DETERMINATION OF THE STRUCTURE AND SEQUENCE OF GAS-PHASE PEPTIDES USING SPECTROSCOPIC AND MASS SPECTROMETRIC METHODS

by

Joshua Lee Fischer

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Chemistry

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL

STATEMENT OF COMMITTEE APPROVAL

Dr. Timothy Zwier

Department of Chemistry, Co-Chair

Dr. Scott McLuckey

Department of Chemistry, Co-Chair

Dr. Paul Wenthold

Department of Chemistry

Dr. Julia Laskin

Department of Chemistry

Approved by:

Dr. Christine Hrycyna

Dedicated to family, for their unyielding emotional support and pride in my work.

ACKNOWLEDGMENTS

I sincerely attribute not only my success in undertaking this doctoral program, but my motivation, to my family foremost. Their unyielding support has been the cornerstone of my drive and has always provided me the motivation to overcome whatever challenges have been placed in front of me. Though I may not have expressed it directly in conversations with my family (to my mother's discontent), the frequent phone calls 'just to check in' always reminded me I have a home full of love supporting me.

Following my family, I was fortunate enough to receive tutelage under some incredible teachers, and would like to acknowledge those who played and essential role in my growth: Mr. McWilliams, my high school chemistry teacher who initially inspired me to pursue science; Dr. Allen, who introduced me how fulfilling academic science can be; Dr. Trimpin, who introduced me to mass spectrometry and its potential, Dr. Brock and Dr. Suits, for showing an unyielding commitment to not only me but all of their students.

I would also like to thank Dr. Zwier and Dr. McLuckey, who oversaw my career at Purdue. Dr. Zwier's passion and pure enjoyment of science was inspirational and his commitment to my success cannot be understated. Dr. McLuckey provided me with mentorship even before I joined the group and encouraged me to pursue the better path, confident in my abilities. Though it may seem trivial, the confidence that I gained from his encouragement cannot be understated. I would also like to thank the teachers not included. I am still coming to understand the role they have played in shaping not only my future.

My lab mates deserve equal acknowledgement. Given the unique lab, the people in it were no exception, and I feel like I truly learned a great deal about science and life from our exchanges.

Additionally, I would like to thank the staff at Purdue. None of this would be possible without support from JAFCI, Chemstores, Procurement (I needed a surprising amount of help from them), the Main Office and Copy Center, and the Business Office.

Last, I would like to thank my closest friends at Purdue for providing me respite from the academic life. Again, the support they offered me may seem trivial to them, but was invaluable to me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SUPPLEMENTAL FIGURES

LIST OF ABBREVIATIONS

AG2P	Ac-γ^2 -Phe-NHMe
AG4P	Ac-γ^4-Phe-NHMe
AIB	α-aminoisobutyric acid
AMBA	α-methylbenzylamine
BSSE	Basis set superposition error
CCDC	Cambridge Crystallography Data Center
CID	Collision induced dissociation
CMR	Covalent modification reaction
Da	Dalton
DDC	Direct dipolar current
DFT	Density Functional Theory
EDC	1-ehtyl-3-(3-dimethylaminopropyl)carbodiimide
ESI	Electrospray Ionization
eV	Electron volt
HOAt	1-hydroxy-7-aza-benzotriazole
HOBt	1-hydroxy-benzotriazole
IR	Infrared
IRIGS	Infrared ion gain spectroscopy
IRPT	Infrared population transfer
IRUVHB	Infrared ultraviolet hole burning
IRUVHF	Infrared ultraviolet hole filling
LD	Laser desorption
MALDI	Matrix-assisted laser desorption/ionization
MCP	Multichannel plate
MPD	Multiphoton dissociation
m/z	Mass-to-charge ratio
MS	Mass spectrometry

NHS	*N*-hydroxysuccinimide
NMR	Nuclear magnetic resonance
PT	Proton transfer
PTR	Proton transfer reaction
R2PI	Resonant two-photon ionization
REX	Rapid energy exchange
RF	Radio frequency
RIDIRS	Resonant ion-depletion infrared spectroscopy
SBA	Sulfobenzoic acid
ToF	Time-of-Flight
UV	Ultraviolet
YAG	Yttreium aluminum garnet
ZP	Zero-point
ZPE	Zero-point energy

ABSTRACT

The function of many biological processes depends on the structure and composition of the biomolecules involved. Both spectroscopy and mass spectrometry provide complimentary information regarding the three-dimensional conformation and the composition, respectively, as well as many other things. Here, double resonance conformer specific spectroscopy coupled with the latest *ab inito* computational methods is used to make structural assignments at the atomic resolution as well obtain information regarding propensities of intramolecular interactions. Additionally, rapid cooling in conjunction with IR excitation to modulate and measure the relative populations of conformers present in the expansion. Two different designer peptide systems are studied, including an achiral acylated α-aminoisobutryic acid dipeptide (Ac-AIB_2-R) with various C-terminal protecting groups (R=NHBn, NHBnF, α-methylbenzylamine) and an acylated γ^4-phenylalanine (Ac-γ^4Phe-NHMe) with the a methyl amine C-terminal protecting group. Mass spectrometry is used to determine the kinetics of gas-phase covalent tagging reactions used to enhance the sequence coverage. The covalent modification reactions utilize click chemistry between NHS or HOBt substituted sulfobenzoic acid tags with nucleophiles present on the residues of the amino acids composing the backbone. Effective temperatures are approximated using the Tolmachev model, which relates the statistical average internal energy of the molecule to a temperature.

CHAPTER 1. INTRODUCTION

1.1 Motivations for Understanding the Conformational Propensities of Natural and Designer Peptides in the Gas-Phase

Peptides are composed of amino acids and are the primary building block of proteins, which provide a significant portion of the machinery for life by catalyzing biochemistry. Many of the chemical mechanisms depend on the three-dimensional conformation of the protein, specifically the interplay between the substituents of the amino acids and the molecule of interest.[1, 2] However, proteins are large systems that are difficult to study sufficiently and, at the current time, peptides provide a convenient intermediate system to study. Studying the conformational preferences of peptides provides a means to mimic and isolate the active site of proteins and understand that chemistry, thereby improving the ability to utilize and develop new biochemical methods. The peptides are also able to adopt multiple three-dimensional conformations, or conformers, and our spectroscopic methods allow us to study a single conformer from the ensemble at a time. By studying the inherent inter and intramolecular forces that drive the conformational preferences of peptides, synthetic molecules that perform similar functions equivalent or exceeding proficiency may be designed, leading to designer peptides.[3, 4] This leads to potential new medicines, making it inherently imperative to understand these interactions at a fundamental level.[4] By studying peptides in the gas-phase, free of solvent, the physical properties of the molecule that drive the peptides conformational propensities can be examined,[5] allowing chemists to 'program' certain motifs. However, studying molecules free of solvated water could provide misinformation regarding the conformational preference in water, where the peptide will inherently perform its function. Water can contribute key stabilizing interactions for some conformers.[6] Gas-phase studies can account for this by adducting a single water molecule at a time, providing insight to the role water plays on the conformational propensities of these molecules.

1.2 Previous Spectroscopic Studies of Peptides in the Gas-Phase

Peptides and proteins have proven difficult to liberate into the gas-phase for interrogation, and a variety of methods have been developed. Here, a brief history will be provided, then in later sections the most relevant methods to this thesis will be disclosed in greater detail.

Studies on neutral peptides date back to the 1960s, where chromatographic methods were used to separate mixtures of amino acids using gas chromatography.[7] Volatilizing peptides composed of two or more amino acids proved difficult, as the temperatures needed to move peptide molecules from the solid-phase to the gas-phase often induced degradation of the peptides, and often peptides condense back to the solid-phase upon interaction with any surface.[8] Additionally, determining the conformation of a molecule in the gas-phase provides certain obstacles not apparent in solid or liquid phase determination, such as the number of molecules available for interrogation. Prior to the recent boom in electron microscopy, the most effective methods for determining protein structure were x-ray crystallography and nuclear magnetic resonance (NMR). Both of these methods are still commonly used today, and are information rich, but have a relatively high density of sample relative to gas-phase studies. This means that in the early years of peptide study, solid- and liquid-state studies enjoyed being able to use millions or billions of molecules to produce signal, where as gas-phase studies struggled to even liberate any molecules. With such a small number of molecules available, sensitive methods are required.

Issue volatilizing peptides have not been overcome to this day. Simply heating the sample to create a vapor-pressure, as described early, provides an excellent method to liberate mono and dipeptides (and, if fortunate, tripeptides) into the gas phase, but larger peptides require more elaborate methods such as laser desorption (LD). LD has been used to free peptide molecules as large as Gramacidin A into the gas-phase.[9] However, LD presents its own suite of problems, as the mechanism and, consequently, the effects that mechanism may play on the conformation of the peptides, remains unclear. Additionally, molecules of similar molecular weight and composition appear to volatilize with completely different efficiencies, meaning every experiment becomes its own adventure. These difficulties aside, LD is still the best method available to liberate larger neutral molecules into the gas-phase.

The study of ions has been greatly assisted by the developments of mass spectrometry (MS. MS introduced electrospray ionization (ESI) and matrix assisted laser-desorption/ionization (MALDI) methods to generate gas-phase ions of peptides in the late 1980's, resulting in a shared noble prize.[10, 11] There are several differences in the mechanisms and utility of these ionization methods, but the most influential difference in relation to this thesis is the amount of charge imparted onto the molecules. In MALDI, primarily singly charged ions are generated, where in ESI primarily multiply charged ions are generated. These ionization methods created a revolution,

leading to gas-phase studies of peptides of a gargantuan quantity and leading MS to become a primary tool used in medicine, pharmaceuticals, quality control, etc. This growth is due to not only these ionization methods providing a facile way to study ions of these biologically relevant molecules, but also the ease with which mass spectrometric methods can sequence a molecule piece by piece, allowing experimentalist to decode the building blocks of analytes.

With a facile ionization method now available to ionize peptides and proteins of seemingly any mass, study into the structure of these biologically relevant molecules began in earnest, with a wealth of information being produced.[12, 13] MS most commonly sequences, or determines the compositional order, of peptides and proteins, though the use of spectroscopic methods combined with quantum chemical computational modelling has allowed several groups to pursue conformational studies as well, though these studies are limited to peptides of relatively small size compared to what may be sequenced. This limitation is a result of spectral density, or the density of spectroscopic peaks within a unit of resolution. However, due to inherent peak broadening through quantum mechanical effects, no amount of available resolution can separate these peaks; the peaks are naturally broadened and therefore overlap, convoluting interpretation of the spectra. As the molecule grows, so do the number of transitions available, which leads to more overlap, further hindering spectral interpretation. Methods exist to help narrow these peaks. The most commonly used method is cooling of the molecule, which exponentially reduces the number of transitions and is the most commonly used method in this these. This will be further discussed in the experimental section 2.2 of this thesis, as well as briefly in each chapter.

1.3 Previous Mass Spectrometric Studies of Peptides

Mass spectrometry, often credited to be pioneered by J.J. Thompson in early 20th century,[14] exploded in popularity with the advent of electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) in the 1990s.[15] The reason for the sudden growth in use: Both methods were readily capable of volatilizing and ionizing large biomolecules, which had resisted any attempts to lift into the gas-phase easily. By generating gas-phase ions of biomolecules, fast methods could be (and have been) developed for separating and identifying these molecules from complex systems such as biological matrices.

Mass spectrometry, in the simplest terms, is a method that separates ions based on their mass to charge ratio (m/z). To oversimplify it, mass spectrometry separates ions based on their m/z and ,through innovative methods, can be used for a multitude of purposes from thermodynamics to analyte identification, and many more methods not described here. Once a molecule is both ionized (i.e. it has an electric charge caused by an inequivalent number of protons and electrons) and volatilized into the gas-phase, electric fields can be used to manipulate the ion. The kinetics of how the ion behaves in the electric field is naturally affected by the mass of the ion, and mass spectrometry uses a variety of methods to identify, separate, and interrogate these ions through these differences in kinetics, which will be discussed in more detail in the experimental section.

1.4 Experimental Approach: Spectroscopic and Mass Spectrometric Methods

Studying molecules in the gas-phase, due to the aforementioned low number density, requires sensitive methods. Fortunately, both spectroscopy and mass spectrometry are known for their high sensitivity and provide an excellent means to study these molecules in the gas-phase, though two very different kinds of insight can be gained through each respective method. These methods are both incredibly flexible and have utility beyond what is discussed here, though a 'general' description will be provided. Here, spectroscopy provides information about the quantum structure of molecules of interest, which can be used to infer information about the structure and three-dimensional conformation of said molecule. Mass spectrometry is used here to separate ions based on their mass-to-charge ratio (m/z), which can be used to isolate ions of particular m/z for subsequent experiments such as fragmentation into its substituent parts or gas-phase ion/ion reactions.

