* structures is then established as a function of b fragment ion size.

Table 3-1. Mass-to-charge ratios of all b ions of interested.

b ions	Precursor peptide	Mass-to-charge ratio (m/z)
b ₂	(Gly) ₃	115
b ₃	(Gly) ₈	173
b ₄	(Gly) ₅	229
b ₅	(Gly) ₅ , (Gly) ₈	287
b ₆	(Gly) ₈	344
b ₇	(Gly) ₈	401
b ₈	(Gly) ₈	458

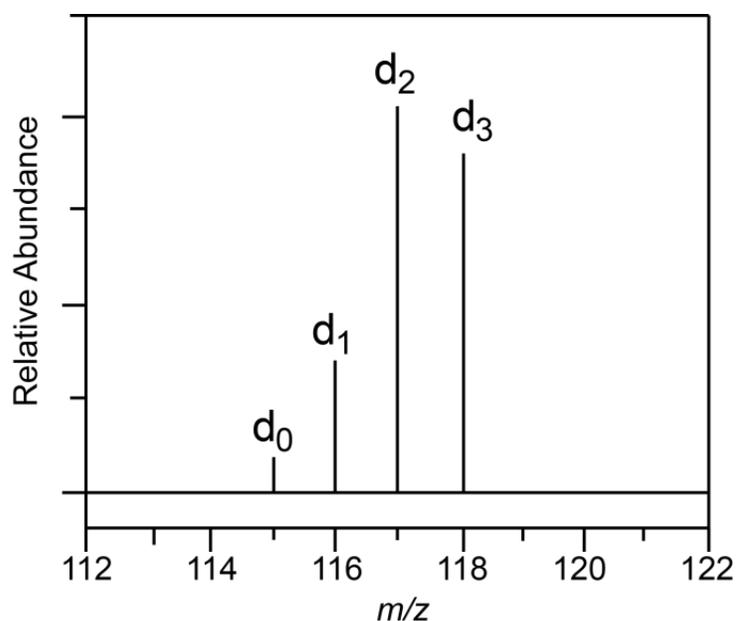


Figure 3-1. HDX mass spectrum with HDX products labeled.

Table 3-2. Photofragments of b₂, protonated cyclo(Gly-Gly) and b₅.

b ions	Photofragments
b ₂ -G ₃	m/z 87 (a ₂)
Cyclo(Gly-Gly)H ⁺	m/z 87 (a ₂) and m/z 59 (a ₂ -H ₂ O)
b ₅ -G ₅	m/z 125 and m/z 154 (b ₃ -H ₂ O)

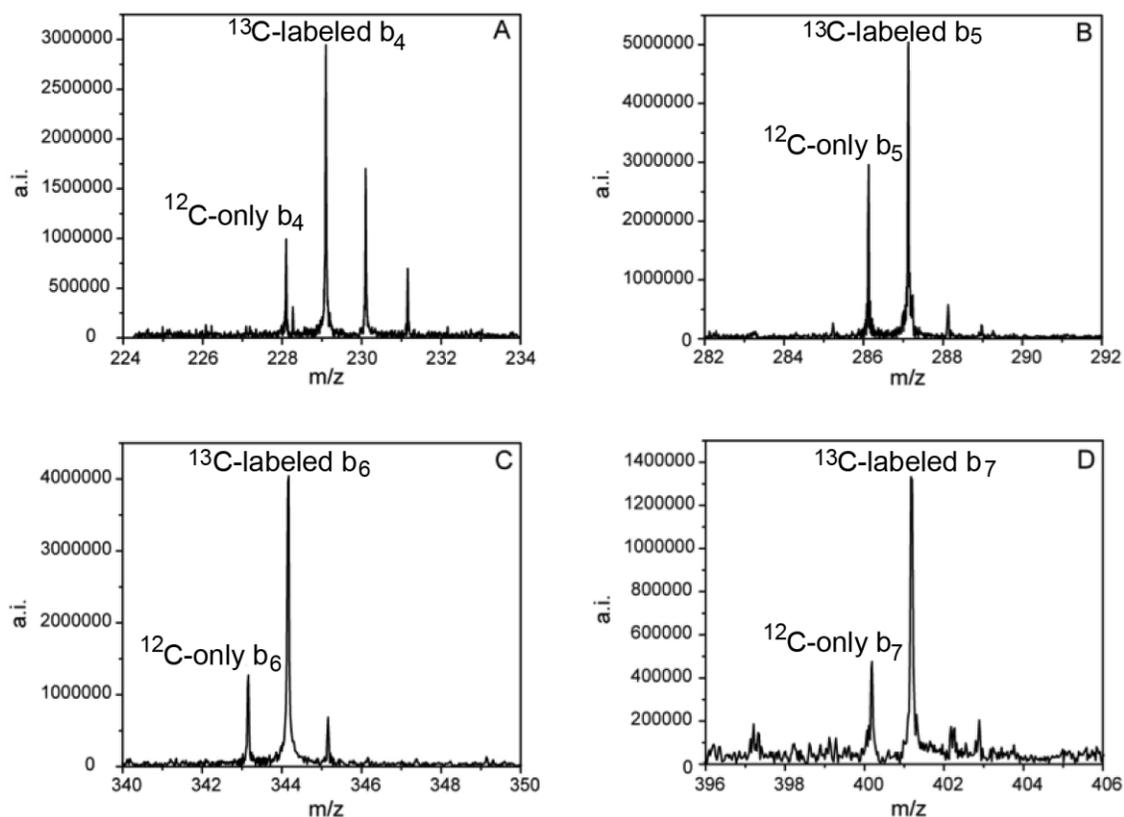


Figure 3-2. Inserts from the nozzle-skimmer CID spectrum of octaglycine labeled with a ^{13}C -Gly as the N-terminal glycine residue, showing the b isotope distributions for (A) b_4 , (B) b_5 , (C) b_6 and (D) b_7 . The ^{13}C -labeled b_n peaks denote b ions that incorporate the ^{13}C -Gly label, whereas ^{12}C -only b_n peaks are entirely composed of ^{12}C -Gly residues.

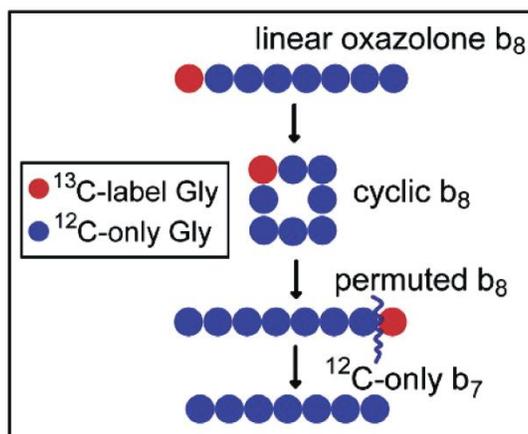


Figure 3-3. The cartoon mechanism rationalizes the appearance of the ^{12}C -only b_7 peak.

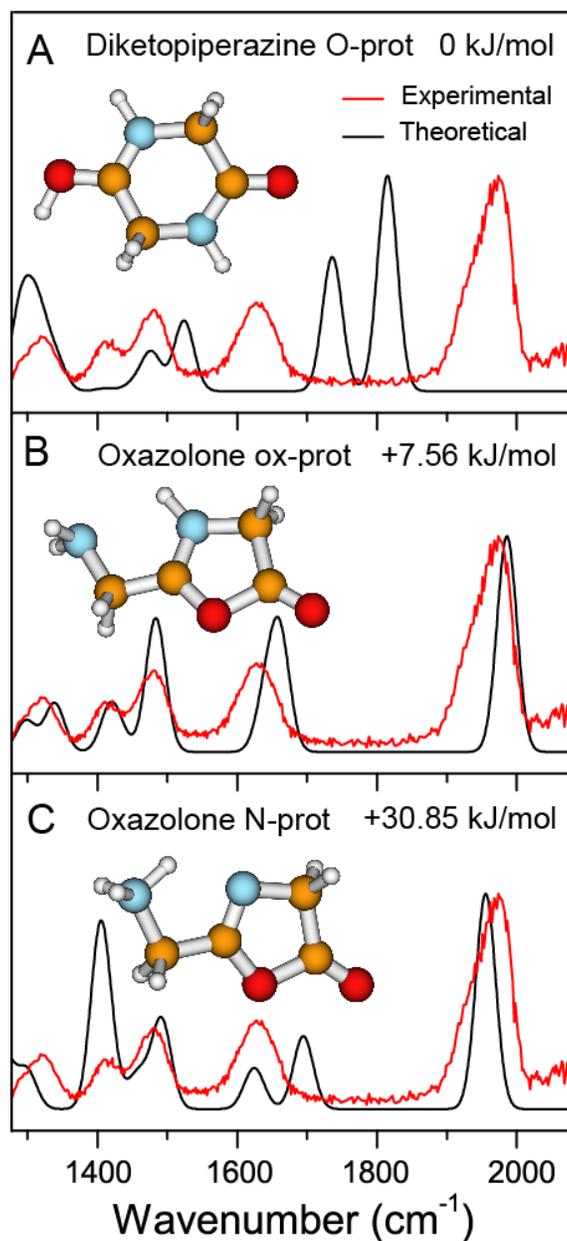


Figure 3-4. IRMPD spectrum of the b_2 fragment generated from Gly-Gly-Gly, compared to computed spectra for (A) diketopiperazine structure protonated on a carbonyl O, (B) *oxazolone* structure protonated on the *oxazolone* ring N, and (C) *oxazolone* structure protonated on the N-terminus.

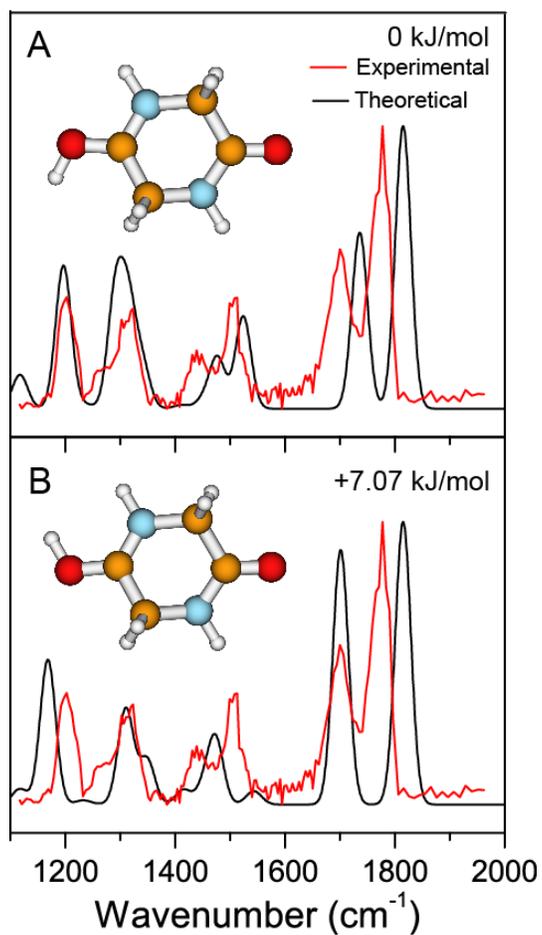


Figure 3-5. IRMPD spectrum of protonated cyclo(Gly-Gly), compared to calculated spectra for (A) *diketopiperazine* with proton pointing to CH₂ group and (B) *diketopiperazine* with proton pointing to the amide N-H group.

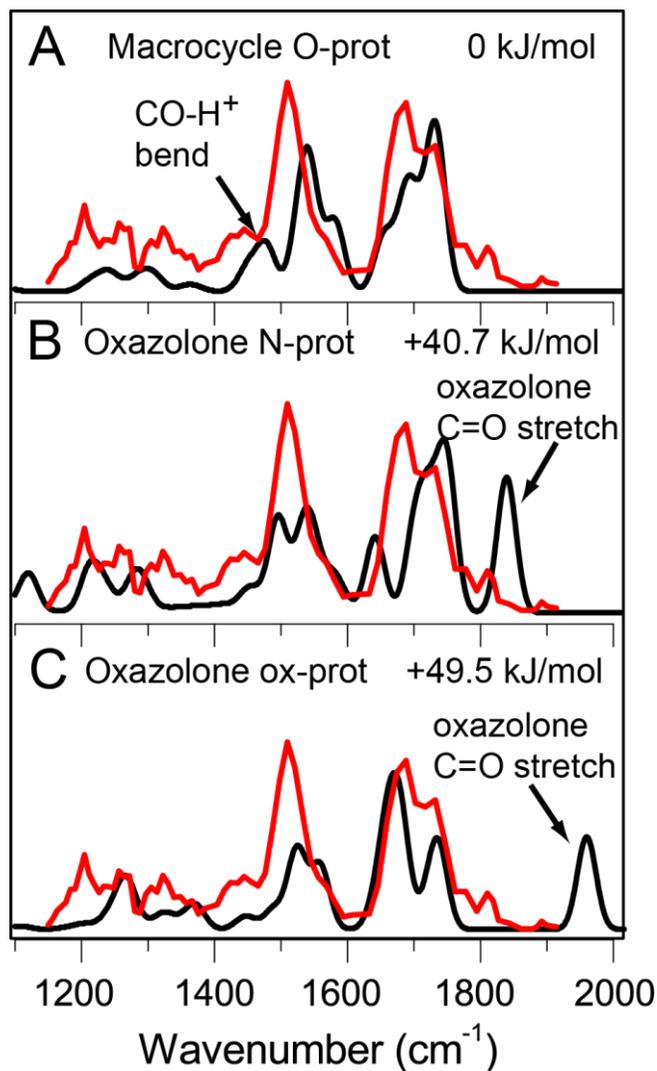


Figure 3-8. IRMPD spectrum of b_5 generated from octaglycine, compared to the two lowest-energy conformers for the various chemical structures: (A) *macrocycle* structure protonated on backbone carbonyl, (B) *oxazolone* structure protonated on N-terminus, (C) *oxazolone* structure protonated on *oxazolone* ring N. The chemical diagnostic bands and the relative energies to the lowest conformer are indicated.

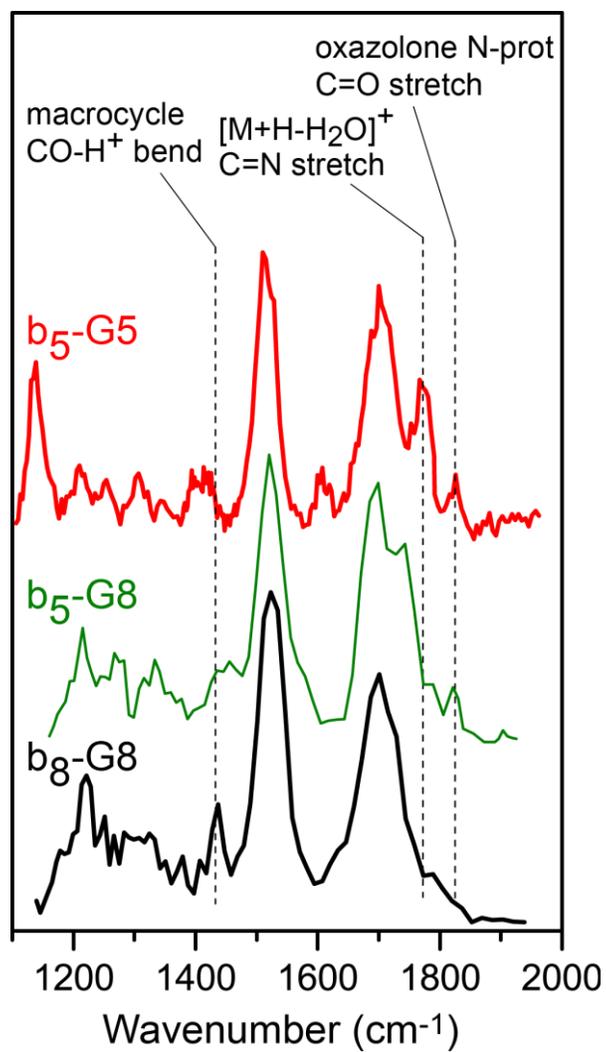


Figure 3-9. Overlaid mid-IRMPD spectra of b₅-G₈, b₅-G₅ and b₈-G₈. The chemically diagnostic modes are indicated.