1.4.1 Spectroscopy

Spectroscopy, the study of the interaction of matter and light, provides a method to interrogate the quantum mechanical structure of molecules, which is sensitive to many properties of the molecule, including the molecules conformation. There are many kinds of spectroscopy, but this thesis will focus on various forms of identifying what wavelength (and therefore energy) of light is absorbed and how that relates to the quantum mechanical structure within a single conformation from the ensemble of conformers that may be present. Modern quantum

computational methods allow us to calculate an approximation of the quantum mechanical structure of a model structure, allowing us to predict what wavelengths of light we would expect the molecule to absorb as well as the relative stability of that models conformation to others. That is to say, if one can accurately calculate the quantum mechanical structure of a hypothesized conformational model of the molecule, they will be able to match that quantum mechanical structure to the molecules physical absorption of light and be able to verify the molecules three-dimensional conformation. Exact calculations of quantum mechanical structure remain impossible at this time, meaning the theory must be supported by experiment and vice-versa. Additionally, molecules are not restricted to adopting just one conformation and therefore multiple structures must be determined and verified against the experimental data, assuming the "best fit" structure is, at the very least, functionally close enough to the experimental structure adopted by the molecule. This process is aided by calculating the relative stability, or energy, of the test conformations. Low energy structures are more likely to be experimentally observed, assuming the theory is accurate. Only when all of these conditions are in agreement can we have confidence that the experiment is accurately described by the theory and assign structures. Structures can be assigned atomic resolution and matched to experimental data, affording a unique view of the three-dimensional structure and insights that may be gained, as well as provide information on the fundamental interactions of the atoms in the molecule with themselves, aided by the same computational support.

1.4.2 Brief Overview of Quantum Mechanics and Spectroscopic Transitions

In order to appreciate the benefits of cooling (as discussed in the review of previous studies outlined in the introduction), the quantum mechanical nature of spectroscopic transitions will be reviewed. This will be by no means exhaustive and is streamlined greatly such that the reader has a general idea of the kinds of the spectroscopic transitions available and how they affect molecular spectra.

Quantum mechanics revolves around the quantized nature of matter at the molecular level. Matter can only exist in 'states' that are made up of quantized energies. Spectroscopically, we are interested in probing the energy of these quantized states with electromagnetic radiation,

specifically light with discreet wavelengths and therefore discreet energies. This energy can be described accurately with Plank's Law:

$$E = \frac{hc}{\lambda} \qquad \text{Equation 1.1}$$

where E is the energy, h is Plank's constant, and λ is the wavelength of the light. We use electromagnetic radiation to add or remove quantized amounts of energy from the molecules.

Molecules, or matter really, exists in quantized, discreet, energetic states, meaning it has internal energies that will be defined by quantized electronic (energy of the electrons and their arrangement), vibrational (the energy of vibrating molecular bonds), and rotational (the energy of the molecule rotating along its three principal axis) states. Counterintuitively to these quantized energetic states, matter cannot be thought of as discrete particles on the quantum scale, but is better represented as a 'probability wave.' That is, the matter isn't necessarily in a specified location, but has a probability to exist in a range of locations. This probability is representable as a wave, and leads to the use wavefunctions. This is of course an oversimplification, but will suffice for this thesis, as going into the effects of superpositions and other quantum mechanical oddities will not be brief and is unnecessary for the bulk of the work of this thesis. The adaptation of these assumption will allow for an adequate discussion of the spectroscopic principles necessary.

As an example, the vibrational energies of a molecular bond can only have discrete energies obtained by solving the Schrödinger equation:

$$\widehat{H}\Psi = i\hbar\frac{\partial}{\partial t}\Psi \qquad \text{Equation 1.2}$$

where $\widehat{H}$ is the Hamiltonian operator, Ψ is the time-dependent wave function, and E is the energy. Ψ can be expressed as the product of separable time-dependent and independent terms,

$$\Psi = \psi e^{-\left(\frac{2\pi iE}{\hbar}\right)t} \qquad \text{Equation 1.3}$$

where ψ is the time-independent wavefunction. Note that ψ has no dependence on time, therefore we can separate ψ from Ψ, and reinsert ψ into $\hat{H}\Psi = i\hbar\frac{\partial}{\partial t}\Psi$E quation 1.2 to create a new expression. This expression can be rearranged such that all of the time-dependent and position dependent terms can be separated, meaning that both sides of the equation must equal some constant. This turns out to be E, and further manipulation yields:

$$\hat{H}\psi = E\psi \qquad \text{Equation 1.4}$$

Further, ψ can be separated into its own electronic, rotational, and vibrational components, meaning we can evaluate only the vibrational energies using

$$\hat{H}\psi_{vib} = E\psi_{vib} \qquad \text{Equation 1.5}$$

where ψ_{vib} is the vibrational component of the time-independent wavefunction.

A molecular vibration consists of the oscillation of a bond, or the oscillation of atomic nuclei. In order to evaluate this expression, $\hat{H}$ needs to be expanded upon.

$$\hat{H} = \hat{T} + \hat{V} \qquad \text{Equation 1.6}$$

where $\hat{T}$ is a kinetic energy operator and $\hat{V}$ is a potential energy operator. The $\hat{T}$ can be expressed using a traditional kinetic energy expression:

$$\hat{T} = \frac{1}{2}mv^2 = \frac{1}{2}mv^2\left(\frac{m}{m}\right) = \frac{\frac{1}{2}(mv)^2}{m} = \frac{p^2}{2m} \qquad \text{Equation 1.7}$$

where m is the mass, v is the velocity, and p is the momentum (that is $p = mv$. When working in quantum mechanics, where particles do not behave classically but as waves, an operator is needed to account for this discrepancy, particularly regarding the phase of the wave.[16]

$$\hat{p} = -(-i\hbar\nabla)$$ Equation 1.8

Therefore,

$$\hat{T} = \frac{\hbar^2}{2m}\nabla^2$$ Equation 1.9

Substituting $\hat{T} = \frac{\hbar^2}{2m}\nabla^2$ Equation 1.9 into $\hat{H} = \hat{T} + \hat{V}$ Equation 1.6 leaves

$$\hat{H} = \frac{\hbar^2}{2m}\nabla^2 + \hat{V}$$ Equation 1.10

The potential term is more complicated, but considering a molecular vibration is an oscillation between two states, a harmonic approximation is often used,[16] which effectively reduces the potential operator to a simple quadratic expression similar to a spring:

$$\hat{V} = \frac{k}{2}r^2 = \frac{m\omega^2}{2}r^2$$ Equation 1.11

where k is the spring constant and ω is the angular frequency. Substituting $\hat{V} = \frac{k}{2}r^2 = \frac{m\omega^2}{2}r^2$ Equation 1.11 into $\hat{H} = \hat{T} + \hat{V}$ Equation 1.6 gives:

$$\hat{H} = \frac{\hbar^2}{2m}\nabla^2 + \frac{m\omega^2}{2}r^2$$ Equation 1.12

Using a harmonic approximation of the vibrational potential, one can solve $\widehat{H} = \frac{\hbar^2}{2m}\nabla^2 + \frac{m\omega^2}{2}r^2$ Equation 1.12 and determine the energies, finding the vibrational energies to follow:[17]

$$E_n = \hbar\omega\left(n + \frac{1}{2}\right) \qquad \text{Equation 1.13}$$

where $E_n$is the energy of the vibration n, $\hbar$ is Plank's constant, ω is the angular frequency of the vibration, and n is the harmonic of the vibration. This equation shows that for the 0th order harmonic, the energy of the vibration will be $\frac{1}{2}\hbar\omega$, for the 1st harmonic the energy will be $\frac{3}{2}\hbar\omega$ and so on. Considering vibrations require atomic movement, energy is stored in these vibrations to facilitate such movement. The amount of internal energy in a molecule determines which quantized vibrational energies of the molecule can be accessed. That is to say, if the individual molecule doesn't have enough internal energy to vibrate at the $n = 3$ harmonic, then that vibrational state won't contribute to the energy of the molecule and that vibrations contribution to the internal energy will be a weighted sum of the n =1 and n =2 vibrational energies, with the weighting factors following a Boltzmann distrubution.

There are two important pieces of information the reader will want to know for this thesis available from $E_n = \hbar\omega\left(n + \frac{1}{2}\right)$ Equation 1.13:

1) The vibration can only contribute energies of $E = \hbar\omega\left(n + \frac{1}{2}\right)$ and that when $n = 0$ the vibration will still contribute energy. No energy between the integer values of n can be contributed (the quantum in quantum mechanics) and that even at $n = 0$, the vibration still contributes energy, known as the zero-point energy.
2) If the experimentalist tries to impart an energy to the vibration that does not equal a harmonic of the vibrational, the molecule has no way to quantum mechanically absorb the energy, so it will not.

Spectroscopy imparts that energy by exposing molecules to electromagnetic radiation, which is effectively an alternating electric potential. Considering vibrational spectroscopy, when the molecule is exposed to electromagnetic radiation equal to the difference in energy between vibrational states, there is a probability that the light will be absorbed and the molecule will

vibrationally "excite" upon absorbing the energy from the light. This can be thought of as the bond vibrating at a higher frequency, therefore having a higher internal vibrational energy (though of course, as is expected from quantum mechanics, this is an oversimplification). By finding the energies of these transitions, and assisted by the latest computation quantum chemistry, spectroscopists can map out the vibrational energies of the molecule. The frequency of these vibrations are sensitive to the environment of the bond, or the conformation of the molecule when considering a molecule isolated in a vacuum, allowing spectroscopists to determine the conformation of the molecule using quantum chemical models to predict what the frequency of these vibrational modes are.

In a general form, these rules apply to rotational and electronic states as well. It will be helpful to have an idea of the energy required for electronic, vibrational, and rotational transitions shortly. Looking back at $E = \frac{hc}{\lambda}$ Equation 1.1, we can see that the energy scales not linearly with wavelength, but as the inverse. To linearize the energy, spectroscopists tend to work with wavenumbers ($\tilde{\nu}$) instead of traditional units of energy such as J or eV.

$$\tilde{\nu} = 1/\lambda \qquad \text{Equation 1.14}$$

Substituting $\tilde{\nu}$ into $E = \frac{hc}{\lambda}$ Equation 1.1 yields the linear:

$$E = hc\tilde{\nu} \qquad \text{Equation 1.15}$$

Electronic transitions are the most energetic, ranging in the tens of thousands of wavenumbers, followed by vibrational transitions, ranging from tens to thousands of wavenumbers, to rotational transitions, ranging from less than a wavenumber to around 100 wavenumbers. Naturally, there are exceptions to these transition rules, and there are just as many exceptions to as cases which obey them. An excellent example is the coupling between rotational, vibrational, and electronic states, known as rovibronic coupling. Consider a molecule with an electronic ground state (denoted S_0) and transition to an electronic excited state (denoted S_1), as well as (oversimply) a rotational state R_1 and transition to rotational state R_2, and vibrational ground state ν_0 to rotational state ν_1. What happens if light is imparted onto a molecule in state S_0 R_1 ν_0 that is equal to the difference of all three transitions? Indeed, such a transition is possible, as well as any combination of these states. So, from the available three transitions described here we

get seven possible transitions. Clearly, in larger molecules which may have many tens or even hundreds of vibrational transitions as well as a multitude of rotational transitions available on the three principle rotational axis, this number grows quite large rather quickly. This compounds with the fact that multiple initial starting states become available as the samples temperature increases. At a higher temperature, the sample has more internal energy and therefore access to more rotational and vibrational states. It should be noted the probability of each combination of transitions varies and often most photon driven transitions are still not allowed. Even with this reduction, the electronic spectra of warm molecules, due to rovibrational coupling that is allowed, tend to resemble a single transition tens of thousands of wavenumbers broad, resolving none of these energy states we seek to identify. These are similar to the spectra typically obtained from room temperature UV-vis experiments.

This is where cooling the molecule to near zero K becomes advantageous. This greatly reduces the number of initial rotational and vibrational states the molecule can be in, exponentially reducing the number of transitions available and reducing the spectral complexity to such a degree that the electronic transitions of individual conformers can be discerned. The energy for electronic transitions tends to be so large that the molecule is often (correctly) assumed to only occupy the lowest energy state, the ground state. The rotational energies are typically not separated in electronic spectroscopy (though with high enough resolution they can be), which results in a broadening of the transition, though when cold the transition is still often only a few wavenumbers wide. The coupling between electronic and vibrational transitions are then separated by tens to thousands of wavenumbers, meaning electronic-vibronic transitions that are only a few wavenumbers wide are separated, creating unique progressions known as Frank-Condon progressions. By cooling the molecule to near 0 K, broad absorptions that would cover tens of thousands of wavenumbers are reduced to a series of discrete peaks that become manageable and allow experimental mapping of the energy states.

1.4.3 Mass Spectrometry

The ability of mass spectrometry to isolate ions in the gas-phase, its incredible sensitivity, and speed of operation provide unique opportunities and experiments. Ion/ion reactions are just one such example (and the focus in this thesis), but many other novel and innovative experiments

exist. Here, the ability of mass spectrometry to separate ions based on their m/z is used to isolate ions of opposite charges and store them within the same trapping device. The natural attraction due to electrostatic forces will draw the ions towards each other, causing them to coalesce and eventually form a cation/anion complex. The potential energy resulting from the electrostatic interaction is converted to internal energy and can result in chemical reactions. Further, the products of the ion/ion reactions can be subjected to additional activation to drive further reactions or fragmentation unique to the ion/ion complex. Here, ion/ion complexes with nucleophiles and reactive leaving group will be stored and then subjected to activation to induce substitution at various energies. The relative energies used to activate the complexes will be used to estimate a temperature and said temperature will be used to determine thermodynamic values for the reaction. These thermodynamic values can be used to inform us of the reaction with different leaving groups and nucleophiles and be used to inform reagent selection and development in the future.

1.5 Thesis Organization

This thesis will be organized into two parts: The conformational spectroscopy of neutral peptides followed by mass spectrometric methods to measure the reaction kinetics of gas-phase ion/ion peptide tagging reactions. In order to understand either of these sections, understanding of the experimental methods is vital, which will be discussed in Chapter 2. The experimental apparatus used in both sections will be discussed as well as some of the underlying theory of the experiments and the necessary conditions. Following will be Chapter 3, which focuses on the conformational spectroscopy of the designer dipeptide aminoisobutryic acid (Aib) with differing functional N-terminal caps. The function of the caps is twofold: to introduce spectroscopic sensitivity and direct conformational preferences. Aib peptides are known to adopt a 3_{10}-helix conformation with no preference between right- or left-handedness. Our goal was to use a chiral N-terminal cap in order to provide a steric directing force to induce preferential right- or left-handedness. Chapter 4 will then look at both the conformational spectroscopy of γ-peptides, as well as a spectroscopic method to induce selected conformers to isomerize into other conformers. γ-peptides have extended the carbon backbone of naturally occurring peptides allowing more flexibility in conformation. Additionally, the effect of moving the functional residue of the peptide to different locations along the backbone will be discussed. Previous studies on γ-Phe-NHMe

found a unique amide stacking motif and we sought to explore the effect of moving the phenyl ring of γ-Phe-NHMe to a different location along the backbone. Lastly, in Chapter 5, mass spectrometry is used to explore the kinetics of gas-phase covalent modification ion/ion reactions between peptides and tags functionalized with reactive leaving groups. The addition of amino acid specific tags to peptides has been done *in solu* to enhance sequence coverage. Control of these reactions can be enhanced even further using mass spectrometry, do to the control afforded by mass isolation, and an understanding of these kinetics will not only enhance tag and leaving group selection, but potentially provide inspiration for reagents for even more efficient reactions.

1.6 References

1. Dill, K. A., Dominant forces in protein folding. *Biochem.* **1990,** *29* (31), 7133-55.

2. Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R., The protein folding problem. *Annu. Rev. Biophys.* **2008,** *37*, 289-316.

3. Guo, L.; Almeida, A. M.; Zhang, W.; Reidenbach, A. G.; Choi, S. H.; Guzei, I. A.; Gellman, S. H., Helix formation in preorganized β/γ-peptide foldamers: hydrogen-bond analogy to the α-helix without α-amino acid residues. *Journal of the American Chemical Society* **2010,** *132* (23), 7868-7869.

4. Sadowsky, J. D.; Fairlie, W. D.; Hadley, E. B.; Lee, H.-S.; Umezawa, N.; Nikolovska-Coleska, Z.; Wang, S.; Huang, D. C. S.; Tomita, Y.; Gellman, S. H., (α/β+ α)-Peptide Antagonists of BH3 Domain/Bcl-xL Recognition: Toward General Strategies for Foldamer-Based Inhibition of Protein− Protein Interactions. *Journal of the American Chemical Society* **2007,** *129* (1), 139-154.

5. Zwier, T. S., Probing Frozen Molecular Embraces. *Science* **2012,** *335* (6069), 668-669.

6. James III, W. H.; Müller, C. W.; Buchanan, E. G.; Nix, M. G.; Guo, L.; Roskop, L.; Gordon, M. S.; Slipchenko, L. V.; Gellman, S. H.; Zwier, T. S., Intramolecular amide stacking and its competition with hydrogen bonding in a small foldamer. *Journal of the American Chemical Society* **2009,** *131* (40), 14243-14245.

7. Biemann, K.; Vetter, W., Separation of peptide derivatives by gas chromatography combined with the mass spectrometric determination of the amino acid sequence. *Biochemical and Biophysical Research Communications* **1960,** *3* (6), 578-584.

8. Posthumus, M. A.; Kistemaker, P. G.; Meuzelaar, H. L. C.; Ten Noever de Brauw, M. C., Laser desorption-mass spectrometry of polar nonvolatile bio-organic molecules. *Analytical chemistry* **1978,** *50* (7), 985-991.

9. Rijs, A. M.; Kabeláč, M.; Abo-Riziq, A.; Hobza, P.; de Vries, M. S., Isolated gramicidin peptides probed by IR spectroscopy. *ChemPhysChem* **2011,** *12* (10), 1816-1821.

10. Fenn, J. B.; Mann, M.; Meng, C. K.; Wong, S. F.; Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science* **1989,** *246*, 64.

11. Karas, M.; Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* **1988,** *60* (20), 2299-2301.

12. El-Aneed, A.; Cohen, A.; Banoub, J., Mass spectrometry, review of the basics: electrospray, MALDI, and commonly used mass analyzers. *Applied Spectroscopy Reviews* **2009,** *44* (3), 210-230.

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

14. Dahl, P. F., *Flash of the Cathode Rays: A History of JJ Thomson's Electron*. CRC Press: 1997.

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

16. Griffiths, D. J.; Schroeter, D. F., *Introduction to quantum mechanics*. Cambridge University Press: 2018.

17. Harris, D. C.; Bertolucci, M. D., *Symmetry and spectroscopy: an introduction to* vibrational *and electronic spectroscopy*. Courier Corporation: 1989.

CHAPTER 2. EXPERIMENTAL

2.1 Introduction

Multiple instruments were used to collect the data contained in this thesis and will be discussed very briefly here and in further detail in the subsequent chapters (each mentioned at the end of the brief description). Additionally, a very generalized and cursory review of spectroscopy has been presented to aide the reader not familiar with quantum mechanics. This is a very distilled interpretation and for a more complete presentation, the reader is encouraged to consult the following texts.[1-3]

The spectroscopic data was acquired using a homebuilt resonant two photon ionization (R2PI) chamber equipped with multiple lasers as well as a heating source and a laser desorption source used to liberate peptides into the gas-phase. Once in the gas-phase, the analyte molecules were entrained in a neutral buffer gas, either Ne or Ar, and introduced into a vacuum chamber via a supersonic expansion, which also cools the molecules to vibrational temperatures of ~10K. The supersonic expansion is sent through a skimmer cone, forming a molecular beam that passes through the extraction region of a Wiley-McLaren time-of-flight (ToF) mass spectrometer. In this region, the analyte molecules are interrogated spectroscopically, leading to the formation of ions through R2PI, which are extracted into the ToF tube and detected at a multi-channel plate (MCP) detector. Modulation of this ion intensity at specific wavelengths is used to generate experimental spectra This chamber will be discussed further in section 2.5.1.

Mass spectrometric data was acquired using multiple method and chambers. Studies using multi-photon dissociation (MPD) were carried out on a homebuilt 3D-ion trap. Ion were generated using nano electrospray ionization (nESI), with two electrodes aligned antiparallel to each other operating at opposite polarities, and collected into a quadrupole ion trap, then transferred into a 3D ion trap, where the ions were interrogated using various dissociation methods. A CO_2 laser was aligned such that the beam was focused to a point that spatially overlaps with the ion cloud within the ion trap. Additionally, waveform generators were used to spatially excite the ions at their secular frequency, known as collisionally induced dissociation (CID) and/or a direct dipolar current (DDC) was placed on the endcap electrodes, physically moving the ions from their region of stability within the center of the ion trap. This chamber will be further discussed in section 2.5.2.

A modified SciEx Qtrap 4000 was used to perform ion-ion reactions. Ions were generated using two nESI sources both aligned such that their sprays enter the same conical orifice of the instrument. Due to ions of opposite polarities requiring opposite voltages, voltage switching throughout the instrument was used to move ions of both polarities into the same quadrupole ion trap, where ion-ion reactions could occur. This instrument and process will be detailed more in section 2.5.3.

2.2 Sample Handling and Pulse Valve Operation for Spectroscopy

Analyte was volatilized using either heating or laser desorption. For heating, the sample is first placed inside of glass wool and then inside of a glass crucible. The glass crucible has a hole drilled in the one side to allow the passage of the carrier gas. The glass wool insulates the sample from the walls of the glass crucible, which slows heat transfer and mitigates thermal decomposition of the sample. The crucible packed with the glass wool is then loaded into a sample holding chamber preceding a pulse valve. When the valve opens, gas will flow through the sample holder before escaping the chamber through the pulse valve into the instrument. The pulse valve, loaded with the sample, is wrapped in a resistive heating rope and placed into the vacuum chamber of the R2PI chamber and the filled with a backing gas, either helium or neon. The pressure of the backing gas depends on the sample, as the degree of cooling effects (such as adduct formation) vary from sample to sample.

A current is run through the heating rope, heating the sample holding chamber and pulse valve to the desired temperature, monitored with a thermocouple wrapped between the heating rope and the sample holding chamber. Due to the high resistance of the heating rope (~60 ohms), passing a current through the rope heats the sample holder and the pulse valve. Note that the pulse valve has a manual fitting at the end that allows for the flow of the pulse valve to be adjusted by controlling how far the pulse valve may open and therefore how quickly gas can escape from the pulse valve into the chamber. This fitting is sealed with an o-ring that must be able to retain its integrity upon heating, and Kalrez o-rings rated for temperatures of up to 325 °C were used for all experiments. Additionally, the valve itself is sealed with a poppet that is moved with a solenoid that also must be able to retain integrity upon heating. A PEEK poppet was used to this end. The valve solenoid is driven by a high voltage driver built by the Jonathan Amy Facility for Chemical

Instrumentation, which provides and initial high voltage pulse to 'drive' the valve open quickly (typically hundreds of volts, into the kilovolt range), followed by a lower voltage that holds the valve open. Typical operation times were ~10 μs for high voltage opening, followed by ~50 μs holding times. Both times can be varied to affect expansion quality and therefore sample cooling efficiency.

2.2.1 Volatilization Methods: Heating and Laser Desorption

Heating the sample simply heats the outside of the pulse valve with the heating rope, which in turn heats sample holder and the backing gas inside the pulse valve and consequently the sample. Upon acquiring enough thermal energy, the sample will change phases from either solid or liquid to gas. Once in the gas phase, the sample is entrained in the backing gas and when the pulse valve opens, sample is carried out of the sample holder into the experimental chamber, which is under vacuum (~$1x10^{-6}$ mtorr).

Laser desorption is a much more complicated ordeal and not as well understood, but nonetheless delivers the desired result: Providing a volatilization method gentle enough to leave the sample intact and is particularly useful for larger samples that are subject to thermal degradation before reaching a high enough energy to volatilize.

To implement laser desorption, a flange orthogonal to the beam of pulse valve had a hole (~3 cm in diameter) bored, surrounded by an O-ring groove on the laboratory side of the flange. A flat, square adapter seals against the O-ring using the pressure differential between the vacuum of the chamber and the ambient lab pressure. This adapter is attached to a load/lock assembly made of a hollow tube equipped with a sealing valve. A metal rod with a sample holder is inserted into the hollow tube, through the valve, through the adapter, and into the chamber where it is positioned ~1–3 mm in front of the pulse valve and directly below the valve orifice. To counteract the leverage against the seal cause by the lever arm inherent to the flange and the load/lock assembly, a stand was placed under the load/lock assembly. The sample holder is a flat bar of aluminum designed to hold a thin, rectangular graphite rod. The rods are made in house by sanding a round graphite rod into a thin strip. Sample is then cut into a fine powder using a razor blade and rubbed into the graphite rode before attaching it to the sample holder with painters' tape at either end. The sample inserts itself into pores and ridges in the graphite, and upon infrared (IR) laser irradiation (1064

nm, 4–6 mJ/pulse, ~1 cm diameter), the graphite absorbs the laser and heats rapidly, also heating the sample and volatilizing it into the gas-phase. The energy is quickly dispersed through the graphite rod, minimizing thermal degradation.

2.2.2 Supersonic Jet-Cooling

Upon volatilization and pulse valve operation, the buffer gas expands from the relatively high-pressure environment inside of the pulse valve into the relatively low-pressure environment inside of the vacuum chamber. The pressure differential causes the buffer gas carrying the sample molecules to expand. However, due to the small orifice that the buffer gas can take to exit the pulse valve and enter the vacuum chamber, the distribution of paths the molecules and carrier gas may take is limited, resulting in the conversion of the random translation motions kinetic energy into a narrow translational distribution in the direction of the expansion.[4, 5] This expansion processes is isentropic, in a perfect system. When observing the resulting expansion from the frame of reference of the flow of the expanding gas (as if the observer were moving along the same translational vector at the same velocity as the flow), the kinetic energy of the expanding gas becomes near zero. The number density of the buffer gas is many orders of magnitude greater than that of the sample. Sample contained in the buffer gas will collide with the buffer gas until it also has near zero kinetic energy from the moving perspective, resulting only gentle collisions between the carrier gas and the sample molecules entrained in the carrier gas. Therefore, any sample entrained in the buffer gas will undergo collisions with the buffer gas until the sample is moving with a similar velocity vector to the buffer gas, resulting in effective kinetic cooling to less than 1 K.[4] Naturally, nothing is perfect, resulting in only the center of the expansion cone leaving the high-pressure region behaving this way.[5] That is, the center of the expansion will behave the closest to ideal, while the edges will contain more random motion and therefore have more energetic collisions and cool less effectively. As the expansion travels, it will continue to expand, even in the ‘well-behaved’ center region, reducing the density and therefore the number of cooling collisions, resulting in decreased cooling efficiency as the expansion travels away from the pulse valve. The cooling efficiency decreases incredibly quickly with distance, meaning the majority of cooling is done only a few millimeters from the pulse valve.