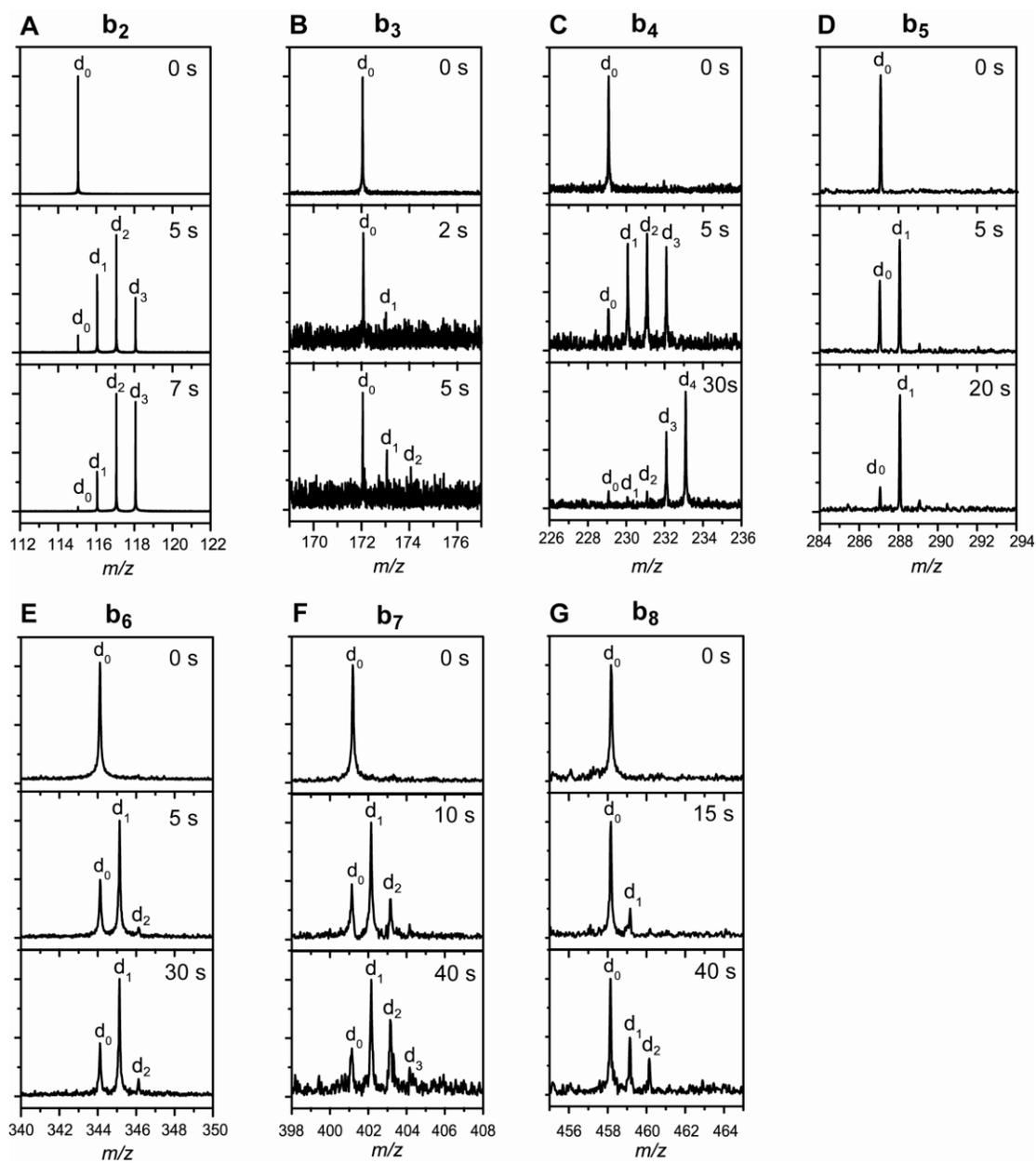


Figure 3-10. Representative H/D exchange (10^{-8} Torr CH_3OD) mass spectra for (A) b_2 generated from tri-glycine, (B) b_3 generated from octa-glycine, (C) b_4 generated from penta-glycine, (D) b_5 generated from octa-glycine, (E) b_6 generated from octa-glycine, (F) b_7 generated from octa-glycine, and (G) b_8 generated from octa-glycine.

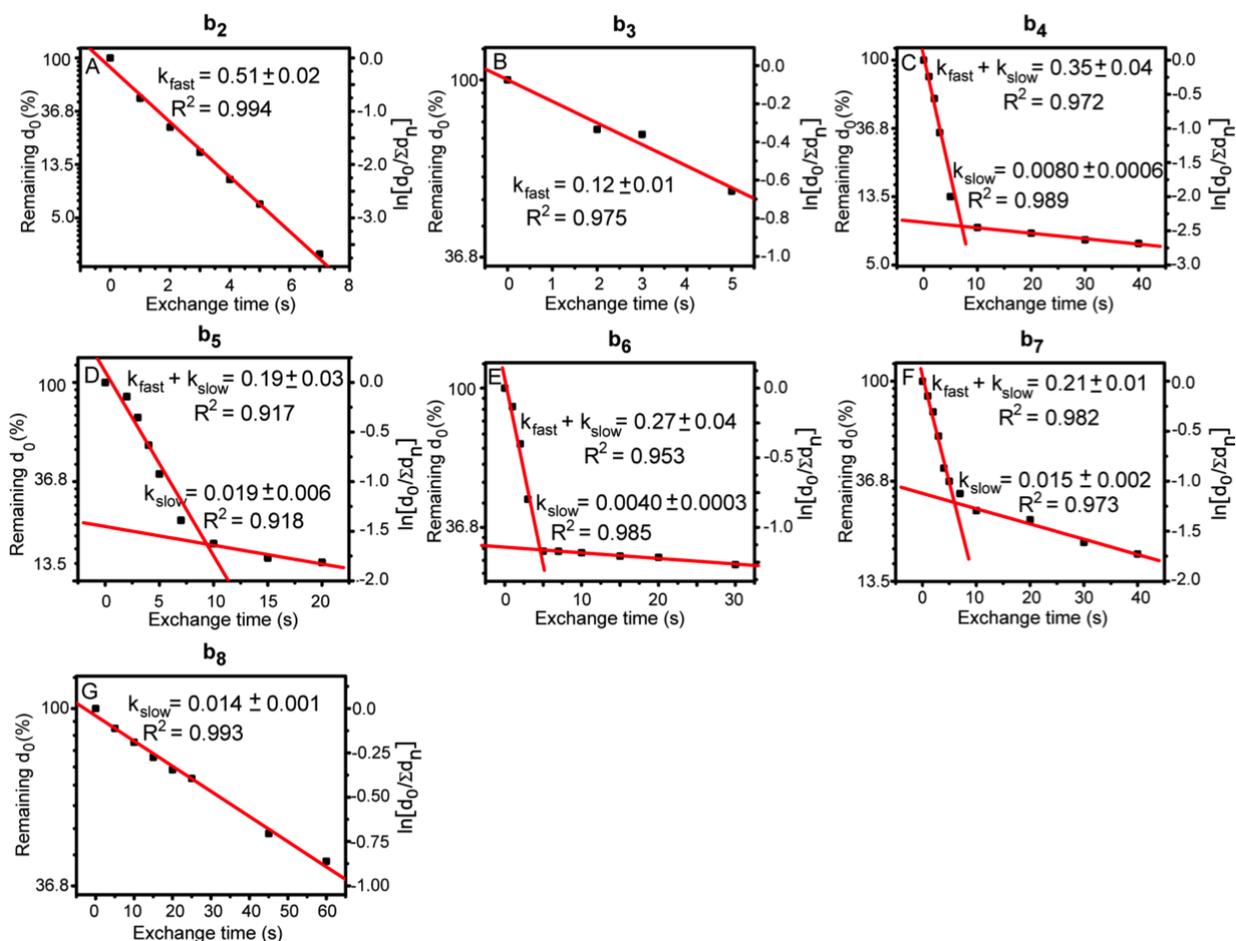


Figure 3-11. Kinetic fitting of the HDX results for glycine-based b fragment ions. (A) b_2 generated from triglycine, (B) b_3 generated from octaglycine, (C) b_4 generated from pentaglycine, (D) b_5 generated from octaglycine, (E) b_6 generated from octaglycine, (F) b_7 generated from octaglycine, and (G) b_8 generated from octaglycine. (H) Relative abundance of “slow”-exchanging structure as a function of b_n fragment size.

Table 3-3. Kinetic fitting results for the $\ln[d_0/\Sigma d_n]$ plots vs. H/D exchange time for the oligoglycine fragments b_2 - b_8 .

	b_n	b_2	b_3	b_4	b_5
Fast -exchanging structure	slope	-0.513	-0.122	-0.350	-0.185
	Error in slope	0.018	0.014	0.042	0.032
	intercept	-0.177	9.53E-4	0.0609	0.109
	Error in intercept	0.070	0.0427	0.0784	0.106
	R	-0.997	-0.987	-0.986	-0.958
	Standard deviation about regression	0.108	0.0499	0.0938	0.124
Slow -exchanging structure	slope			-0.00795	-0.0189
	Error in slope			6E-4	0.0056
	intercept			-2.38	-1.45
	Error in intercept			0.016	0.09
	R			-0.994	-0.958
	Standard deviation about regression			0.0134	0.0400

	b_n	b_6	b_7	b_8
Fast -exchanging structure	slope	-0.266	-0.212	
	Error in slope	0.042	0.014	
	intercept	0.0667	0.0506	
	Error in intercept	0.0785	0.0426	
	R	-0.976	-0.991	
	Standard deviation about regression	0.0938	0.0589	
Slow -exchanging structure	slope	-0.00403	-0.986	-0.0142
	Error in slope	3E-4	0.002	5E-4
	intercept	-1.14	-1.12	-0.0409
	Error in intercept	0.01	0.05	0.0145
	R	-0.992	-0.986	-0.997
	Standard deviation about regression	0.00524	0.0400	0.0266

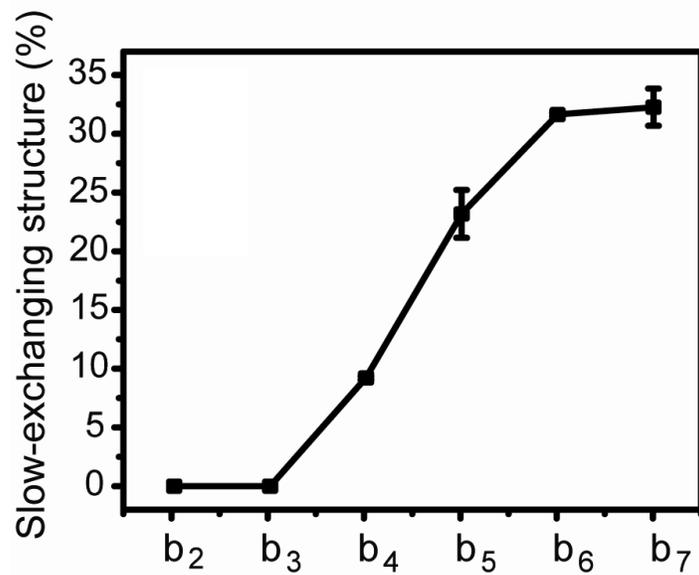


Figure 3-12. The relative abundance of slow-exchanging structure as a function of b fragment size.

CHAPTER 4
STRUCTURAL ANALYSIS FOR LEU-ENKEPHALIN B₂-B₄ WITH INFRARED
MULTIPLE-PHOTON DISSOCIATION SPECTROSCOPY AND GAS-PHASE
HYDROGEN/DEUTERIUM EXCHANGE*

Background

In the last chapter, the dependence of *macrocycle* structure formation on the chain length of b fragment ions has been studied, and the system used is oligoglycines where all the amino acid residues are the same. In this chapter, another peptide, Leu-enkephalin (Try-Gly-Gly-Phe-Leu), will be investigated. Leu-enkephalin is a pentapeptide that has been investigated in numerous studies. The first direct evidence for a mixture of *oxazolone* and *macrocycle* structures comes from infrared multiple-photon dissociation (IRMPD) spectroscopy by Polfer et al. on b₄ generated from YGGFL.³⁹ *Oxazolone* and *macrocycle* structures are identified based on diagnostic vibrations, involving the *oxazolone* ring C=O stretch and *macrocycle* CO-H⁺ bending. Rearrangement processes were also seen for an a-type fragment in Leu-enkephalin by Vachet and Glish.¹⁶⁵ Structure of Try-Gly-Gly-Phe-Leu is as shown in Figure 4-1. The combined approach of IRMPD spectroscopy and gas-phase HDX is hence a suitable solution to this problem, as shown in the previous chapter. We apply this complementary methodology here to study the b₂-b₄ CID products of Leu-enkephalin. Both techniques are found to be highly complementary, as they cancel out each others' weaknesses. While IRMPD spectroscopy confirms the chemical structures that are formed, HDX allows their quantification.

* Reproduced in part with permission from Chen, X.; Steill, J.D.; Oomens, J.; Polfer, N.C. Oxazolone vs. Macrocycle Structures for Leu-Enkephalin b₂-b₄: Insights from Infrared Multiple-Photon Dissociation Spectroscopy and Gas-Phase Hydrogen/Deuterium Exchange, *J. Am. Soc. Mass Spectrom.* 2010, 21, 1313-1321. Copyright 2010 American Society for Mass Spectrometry.

Experimental

Sample Preparation

The pentapeptide Leu-enkephalin (Tyr-Gly-Gly-Phe-Leu) (Sigma Aldrich, St. Louis, MO) and deuterating reagent CH₃OD (Sigma Aldrich, St. Louis, MO) were employed without further purification. Leu-enkephalin was used as 20 μM solution in 50:50 (vol:vol) water/methanol with 2% acetic acid to aid protonation.

Mass Spectrometry and Hydrogen/Deuterium Exchange (HDX)

The hydrogen/deuterium exchange (HDX) experiments were carried out at the University of Florida using a commercially available 4.7 Tesla actively shielded Bruker Bioapex II Fourier transform ion cyclotron resonance (FT-ICR) mass spectrometer equipped with an Apollo API 100 source (Bruker Daltonics, Billerica, MA) at Dr. David Powell's lab, as reported in Chapter 3. The singly protonated Leu-enkephalin (m/z 556) was ionized by electrospray ionization (ESI). "Nozzle-skimmer" dissociation was applied in the ESI source region to generate CID product ions. The produced ions were then accumulated in the hexapole for 3s, prior to transfer to the ICR cell. A frequency sweep was performed in the FT-ICR experimental sequence to eject other ions and to mass-isolate each desired b product ion (b_2 - b_4). The mass-to-charge ratios of b_2 , b_3 and b_4 are 221, 278, and 423, respectively. The monoisotopic peak of each species was then subjected to gas-phase hydrogen/deuterium exchange (HDX) with CH₃OD in the ICR cell for different time periods. CH₃OD was leaked into the vacuum chamber using a Varian leak valve to attain a constant pressure of 1×10^{-8} Torr. Note that CH₃OD was degassed using several freeze-thaw cycles before introduction into the mass spectrometer.

Mass spectra with different HDX times were recorded. The data are represented here by plotting the natural logarithm of the ratio of d_0 divided by the sum of all ions, $\ln[d_0/\sum d_n]$, as a function of the HDX time, as described previously. The depletion of d_0 is also represented as a percentage, which is plotted on a natural logarithm scale. Because the exchange reagent is considered to be in great excess of the analyte, the H/D exchange reaction can be approximated as a first-order reaction in the analyte concentration. Linear fits were performed to determine the reaction kinetics.

Mass Spectrometry and Infrared Photodissociation Spectroscopy

The infrared photodissociation experiments were performed at the FOM institute for Plasma Physics “Rijnhuizen” using the free electron laser FELIX. A laboratory-built FT-ICR mass spectrometer described in detail previously was employed for the mass spectrometry measurements. The fragment ions (b_2 , b_3) were generated by “nozzle-skimmer” dissociation. The infrared spectrum for b_4 is taken from a previously published study³⁹. The fragment ion was accumulated in the hexapole, and then transferred to the ICR cell. The ion of interest was mass-selected and irradiated with the tunable output from the free electron laser.

The IR photodissociation spectrum was recorded by monitoring the infrared multiple-photon dissociation (IRMPD) yield as a function of the wavelength (here 1300–1975 cm^{-1}). This yield is represented with the equation (3-2). The yield is further normalized linearly with FELIX laser power at each wavelength step. The main photodissociation product of b_3 (m/z 278) was found to be m/z 221 (b_2), as well as a minor fragment at m/z 193 (a_2). For b_2 (m/z 221), the m/z 193 (a_2) photofragment was most abundant.

Computations

Calculations were carried out at the High-Performance Computing (HPC) Center at the University of Florida using the AMBER force field¹⁶⁶ and the Gaussian03 package.¹⁶⁷

Candidate structures for b_2 fragments were generated using an in-house developed method, involving conformational searching by molecular mechanics and frequency calculations by density functional theory (DFT). The chemical structures (*oxazolone* and *diketopiperazine*) were initially built and optimized using semi-empirical methods (AM1) in HyperChem (Hypercube Inc., Gainesville, FL). For *oxazolones*, two proton attachment sites were considered, at the N-terminus and *oxazolone* ring nitrogen. For the *diketopiperazine* structure, merely the backbone carbonyl oxygen was considered.

The chemical input structures for the molecular mechanics calculations required a parameterization procedure. The AM1 structures were optimized using DFT (B3LYP/6-31 g*), followed by an electrostatic potential fitting with ab initio methods (HF/6-31 g*) to derive the atomic point charges. The geometry and ESP-derived charges were imported into the AMBER suite of programs, where a restrained electrostatic potential (RESP) fitting was performed.¹⁶⁸ Each chemical structure was parameterized separately in AMBER, followed by a conformational search using simulated annealing cycles. Two separate runs with starting temperatures at 300 and 500 K were carried out, resulting in 200 candidate structures per dynamics simulation.

All output structures from AMBER were then optimized again at the DFT level, initially using B3LYP/3-21. Further optimization was performed at the B3LYP/6-31G(d). Merely the 20 unique lowest-energy conformers were then optimized at B3LYP/6-

31G+(d,p), followed by a single-point MP2/6-31G+(d,p) calculation. The MP2 electronic energy for each conformer was corrected for the zero-point energy (ZPE) from B3LYP/6-31G+(d,p) of theory to yield the final ZPE-corrected energies. Note that MP2 energies were considered to account for the dispersion interaction involving the tyrosine side-chain. All energies are presented here are relative to the lowest-energy conformer. The frequency spectra of the lowest-energy structures at the B3LYP/6-31G+(d,p) level were scaled by 0.965.¹⁶⁹ Stick spectra were convoluted using a 20 cm⁻¹ full width at half-maximum (fwhm) Gaussian profile to allow easier comparison with the recorded IR photodissociation spectra.

Results and Discussion

Infrared Spectroscopy Results

IRMPD spectrum of the b₂ fragment (*m/z* 221), generated by “nozzle-skimmer” CID from protonated Leu-enkephalin, was recorded in the mid-IR range (1300-1975 cm⁻¹). To structurally analyze the ion, the experimental spectrum is compared to theoretical spectra of a number of candidate chemical structures. For theoretical study, the chemical structures considered include an *oxazolone* structure protonated on the N-terminus (*oxazolone* N-prot), an *oxazolone* structure protonated on the *oxazolone* ring N (*oxazolone* ox-prot), and a *diketopiperazine* structure protonated on a carbonyl O (*diketopiperazine* O-prot). Figure 4-2 displays a comparison of the experimental results with the calculated spectra of the lowest-energy conformer for each chemical structure, along with their structures and relative energies. As explained in the experimental section, the electronic energies (MP2) are corrected for the zero-point energy (ZPE) at

B3LYP/6-31g+(d,p). Both stick spectra and convoluted Gaussian profiles are shown for easier comparison. The results from theoretical study are shown in Table 4-1.