Additionally, from the moving frame of refence of the buffer gas, any collisions between the sample and the buffer gas will be sufficiently gentle to cool the rotational and vibrational energies of the molecule. That is, if the molecule collides gently enough with a gas molecule, the energy of the molecular rotation or vibration will be imparted into the buffer gas molecule, removing internal energy from the molecule and transferring it to the buffer gas. Accordingly, the center of the expansion is separated from the overall expansion using a skimmer cone, as this region has the most uniform distribution and therefore cooling potential.

As described earlier, rotational and vibration cooling is substantial, resulting in rotational and vibrational temperatures on the order of 2–5 K and 10–15 K, respectively. Additionally, this cooling funnels conformers into low-energy conformations, reducing the conformational complexity and therefore the number of spectral peaks, allowing for conformer specific spectroscopy. If the cooling is fast relative to the rate of isomerization, conformers will be trapped in that conformational family, known as kinetic trapping. Small variations in conformation within that conformational family will typically be cooled, resulting in the lowest energy conformer of that family being observed. If the cooling is slow relative to the isomerization rate of the sample, only the lowest energy conformers will be observed.

Clearly, this cooling results in a reduction of the total number of conformers observed in the expansion, which is beneficial to the spectroscopist. Having hundreds of conformers available will result in hundreds of sets of spectroscopic peaks, creating spectral overlap that is unmanageable. The cooling is so complete and efficient that typically only around one to four conformers are observed in the expansion, greatly reducing the spectral complexity. Additionally, due to the rotational and vibrational cooling to zero-point levels, the number of transitions available is greatly reduced, affording incredibly narrow electronic transitions (on the order of ~10 $\tilde{\nu}$), compared to the transitions that can be tens of thousands of $\tilde{\nu}$ wide in analogous UV-Vis methods done at room temperature.

2.3 Double Resonance Spectroscopic Methods

Conformation specific spectroscopy requires the use of at least two lasers. In this thesis, no more than two lasers were used to probe the sample, though more elaborate experiments exist in which more lasers can be used to obtain complimentary data.[6] These lasers fire at rates of *f* and

f/2, where in this case *f* = 20 Hz and *f*/2 = 10 Hz. Data points are acquired at a frequency of 20 Hz, meaning every other shot can have sample interacting with both lasers. One laser is used as a "pump" laser, fixed at a single resonant wavelength, while the other laser is tunable, allowing for scans as functions of the tunable lasers wavelength to be obtained, known as a "probe" laser. The laser fired at 20 Hz is applied in every data point while the laser fired at 10 Hz is only applied to every other data point. One of the two lasers must be at a wavelength that is absorbed by the sample. The difference of the both data points, one containing the 20 Hz laser and the other containing both the 20 Hz and the 10 Hz laser, is taken, which is recorded as a single data point. If the light the one laser is not absorbed, no change will be observed between the two data points, resulting in a *zero* value if the difference of the pump/probe spectra is taken. If the light of the both lasers is absorbed by the sample, effectively a hole is 'burned' in the population and signal produced is reduced after taking the difference, resulting in a negative value when the difference of the pump/probe signals is taken. That is, this method actively baseline subtracts and generates a zero-baseline! The difference between these two data points is what the final plot is composed of. For example, the 20 Hz laser can be fixed on a conformer specific transition, even if other conformers are present in the expansion, signal depletion will only been seen when the 10 Hz laser is absorbed by the same conformer (and not any of the others, assuming the absorption wavelength is unique to that conformer), which is what allows conformer specific spectra to be generated. This scheme can also be switched, yielding different spectra. These methods will be discussed in greater detail below.

2.3.1 Spectroscopic Methods: Resonant Two-Photon Ionization (R2PI)

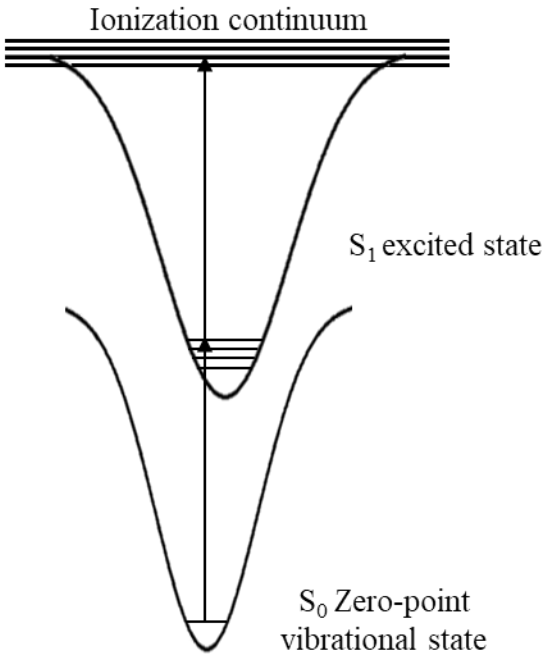


Figure 2.1. Schematic of resonant two-photon ionization.

R2PI involves the absorption of two subsequent photons of electromagnetic radiation, typically in the ultraviolet (UV) region of the electromagnetic spectrum. Considering the molecules are cooled to their zero-point vibrational levels, they all start in the same vibrational state. A photon is absorbed, promoting an electron to an excited electronic state and then, upon absorption of the subsequent second photon, enough internal energy is acquired by the electron to reach the ionization continuum, in which the excited electron enters and unbound state and the electron is ejected, resulting in cation formation. If the frequency of the laser is not resonant with a transition from the ground electronic zero-point vibrational state to some electronic excited state, the probability of absorption of the first photon of light is low and ions are not generated.

The ions are generated in the extraction region of a Wiley-McLaren time-of-flight mass spectrometer and thereby extracted to the detector upon creation, generating an ion current. The ion current is measured as a function of the frequency of the laser, generating UV spectra. Additionally, this ion current is used to monitor the response in double resonance experiments.

2.3.2 Spectroscopic Methods: Resonant Ion Depletion Infrared Spectroscopy

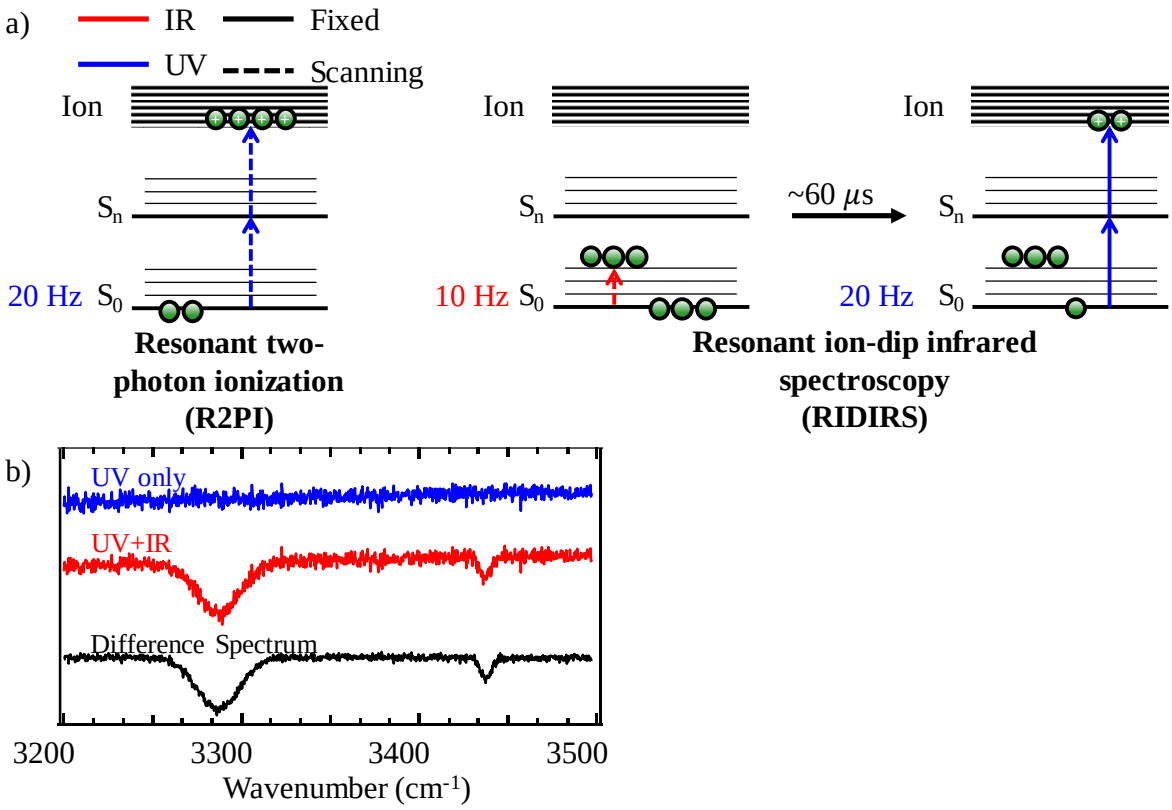


Figure 2.2. Schematic of a) Resonant two-photon ionization and Resonant ion-dip infrared spectroscopy methods and b) cartoon illustrating the effectiveness of active background subtraction methods used in both the above mentioned methods.

Infrared spectra were acquired using resonant ion depletion infrared spectroscopy (RIDIRS). RIDIRS requires an initial ion current to monitor, which is generated using R2PI fixed at a single conformer specific UV frequency. Every other laser pulse an IR laser is introduced temporally preceding and spatially overlapping the UV laser. If the frequency of the IR laser is resonant with a vibrational transition, population will be removed from the ground electronic zero-point vibrational state and into the ground electronic excited vibrational state before R2PI occurs. This leaves less population to be ionized by the UV laser and consequently results in a lower ion current, or ion depletion. This depletion is monitored as a function of the frequency of the IR laser, resulting in a conformer specific infrared spectra selected by the frequency of the UV laser. By utilizing active baseline subtraction, the active subtraction of datapoint n-1 from data point n, not only is this difference monitored, but systematic noise is mitigated.

2.3.3 Spectroscopic Methods: Infrared-Ultraviolet Hole Burning

The Infrared-Ultraviolet Hole Burning (IRUVHB) method uses a similar mechanism to RIDIRS, with the primary difference being which laser is fixed and which laser is held constant. In RIDIRS, the UV laser is fixed onto a conformer specific transition, while the IR laser is scanned, firing at frequency f and $f/2$ respectively. The IR laser still precedes the UV laser temporally, and the UV laser is still fired at frequency f and the IR laser at frequency $f/2$. In IRUVHB, the UV laser is scanned while the IR laser is fixed on a conformer specific transition. The mechanism is exactly the same at this point: When the UV laser is scanned over a resonant electronic transition, population is removed from the zero-point vibrational level. If the UV laser scans over an electronic transition of any conformer present in the expansion, signal is generated. However, considering the IR laser precedes the UV laser, the relative population of that conformer is depleted, resulting in a depletion of signal for all UV transitions belonging to that conformer. The difference spectrum produced will therefore only contain electronic transitions that share a transition with the IR transition selected. That is, by scanning the UV laser while holding the IR laser at a fixed frequency, conformer specific electronic spectra are generated.

2.3.4 Spectroscopic Methods: Infrared Population Transfer

Infrared Population Transfer (IRPT) is a double-resonance technique that allows for quantitative determination of the relative populations of different conformers present in the gas expansion. The UV laser is fixed on a conformer specific transition, generating ion signal via R2PI, while the IR laser is scanned across the range of interest. IRPT requires a three-step cooling process in which the sample is initially collisionally cooled upon exiting the pulse valve, heated with the IR laser as immediately as possible after the initial cooling (such that the expansion remains dense enough for cooling collisions to occur), and then cooled a second time. If the IR heating imparts enough energy to overcome isomerization barriers within the cooling timescale, the molecule will isomerize, depleting the initial population of that conformer, and increasing the population of the conformer(s) that the excited molecule isomerizes into. If the conformer that absorbs the IR radiation shares a transition with the conformer selected by the UV laser, decrease in ion signal generated from the UV laser will be observed. Conversely, if the conformer absorbing the IR radiation does not share the electronic transition selected by the UV laser, there will be a gain in

population caused by the unshared conformer isomerizing into the shared conformer, and therefore a gain in ion signal.

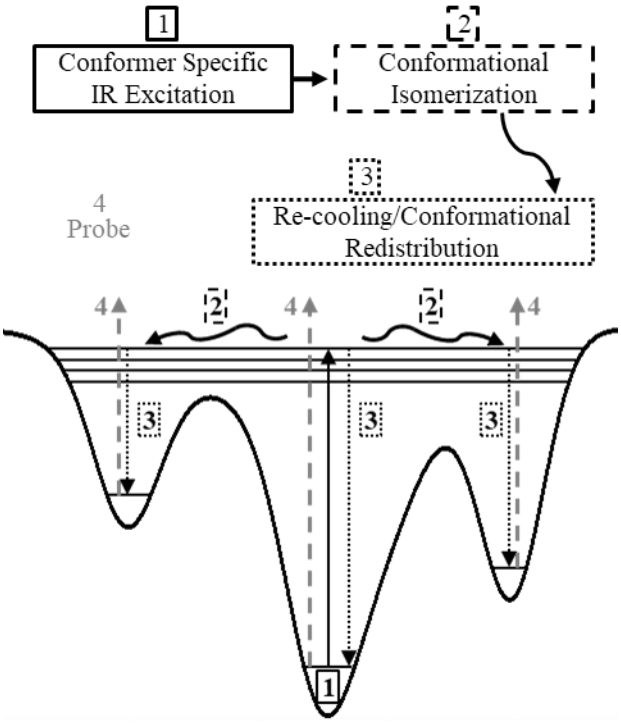


Figure 2.3. (*Top*) Flow-chart of IRPT steps. Borders around text and connecting arrows are coded to match the steps in the potential energy surface presented below. (*Bottom*) Cartoon depicting a mock potential energy surface of an IRPT process, with probed molecules starting from the 1 box and progressing following the attached arrows.