It is clear that there is no match between the experimental spectrum and the diagnostic modes of protonated *diketopiperazine* structure, despite the fact that this structure is lowest in energy. The prominent C=O (1810 cm^{-1}) and C-N (1715 cm^{-1}) stretches are clearly not observed. This result suggests that no *diketopiperazine* structure is present in b_2 of Try-Gly, and this is similar to the result from Chapter 3 that b_2 made from triglycine adopts exclusively *oxazolone* structure. Similar results have been observed by others for b_2 fragments in different systems, Ala-Ala by Oomens and Van Stipdonk and co-workers,⁵¹ and Ala-Gly by Wysocki and co-workers,⁵² and a number of b_2 fragments from tryptic digest peptides by and co-workers⁵³. This indicates that the kinetic barrier to a nucleophilic attack from the N-terminus is too high, and hence the *oxazolone* formation pathway is favored for kinetic reasons, as predicted by Paizs.¹⁷⁰

The broad band at around 1900 cm^{-1} clearly shows one main peak and a prominent shoulder, which is consistent with the presence of *oxazolone* N-prot and ox-prot structures. The peak at lower-frequency is assigned to the *oxazolone* C=O stretch of the N-prot structure, whereas the higher-frequency shoulder matches the corresponding band of the *oxazolone* ox-prot structure. The predicted energy gap, calculated at the MP2 (Møller-Plesset) level ($\sim 11\text{ kJ mol}^{-1}$) is somewhat large to account for the presence of both structures, and hence this may be due to slight inaccuracies in the calculation. Nonetheless, the computations favor the N-prot *oxazolone*, which also seems to be validated by the significantly more intense band for

this structure in the IRMPD spectrum. This is in strong contrast to the b_2 fragment involving primary structures Gly-Gly, Ala-Gly and Ala-Ala, where the ox-prot structure mainly accounts for the structures that are observed. The structures indicate that the tyrosine side-chain in b_2 Try-Gly plays an important role in stabilizing the N-prot *oxazolone*, due to interaction between the proton and the tyrosine π -cloud. In order to account for these interactions more accurately, the energies are computed at the MP2 level, which includes dispersion interactions.

The overlaid experimental mid-IRMPD ($1300\text{-}1975\text{ cm}^{-1}$) spectra for b_2 (m/z 221), b_3 (m/z 278) and b_4 (m/z 425) are shown in Figure 4-3 for comparison. Note that only b_2 and b_3 were recorded in this study, whereas b_4 originates from the study by Polfer et al.³⁹ It is clear that there are some differences between these spectra. The chemical interpretations of the spectral bands are indicated and are based on comparisons with DFT-calculated spectra for b_2 and b_4 . The bands in the region $1780\text{-}1940\text{ cm}^{-1}$ are chemically diagnostic modes for *oxazolone* C=O stretch modes, as these consistently appear in a higher frequency region than the amide C=O stretch. It can be seen that the position of the *oxazolone* C=O stretch is progressively red-shifted in larger b fragments, due to an increase in hydrogen-bonding interactions. b_2 and b_4 exhibit multiple *oxazolone* bands due to different protonation sites, while b_3 merely appears to show one band, and this suggests that only one site of proton attachment is favored in b_3 . For identification of *macrocycle* structures, the *macrocycle* CO-H⁺ bending mode is typically observed at $\sim 1445\text{ cm}^{-1}$.³⁹ This is rarely a pure mode, exclusively due CO-H⁺ bending of a proton shared between two carbonyls. In the IRMPD spectrum of b_2 , where such a CO-H⁺ bending mode is not possible the band at 1445 cm^{-1} corresponds to CH₂ bending

on the ring. Nonetheless, the proton bending mode lends brightness to the band at this position. While this band is found to be relatively weak in b_3 , its relative intensity is much enhanced in b_4 . This is consistent with the hypothesis that b_4 contains a higher proportion of *macrocycle* structures than b_3 .

In summary, the IR spectra provide strong evidence that b_2 is exclusively composed of *oxazolone* structures, whereas a mixture of *oxazolone* and *macrocycle* structures are observed for b_4 . For b_3 , the identification of *oxazolone* is unambiguous, whereas conclusion for *macrocycle* is difficult to make due to the weak band at 1450 cm^{-1} and possible overlap between CO-H⁺ bending and *oxazolone* ring CH₂ bending. The complementary technique of HDX will be employed to shed more information on the chemical structures that are present.

Hydrogen-Deuterium Exchange (HDX) Experiments

The mass spectral distributions for different HDX times for the b_2 , b_3 and b_4 are shown in Figure 4-4. There are 4 exchangeable hydrogens in b_2 , as illustrated in Figure 4-5. It can be seen that only 3 exchanges occurred in b_2 . For the *oxazolone* N-prot and ox-prot structures, this suggests that the nitrogen-bound hydrogens are exchanged, whereas the side-chain tyrosine OH is not. Similar phenomena are also found for b_3 and b_4 , which show one less exchange than the number of exchangeable hydrogens. These results are consistent with the picture that the tyrosine OH is much less labile to HDX with CH₃OD. This also supports the hypothesis that the protonated site is the “site of entry” for the deuterium, as proposed by the “relay” mechanism.^{128,171} A full deuteration of the peptide fragment would then take place via a “mobile” deuterium, which shuttles to NH amide sites on the backbone. Within this model, the deuterium does not exchange for

the tyrosine OH, and hence HDX is primarily sensitive to the chemical structure of the peptide fragment (i.e, sites of proton attachment).

Similarly to previous HDX studies,^{71,136} bimodal distributions are observed for b_3 and b_4 . On the other hand, b_2 merely displays one distribution. These results are consistent with the findings from IRMPD above, where b_2 is exclusively composed of *oxazolones*, whereas b_4 is made up of a mixture of *oxazolone* and *macrocycle*. To calculate the relative abundances of these structures, kinetic fitting for these HDX data is required.

Kinetic Fitting of HDX Data

To determine pseudo-first-order HDX kinetics, the natural logarithm of the relative depletion of the undeuterated peak, $\ln[d_0/\sum d_n]$, is plotted against the HDX time for b_2 , b_3 and b_4 , as shown in Figure 4-6. A single kinetic rate ($k = 0.43 \text{ s}^{-1}$) is observed for b_2 , while two distinct kinetic rates are required to fit the data for b_3 and b_4 . A least-squares linear regression fit is employed to determine both the pseudo-first-order rate constant and the intercept. As introduced in the last chapter, the rate of the “slow”-exchanging population can be determined accurately, assuming that the “fast”-exchanging population is fully depleted at longer times. The relative abundances of these “slow”-exchanging structures can be approximated by the intercept of the “slow”-exchanging reaction.

The rates and abundances from the HDX kinetic fitting results are summarized in Table 4-2, and Table 4-3 summarize the kinetic fitting results. In the case of b_3 , the intercept equates to a relative abundance of ~6% for the “slow”-exchanging structure at the beginning of the experiment. Note that the large error bars in this case ($\pm 34\%$) are

due to the low ion abundance for b_3 . This implies that the “fast”-exchanging structure accounts for the remainder (i.e., 94%). Note that the higher rate at shorter HDX times in Figure 4-6 corresponds to the combined depletion rates of the “fast”- and “slow”-exchanging structures. Both “fast” and “slow” rates can be distinguished, as the difference in rate constant is more than an order of magnitude: $k_{\text{slow}} = 0.019 \text{ s}^{-1}$ vs $k_{\text{fast}} = 0.40 \text{ s}^{-1}$ ($= 0.42 - k_{\text{slow}}$). Similarly, the abundance of the “slow” structure in b_4 is approximated at ~31%, compared to 69% for the “fast” structure.

The exclusive presence of *oxazolone* structures for b_2 correlates well with a single HDX rate. Conversely, the presence of “fast” and “slow” rates for b_3 and b_4 suggest the presence of two distinct chemical structures. Given the unambiguous identification of the *oxazolone* structure for b_3 , but more tenuous identification of the *macrocycle*, this strongly suggests that the majority “fast” structure corresponds to the *oxazolone*, whereas the “slow” structure is related to the *macrocycle*. In fact, the “fast” rate for b_2 and b_3 are nearly identical (0.43 vs. 0.42), basically confirming that the “fast” structure corresponds to the *oxazolone*. For the “slow” structure, there is a significant increase from b_3 (6%) to b_4 (31%). This trend is in agreement with the increase in intensity of the 1440 cm^{-1} band from b_3 to b_4 . All of these results support the hypothesis that the “slow”-exchanging structure corresponds to the *macrocycle*, while the “fast”-exchanging structure is the *oxazolone*. The same trends were observed in the study of size effect in Chapter 3. Moreover, in a recent study by Wysocki and co-workers on His-Ala b_2 , *oxazolone* and *diketopiperazine* structures were considered.¹⁴⁷

Comparison with Oligoglycines Study

There are some differences between the present study and the oligoglycine b fragment study. For oligoglycine b_2 - b_8 , a general categorization into “fast”- and “slow”-

exchanging structures could be established. Such a categorization is less straightforward here, as the magnitude for k_{fast} and k_{slow} drop by a factor of ~ 10 from b_3 to b_4 (see Table 4-2). The slower kinetics for b_4 are possibly due to the bulky phenylalanine side-chain, which affects the H/D exchange as a result of steric hindering. Conversely, for oligoglycine b fragments, no such side-chain effects are expected. Fortunately, the relative ratio in HDX rates between “fast” and “slow” remains similar for b_3 and b_4 , thus the separation of b_3 and b_4 fragment ions can be done through the kinetic analysis.

Another difference between both studies relates to the abundance of the “slow”-exchanging structure (i.e., *macrocycle*) for b_4 , which appears to be considerably enhanced in Leu-enkephalin compared to pentaglycine (31% vs. 9%). This difference must be related to the subtle primary structure differences between Tyr-Gly-Gly-Phe and Gly-Gly-Gly-Gly, which shows that the primary structure affects the propensity for “head-to-tail” cyclization. And this leads to the investigation of primary structure effect, which will be discussed in Chapter 5.

An important realization in the comparison between the IRMPD and HDX results for Leu-enkephalin is that the appearance of two distinct kinetics rates in the HDX experiments is only related to the presence of two considerably different chemical structures (i.e., *oxazolone* and *macrocycle*), not to the presence of different protonation sites of the same chemical structure. For b_2 , IRMPD measurements confirm the presence of N-prot and ox-prot *oxazolones*; this, however, only results in a single kinetic HDX rate. On the other hand, the presence of a mixture of *oxazolone* and *macrocycle* in b_3 and b_4 does result in the observation of two distinct HDX rates. This suggests that in the structural characterization of these b fragments, HDX is not sensitive to the site of

proton attachment as such, but rather the chemical structure. This observation is consistent with the “relay” mechanism in HDX of low-basicity deuterating reagents (such as CH_3OD ¹²⁸). As previously discussed, the transition state for HDX is likely to be higher for a *macrocycle* compared to an *oxazolone* structure, due to the ring strain in the *macrocycle* structure to allow this mechanism to take place.

Recently, Solouki and co-workers employed the more basic deuterating reagent ND_3 to characterize b fragments, where they also distinguished “fast”- and “slow”-exchanging structures.⁷² This observation is intriguing, since the mechanism of HDX for higher-basicity HDX reagents, such as ND_3 , is thought to be different (‘onium’ mechanism). This shows that a number of approaches are possible in the HDX characterization of b fragments, even if it remains to be seen whether the interpretation is always unambiguous.

Summary

In this study, we have applied IRMPD spectroscopy and gas-phase H/D exchange to the characterization of b_2 - b_4 from Leu-enkephalin. IRMPD was used to qualitatively identify structures, whereas HDX was employed to quantify the structures. For b_2 , by comparing the measured IRMPD spectrum (1300-1975 cm^{-1}) to theoretical spectra, the *diketopiperazine* structure was excluded based on the fact that the prominent C=O (1810 cm^{-1}) and C-N (1715 cm^{-1}) stretches were not observed. Conversely, the characteristic *oxazolone* C=O stretch modes at $\sim 1900 \text{ cm}^{-1}$ allowed identification of N-prot and ox-prot *oxazolones*. The exclusive presence of one chemical structure (i.e., *oxazolone*) correlates well with the presence one rate constant in the HDX measurements of b_2 , which happens to be “fast”. An overlay of IRMPD spectra of b_2 , b_3 and b_4 shows evidence for both *macrocycle* and *oxazolone* structures in b_3 and b_4

based on vibrations at 1440 cm^{-1} (CO-H⁺ bending of *macrocycle*) and $>1770\text{ cm}^{-1}$ (*oxazolone* C=O). This is in agreement with the presence of two distinct HDX rates and bimodal distributions in the corresponding HDX mass spectra. Using a recently the kinetic fitting analysis, as shown in Chapter 3 for oligoglycine b CID products, the relative abundances of the "slow"- and "fast"- exchanging structures were inferred from kinetic fitting of the HDX data. Similarly to the more extensive study on oligoglycine b₂-b₈, Leu-enkephalin b₂-b₄ show an increase in the relative abundance of the *macrocycle* structure with fragment size. In fact, the relative abundance for *macrocycle* b₄ for the sequence Try-Gly-Gly-Phe is considerably larger than for Gly-Gly-Gly-Gly (31% vs. 9%). This shows that apart from the chain length, the primary structure also plays a key role in the relative propensity in forming *oxazolone* vs. *macrocycle* structure. While more complementary IRMPD/HDX studies are required to validate the hypothesis that *macrocycle* structures exhibit slower HDX kinetics compared to *oxazolones*, these results show that HDX is a promising technique in quantifying both structures. In fact, HDX is much sensitive relative to IRMPD in detecting low-abundance structures, such as the *macrocycle* for b₃.

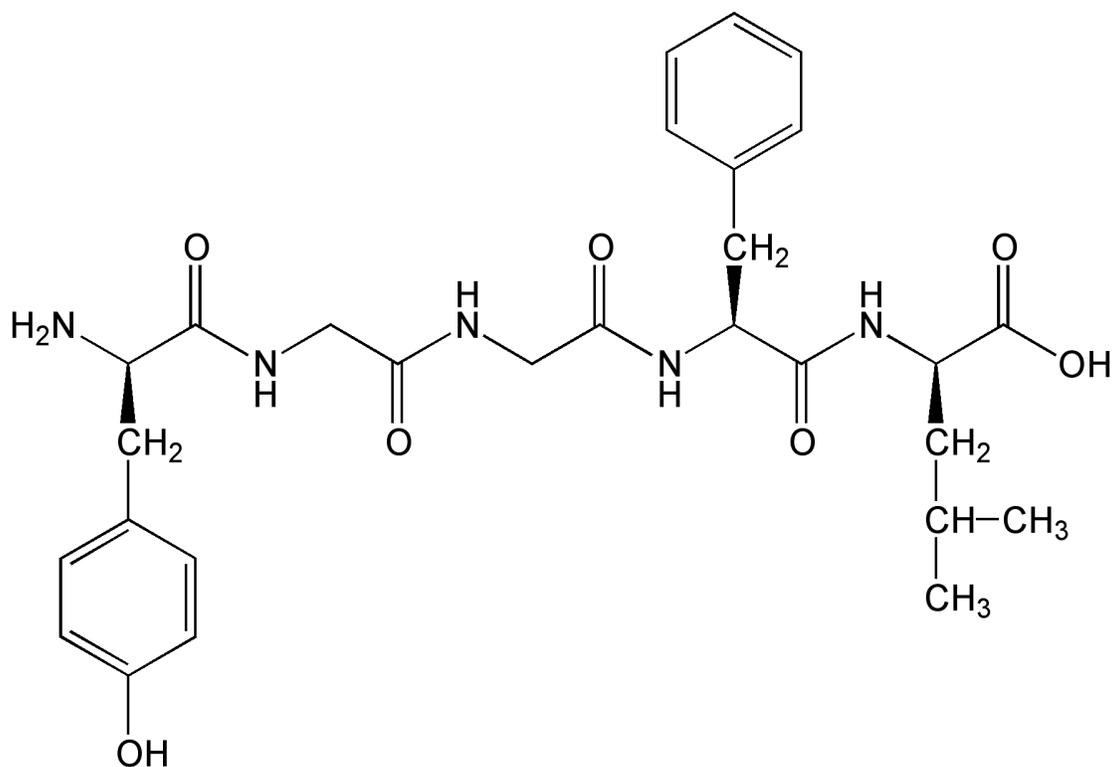


Figure 4-1. Chemical structure of Leu-enkephalin.

Table 4-1. Energies for the lowest-energy conformers of each chemical structure of b_2 . The electronic energy for each conformer at the MP2/6-31G+(d,p) level was corrected for the zero-point energy (ZPE) derived at the B3LYP/6-31G+(d,p) level.

Structure	Electronic energy / Hartrees	ZPE-corrected energy / Hartrees
<i>Diketopiperazine</i> O-prot	-759.757356	-759.520926
<i>Oxazolone</i> N-prot	-759.7521904	-759.5155784
<i>Oxazolone</i> ox-prot	-759.7469593	-759.5114213

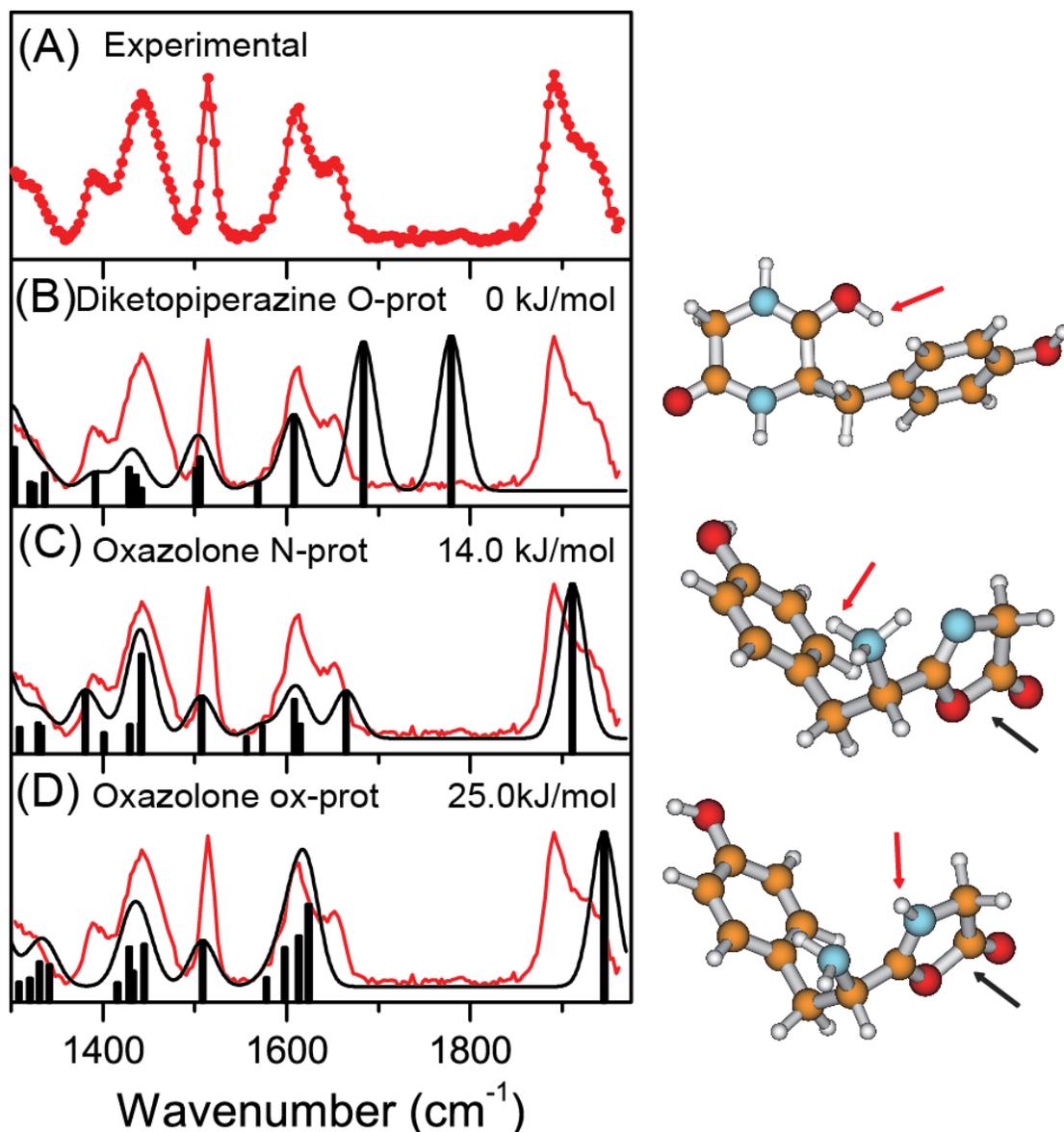


Figure 4-2. Experimental IRMPD spectrum of b_2 compared with theoretical spectra by computational studies. (A) IRMPD spectrum of the b_2 fragment generated from protonated Leu-enkephalin, (B) *diketopiperazine* structure protonated on a carbonyl O, (C) *oxazolone* structure protonated on the on the N-terminus, and (D) *oxazolone* structure protonated on the *oxazolone* ring N. The scaling factor is 0.965. Corresponding structures are presented on the right. The site of proton attachment (red arrow) and *oxazolone* rings (black arrow) are indicated.

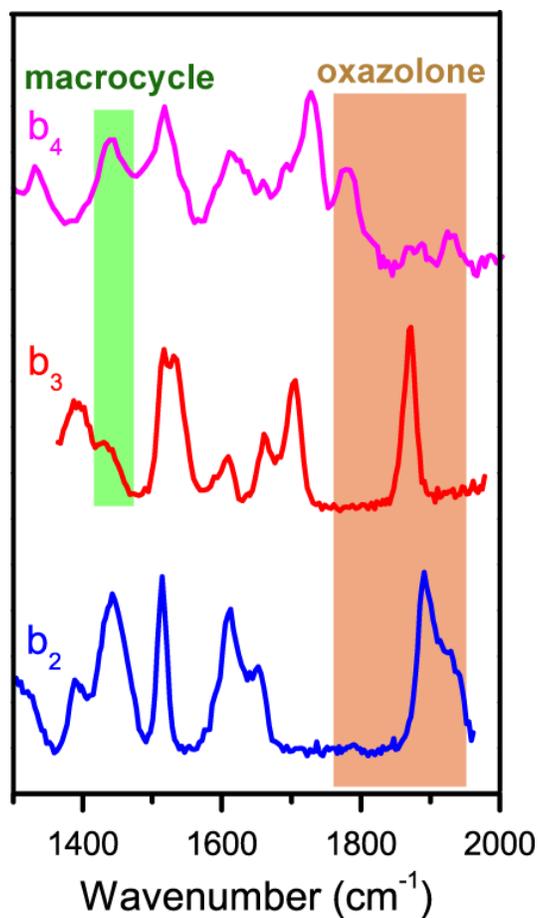


Figure 4-3. Overlaid mid-IRMPD spectra of b_2 , b_3 and b_4 . The spectrum of b_4 is adapted from previous publication of Polfer et al.³⁹. The chemically diagnostic modes are indicated.

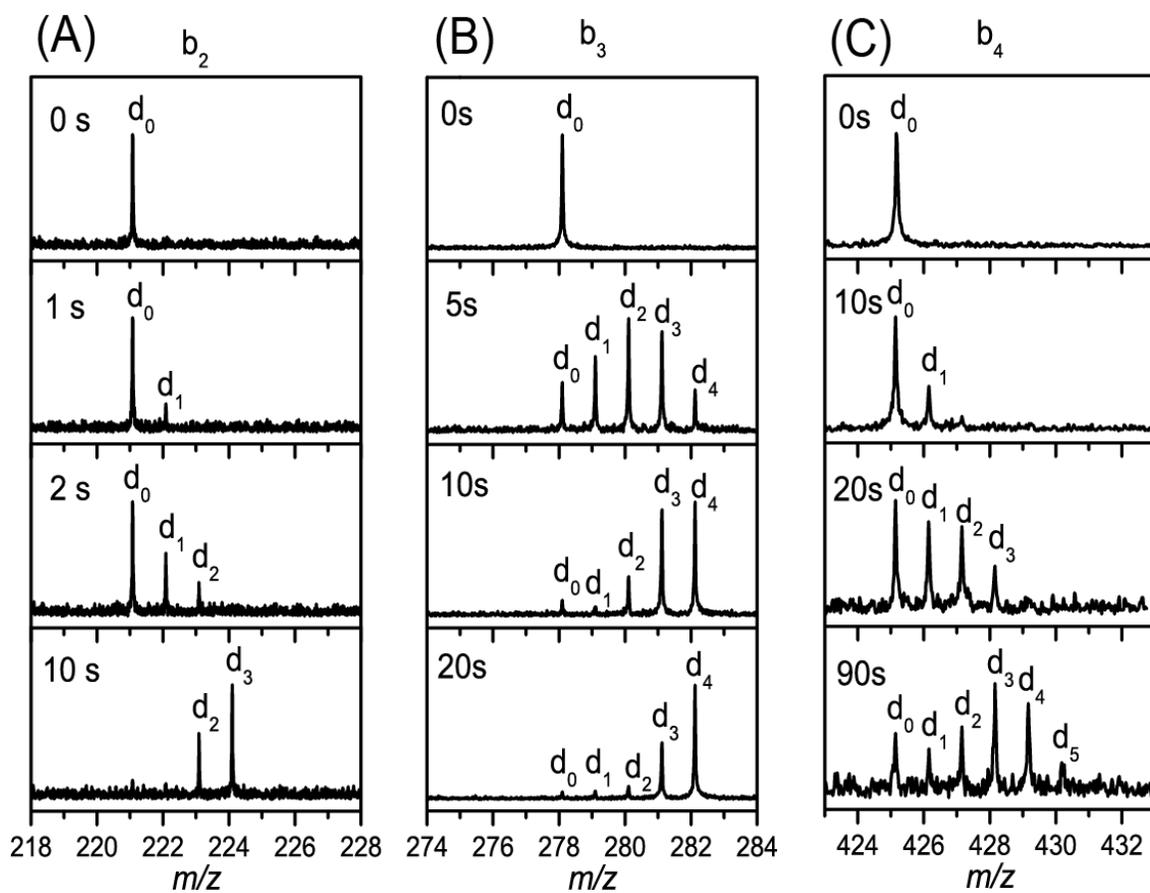


Figure 4-4. H/D exchange (10^{-8} Torr CH₃OD) mass spectra for (A) b₂, (B) b₃, and (C) b₄ for different exchange times.

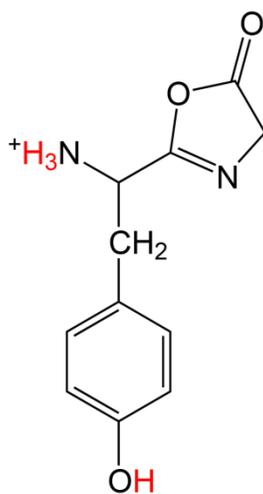


Figure 4-5. Structure of b₂ of Try-Gly of an oxazolone structure protonated at the N-terminus. Exchangeable hydrogens are shown in red.

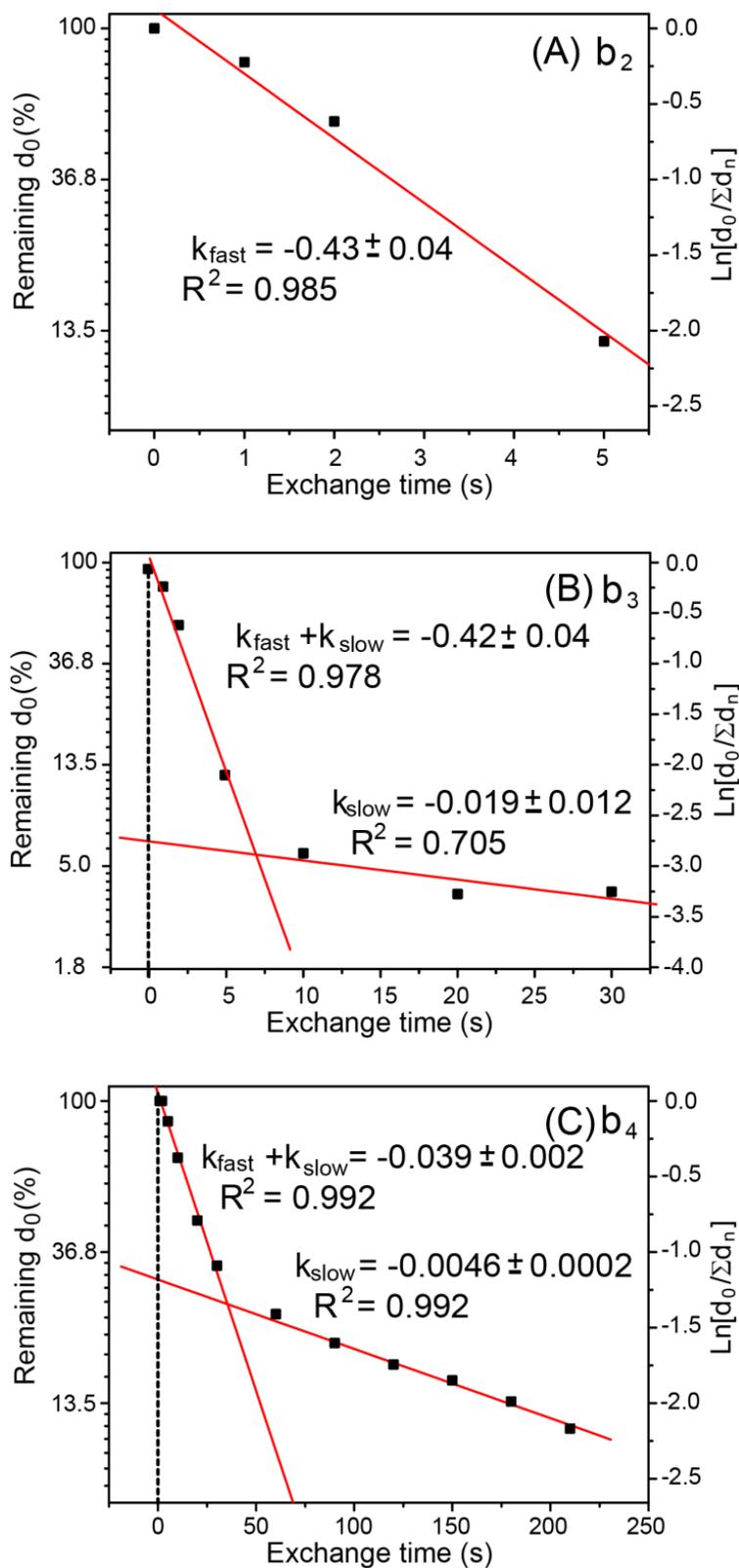


Figure 4-6. Kinetic fitting of the HDX results for (A) b_2 , (B) b_3 , and (C) b_4 .

Table 4-2. Exchange rates and relative abundances of “fast” and “slow”-exchanging structures for the Leu-enkephalin fragments b₂-b₄.

		Fast-exchanging structure	Slow-exchanging structure
b ₂	Rate	0.43±0.04	
	Abundance (%)	100	
b ₃	Rate	0.40±0.03	0.019±0.012
	Abundance (%)	93.6±34.5	6.4±34.5
b ₄	Rate	0.034±0.002	0.0046±0.0002
	Abundance (%)	69.4±3.9	30.6±3.9

Table 4-3. Kinetic fitting results for the ln[d₀/Σd_n] plots vs. H/D exchange time for the Leu-enkephalin fragments b₂-b₄.

b_n		b₂	b₃	b₄
Fast-exchanging structure	slope	-0.428	-0.424	-0.0393
	Error in slope	0.0370	0.0451	0.00205
	intercept	0.128	0.156	0.0486
	Error in intercept	0.101	0.124	0.0346
	R	-0.993	-0.989	-0.996
	Standard	0.138	0.169	0.0472
Slow-exchanging structure	slope		-0.0189	-0.00459
	Error in slope		0.0122	2.42E-4
	intercept		-2.75	-1.18
	Error in intercept		0.264	0.0378
	R		-0.839	-0.996
	Standard		0.173	0.0230

CHAPTER 5 DEPENDENCE OF HEAD-TO-TAIL CYCLIZATION ON PRIMARY STRUCTURE OF PEPTIDES IN COLLISION-INDUCED DISSOCIATION

Background

The influence of the backbone length on the formation of *macrocycle* structure has been shown in Chapters 3 and 4, based on the results for series of b ions from oligoglycines and from YGGFL. There are some differences in the relative abundances of *oxazolone/macrocycle* structures in both of those datasets, which are probably due to the different amino acid compositions of the ions. This suggests that the primary structure of the peptide also affects the formation of *macrocycle* structures.

The effect of certain amino acids on scrambling has been investigated by several groups. Van Stipdonk and co-workers studied the influence of some amino acids with multiple-stage tandem mass spectrometry experiments. It was found that with the presence of arginine, only direct sequence ions were observed. This was interpreted in the sense that arginine could inhibit the formation of macrocyclic b ions regardless of its sequence position.⁶¹ Another study by the same group showed that the reopening of *macrocycle* structures is influenced by the amino acid side chains. For a series of permuted isomers with glutamine, b_5^+ ions showed nearly identical MS^n spectra, which suggested that *macrocycle* structures tend to reopen at the position of glutamine.⁶² Histidine residue was investigated by Paizs, Harrison and co-workers by tandem mass spectrometry and theoretical studies on the singly-protonated peptides containing His residue. The results indicate that cyclization/reopening is less active for b_n ions containing His residue than for those with only aliphatic residues.¹⁷²

It has been seen in the previous IRMPD results that unambiguous results can be made for *oxazolone* structures based on the chemically diagnostic bands in 1750-1950

cm^{-1} . However, since the macrocyclic structures have chemically diagnostic bands located in a very congested region, it is very difficult to identify *macrocyclic* structures from IRMPD spectra. In Chapter 3, for the study of b_2 from triglycine, cyclo(Gly-Gly) was utilized as a reference system to exclude the diketopiperazine structure for b_2 . A direct comparison of IRMPD spectra of b fragments to IRMPD spectra of synthetic cyclic peptides is a useful and straightforward approach to identify *macrocyclic* structures.

In this chapter, a systematic study of the effect of specific residues on head-to-tail cyclization is presented for the b_6 sequence motif QWFGLM, where all amino acid residues are different, thus in complete contrast to the oligoglycine studies. In addition, this peptide contains no basic amino acids (i.e., arginine, lysine, or histidine), and hence no proton will be located at the side chain. The experimental methods employed include IRMPD and gas-phase HDX. Commercially available cyclo(QWFGLM) is a reference peptide in the study of the b_6 motif QWFGLM. For the investigation of the proline effect on the head-to-tail cyclization, a synthetically made peptide, cyclo(QPFGLM) is utilized.

Experimental

Materials

All resins were purchased from Advanced ChemTech (Louisville, KY). Fmoc protected amino acids and 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were obtained from AnaSpec (Fremont, CA). Trifluoroacetic acid (TFA, 99%), N,N-Diisopropylethylamine (DIPEA), triisopropylsilane (TIS), N,N-dimethylformamide (DMF), dichloromethane (DCM), piperidine (>99.5%), water, methanol, and formic acid were available from Sigma Aldrich (St. Louis, MO). Cyclo(QWFGLM) was purchased from Bachem (Torrance, CA).

Preparation of Peptides

Fmoc synthesis of linear peptides

Most linear peptides used in this project were prepared using solid-phase synthesis techniques using 9-fluorenylmethoxycarbonyl (Fmoc) amino acid loaded Wang resins³ and Fmoc protected amino acids. After the final deprotection, peptides were cleaved from the resins were achieved utilizing a cleavage cocktail with 95:2.5:2.5 (vol:vol:vol) TFA:H₂O:TIS. The synthesized peptides that were dissolved in the cleavage solution were precipitated by addition of diethyl ether. The products were collected by centrifugation and washed several times with diethyl ether. The solids were then dried in air at ambient temperature.