By monitoring the gain and depletion of various conformers selected by the UV laser, the relative populations of each conformer can be obtained according to the following equation:

$$0 = \sum_i F_i * I_i^{PT}(\tilde{\nu})$$

Equation 2.1

where F_i is the fractional abundance of conformer i, and $I_i^{PT}(\tilde{\nu})$ is the measured intensity of conformer i from the experiment. This equation is essentially a conservation of mass relationship: If a conformer isomerizes, the depletion of that conformers population requires a gain in the population of other conformers. By satisfying this equation along every point in the spectra,

stringent conditions are met affording high confidence that the all of the conformers are accounted for and their fractional abundances are accurate.[7-9]

2.4 Mass Spectrometry

2.4.1 Overview

Mass spectrometry uses electric fields to separate ions based on their mass-to-charge ratio (m/z). Various methods exist to accomplish ion separation, as well as various techniques to obtain information about the ions. Here, general information about the methods used in this thesis will be covered, mainly quadrupole ion trapping. First, ion trapping and the methods made available through the storage of ions will be covered. Further, due to the importance of ion/ion reactions to this work, the storage of opposite polarity ions and the resulting ion/ion reactions will be covered separately. Details of the experimental chambers used will be covered in the **Experimental Chamber** section.

2.4.2 Quadrupole Ion Trapping and Methods

Ion Trapping

The ion trapping methods used in this thesis were all performed within a quadrupole ion trap, therefore the theory of quadrupole ion trapping will be discussed briefly here. A quadrupole ion trap consists of a set of four rods, evenly spaced from each other radially. An RF voltage is applied to these rods, with any rod paired to the rod 180° apart, radially. Rods 90° apart are of opposite polarity, creating an alternating polarity of rods and the quadrupolar field. An auxiliary DC potential can also be applied to all the rods to apply the potential to the ions stored within the ion trap. Detailed derivation of the interactions of ions trapped within ideal quadrupolar field can be found here.[10] Briefly, the resulting potentials result in differential equations resembling the Mathieu equation, leading to the derivation of the unitless parameters a and q, represented by

$$a = \frac{8zeU_{DC}}{mr_0^2\Omega^2} \qquad \text{Equation 2.2}$$

$$q = \frac{4zeV_{RF}}{mr_o^2\Omega^2}$$ Equation 2.3

where z is the charge of the ion, e is the elementary charge, U_{DC} is the DC voltage applied to the rods, V_{RF} is the 0-peak voltage of the RF potential applied to the rods, m is the mass of the ion, $r_0$is the distance from the center of the ion trap to a rod, and Ω is the RF drive frequency applied to the rods. Note that both a and q are inversely dependent on the m/z of the ions.

Solving the Mathieu equation with these parameters leads to regions of stability found between the select combinations of a and q values, illustrated by the Mathieu stability diagram ($q = \frac{4zeV_{RF}}{mr_o^2\Omega^2}$ Equation 2.3), shown below. It is important to note that the a and q values are dependent on the m/z of the ions. Ions alternating in an RF field will have a natural secular frequency. The a and q values of the ions are determined by the potentials applied to the ions as well as their m/z. Ions are brought into the quadrupole region through a potential gradient, after which large common polarity is applied to trapping lenses at the ends of the quadrupole. The alternating RF voltage confines the ions movement and focuses the ions into the center of the trap radially. Axially, along the path the ions entered the trap, the ions are confined by the relatively high potential applied to the trapping lenses. Typically, the DC potential, U, is set to zero which results in the a value of the ions also being zero, which allows for the largest mass range of ions to be stored within the trap.

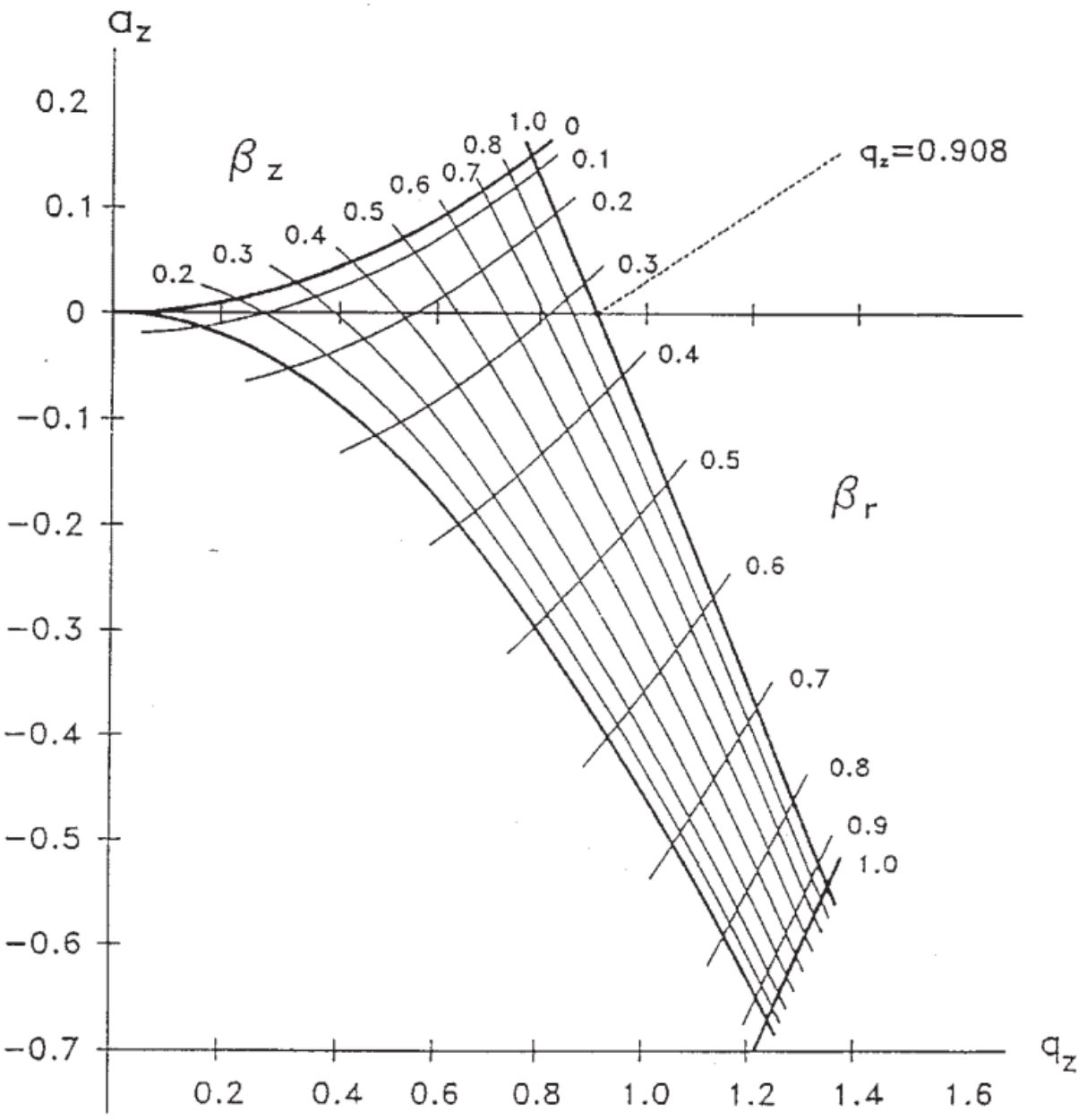


Figure 2.4. Mathieu stability diagram taken from March et al.[10]

These equations are for ideal ions that are 'born' into the quadrupole ion trap and undergo no perturbing forces, such as collisions with neutral gas molecules or the repelling electrostatic forces of other ions within the trap. However, in practice it is found that ions do not 'behave' well in such conditions and are difficult to trap, especially upon injection. This is mitigated by the introduction of a backing gas. The backing gas collisionally cools the kinetic energy of ions both as they enter the trap and as they are electrostatically repelled by other ions.

Ion Isolation

Once ions have been successfully trapped within the quadrupole, ions of particular m/z can be isolated by taking advantage of the Mathieu stability region. Two common methods are employed throughout this thesis and will be discussed: Apex isolation and Sequential Resonance Ejection Ramping.

Apex isolation is achieved by setting the a and q values such that ions of the desired m/z sit at the apex of the stability region. That is, V_{RF} and U are adjusted such that the ion of interest has an a value of 0.237 and a q value of 0.706. Ions of higher or lower m/z than the ion of interest will exist outside of the stability diagram, and either be ejected from the trap or neutralize by colliding with the rods of the quadrupole. In this work, this method is most commonly used by applying the required voltages to a transmission quadrupole, effectively converting the quadrupole into a mass filter. Only ions of the desired m/z can have a stable trajectory as they travel through the quadrupole, with ions of all other m/z being ejected from the quadrupole as they travel its length.

Sequential Resonance Ejection Ramping

Sequential resonance ejection ramping is accomplished through the secular frequency of the ion, indicated by the β lines shown in Figure 2.4. An auxiliary RF voltage is applied to the rods of the quadrupole. If the auxiliary RF frequency is equal to the secular frequency of an ion, it gains kinetic energy and travels along longer trajectories. If enough kinetic energy is gained, the ion no longer becomes stable within the trap and is either ejected or neutralized. As the typical value of U is zero, by ramping the value of V_{RF}, ions (a, q) coordinate will be $(0,q)$, meaning that ramping the value of V_{RF} will move ions along the $a = 0$ coordinate of the stability region. When

an ions q value falls upon the β line created by the auxiliary field, it will be excited, and with a large enough amplitude ejected. By using a ramping function for V_{RF} that contains all voltages except those that would successfully trap the ion of interest, all other ions are ejected.

Activation Methods: Collision Induced Dissociation and Direct Dipolar Current Activation

Once ions are stored in the ion trap, various methods can be used to impart energy into the ions, which induces fragmentation. Fragments that retain charge and still have a and q values within the stability region remain trapped. These fragment ions can then be mass resolved to generate a mass spectrum, known as MS/MS, or used for various purposes, such as determining the constituent sequence of the ion activated or used for further experiments. The use of fragments for further experiments is referred to as MS^n, where n refers to the number of times new species of ions are generated.

Collision Induced Dissociation

CID operates on the same principles as sequential resonance ejection ramping, with the only difference being a single value of V_{RF} is used, and at an amplitude that is not great enough to eject ions from within the trap. By selecting a single value for V_{RF}, only the m/z of interest is excited. By using a voltage low enough for the ion to remain trapped, the imparted kinetic energy is converted into internal energy through collisions with the backing gas of the trap. Energy imparted into the excited ion is redistributed through all of its vibrational modes with equal probability. If the ion gains enough internal energy that a vibrational mode is excited beyond the bond dissociation limit, the ion will fragment along said vibration. The resulting fragments that retain charge will have different m/z values, and therefore different q values, and be trapped within the quadrupole.

Direct Dipolar Current Activation

DDC activation activates ions through collisions, similar to CID. However, rather than exciting an ion of a single m/z, DDC is a broadband method that excites all ions stored within the trap. This is accomplished through applying an auxiliary DC potential to two opposing rods. This potential along one axis offsets the trapping potential created by the quadrupolar field, which

confines ions to the center. After adding the DC potential along the desired axis, the minimum of the trapping potential lies off center. When ions are stored away from the center of the trap, the trajectories become less stable, resulting in energetic collisions with the backing gas and consequential activation of all the ions stored within the quadrupole.

2.4.3 Ion/Ion Reactions

Ion/ion reactions are accomplished by storing both cations and anions within the same ion trap. The electrostatic attraction between the ions will naturally lead to ions coalescing into a cation/anion complex. The ions coalesce through a collapsing orbit, with the angular momentum removed through collisions with the backing gas. As the angular momentum is removed, the force maintaining the radius of the orbit is diminished, resulting in the orbit shrinking with each collision. Eventually, and remarkably quickly, the orbit will collapse completely, resulting in a cation/anion complex. During the orbital collapse, chemistry may occur depending on the ions involved, including proton transfer and electron transfer, for example.

In order to perform ion/ion reactions within a quadrupole ion trap, the trapping lenses must be modified from having the typical DC voltage applied to having an RF voltage. Typically, in order to store single polarity ions, relatively high voltages, of the same polarity as the ion, relative to the rods of the quadrupole, are applied to the trapping lenses. If negatively charged ions were stored within the same voltage parameters used to trap positive ions, the negatively charged ions would move towards the high positive voltages, displacing the ions from the center of the trap where the positive ions are confined, where they would eventually neutralize. By applying a RF voltage, ions of both polarities are focused towards the center of the trap.

2.5 Experimental Chambers

2.5.1 Resonant Two-Photon Ionization Chamber (R2PI)

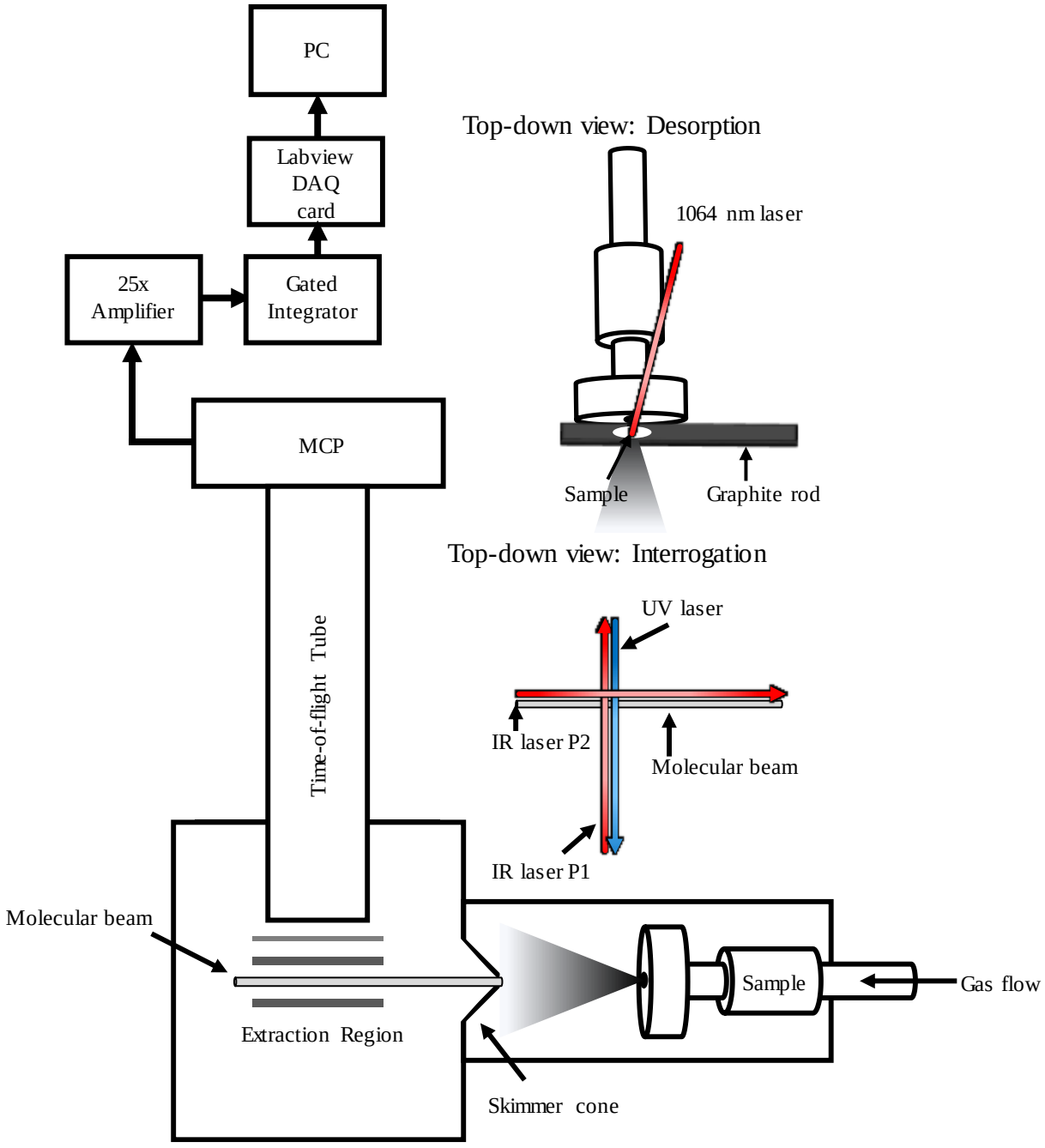


Figure 2.5. Schematic of the double resonance R2PI spectroscopy chamber.

The R2PI chamber is composed of three separate units: A sample injection chamber, an interrogation chamber, and a detection chamber, as well as exterior lasers. The sample injection chamber is shown to the right of Figure 2.5. This is where the sample is volatilized, entrained in the supersonic expansion, and cooled. To transfer from the sample from the injection chamber into the interrogation chamber, the resulting expansion passes through a skimmer cone, effectively creating a molecular beam. This molecular beam passes between the extraction plates of a Wiley-McLaren ToF mass spectrometer.

While in the extraction region, two different interrogation methods were used, depending on the experimental setup. The UV laser was always in the same configuration, propagating perpendicular to both the ToF chamber as well as the molecular beam, as shown in Figure 2.1. The IR laser can either counter propagate the UV laser (IR laser P1) or the molecular beam (IR laser P2). For RIDIRS experiments, P1 or P2 may be used. For IRPT experiments, only P2 may be used. Using P1 requires tighter alignment to achieve the required spatial overlap of the UV and IR laser, but due to being able to the see the laser path as it both enters and exits the instrument, alignments are actually easier. Using P2, because the laser counter-propagates the molecular beam, requires zero spatial overlap with the UV laser, but is much more difficult do to only being able to see where the beam enters the chamber; there is no way to know whether you're counter propagating the molecular beam or not once the chamber is sealed and under vacuum. However, for IRPT, the laser timing can be such that cooled molecules can be irradiated while still in a collision dense region, resulting in the second cooling necessary for IRPT.

The UV laser is used to generate ions through R2PI through the ejection of an electron. A description of R2PI can be found above, in section 2.3.1. The R2PI event occurs between the extraction plates of a Wiley-McLaren ToF. Once the analyte has acquired a charge, the high voltage of the plates (~4 kV) extracts the ions down the ToF tube, converting the potential energy imparted by the extraction plates into kinetic energy, where the ions will eventually collide with a multi-channel plate detector, generating signal. The amount of time it takes from the firing of the laser to when the ions reach the detector can be converted into the m/z using the following relationship:

$$t = \frac{d}{2U}\sqrt{\frac{m}{z}} \qquad \text{Equation 2.4}$$

where t is the arrival time, U is the voltage applied, and d is the length of the ToF tube. Due to the reliance on the UV laser to generate the ions, a photodiode that detected each laser pulse was used to trigger an oscilloscope used to collect data.

Signal generated by the MCP was sent through a 25x preamplifier before entering a gated integrator. When taking R2PI spectra, a bypass through the gated integrator was used to send the signal directly to the oscilloscope, where data was recorded directly from the oscilloscope using an in house LabView code. When taking RIDIRS or IRPT spectra, the gated integrator was used to perform active baseline subtraction to monitor any modulation of signal induced by the IR laser. The output of the active baseline subtraction through the gated integrator was then sent to the same oscilloscope and data was recorded using a second in house LabView code.

The lasers used with the R2PI chamber include two different Nd:YAG pumped dye lasers as well as a Nd:YAG pumped KTP/KTA optical parametric converter (LaserVision). The dye lasers were used to generate UV photons, while the optical parametric converter was used to generate IR photons. The primary UV laser used was a Radiant Labs Narrowscan dye laser using Coumarin 540A as the dye. In an auxiliary capacity, a Scanmate Pro dye laser using Coumarin 540A. Both dye lasers were pumped with Surelite II lasers (Continuum). The IR laser was pumped with a Surelite III (Continuum).

2.5.2 Modified SciEx 5600 QToF

For all mass spectrometry experiments discussed here, a modified SciEx 5600 QToF was used. A schematic for the instrument can be found in Figure 2.6. The instrument consists of quadrupole ion filter guide in front of a q-Q-q-ToF, and the ion optics are capable of polarity switching within the same scan function. Lower case q represents higher vacuum regions (<1mtorr), where capital Q represents lower vacuum regions (>1mtorr). The ionization source has been modified so that sequential injection of positively and negatively charged ions can be accomplished by utilizing separate power supplies. After ion injections, ions travel through a quadrupole ion guide (QJet), then through q0. Ions can be stored in q0, but in this thesis it was used as a mass filter. Mass filtering was accomplished by adjusting the a and q values using U_{DC} and V_{RF} such that the m/z of interest sits at the apex of the Mathieu stability diagram ($a = 2.300$,

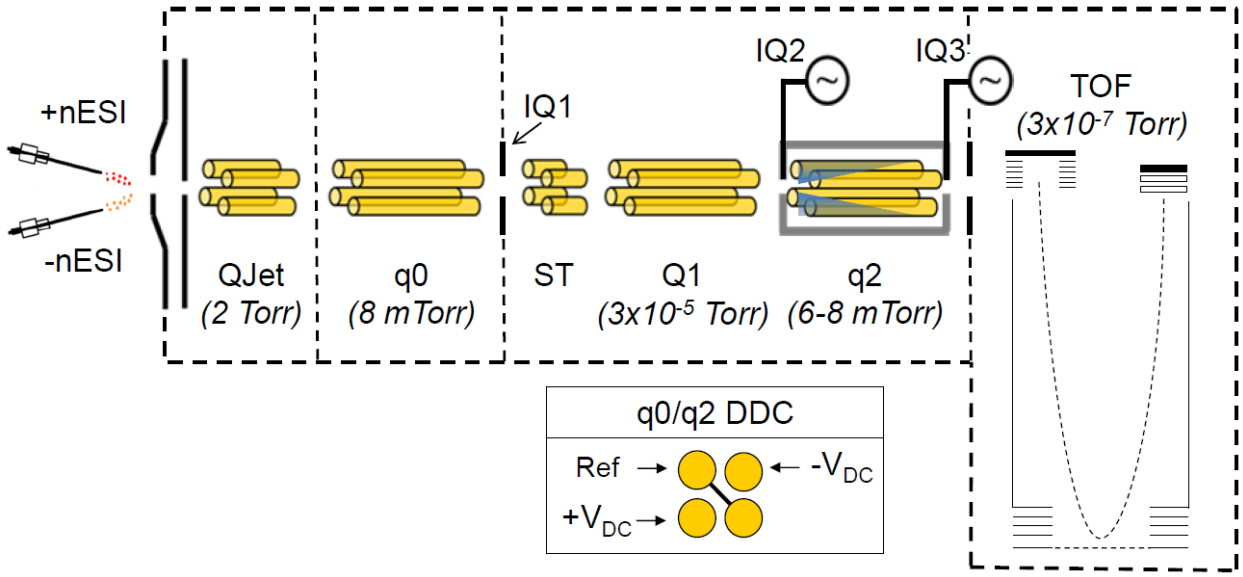


Figure 2.6. Schematic of the modified SciEx 5600 QToF instrument.

$q = 0.708$), destabilizing the path of all ions that are not of interest. Selected ions then travel through Q1, which is capable of storing ions of both polarities, and into q2, where they are eventually stored. Modifications to the trapping lenses surrounding Q1 and q2 allow the application of RF voltages, allowing the trapping of multiply charged ions in either quadrupole; in this thesis ions were always trapped in the higher pressure q2 for collisional activation purposes, which utilizes the higher collision frequency with the backing gas resulting from the higher pressure. Modifications to the rods of q0/q2 include the ability to apply voltages to a pair of opposing rods for DDC activation. The high vacuum in Q1 reduces the trapping and activation efficiency resulting from collisional cooling and activation, respectively. After the desired experiment is completed, a LINAC is used to focus the ions into a tighter packet, and injected through further focusing optics into a reflectron ToF for mass analysis.

2.6 Theoretical Techniques

The spectroscopic studies covered in this thesis were all supplemented with computational chemistry using either molecular force fields or density functional theory (DFT), often using both in conjunction. A summary of how the computational methods were used and applied is covered here.

Due to the conformational flexibility of the molecules studied, the conformational landscape is complex with a vast number of local minimum (with each local minimum resulting in a conformer). Isomerization of each conformer into another conformer is theoretically possible, but activation barriers must be overcome, and some conformers lie at such high relative energies relative to the global energetic minimum, the chance of their observation in the experiment is unlikely. Due to the high number of possible conformers, using high levels of theory to explore this potential energy surface is unreasonable, and a computationally cheaper method is used instead. To that end, the conformational landscape is explored using molecular force fields, which use semi-empirical non-quantum methods to determine the relative energies of populations. The lowest energy conformers found using the force field method are then subjected to further optimization using quantum DFT methods, which are used to determine more realistic relative conformer energies.

2.6.1 Exploring the Conformational Potential Energy Surface Using Molecular Force Fields

In order to use computational chemistry to assist our assignment of experimental spectra to conformers, a population of low energy conformers with predicted spectra is needed. As described above, the conformational potential energy surface is complex, and increases in complexity exponential as the molecule's atom count grows. A systematic exploration of the potential energy surface is impractical due to the computational cost. To that end, the MacroModel software (part of the Schrödinger software suite) was used to apply a Monte Carlo algorithm to randomize the dihedral angle of bonds in the submitted molecule. From there, either the OPLS3 or AMBER* force fields were used to minimize the energy of the resulting structure. It was empirically found that different conformers can be generated using OPLS3 and AMBER*, though most of the conformers generated are the same. However, if the experimentally observed conformer is not generated, our efforts to assign a structure to the experimental data will fail or be incorrect, meaning every possible low energy conformer is worth considering.

The Monte Carlo dihedral randomization method was then applied up to 500 times to the resulting minimized conformer, generating up to 500 new conformers. Naturally, some of the conformers randomized could be the same upon minimization, so a redundant conformer elimination was performed to remove duplicate conformers. The redundant conformer elimination

used the root-mean-square-error of the positions of the heavy atoms in the conformer, with a threshold of the heavy atoms in the conformer, with a threshold of the heavy atoms in the conformer, with a threshold of ±0.01 Å. Each conformer in the resulting list was then individually subjected to the Monte Carlo dihedral randomization process, minimized, and compared against the initial list to remove redundant conformers. This process was repeated with the resulting list until no new conformers were found.