N-terminal acetylated peptides were obtained by mixing 50 molar excess of acetic anhydride and DIPEA with peptides in DMF for 30 min prior to the last deprotection step. After the acetylation, the peptides underwent deprotection, cleavage, and purification as other linear peptides.

Q(Trt)PFGLM was made with Fmoc-protected resins, preloaded with 2-chlorotrityl. The cleavage was done with 5% TFA and 2.5% TIS in DCM. The trityl protecting group on the side chain of Gln can be retained after the cleavage. Following cleavage, the peptide was dissolved in acetonitrile and freeze-dried.

Synthesis of head-to-tail cyclic peptides

The synthesis of cyclic peptides were done in Synthetic Organic Chemistry group of the University of Amsterdam with the help from Jochem Rutters and Dr. Jan Maarseveen. Linear peptide Q(Trt)PFGLM, made with solid phase synthesis, were dissolved in THF at concentrations of less than 10⁻³ mol/L, with 4.4 equivalent of DIPEA. 2.2 equivalent of HATU and 3H-1,2,3-Triazolo[4,5-b]pyridin-3-ol, were added into the

solution. The reaction solution was kept at room temperature while stirring for 24h. The aliquot from the reaction solution was characterized by LC/MS to make sure that the reaction is complete.

Purification of synthesized head-to-tail cyclic peptides

The THF was evaporated off from the reaction solution, and the remaining solid was then dissolved in ethyl acetate. 1M KHSO₄ aqueous solution was added to dissolve the unreacted coupling reagents. The organic phase was collected and solid Na₂SO₄ was added to remove the remaining water. The organic phase was lyophilized to obtain solid products. The trityl protecting group at the glutamine side chain was then removed by addition of 95%TFA. Crude peptides was purified by reversed-phase HPLC (RF-HPLC) on a C₁₈ column using a gradient of 0-80% B (Buffer A: water/0.05%TFA; Buffer B: 90% acetonitrile/10% water/0.045% TFA) over 30 min.

Mass Spectrometry

All sample solutions were prepared by dissolving peptides in 49:29:2 (vol:vol:vol) MeOH:H₂O:formic acid at concentrations of 100μM.

Infrared multiple photon dissociation (IRMPD) experiments were performed with the same set-up as described in Chapter 3, using the free electron laser FELIX at FOM institute “Rijnhuizen” on the 4.7 T FTICR-MS. The mass-to-charge ratios (*m/z*) of the ions of interested generated by “nozzle-skimmer” dissociation in the ESI source region are listed in Table 5-1. In the photodissociation experiments presented here, the wavelength was scanned from 1200–2000 cm⁻¹.

Gas-phase hydrogen/deuterium exchange experiments were performed on cyclo(QWFGLM)H⁺, QWFGLMPG b₆ and Ac-QWFGLMPG b₆ , using the 4.7 T FTICR instrument at the University of Florida and the same procedure described in Chapter 3.

Again, the deuterating reagent used was CH₃OD and the background pressure was kept constant at 10⁻⁸ Torr.

Results and Discussion

Commercial Cyclo(QWFGLM) as Direct Reference

IRMPD of cyclo(QWFGLM) and analogs

The IRMPD spectrum of protonated cyclo(QWFGLM) is contrasted with those of b₆ from QWFGLMPG and Ac-QWFGLMPG in Figure 5-1. The clear absence of *oxazolone* bands in the spectrum of b₆ from the linear peptide QWFGLMPG suggests that this ion does not adopt an *oxazolone* structure.

To confirm that this ion is exclusively macrocyclic structure, its IRMPD spectrum is compared with that of commercial cyclo(QWFGLM). Both spectra are close to identical, clearly reproducing all of the main features. The only discrepancy is seen in the range from 1600-1650 cm⁻¹, a range that includes NH₂ scissoring and C=O stretching modes. It is conceivable that these minor differences are due to variations in subpopulations that are formed, although this is beyond the information that can be obtained from this IRMPD spectrum.

Note that, to promote isomerization from *cyclic* to *oxazolone* structures for the cyclic peptide, harsh source conditions were employed, where some of the protonated cyclo(QWFGLM) precursor ions were fragmented by nozzle-skimmer CID. Nonetheless, the absence of bands at higher than 1700 cm⁻¹ clearly indicates that there is no *oxazolone* structure present in the protonated cyclo(QWFGLM) in the gas phase, even under these energetic conditions.

For N-terminally acetylated Ac-QWFGLMPG, the N-terminus is not a nucleophile, thus precluding head-to-tail cyclization. In fact, the IRMPD spectrum for b₆ from Ac-

QWFGLMPG displays a prominent band at 1900 cm^{-1} , which can be unambiguously assigned to the C=O stretch of a protonated *oxazolone* moiety. This confirms that N-terminal acetylation results in exclusive presence of an *oxazolone* structure for this b_6 ion. Compared to the other two spectra, the amide C=O stretch band is much broader, which is consistent with a more dynamic *oxazolone* structure. The band at 1440 cm^{-1} , which had been assigned to the CO-H⁺ bending mode of *macrocycle* structures, and hence is used as the only diagnostic band for *macrocycle* structures in the mid-IR range. Ac-QWFGLMPG b_6 , which is exclusively *oxazolone*, also displays a band at this position, albeit at a lower intensity. This confirms that the 1440 cm^{-1} region is congested with other bands, and that identification of *macrocycle* structures is not unambiguous based on this feature. In principle, however, the higher band intensities at 1440 cm^{-1} for QWFGLMPG b_6 and cyclo(QWFGLM)H^+ supports the case for the presence of the CO-H⁺ bending mode at 1440 cm^{-1} .

HDX of *cyclo(QWFGLM)* and analogs

The three ions above were also subjected to gas-phase hydrogen/deuterium exchange (HDX) with deuterated methanol (CH₃OD) in the ICR cell of an FTICR mass spectrometer. As described in Chapter 3, the kinetic analysis of the data was done by plotting the relative abundance of the undeuterated peak, d_0 , as a function of exchange time.

The kinetic plots of HDX for protonated cyclo(QWFGLM) and QWFGLMPG b_6 are shown in Figure 5-2. Two distinct rates of exchange can be resolved in both plots. In addition, the kinetic rates are very close. Since both IRMPD spectra are nearly identical, it is not surprising that they display close values for the kinetic rates. A closer examination of the magnitudes of the rate constants for cyclo(QWFGLM)H^+ shows that

they are both relatively “slow” compared to the “fast”-exchanging structure for **b**₅ ($k_1 = 0.027 \text{ s}^{-1}$ and $k_2 = 0.009 \text{ s}^{-1}$ vs $k_{fast} = 0.17 \text{ s}^{-1}$). In fact, both of these rate constants are much more in accordance with the “slow”-exchanging structure for **b**₅ (0.02 s^{-1}). The absence of an *oxazolone* structure for cyclo(QWFGLM)H⁺, as confirmed by IRMPD spectroscopy, hence correlates well with the absence of a “fast”-exchanging structure in the HDX results. Similarly, both structures found in **b**₆ from QWFGLMPG are also “slow”-exchanging structures, as $k_1 = 0.037 \text{ s}^{-1}$ and $k_2 = 0.0069 \text{ s}^{-1}$. The appearance of two kinetic rates indicates that HDX can separate the subpopulations of isomers, which is compatible with the differences in the 1600-1650 cm^{-1} region of IRMPD spectra.

So far, it is not clear what isomers account for the subpopulations in macrocyclic QWFGLM. It is possible that different protonation sites have to be considered for this peptide, such as the tryptophan side chain for instance, or a number of backbone carbonyls. If the structures do not interconvert (i.e., no proton transfer), this may result in multiple exchange kinetics. Alternatively, cyclo(QWFGLM)H⁺ is made up of different structural variants, such as cis–trans isomers for example.

HDX for Ac-QWFGLMPG **b**₆ shows an extremely slow exchange with CH₃OD, as no exchange is observed even after 90 s (see Figure 5-3). This can be rationalized in the framework of the “relay” mechanism that was discussed in Chapter 2.¹²⁸ Figure 5-4 illustrates the exchange progress with CH₃OD. An *oxazolone* with a free N-terminus exchanges with CH₃OD through a concerted movement of the *oxazolone* proton to the deuterating molecule, and movement of a deuteron to the N-terminus. Both the N-terminus and the *oxazolone* ring N have similar proton affinities, and hence there is no energetic penalty for substituting the proton from the *oxazolone* ring for a deuteron on

the N-terminus, or vice-versa. In Ac-QWFGLMPG b_6 , the N-terminus is acetylated and hence the N-terminus is no longer a basic site. Instead, the second most basic site in the molecule may be at a carbonyl oxygen, or the side chain amine. Since these sites are much less basic than the N-terminus, it is energetically unfavorable to substitute the proton from the *oxazolone* ring for a deuteron on a less basic site.

The Effect of Proline on Head-to-Tail Cyclization

Charaterization of synthetic peptides

The LC/MS spectrum of purified linear Q(Trt)PFGLM is shown in Figure 5-5. As the HPLC spectrum shows, the product from the solid phase synthesis has high purity. This peptide was then subjected to head-to-tail cyclization in solution. The trityl protecting group on the side chain of glutamine is kept to prevent the cyclization from the side chain amine. Figure 5-6 shows the LC/MS spectrum of the aliquot from the cyclization reaction solution after 24h. The product with retention time of 6.15 min has the mass of 674, which is the mass of cyclo(QPFGLM). The product with retention of 6.95 min shows to intense peaks in the mass spectrum, 674 and 1347, and this indicates that cyclo(QPFGLM)₂ is also synthesized. Both of these two products were then purified and collected with HPLC.

IRMPD results

The IRMPD spectra of QWFGLMPG b_6 , QPFGLMPG b_6 and QWPFGLMPG b_7 are compared in Figure 5-7, where the *oxazolone* region (1750-1930 cm^{-1}) is highlighted in pink. It is clear that the b_6 ion generated from QPFGLMPG displays a band at 1840 cm^{-1} , which is likely assigned to the C=O stretch mode of an *oxazolone* structure protonated at the N-terminus. The spectrum of QWPFGLMPG b_7 ion, where the proline residue is inserted between the tryptophan and phenylalanine, displays an even more intense

oxazolone band. These results indicate that the presence of proline in these peptide play a role in favoring *oxazolone* structure formation.

To confirm that the 1840 cm^{-1} band in fact corresponds to an *oxazolone* structure, the IRMPD spectrum of synthetically made cyclo(QPFGLM) is used as a direct reference. The overlaid IRMPD spectra of these two ions are shown in Figure 5-8. The spectrum of protonated cyclo(QPFGLM) shows no band in the *oxazolone* region, whereas the band for QPFGLMPG b_6 is apparent. This lends further credence to the claim that *oxazolone* structures are formed for QPFGLMPG b_6 and QWPFGLMPG b_7 . As noted for QWFGLM systems, there are also differences in the range from $1600\text{-}1650\text{ cm}^{-1}$, where NH_2 scissoring and $\text{C}=\text{O}$ stretching modes are located, although these differences are harder to explain from a structural point of view.

The effect of the proline residue on the formation of *macrocycle* structure may be due to its chemical structure. Of all the natural amino acids, proline is unique in that it contains a secondary amine involving a cyclic side chain. This feature makes the dihedral angle of proline particularly rigid ($\sim 60^\circ$), often leading to a 'kink' in the secondary structure of peptides at this position. This rigidity constrains the flexibility of the backbone, and is hence compatible with limiting nucleophilic attack from the N-terminus.

The Effect of Glutamine Side Chain on the Head-to-Tail Cyclization

There are two amines in the structure of glutamine: one at the N-terminus and the other at the side chain. In the gas phase, the head-to-tail cyclization is done through the nucleophilic attack from the N-terminus. It is also possible that a competing nucleophilic attack can occur from the amine at the glutamine side chain, forming different kind of *macrocycle* structure. To investigate whether the side chain of glutamine is involved in

gas-phase cyclization reactions during CID, a systematic study was performed using the QWFGLG motif. Four peptides were prepared: QWFGLGPG, Ac-QWFGLGPG, Q(N- γ -ethyl)WFGLGPG, and Ac-Q(N- γ -ethyl)WFGLGPG. The structures the *oxazolone* structure of b_6 ions made from the four peptides are illustrated in Figure 5-9. For QWFGLGPG b_6 , attack from both the N-terminus and side chain are possible. For Ac-QWFGLGPG b_6 , only the attack from the glutamine side chain can occur, while for Q(N- γ -ethyl)WFGLGPG b_6 , only the N-terminus can attack. No cyclization can happen for Ac-Q(N- γ -ethyl)WFGLGPG b_6 since both amines are protected, and hence only *oxazolone* structures can be generated.

The IRMPD spectra of these four b_6 ions are shown in Figure 5-10, and the spectral range for *oxazolone* bands is again highlighted in pink. There is a very weak *oxazolone* band observed in the spectrum of b_6 from QWFGLGPG $\sim 1800\text{ cm}^{-1}$, indicating that a small abundance of *oxazolone* structures are present. For Q(N- γ -ethyl)WFGLGPG b_6 , no band is seen in this region suggesting that this ion is exclusively *macrocycle*. Conversely, two bands are observed in the spectra of Ac-QWFGLGPG b_6 and Ac-Q(N- γ -ethyl)WFGLGPG b_6 at 1830 cm^{-1} and 1920 cm^{-1} , assigned to *oxazolone* structures protonated at the *oxazolone* ring N and protonated at the N-terminus, respectively. It is apparent that only when the N-terminus is blocked, do the *oxazolone* bands become prominent. These results indicate that in the gas phase, for peptides with glutamine at the N-terminus, only the head-to-tail cyclization has to be considered, while cyclization from the glutamine side chain can be neglected.

Summary

In this chapter, it has been shown that apart from the size effect, the primary structure has an influence on the formation of *macrocycle* structures. Synthetically made cyclic peptides can provide a direct comparison for IRMPD spectra and HDX, in order to identify *macrocycle* structures. The IRMPD spectra of QWFGLMPG b₆ and cyclo(QWFGLM) are nearly identical, which indicates that QWFGLMPG b₆ exclusively adopts a *macrocycle* structure. This is confirmed by the gas-phase HDX kinetic study, where both ions display “slow”-exchanging rates that are close to each other. The influence of the proline residue on the head-to-tail cyclization is investigated. An *oxazolone* band is seen in the IRMPD spectra of both QPFGLMPG b₆ and QWPFGLMPG b₇, suggesting that the presence of proline reduces the propensity for the formation of *macrocycle* structures. A systematic chemical protection study was performed on the QWFGLG motif, to investigate nucleophilic attacks from the glutamine side-chain. Based on these results, it appears that such a side-chain attack does not compete with the head-to-tail cyclization reaction.

Table 5-1. Mass-to-charge ratios (m/z) of the ions of interest in this project.

Ions of interest	Mass-to-charge ratio (m/z)
Cyclo(QWFGLM)H ⁺	763
Cyclo(QPFGLM)H ⁺	674
QWFGLMPG b ₆	763
QWFGLGPG b ₆	689
QPFGLMPG b ₆	674
QWPFGLMPG b ₇	860
Ac-QWFGLMPG b ₆	805
Ac-QWFGLGPG b ₆	731
Q(ethyl)WFGLGPG b ₆	717
Ac-Q(ethyl)WFGLGPG b ₆	759

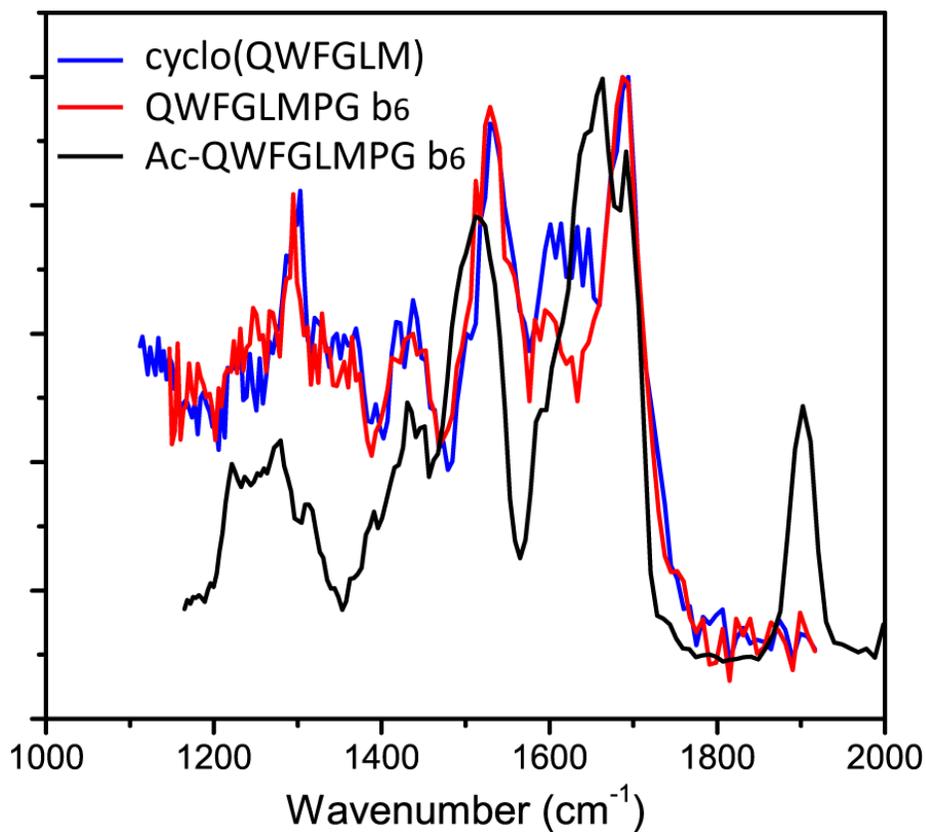


Figure 5-1. Overlaid IRMPD spectra of protonated cyclo(QWFGLM) and b₆ from QWFGLMPG and Ac-QWFGLMPG.