2.6.2 Refining the Conformational Potential Energies using DFT

From the resulting list obtained using the molecular force field method, the lowest 100 energy conformers were submitted to further minimization using DFT using either the B3LYP with Grimme's dispersion correction (GD3BJ) or M05-2x functional with the 6-31(g)+d basis set using Gaussian 16 software. The vibrational frequencies of the resulting minimized structures were then calculated, and the resulting energies were used to determine the zero-point (ZP) energy of each conformer. The relative ZP energy of each conformer provides a rough account of the probability of seeing each conformer experimentally, and the vibrational energies are used to generate the predicted IR spectra of the conformers. It should be noted that DFT is an approximation method, meaning some density functionals will represent intramolecular interactions better than others.

2.7 References

1. Atkins, P. W.; De Paula, J., Physical chemistry. Oxford university press, Oxford UK: 1998.

2. Griffiths, D. J.; Schroeter, D. F., *Introduction to quantum mechanics*. Cambridge University Press: 2018.

3. Harris, D. C.; Bertolucci, M. D., *Symmetry and spectroscopy: an introduction to vibrational and electronic spectroscopy*. Courier Corporation: 1989.

4. Smalley, R. E.; Wharton, L.; Levy, D. H., Molecular optical spectroscopy with supersonic beams and jets. *Accounts of Chemical Research* **1977,** *10* (4), 139-145.

5. Lubman, D. M.; Rettner, C. T.; Zare, R. N., How isolated are molecules in a molecular beam? *The Journal of Physical Chemistry* **1982,** *86* (7), 1129-1135.

6. Hamilton, C. E.; Kinsey, J. L.; Field, R. W., Stimulated emission pumping: new methods in spectroscopy and molecular dynamics. *Annual Review of Physical Chemistry* **1986,** *37* (1), 493-524.

7. Buchanan, E. G.; James III, W. H.; Gutberlet, A.; Dean, J. C.; Guo, L.; Gellman, S. H.; Zwier, T. S., Single-conformation spectroscopy and population analysis of model γ-peptides: New tests of amide stacking. *Faraday discussions* **2011,** *150*, 209-226.

8. Dian, B. C.; Longarte, A.; Winter, P. R.; Zwier, T. S., The dynamics of conformational isomerization in flexible biomolecules. I. Hole-filling spectroscopy of N-acetyl tryptophan methyl amide and N-acetyl tryptophan amide. *The Journal of chemical physics* **2004,** *120* (1), 133-147.

9. Dian, B. C.; Longarte, A.; Zwier, T. S., Conformational dynamics in a dipeptide after single-mode vibrational excitation. *Science* **2002,** *296* (5577), 2369-2373.

10. March, R. E., An introduction to quadrupole ion trap mass spectrometry. *Journal of mass* spectrometry **1997,** *32* (4), 351-369.

CHAPTER 3. SINGLE-CONFORMATION SPECTROSCOPY OF CAPPED AMINOISOBUTYRIC ACID DIPEPTIDES: THE EFFECT OF C-TERMINAL CAP CHROMOPHORE ON CONFORMATION

3.1 Introduction

The forces giving rise to well-structured peptides and proteins have been of interest for decades.[1–3] In the 1950's, Pauling and co-workers emphasized hydrogen-bonding interactions in discerning alpha helices and beta sheets, the two dominant secondary structural motifs of proteins.[4] Donohue soon followed Pauling's alpha helix by proposing the 3_{10} helix.[5] These two helical forms, the alpha helix and the 3_{10} helix, are regularly observed in protein structures and are supported by intra-backbone hydrogen bonds closing 13- and 10-membered rings (C13 and C10), respectively, requiring only slight variations in ϕ/ψ backbone angles. Additional C10 structures, β–turns requiring only four successive residues, were first reported by Venkatachalum.[6] While the repeating backbone ϕ/ψ angles of the 3_{10}-helix (-60°/-30°) comprise Type III turns, Type I and Type II C10 structures require non-repeating ϕ/ψ angles along the peptide backbone chain to "turn" its direction.[6]

Despite early emphasis on hydrogen bonding patterns, the modest stability of folded proteins, relative to their unfolded configurations, means that a variety of forces, including short range van der Waals forces, as well as hydrogen bonding, are important for the unique folds of proteins.[1, 2] For example, single methylation and double methylation of glycine's alpha carbon (e.g., Gly vs Ala vs Aib) provides dramatic increases in the residue's helicogenicity.[7]

In biological settings, most proteinogenic helices are right handed α-helices having 3.6 residues/turn, with ϕ/ψ = -60°/-45°, supported by NH···O=C hydrogen bonds that close 13-membered rings (referred to as C13 hydrogen bonds).[8, 9] The 3_{10}-helix comprises about 10% of protein helices and is wound more tightly, with 3.0 residues/turn, $\phi/\psi \cong$ -57°/-30°, forming a series of C10 hydrogen bonds.[5, 9] Interestingly, in room temperature solution, the ϕ/ψ angles for a 3_{10}-helix are rather loosely defined due to the steric strain resulting from the necessary residue overlap of a helix having an integer number of residues/turn.[10] In some cases, α- and 3_{10}-helices can interconvert and it has been suggested that the 3_{10}-helix could be an intermediate state for the formation of α-helices.[10, 11] Indeed, the balance between α-helix and 3_{10}-helix can be tipped by the

presence of even one amino acid that directs formation of one helix over the other,[12, 13] suggesting that the C10 vs C13 hydrogen bonding pattern is not by itself determinative.

Aminoisobutryic acid (Aib) is a non-proteinogenic, naturally occurring amino acid with geminal methyl groups substituted on the α-carbon, which is known to favor 3_{10}-helix formation.[14, 15] The symmetric methyl substitution renders Aib achiral, removing the chiral influence directing helix formation and rendering left- and right-handed helices of poly-Aib peptide enantiomers equivalent in energy. Aib is most commonly found naturally in the peptaibol family of antimicrobial peptides,[15, 16] which have been used in foldamer studies to generate transmembrane pores.[17] The heliogenic character imparted by Aib is believed to drive the formation of these transmembrane pores, which result in antimicrobial activity.[17] Vibrational modes of the polar amide groups inherent to peptides are highly sensitive to their local environment, including their hydrogen-bonding network, and consequently are diagnostic of the secondary structures adopted by peptides and proteins. The NH stretch frequencies found in the Amide A region shift to lower frequencies with increasing hydrogen bond strength, providing a particularly useful diagnostic. However, in aqueous solution, these absorptions occur in the midst of the broad OH stretch band, masking their presence. Using nonpolar $CDCl_3$ to minimize spectral interference, Toniolo and co-workers have studied a set of Z-(Aib)$_n$-O*t*Bu (n=1–12) peptides in the Amide A region that pointed to these peptides undergoing 3_{10}-helix formation in non-polar solvents. [14] These authors identified two NH stretch fundamentals assigned to "free" and "intramolecularly hydrogen bound" NH groups. As n increases, the "hydrogen bound" band grew at the expense of the free NH stretch band. Decades later, Maekawa *et al.* used 2D-IR spectroscopy in the Amide I region (n = 3, 5, 8, 10) to investigate the C═O stretch vibrations, and showed that the onset of 3_{10}-helix formation occurs with $n > 5$.[18] More recently, infrared spectra from some of us, using ^{13}C-isotopologues of Z-(Aib)$_6$-O*t*Bu, showed isotope shifted localized ^{13}C═O Amide I modes that confirmed 3_{10}-helix formation.[19] Although isotope editing singles out the local mode absorptions of individual ^{13}C=O sub-units by shifting their absorptions to lower frequencies, these spectra are still influenced by solvent effects.

Studying the intramolecular interactions of peptides under jet-cooled conditions in the gas-phase provides the unique opportunity to interrogate the peptides without solvent influence.[2040] Often, gas-phase studies of peptides are difficult because peptides are nonvolatile, necessitating laser desorption to bring them into the gas phase. Furthermore, most do not have a natural

chromophore to sensitize their detection. Previous work has used phenyl ring caps as chromophores at either terminus of the peptides.[25–33] However, weak chromophore absorption, low number densities and shot-to-shot fluctuations associated with laser desorption, indicate that further experimental improvements would be welcome. Even with these shortcomings, desorption experiments have been used successfully to interrogate a range of peptides in the gas-phase.[22–40] The hydrogen-bonded networks present in these peptides can be probed in exquisite detail through the NH stretch and amide I/II regions, which reflect the hydrogen-bonded network present in the peptide. Comparisons to high-quality quantum chemical calculations routinely confirm the spectral assignments.

The present study on capped Aib peptides has as one point of comparison an earlier study from our group on the gas phase conformational preferences of Z-capped oligo-glycines up to n = 5. [28] Since glycine is the only naturally-occurring achiral amino acid, with its two hydrogen atoms on C(α), it is a seemingly close analog of the Aib peptides studied here. Interestingly, Z-$(Gly)_5$ folds not into an α- or 3_{10}-helix, but instead into an emergent form of a 14/16-mixed helix, with amide-amide hydrogen bonds that alternate direction, N→C and C→N.

Mons and co-workers have studied the single-conformation spectroscopy of a capped tripeptide, Ac-Aib-Phe-Aib-NH_2, under jet-cooled conditions.[24] In that case, the Amide NH stretch spectrum of the main conformer was assigned to a structure containing two C10 hydrogen bonds, indicating incipient 3_{10}-helix formation, with C10 NH stretch fundamentals at 3374 cm^{-1} and 3405 cm^{-1}. Expanding upon this work, Gord and co-workers studied a series of Aib-oligomers with benzoylcarboxy (Z) and methyl ester (OMe) caps at the N- and C-termini, respectively: (Z)-Aib_n-OMe (n = 1, 2, 4).[31] The smaller n = 1,2 peptides formed conformers that exhibited Ramachandran angles similar to those in a 3_{10}-helix, while the n = 4 peptide had sufficient length to fold into an incipient 3_{10}-helix, identified by two characteristic C10 hydrogen bonds with NH stretch fundamental at 3383 cm^{-1} and 3406 cm^{-1}, and two free NH stretch transitions.

In the current work, we study three short peptides of the form Ac-Aib-Aib-R, where the C-terminus is a benzyl amide or a substituted benzyl amide. These molecules are *capable* of forming C10 structures, including a single Type III beta turn, which has the same Ramachandran angles as those present in 3_{10}-helix formation. We employ single-conformation gas phase spectroscopic techniques, coupled with quantum chemical computations, to determine gas phase conformations of these three molecules, with R being a benzylamide, a *para*-fluoro-benzylamide, and a chiral

(S)-alpha-methyl benzyl amide. These molecules are part of a longer-term research project to study peptides capable of forming both left- and right-handed helices. While these structures themselves are too small to produce fully-formed 3_{10}-helices, their elongated versions of the benzyl amide and the fluorobenzyl amide would have left- and right-handed 3_{10}-helices that are equal in energy. By adding a chiral alpha-methyl (S) chromophore cap, we introduce a slight steric asymmetry in order to probe its effect on the structure adopted by the parent peptides, and probe the right- and left-handed structures that emerge. Lastly, we test the fluorinated phenyl ring as chromophore cap to see the effects of its stronger absorption on signal size and whether fluorination changes the conformational preferences.

3.1.1 Methods

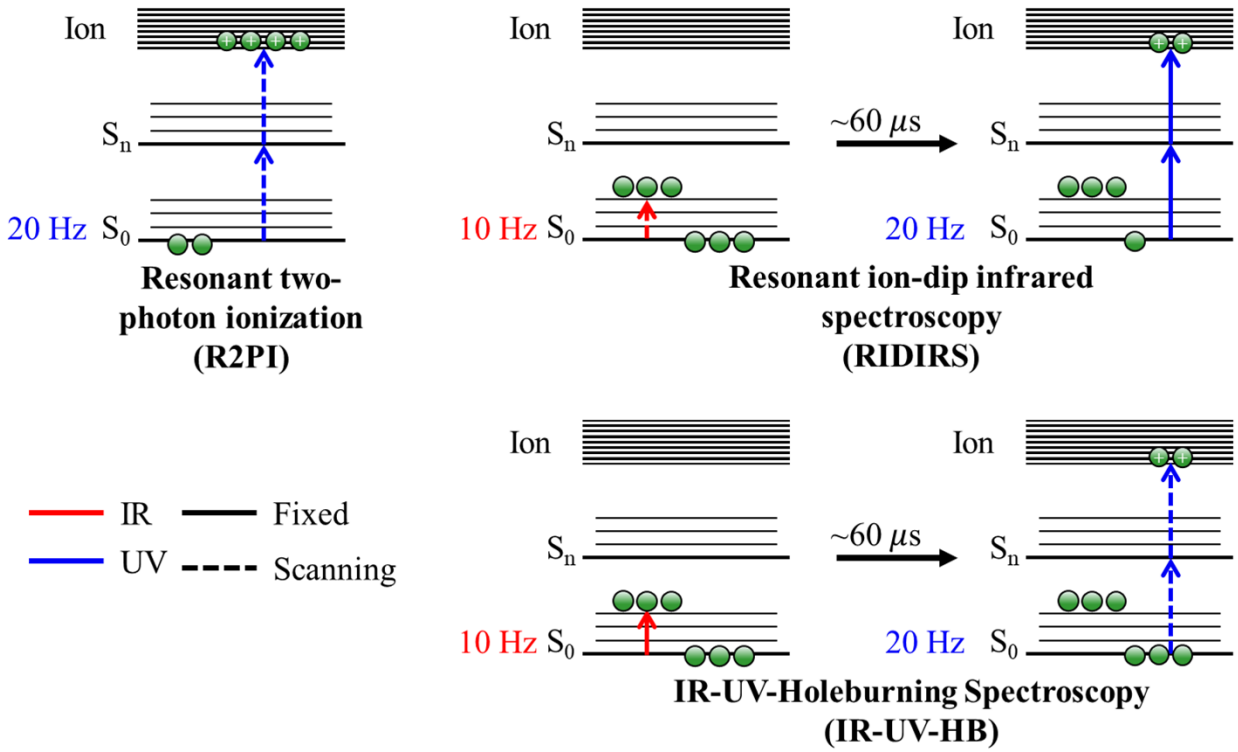


Figure 3.1 Schematic for double resonance techniques using a 60 μs delay between the first and second laser .