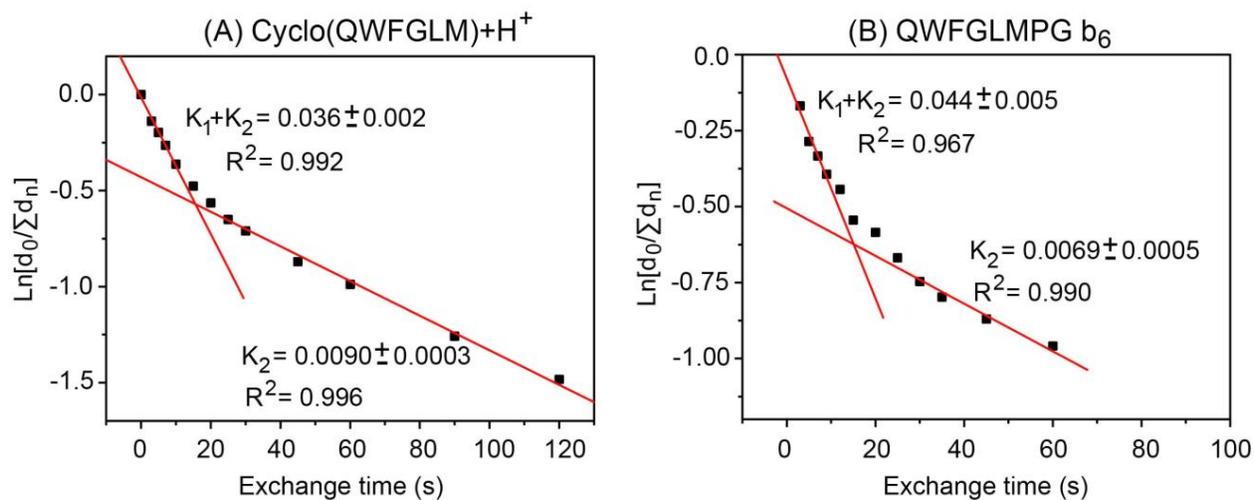


Figure 5-2. Natural logarithm of relative d_0 depletion, $\ln[d_0/\sum d_n]$, as a function of time for (A) protonated cyclo(Gln-Trp-Phe-Gly-Leu-Met), and (B) b_6 generated from protonated linear QWFGLMPG.

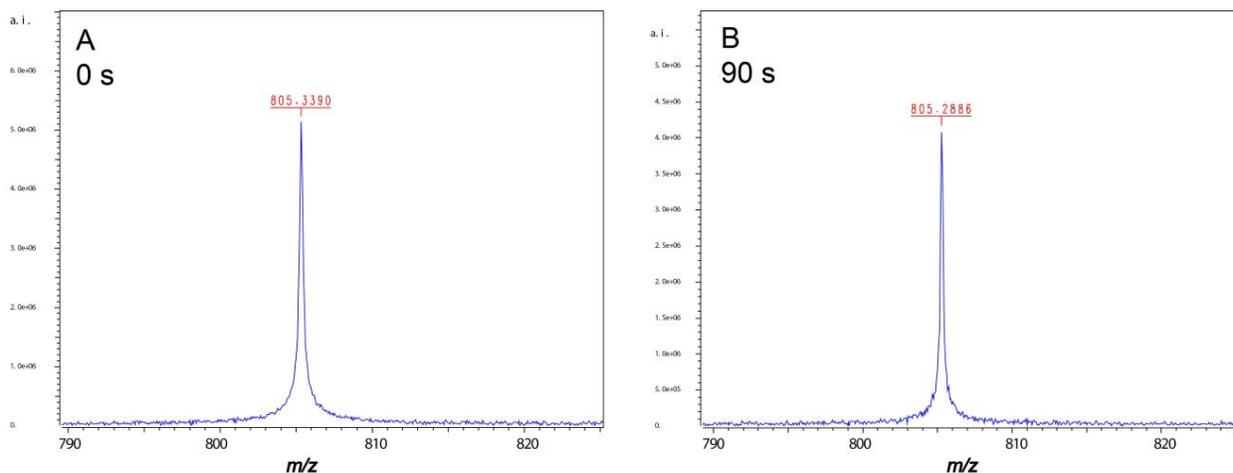


Figure 5-3. Mass spectra of Ac-QWFGLMPG b_6 ion after exchanging with CH_3OD for (A) 0s and (B) 90s. No exchange is seen after 90s of exchange, indicating that such reaction is extremely slow.

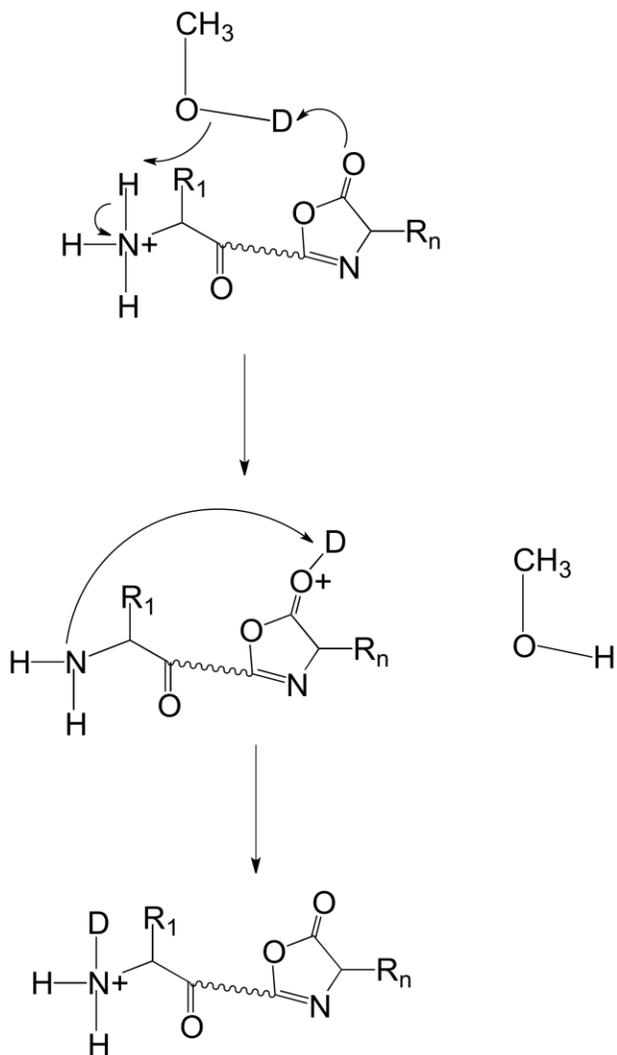


Figure 5-4. Schematic presentation of an *oxazolone* exchanges with CH_3OD .

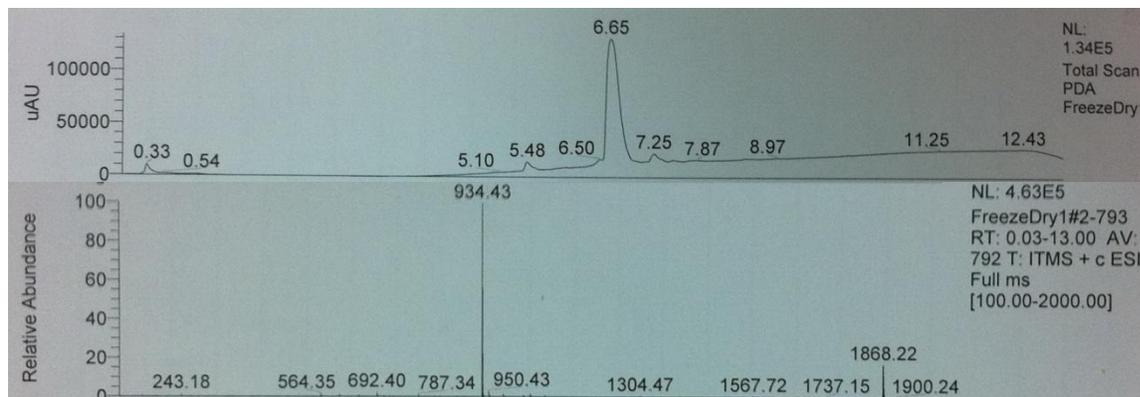


Figure 5-5. LC/MS spectrum of linear Q(Trt)PFGLM after purification.

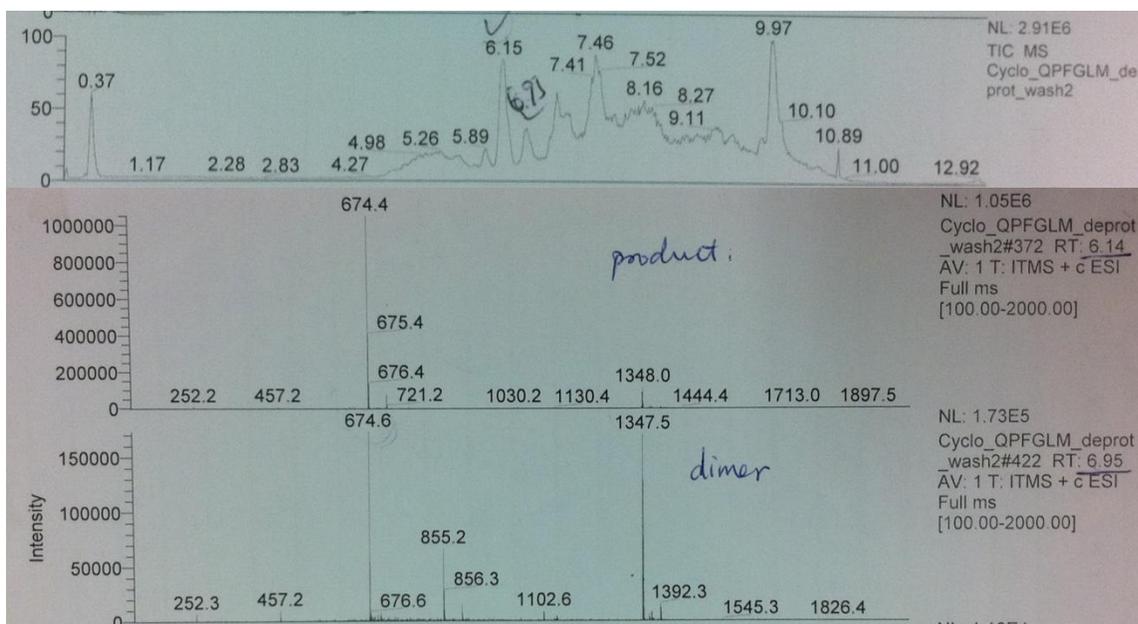


Figure 5-6. LC/MS spectrum of aliquot taken from cyclization reaction solution. The top figure is the HPLC spectrum of the aliquot. The middle figure is the MS spectrum of the synthetic cyclo(QPFGLM) with a retention time of 6.14 min. The bottom figure is the MS spectrum of the synthetic cyclo(QPFGLM)₂ with a retention time of 6.95 min.

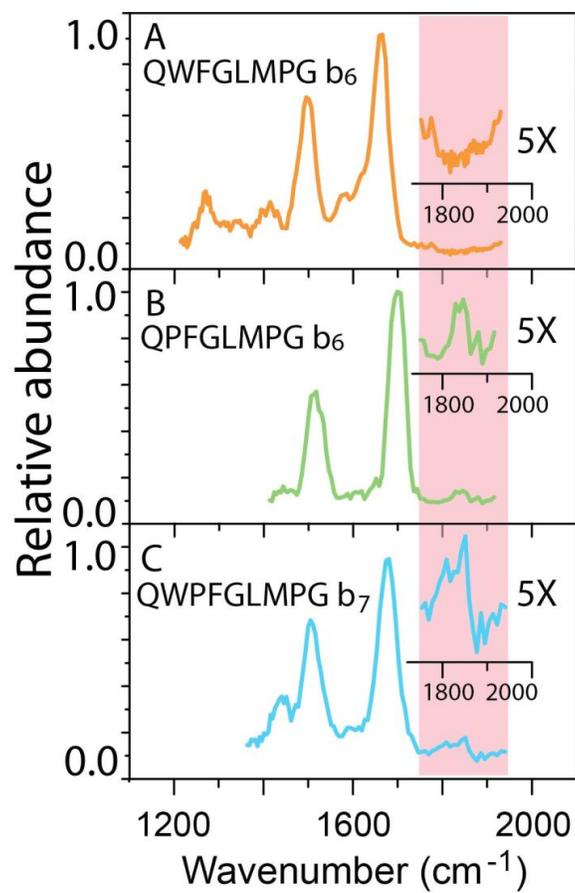


Figure 5-7. IRMPD spectra of (A) QWFGMLPG b₆, (A) QPFGLMPG b₆, and (A) QWPFGLMPG b₇. In the inserts, the intensities are magnified by 5x in the range of 1750-1930 cm⁻¹.

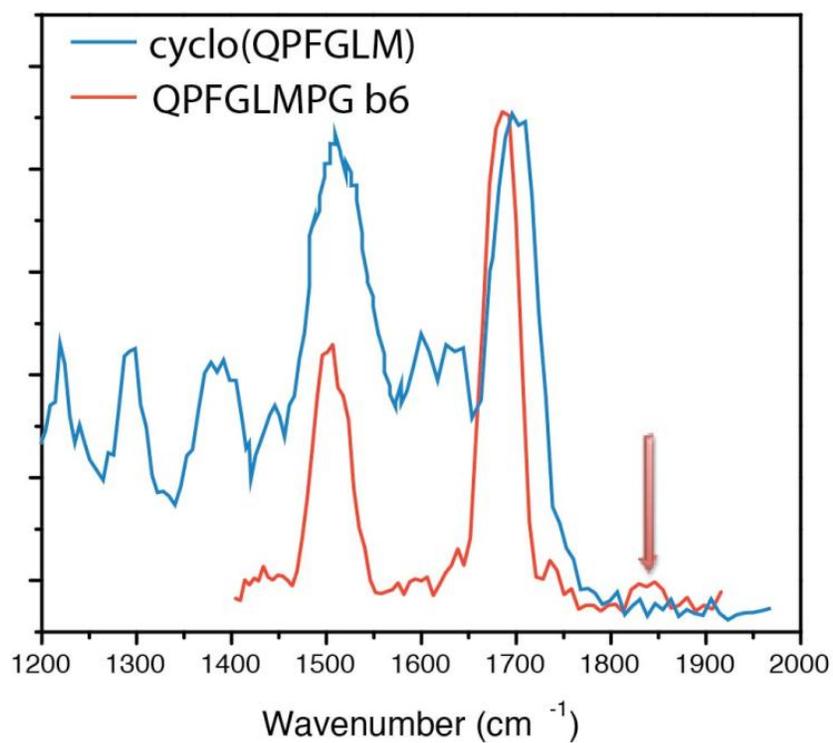


Figure 5-8. Overlaid IRMPD spectra of cyclo(QPFGLM) and QPFGLMPG b₆. The *oxazolone* C=O stretch band is pointed by the red arrow.

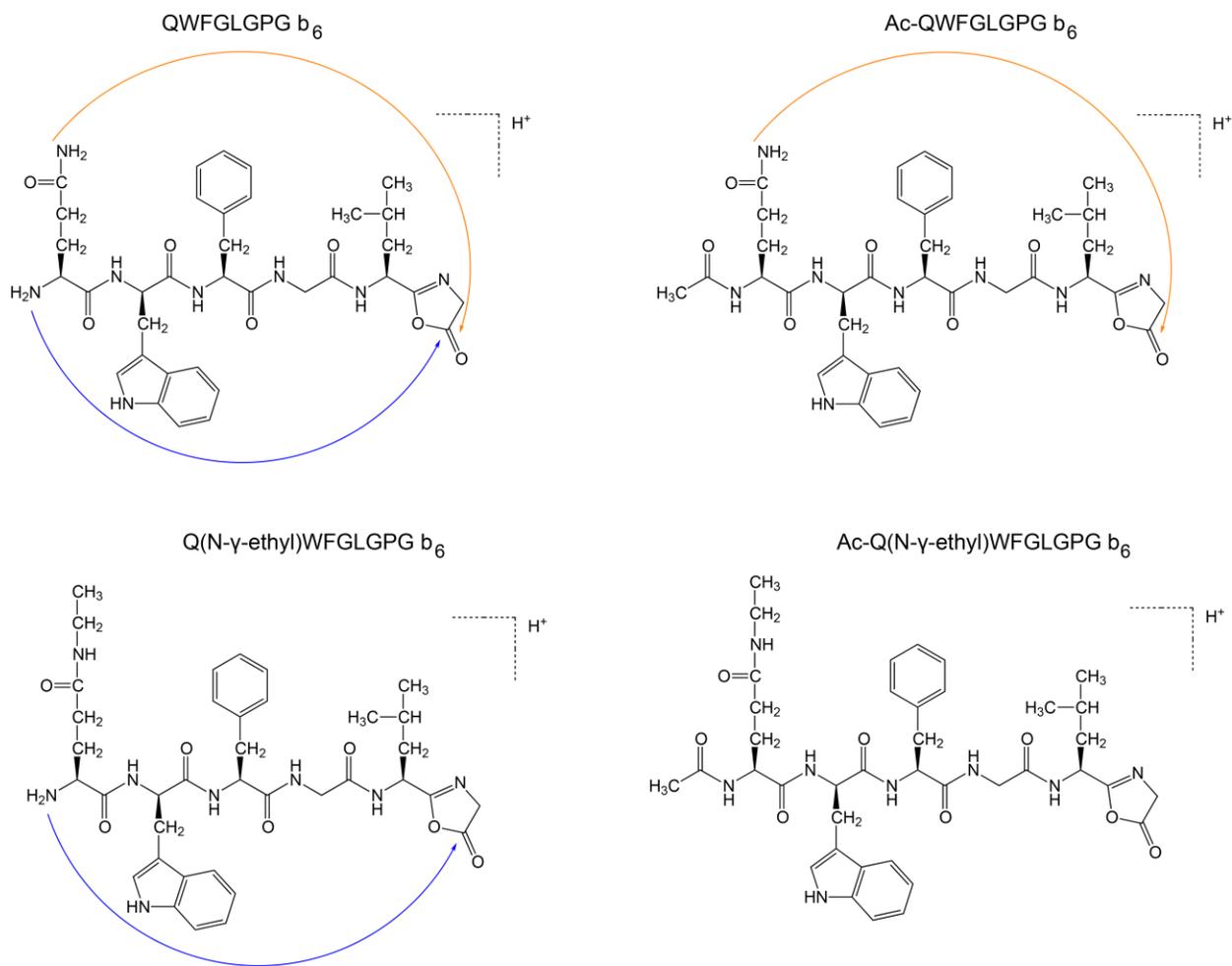


Figure 5-9. The structures of the b_6 ions from four peptides of QWFGLG system. The possible attack from the N-terminus is labeled in blue, and the one from the glutamine side-chain is labeled in orange.

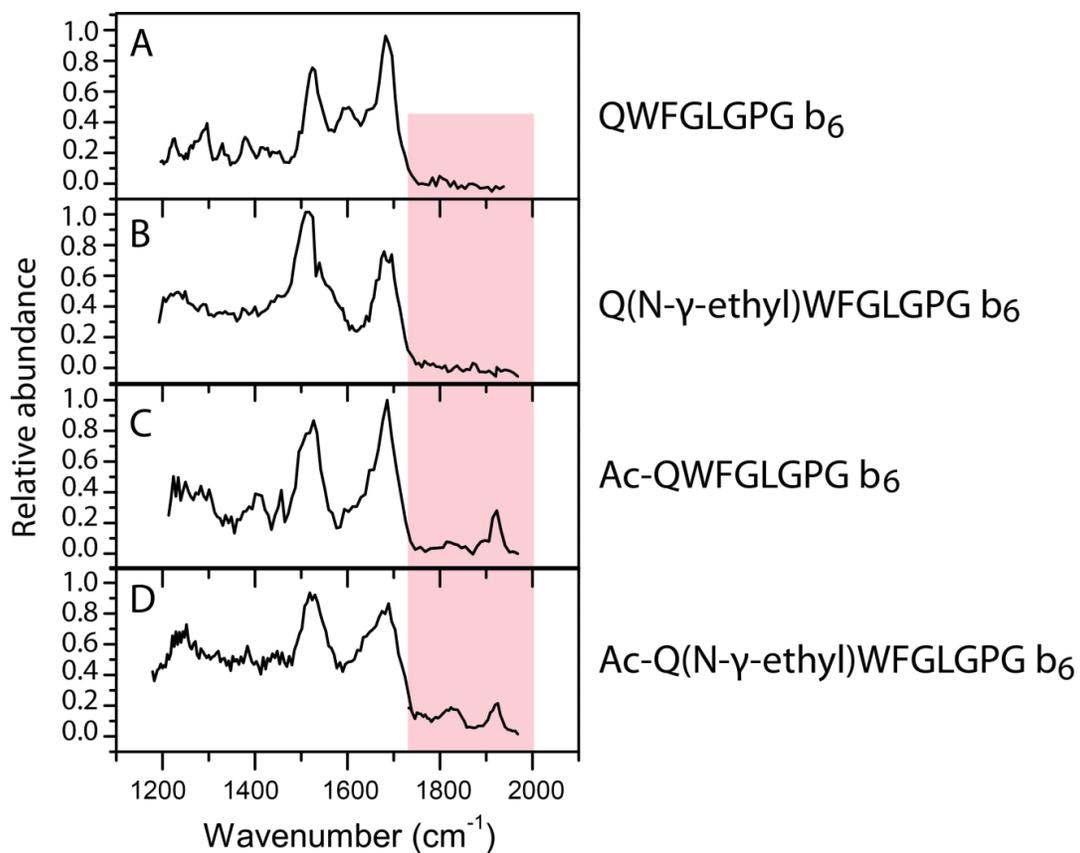


Figure 5-10. IRMPD spectra of b₆ ions generated from (A) QWFGLGPG, (B) Ac-QWFGLGPG, (C) Q(N-γ-ethyl)WFGLGPG, and (D) Ac-Q(N-γ-ethyl)WFGLGPG. The range for *oxazolone* bands is highlighted in pink.

CHAPTER 6 CONCLUSIONS AND FUTURE DIRECTIONS

The results in this thesis have given insights into head-to-tail cyclizations of “b” ions in collision-induced dissociation (CID). This process rationalizes scrambling of the original amino acid sequence, and hence it is important to understand the chemistry of b ions in the gas phase.

It has been shown that combining infrared multiple photon dissociation (IRMPD) spectroscopy and gas-phase hydrogen/deuterium exchange (HDX) is a powerful approach for structurally elucidating b ions. IRMPD spectra of b ions can be used to identify the structures based on the chemically diagnostic bands. In particular, *oxazolone* structures can be identified by the bands associated to the *oxazolone* C=O stretch modes in the range of 1750-1950 cm^{-1} . The diagnostic bands of *macrocycle* structures are located at around 1440 cm^{-1} , which are assigned to CO-H⁺ bending mode. However, as this region is congested with other vibrational modes, unambiguous identification is often more challenging. A further limitation is that IRMPD spectroscopy cannot provide information about relative abundances in mixtures, since the intensity of the IRMPD band does not linearly relate to the abundance of one structure. On the other hand, the relative abundances of *oxazolone/macrocycle* structures can be approximated with gas-phase HDX. The exchange reaction in the ICR cell is assumed to be a first-order reaction, and hence the kinetic plotting can be used to calculate the exchange rates and relative abundances of isomers. It has been found that the exchange rates of *oxazolones* are around 10 times faster than *macrocycle* structures.

The first project on this dissertation is about the influence of the peptide backbone length on b ion structures. A series of b ions, b₂-b₈, made from oligoglycines were

investigated by IRMPD spectroscopy and gas-phase HDX. The glycine residue is the simplest amino acid, devoid of a nucleophilic side chain, and hence is an excellent candidate for the size effect study. It was found that the formation of *macrocycle* structures is favored for larger b fragments. Smaller b ions, b_2 and b_3 , were identified as exclusively *oxazolones*, whereas mid-size b ions, b_4 - b_7 , were found to be a mixture of *oxazolone* and *macrocycle* structures. With the kinetic fitting of HDX data, the relative abundances of *macrocycle* structures were calculated to be 7%, 21%, 31% and 32%, respectively for b_4 - b_7 . Thus, *oxazolone* structures are still the majority of the population. The largest b ion, b_8 , was found to exclusively adopt a *macrocycle* structure, as confirmed by the absence of an *oxazolone* band in the IRMPD spectrum and a single HDX exchange rate.

CID product ions from Leu-enkephalin (YGGFL) were studied with the same techniques as above. A similar dependency on the size of backbone was observed in this peptide. b_2 was found to adopt only *oxazolone* structures, as shown by the good match with the theoretical *oxazolone* spectra and a single HDX exchange rate. Both *oxazolone* and *macrocycle* structures were confirmed in b_3 and b_4 ions, and the relative abundances were calculated to be 6% and 31%, respectively. The relative ratios of *oxazolone/macrocycle* for b_3 of GGG and YGG, and b_4 of GGGG and YGGF are different, which indicates that besides the size effect, the primary sequence plays a role in the formation of *macrocycle* structures.

Thus, a third project was conducted to study how the primary structure affects the formation of *macrocycle* structure for the sequence motif QWFGLM. In addition to IRMPD and HDX methods, synthetic cyclic peptides were used to provide a direct

comparison, to verify *macrocycle* identification. By comparing IRMPD spectrum of QWFGLMPG b_6 ion with those of cyclo(QWFGLM) and Ac-QWFGLMPG b_6 , it was found that b_6 of QWFGLM was exclusively *macrocycle*. The effect of proline on the head-to-tail cyclization was investigated by measuring IRMPD spectra of synthetic cyclo(QPFGLM), QPFGLMPG b_6 and QWPFGLMPG b_7 . Results indicated that the presence of proline residue in the primary sequence could reduce the propensity of the formation of *macrocycle* structures, consistent with the limited flexibility of the backbone due to proline. Lastly, studies of b_6 ions of QWFGLG, Ac-QWFGLG, Q(N- γ -ethyl)WFGLG, and Ac-Q(N- γ -ethyl)WFGLG revealed that no cyclization from the amino group on the glutamine side chain takes place.

Several avenues for future work are proposed. First, the size effect study has so far only been carried out for b ions up to b_8 . It is not clear yet whether larger b_n fragments, where $n > 8$, also exclusively adopt *macrocycle* structures. While the propensity for head-to-tail cyclization increases with chain length for b_2 - b_8 , very large *macrocycle* are expected to be energetically disfavored due to entropic effects. In other words, the degrees of freedom of a *macrocycle* structure is reduced vis-à-vis a linear *oxazolone* structure. One might hence expect an upper limit for *macrocycle* structures, which remains to be confirmed.

Much work remains to be done on the correlation between primary structure and propensity for *macrocycle* formation. The presence of proline appears to reduce this tendency, possibly due to kinetic effects. Other residues (e.g. arginine, lysine) might offer competing nucleophilic attacks from their side-chain groups. This was not confirmed for glutamine, but may yet be confirmed for other amino acid residues.

Lastly, although the scrambling phenomenon has been observed in a number of model peptides, there is no study showing the extent of scrambling in proteomics studies. Such an analysis is on-going in our laboratory, involving high-resolution and high-accuracy MS/MS data of a “well-defined” (i.e., with a low number of known proteins) proteomics study.

LIST OF REFERENCES

- (1) Garrett, R.; Grisham, C. M. *Biochemistry*, 2nd ed.; Saunders College Pub.: Fort Worth, 1999.
- (2) Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.
- (3) Chan, W. C.; White, P. D. *Fmoc solid phase peptide synthesis: a practical approach*; Oxford University Press: New York, 2000.
- (4) Kent, S. B. *Annu Rev Biochem* **1988**, *57*, 957-89.
- (5) Davies, J. S. *J Pept Sci* **2003**, *9*, 471-501.
- (6) Domon, B.; Aebersold, R. *Science* **2006**, *312*, 212-217.
- (7) Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M. *Science* **1989**, *246*, 64-71.
- (8) Hillenkamp, F.; Karas, M.; Beavis, R. C.; Chait, B. T. *Anal. Chem.* **1991**, *63*, A1193-A1202.
- (9) Karas, M.; Hillenkamp, F. *Anal. Chem.* **1988**, *60*, 2299-2301.
- (10) Steen, H.; Mann, M. *Nature Rev. Mol. Cell Biol.* **2004**, *5*, 699-711.
- (11) Niall, H. D. *Methods Enzymol* **1973**, *27*, 942-1010.
- (12) Berthod, A.; Crank, J. A.; Rundlett, K. L.; Armstrong, D. W. *Rapid Comm Mass Spectrom* **2009**, *23*, 3409-3422.
- (13) Standing, K. G. *Curr. Opin. Struct. Biol.* **2003**, *13*, 595-601.
- (14) Horn, D. M.; Zubarev, R. A.; McLafferty, F. W. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 10313-10317.
- (15) Zubarev, R. A.; Horn, D. M.; Fridriksson, E. K.; Kelleher, N. L.; Kruger, N. A.; Lewis, M. A.; Carpenter, B. K.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 563-573.
- (16) Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J. H.; Begley, T. P.; McLafferty, F. W. *J. Am. Chem. Soc.* **2002**, *124*, 672-678.
- (17) Ge, Y.; ElNaggar, M.; Sze, S. K.; Bin Oh, H.; Begley, T. P.; McLafferty, F. W.; Boshoff, H.; Barry, C. E. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 253-261.
- (18) Sze, S. K.; Ge, Y.; Oh, H. B.; McLafferty, F. W. *Anal. Chem.* **2003**, *75*, 1599-1603.
- (19) Skoog, D. A.; Holler, F. J.; Crouch, S. R. *Principles of instrumental analysis*; 6th ed.; Thomson Brooks/Cole: Belmont, CA, 2007.

- (20) Hoffmann, E. d.; Stroobant, V. *Mass spectrometry: principles and applications*; 2nd ed.; Wiley: Chichester ; New York, 2001.
- (21) Rockwood, A. L.; Busman, M.; Udseth, H. R.; Smith, R. D. *Rapid Comm Mass Spectrom* **1991**, *5*, 582-585.
- (22) Cech, N. B.; Krone, J. R.; Enke, C. G. *Anal. Chem.***2001**, *73*, 208-213.
- (23) Iribarne, J. V.; Thomson, B. A. *J. Chem. Phys.* **1976**, *64*, 2287-2294.
- (24) de la Mora, J. F. *Anal. Chim. Acta* **2000**, *406*, 93-104.
- (25) Roepstorff, P.; Fohlmann, J. *J. Biomed. Mass Spectrom.* **1984**, *11*, 601.
- (26) Horn, D. M.; Ge, Y.; McLafferty, F. W. *Anal. Chem.* **2000**, *72*, 4778-4784.
- (27) Stensballe, A.; Jensen, O. N.; Olsen, J. V.; Haselmann, K. F.; Zubarev, R. A. *Rapid Commun. Mass Spectrom.* **2000**, *14*, 1793-1800.
- (28) Hakansson, K.; Cooper, H. J.; Emmett, M. R.; Costello, C. E.; Marshall, A. G.; Nilsson, C. L. *Anal. Chem.* **2001**, *73*, 4530-4536.
- (29) Zubarev, R. A.; Kelleher, N. L.; McLafferty, F. W. *J. Am. Chem. Soc.***1998**, *120*, 3265-3266.
- (30) Schnier, P. D.; Price, W. D.; Strittmatter, E. F.; Williams, E. R. *J. Am. Soc. Mass Spectrom.* **1997**, *8*, 771-780.
- (31) Mabud, M. D. A.; Dekrey, M. J.; Cooks, R. G. *Int. J. Mass Spectrom.* **1985**, *67*, 285-294.
- (32) Coon, J. J.; Syka, J. E. P.; Schroeder, M. J.; Shabanovitz, J.; Hunt, D. F. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9528-9533.
- (33) Wells, J. M.; McLuckey, S. A. *Biol. Mass Spectrom.* **2005**, *402*, 148-185.
- (34) Sleno, L.; Volmer, D. A. *J. Mass Spectrom.* **2004**, *39*, 1091-1112.
- (35) Little, D. P.; Speir, J. P.; Senko, M. W.; Oconnor, P. B.; Mclafferty, F. W. *Anal. Chem.***1994**, *66*, 2809-2815.
- (36) Biemann, K.; Scoble, H. A. *Science* **1987**, *237*, 992-998.
- (37) Chait, B. T.; Wang, R.; Beavis, R. C.; Kent, S. B. *Science* **1993**, *262*, 89-92.
- (38) Little, D. P.; Speir, J. P.; Senko, M. W.; O'Connor, P. B.; McLafferty, F. W. *Anal. Chem.* **1994**, *66*, 2809-2815.

- (39) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2007**, *129*, 5887-5897.
- (40) Dongre, A. R.; Jones, J. L.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **1996**, *118*, 8365-8374.
- (41) Burlet, O.; Yang, C. Y.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 337-344.
- (42) Paizs, B.; Suhai, S. *Mass Spectrom. Rev.* **2005**, *24*, 508-548.
- (43) Boyd, R.; Somogyi, A. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1275-1278.
- (44) Biemann, K. *Biomed. Environ. Mass Spectrom.* **1988**, *16*, 99-111.
- (45) Yalcin, T.; Khouw, C.; Csizmadia, I. G.; Peterson, M. R.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **1995**, *6*, 1165-1174.
- (46) Harrison, A. G. *Mass Spectrom. Rev.* **2009**, *28*, 640-654.
- (47) Yalcin, T.; Csizmadia, I. G.; Peterson, M. B.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **1996**, *7*, 233-242.
- (48) Nold, M. J.; Wesdemiotis, C.; Yalcin, T.; Harrison, A. G. *Int. J. Mass Spectrom. Ion Proc.* **1997**, *164*, 137-153.
- (49) Paizs, B.; Lendvay, G.; Vekey, K.; Suhai, S. *Rapid Commun. Mass Spectrom.* **1999**, *13*, 525-533.
- (50) Rodriguez, C. F.; Shoeib, T.; Chu, I. K.; Siu, K. W. M.; Hopkinson, A. C. *J. Phys. Chem. A* **2000**, *104*, 5335-5342.
- (51) Oomens, J.; Young, S.; Molesworth, S.; van Stipdonk, M. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 334-339.
- (52) Yoon, S.; Chamot-Rooke, J.; Perkins, B.; Hilderbrand, A. E.; Poutsma, J.; Wysocki, V. H. *J. Am. Chem. Soc.* **2008**, *130*, 17644-17645.
- (53) Bythell, B.; Erlekam, U.; Paizs, B.; Maitre, P. *Chem. Phys. Chem.* **2009**, *10*, 883-885.
- (54) Chen, X.; Yu, L.; Steill, J. D.; Oomens, J.; Polfer, N. *J. Am. Chem. Soc.* **2009**, *131*, 18272-18282.
- (55) Chen, X.; Steill, J. D.; Oomens, J.; Polfer, N. C. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1313-1321.
- (56) Bleiholder, C.; Osburn, S.; Williams, T.; Suhai, S.; van Stipdonk, M. J.; Harrison, A. G.; Paizs, B. *J. Am. Chem. Soc.* **2008**, *130*, 17774-17789.

- (57) Harrison, A. G.; Young, A. B.; Bleiholder, B.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128*, 10364-10365.
- (58) Erlekam, U.; Bythell, B.; Scuderi, D.; Van Stipdonk, M. J.; Paizs, B.; Maitre, P. *J. Am. Chem. Soc.* **2009**, *131*, 11503-11508.
- (59) Harrison, A. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2248-2253.
- (60) Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1776-1780.
- (61) Molesworth, S. P.; Van Stipdonk, M. J. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1322-1328.
- (62) Molesworth, S.; Osburn, S.; Van Stipdonk, M. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1028-1036.
- (63) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2005**, *127*, 17154-17155.
- (64) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. *J. Am. Chem. Soc.* **2008**, *130*, 17644-17645.
- (65) Chen, X.; Yu, L.; Oomens, J.; Steill, J. D.; Powell, D. H.; Polfer, N. In *Proceedings from the 57th ASMS Conference on Mass Spectrometry and Allied Topics* Philadelphia, PA, 2009, p WOAp3:30.
- (66) Polfer, N. C.; Bohrer, B. C.; Plasencia, M. D.; Paizs, B.; Clemmer, D. E. *J. Phys. Chem. A* **2008**, *112*, 1286-1293.
- (67) Garcia, I.; Giles, K.; Bateman, R. H.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 609-613.
- (68) Chen, X.; Powell, D. H.; Polfer, N. C. In *Proceedings from the 56th American Society for Mass Spectrometry Conference* Denver, CO, 2008, p MPMM 334.
- (69) Gucinski, A. C.; Somogyi, A.; Chamot-Rooke, J.; Wysocki, V. H. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1329-1338.
- (70) Bythell, B.; Somogyi, A.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 618-624.
- (71) Somogyi, A. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1771-1775.
- (72) Fattahia, A.; Zekavata, B.; Solouki, T. *J. Am. Soc. Mass Spectrom.* **2009**, *21*, 358-369.
- (73) Aebersold, R.; Goodlett, D. R. *Chem. Rev.* **2001**, *101*, 269-295.

- (74) Fernandez-Lima, F. A.; Becker, C.; McKenna, A. M.; Rodgers, R. P.; Marshall, A. G.; Russell, D. H. *Anal. Chem.* **2009**, *81*, 9941-9947.
- (75) Stokes, A. A.; Clarke, D. J.; Weidt, S.; Langridge-Smith, P.; Mackay, C. L. *Int. J. Mass Spectrom.* **2010**, *289*, 54-57.
- (76) Strohmalm, M.; Novak, P.; Pompach, P.; Man, P.; Kavan, D.; Witt, M.; Dzubak, P.; Hajduch, M.; Havlicek, V. *J. Mass Spectrom.* **2009**, *44*, 1565-1570.
- (77) Tipton, J. D.; Carter, J. D.; Mathias, J. D.; Emmett, M. R.; Fanucci, G. E.; Marshall, A. G. *Anal. Chem.* **2009**, *81*, 7611-7617.
- (78) Li, L. J.; Masselon, C. D.; Anderson, G. A.; Pasa-Tolic, L.; Lee, S. W.; Shen, Y. F.; Zhao, R.; Lipton, M. S.; Conrads, T. P.; Tolic, N.; Smith, R. D. *Anal. Chem.* **2001**, *73*, 3312-3322.
- (79) Patrie, S. M.; Robinson, D. E.; Meng, F. Y.; Du, Y.; Kelleher, N. L. *Int. J. Mass Spectrom.* **2004**, *234*, 175-184.
- (80) Marshall, A. G.; Hendrickson, C. L.; Jackson, G. S. *Mass Spectrom. Rev.* **1998**, *17*, 1-35.
- (81) Marshall, A. G. *Int. J. Mass Spectrom.* **2000**, *200*, 331-356.
- (82) Amster, I. J. *J. Mass Spectrom.* **1996**, *31*, 1325-1337.
- (83) Lawrence, E. O.; Livingston, M. S. *Phys. Rev. Lett.* **1932**, *40*, 555.
- (84) Sommers, H.; Thomas, H. A.; Hipple, J. A. *Phys. Rev.* **1951**, *82*, 697-702.
- (85) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *26*, 489-490.
- (86) Comisarow, M. B.; Marshall, A. G. *Chem. Phys. Lett.* **1974**, *25*, 282-283.
- (87) Dehmelt, H. *Reviews of Modern Physics* **1990**, *62*, 525-530.
- (88) Schaub, T. M.; Hendrickson, C. L.; Horning, S.; Quinn, J. P.; Senko, M. W.; Marshall, A. G. *Anal. Chem.* **2008**, *80*, 3985-3990.
- (89) Comisarow, M. B. *J. Chem. Phys.* **1978**, *69*, 4097-4104.
- (90) Marshall, A. G.; Verdun, F. R. *Fourier Transforms in NMR, Optical, and Mass Spectrometry: A User's Handbook*; First Edition ed.; Elsevier: Amsterdam, 1990.
- (91) Barrow, M. P.; Burkitt, W. I.; Derrick, P. J. *Analyst* **2005**, *130*, 18-28.
- (92) Feng, X. D.; Clipston, N.; Brown, T.; Cooper, H.; Reuther, U.; Hirsch, A.; Derrick, P. J.; Drewello, T. *Rapid Comm Mass Spectrom* **2000**, *14*, 368-370.

- (93) Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardmanc, M.; Cooks, R. G. *J. Mass Spectrom.* **2005**, *40*, 430-443.
- (94) Ledford, E. B.; Rempel, D. L.; Gross, M. L. *Anal. Chem.* **1984**, *56*, 2744-2748.
- (95) Jeffries, J. B.; Barlow, S. E.; Dunn, G. H. *Int. J. Mass Spectrom.* **1983**, *54*, 169-187.
- (96) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 3248-3250.
- (97) Bomse, D. S.; Woodin, R. L.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1979**, *101*, 5503-5512.
- (98) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. *Chem. Phys. Lett.* **1979**, *63*, 630-636.
- (99) Grant, E. R.; Schulz, P. A.; Sudbo, A. S.; Shen, Y. R.; Lee, Y. T. *Phys. Rev. Lett.* **1978**, *40*, 115-118.
- (100) Lehmann, K. K.; Scoles, G.; Pate, B. H. *Annual Review of Physical Chemistry* **1994**, *45*, 241-274.
- (101) Black, J. G.; Yablonovitch, E.; Bloembergen, N.; Mukamel, S. *Phys. Rev. Lett.* **1977**, *38*, 1131-1134.
- (102) Bagratashvili, V. N. *Multiple photon infrared laser photophysics and photochemistry*; Harwood Academic Publishers: Chur [Switzerland] ; New York, NY, 1985.
- (103) Polfer, N. C.; Oomens, J. *Mass Spectrom. Rev.* **2009**, *28*, 468-494.
- (104) Eyler, J. R. *Mass Spectrom. Rev.* **2009**, *28*, 448-467.
- (105) Shin, S. K.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1990**, *112*, 2057-2066.
- (106) Peiris, D. M.; Cheeseman, M. A.; Ramanathan, R.; Eyler, J. R. *J. Phys. Chem.* **1993**, *97*, 7839-7843.
- (107) Atrill, S. R.; Stace, A. *Phys. Chem. Chem. Phys.* **2000**, *2*, 823-830.
- (108) Cagmat, E. B.; Szczepanski, J.; Pearson, W. L.; Powell, D. H.; Eyler, J. R.; Polfer, N. C. *Phys. Chem. Chem. Phys.* **2010**, *12*, 3474-3479.
- (109) Mino Jr., W. K. S., J.; Pearson, W.L.; Powell, D.H.; Dunbar, R.C.; Eyler, J.R.; Polfer, N.C. *Int. J. Mass Spectrom.* In press
- (110) Englander, S. W.; Sosnick, T. R.; Englander, J. J.; Mayne, L. *Current Opinion in Structural Biology* **1996**, *6*, 18-23.

- (111) Anderegg, R. J.; Wagner, D. S.; Stevenson, C. L.; Borchardt, R. T. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 425-433.
- (112) Dharmasiri, K.; Smith, D. L. *Anal. Chem.* **1996**, *68*, 2340-2344.
- (113) Wagner, D. S.; Anderegg, R. J. *Anal. Chem.* **1994**, *66*, 706-711.
- (114) Englander, S. W. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1481-1489.
- (115) Engen, J. R.; Smith, D. L. *Anal. Chem.* **2001**, *73*, 256A-265A.
- (116) Busenlehner, L. S.; Armstrong, R. N. *Archives of Biochemistry and Biophysics* **2005**, *433*, 34-46.
- (117) Taylor, S. S.; Yang, J.; Wu, J.; Haste, N. M.; Radzio-Andzelm, E.; Anand, G. *Biochim. Biophys. Acta, Proteins Proteomics* **2004**, *1697*, 259-269.
- (118) Freitas, M. A.; Hendrickson, C. L.; Emmett, M. R.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **1998**, *9*, 1012-1019.
- (119) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 9765-9766.
- (120) Gard, E.; Green, M. K.; Bregar, J.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 623-631.
- (121) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53-71.
- (122) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9-14.
- (123) Chipuk, J. E.; Brodbelt, J. S. *J. Am. Soc. Mass Spectrom.* **2007**, *18*, 724-736.
- (124) Crestoni, M. E.; Fornarini, S. *J. Mass Spectrom.* **2003**, *38*, 854-861.
- (125) Freitas, M. A.; Marshall, A. G. *J. Am. Soc. Mass Spectrom.* **2001**, *12*, 780-785.
- (126) Mo, J. J.; Hakansson, K. *Anal. Chem.* **2007**, *79*, 7893-7898.
- (127) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1994**, *116*, 9765-9766.
- (128) Campbell, S.; Rodgers, M. T.; Marzluff, E. M.; Beauchamp, J. L. *J. Am. Chem. Soc.* **1995**, *117*, 12840-12854.
- (129) Harrison, A. G. *Mass Spectrom. Rev.* **1997**, *16*, 201-217.
- (130) Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1275-1281.

- (131) Wu, Z. C.; Fenselau, C. *J. Am. Soc. Mass Spectrom.* **1992**, *3*, 863-866.
- (132) Wu, J. Y.; Lebrilla, C. B. *J. Am. Chem. Soc.* **1993**, *115*, 3270-3275.
- (133) Zhang, K.; Zimmerman, D. M.; Chungphillips, A.; Cassady, C. J. *J. Am. Chem. Soc.* **1993**, *115*, 10812-10822.
- (134) Gronert, S.; Simpson, D. C.; Conner, K. M. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2116-2123.
- (135) Reid, G. E.; Simpson, R. J.; O'Hair, R. A. J. *Int. J. Mass Spectrom.* **1999**, *191*, 209-230.
- (136) Herrmann, K. A.; Kuppanan, K.; Wysocki, V. H. *Int. J. Mass Spectrom.* **2006**, *249/250*, 93-105.
- (137) Garcia, I.; Giles, K.; Bateman, R. H.; Gaskell, S. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1781-1787.
- (138) O'Hair, R. A. J.; Reid, G. E. *Eur. Mass Spectrom.* **1999**, *5*, 325-334.
- (139) Harrison, A. G. *Int. J. Mass Spectrom.* **2001**, *210*, 361-370.
- (140) Cooper, T.; Talaty, E.; Grove, J.; Van Stipdonk, M. J.; Suhai, S.; Paizs, B. *J. Am. Soc. Mass Spectrom.* **2006**, 1654-1664.
- (141) Bythell, B.; Molesworth, S.; Osburn, S.; Cooper, T.; Paizs, B.; Van Stipdonk, M. J. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1788-1798.
- (142) Molesworth, S.; Osburn, S.; van Stipdonk, M. J. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 2174-2181.
- (143) Oepts, D.; van der Meer, A. F. G.; van Amersfoort, P. W. *Infrared phys. technol.* **1995**, *36*, 297-308.
- (144) Valle, J. J.; Eyler, J. R.; Oomens, J.; Moore, D. T.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Hendrickson, C. L.; Marshall, A. G.; Blakney, G. T. *Rev. Sci. Instrum.* **2005**, *76*, 23103-23109.
- (145) Polfer, N. C.; Oomens, J.; Moore, D. T.; von Helden, G.; Meijer, G.; Dunbar, R. C. *J. Am. Chem. Soc.* **2006**, *128*, 517-525.
- (146) Bagratashvili, V. N.; Letokov, V. S.; Makarov, A. A.; Ryabov, E. A. *Multiple Photon Infrared Laser Photophysics and Photochemistry* Harwood, Chur., 1985.
- (147) Perkins, B.; Chamot-Rooke, J.; Yoon, S.; Gucinski, A.; Somogyi, A.; Wysocki, V. H. *J. Am. Chem. Soc.* **2009**, *131* 17528– 17529.

- (148) 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*, 1765-1769.
- (149) Roscioli, J. R.; McCunn, L. R.; Johnson, M. A. *Science* **2007**, *316*, 249-254.
- (150) Asmis, K. R.; Pivonka, N. L.; Santambrogio, G.; Brummer, M.; Kaposta, C.; Neumark, D. M.; Woste, L. *Science* **2003**, *299*, 1375-1377.
- (151) Fridgen, T. D.; McMahon, T. B.; MacAleese, L.; Lemaire, J.; Maitre, P. *J. Phys. Chem. A* **2004**, *108*, 9008-9010.
- (152) Moore, D. T.; Oomens, J.; van der Meer, A. F. G.; von Helden, G.; Meijer, G.; Valle, J. J.; Marshall, A. G.; Eyley, J. R. *Chem. Phys. Chem.* **2004**, *5*, 740-743.
- (153) Fridgen, T. D.; MacAleese, L.; Maitre, P.; McMahon, T. B.; Boissel, P.; Lemaire, J. *J. Phys. Chem. Chem. Phys.* **2005**, *7*, 2747-2755.
- (154) Chin, W.; Compagnon, I.; Dognon, J.-P.; Canuel, C.; Piuzzi, F.; Dimicoli, I.; von Helden, G.; Meijer, G.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 1388-1389.
- (155) Inokuchi, Y.; Kobayashi, Y.; Ito, T.; Ebata, T. *J. Phys. Chem. A* **2007**, *11*, 3209-3215.
- (156) Bleiholder, C.; Osburn, S.; Williams, T. D.; Suhai, S.; Van Stipdonk, M.; Harrison, A. G.; Paizs, B. *J. Am. Chem. Soc.* **2008**, *130*, 17774-17789.
- (157) Green, M. K.; Lebrilla, C. B. *Mass Spectrom. Rev.* **1997**, *16*, 53-71.
- (158) Gard, E.; Green, M. K.; Bregar, J.; Lebrilla, C. B. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 623-631.
- (159) Harrison, A. G.; Young, A. B.; Bleiholder, C.; Suhai, S.; Paizs, B. *J. Am. Chem. Soc.* **2006**, *128*, 10364-10365.
- (160) Polfer, N.; Paizs, B.; Snoek, L. C.; Compagnon, I.; Suhai, S.; Meijer, G.; von Helden, G.; Oomens, J. *J. Am. Chem. Soc.* **2005**, *127*, 8571-8579.
- (161) Frison, G.; van der Rest, G.; Turecek, F.; Besson, T.; Lemaire, J.; Maitre, P.; Chamot-Rooke, J. *J. Am. Chem. Soc.* **2008**, *130*, 14916-14917.
- (162) Lucas, B.; Gregoire, G.; Lemaire, J.; Maitre, P.; Glotin, F.; Schermann, J.-P.; Desfrancois, C. *Int. J. Mass Spectrom.* **2005**, *243*, 97-105.
- (163) Lucas, B.; Gregoire, G.; Lemaire, J.; Maitre, P.; Ortega, J.-M.; Rupenyan, A.; Reimann, B.; Schermann, J.-P.; Desfrancois, C. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2659-2663.

- (164) Simon, A.; MacAleese, L.; Maitre, P.; Lemaire, J.; McMahon, T. B. *J. Am. Chem. Soc.* **2007**, *129*, 2829-2840.
- (165) Vachet, R. W.; Bishop, B. M.; Erickson, B. W.; Glish, G. L. *J. Am. Chem. Soc.* **1997**, *119*, 5481-5488.
- (166) Case, D. A.; Pearlman, D. A.; Caldwell, J. W.; Cheatham III, T. E.; Ross, W. S.; Simmerling, C. L.; Darden, T. A.; Merz, K., M.; Stanton, R. V.; Cheng, A. L.; Vincent, J. J.; Crowley, M.; Tsui, V.; Radmer, R.; Duan, Y.; Pitera, J.; Massova, I. G.; Seibel, G. L.; Singh, U. C.; Weiner, P. J.; Kollmann, P. A.; University of California: San Francisco.
- (167) Gaussian 03, Revision C.02, Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A.; Gaussian, Inc.: Wallingford, CT, 2004.
- (168) Cieplak, P.; Cornell, W. D.; Bayly, C. I.; Kollmann, P. A. *Computat. Chem.* **1995**, *16*, 1357-1377.
- (169) Merrick, J. P.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683-11700.
- (170) Paizs, B.; Suhai, S. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 103-113.
- (171) Wyttenbach, T.; Bowers, M. T. *J. Am. Soc. Mass Spectrom.* **1999**, *10*, 9-14.
- (172) Bythell, B. J.; Knapp-Mohammady, M.; Paizs, B.; Harrison, A. G. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1352-1363.

BIOGRAPHICAL SKETCH

Xian Chen was born to Shanfeng Chen and Suyun Liao in Xiamen, a city on the southeast coast of China. She was the only child in the family. In 1999, she graduated from Xiamen No.1 Middle School after studying for six years. She then moved to Hefei, Anhui, to attend the University of Science and Technology of China (USTC), where she enrolled in the Department of Polymer Science and Engineering. She joined the group of Dr. Guangzhao Zhang and worked on her undergraduate research thesis entitled *Charaterization of Surfactants using Quartz Crystal Microbalance (QCM)*. In July 2004, she received a Bachelor of Engineering degree. A month after graduation, she flew overseas to Gainesville, Florida, to begin her graduate studies in analytical chemistry in the Department of Chemistry at the University of Florida. She joined Dr. Charles Cao's research group and got a masters degree with a thesis entitled *Synthesis of Semiconductor Nanocrystals using Selenium Dioxide* in 2007. She then joined Dr. Nicolas Polfer's research group, and began her work on investigating the influences on structures of peptide *b* fragment ions with infrared multiple photon dissociation, gas-phase hydrogen/deuterium exchange, and Fourier transform ion cyclotron resonance mass spectrometry. She received her Doctor of Philosophy from the University of Florida in December 2010.