3.1.2 Experimental

Solid sample was brought into the gas-phase using laser desorption by crushing sample into a fine powder, mixing the powder into the surface of a flat graphite bar, and then desorbing the sample into the gas-phase using the 1064 nm fundamental of a Nd:YAG laser operating at 20 Hz (Continuum Minilite II, ~4 mJ/pulse, 2 mm beam diameter). The resulting desorption plume was entrained orthogonally into the high-collision regime of a supersonic expansion to cool conformers to their zero-point vibrational levels. The expansion was generated with a pulsed valve (Parker, Series 9, 1 mm dia. nozzle, 4 bar backing pressure Ar, 20 Hz, 500 μs pulse duration). The resulting free jet was skimmed through a 3-mm conical skimmer ~2–3 cm downstream from the pulsed valve orifice, forming a molecular beam that entered the extraction region of a Wiley-McLaren time-of-flight mass spectrometer, where single- and double-resonance laser spectroscopy techniques were used to interrogate the sample.

Figure 3.1 illustrates the set of laser-based methods used in this study. The UV spectrum of the three molecules in the series were recorded in the S0–S1 origin region using one-color, resonant two-photon ionization (R2PI). Ions were generated by introducing the frequency doubled output of a tunable dye laser into the ion source region of the TOFMS. A Radiant Dyes Narrowscan dye laser (Coumarin 540A) pumped with the third-harmonic from a Nd:YAG laser (Continuum, Surelite II) was used for this purpose. The R2PI spectrum in the UV is recorded by monitoring the ion intensity in the parent mass channel as a function of the UV frequency in the region of the S_0-S_1 origin for each unique aromatic cap (R=NHBn or AMBA, 37400–37800 cm-1; R = NHBn-F, 36900–37200 cm^{-1}). By cooling the conformers to their zero-point vibrational levels, there is no contribution to the spectrum from vibrational hot bands, reducing spectral complexity to the point that electronic transitions for specific conformers can be resolved.

Double resonance techniques were used to record single-conformation IR and UV spectra. These methods depend on depleting the population of a single conformer in its zero-point vibrational level, being monitored via R2PI. In resonant-ion dip infrared spectroscopy (RIDIRS), the UV laser (20 Hz) has its wavelength fixed on a conformer-specific UV transition, generating a 20 Hz ion signal in the parent mass channel. The IR laser (10 Hz) was aligned antiparallel to the molecular beam and set to temporally precede the UV laser such that the laser pulse intersects the molecular beam as it passes through the skimmer cone, ~60 μs before the UV laser. This configuration ensures the maximum number of skimmed, jet cooled molecules are exposed to the

IR laser. The IR laser is scanned in wavelength and when the IR laser is resonant with a vibrational transition of the same conformation being monitored by the UV laser, population is removed from the zero-point vibrational level to an excited vibrational state, leading to a depletion in the R2PI signal of the monitored zero-point level. The repetition rates of the lasers create an alternating IR-on/IR-off scheme. The difference signal from successive UV laser pulses (one with IR, one without) is recorded using a gated integrator (Stanford Research Systems, SR 250) in active baseline subtraction mode, which outputs the difference of IR-off and IR-on ion signal as a function of the tuned frequency of the IR laser.

To produce conformer specific UV spectra via IR-UV holeburn spectroscopy (IR-UV HB), the same laser scheme as RIDIRS is used, except the wavelength of the IR laser is fixed on a conformer specific IR transition while the UV laser is tuned across a frequency range. As in RIDIRS, IR-UV HB spectra were generated by plotting the difference of the IR-off and IR-on ion signals as a function of UV frequency. In the fluorinated, NHBn-F capped molecule, the larger oscillator strength of the electronic transition made it possible to record a conformer specific UV spectrum using UV-UV hole-burning. In this case, a UV hole-burn laser is fixed on a UV transition, partially depleting its ground state zero-point level population.

Conformation specific IR-spectra were acquired using a Nd:YAG pumped KTP/KTA optical parametric converter (LaserVision), generating light in the Amide A (NH Stretch) region (3200–3500 cm-1, 20–25 mJ/pulse). Light for the amide I/II regions was generated by difference frequency mixing the output from the optical parametric converter in a $AgGaSe_2$ crystal (1400–1800 cm-1, 0.75–1.50 mJ/pulse), which was then focused using a 500 mm CaF_2 lens.

Synthesis of the molecules was carried out at Fairfield University. The most direct synthesis of the three peptide samples was to couple the oxazolone of Ac-Aib-Aib-OH to the three different amine caps in dry MeCN (overnight reflux).[41, 42] After isolation, the peptides were purified using flash silica gel chromatography. Crystals for X-ray analysis were grown via diffusion of diethyl ether into a solution of methylene chloride.

3.1.3 Computational

A conformational search of the potential energy surface was performed using the Amber* force field in the Macromodel software suite.[43] The 100 lowest energy structures within 50 kJ/mol

of the global minimum were submitted to geometry optimization (tight optimization) and frequency calculations using Density Functional theory calculations that employed the B3LYP functional with a Grimme dispersion correction (DFT B3LYP-D3BJ) with a 6-31+G(d) basis set, using the Gaussian 09 software package.[44] Tight optimization conditions were employed. Harmonic vibrational frequencies were calculated at the same level of theory, and were scaled to account for anharmonicity using the accepted values of 0.958 for the NH stretch region, 0.981 for the amide I region, and 0.970 for the amide II region.[28–31, 45]

3.1.4 Crystallography

A single crystal of AMBA-F was coated with Fomblin oil and transferred to the goniometer head of a Bruker Quest diffractometer with kappa geometry, an I-μ-S microsource X-ray tube, laterally graded multilayer (Goebel) mirror single crystal for monochromatization, a Photon2 CMOS area detector and an Oxford Cryosystems low temperature device. Examination and data collection were performed with Cu Kα radiation (λ = 1.54178 Å) at 150 K. Data were collected, reflections were indexed and processed, and the files scaled and corrected for absorption using APEX3.[46] The space group was assigned and the structures were solved by direct methods using XPREP within the SHELXTL suite of programs[47] and refined by full matrix least squares against F^2 with all reflections with Shelxl2018 using the graphical interface Shelxle.[48, 49] H atoms attached to carbon atoms were positioned geometrically and constrained to ride on their parent atoms. C-H bond distances were constrained to 0.95 Å for aromatic C-H moieties, and to 1.00 and 0.98 Å for aliphatic C-H and CH_3 moieties, respectively. Water and amide H atom positions were refined and O-H and N-H distances were restrained to 0.84(2) and 0.88(2) Å, respectively. U_{iso}(H) values were set to a multiple of U_{eq}(C) with 1.5 for CH_3 and OH, and 1.2 for C-H and N-H units, respectively. The water molecule site was found to be partially occupied. The occupancy refined to 94.0(7) %. Additional data collection and refinement details in table format can be found in the Supporting Information. Complete crystallographic data, in CIF format, have been deposited with the Cambridge Crystallographic Data Centre. CCDC 1898412contains the supplementary crystallographic data for this paper. These data can be obtained free of charge from The Cambridge Crystallographic Data Centre via www.ccdc.cam.ac.uk/data_request/cif.

3.2 Results

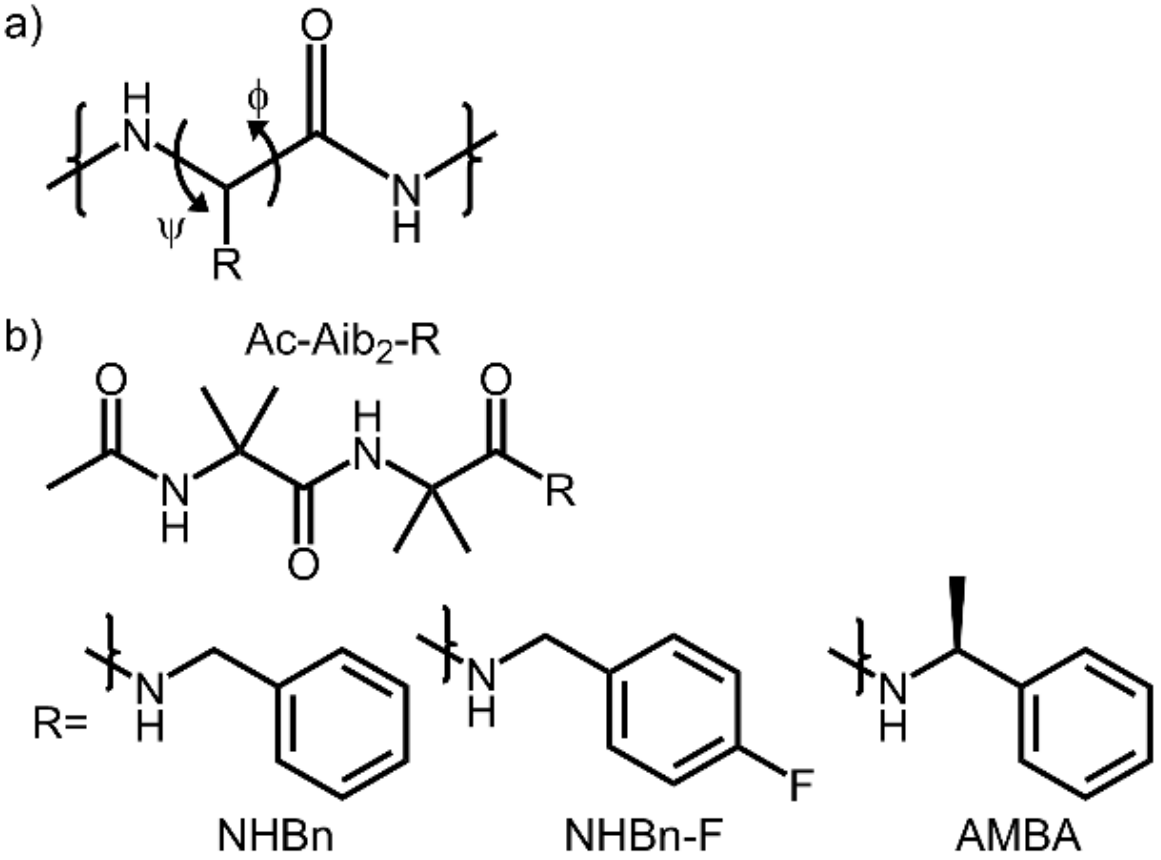


Figure 3.2 (a) Diagram depicting the Ramachandran ϕ/ψ angles of a peptide, which determine the secondary structure according to the Ramachandran plot. (b) The Aib backbone structure with the three R caps used in this study. The shorthand used throughout the text is the text beneath the respective R group.

3.2.1 Overview

Three peptide caps were chosen for this study, shown in Figure 3.2. In what follows, the 'R' labels in the figure will serve as shorthand designations for each Ac-Aib_2-R molecule. The Aib dipeptide with an NHBn cap serves as a reference, as our group most commonly studies peptides with this cap,[29–33] The NHBn-F cap is designed to increase the oscillator strength of the $S_0 \rightarrow S_1$ transition, the region where we carry out our UV spectroscopy. This cap may also weaken NH••π interactions by withdrawing electron density from the π cloud on the phenyl ring. Lastly, the AMBA cap will be used to introduce a chiral center outside of the folding region of the peptide,

converting any left- or right-handed enantiomeric secondary structures found in NHBn into diastereomers that can be spectroscopically distinguished.

3.2.2 Nomenclature.

We label an intramolecular hydrogen bond using the symbol 'Cn', where n denotes the size of the ring the hydrogen bond closes. For example, a structure that has a NH···O=C hydrogen bond that closes off a 10-membered ring, is labeled a C10 hydrogen bond. As a shorthand method for naming the capped dipeptide structures, we adopt a nomenclature used in previous work:[31] Starting from the N-terminus and proceeding to the C-terminus, each successive amide NH group in the structure is characterized by the type of hydrogen bond in which this NH group is involved. A free NH is denoted by 'F', a Cn hydrogen bond by 'n', or an NH••π-interaction by 'π'. For example, the assigned β-turn for NHBn contains two free NH bonds a NH(1) and NH(2), with NH(3) engaging in a C10 hydrogen bond, thereby labeled as F-F-10. Throughout this work, right- and left-handed turns are chiral secondary structural elements that are at times referred to as enantiomers of one another; similarly, the presence of a single chiral center in conjunction with a chiral structural element is referred to as a diastereomer. The left-/right-handed forms of the β-turns are designated as II/II', III/III', etc.

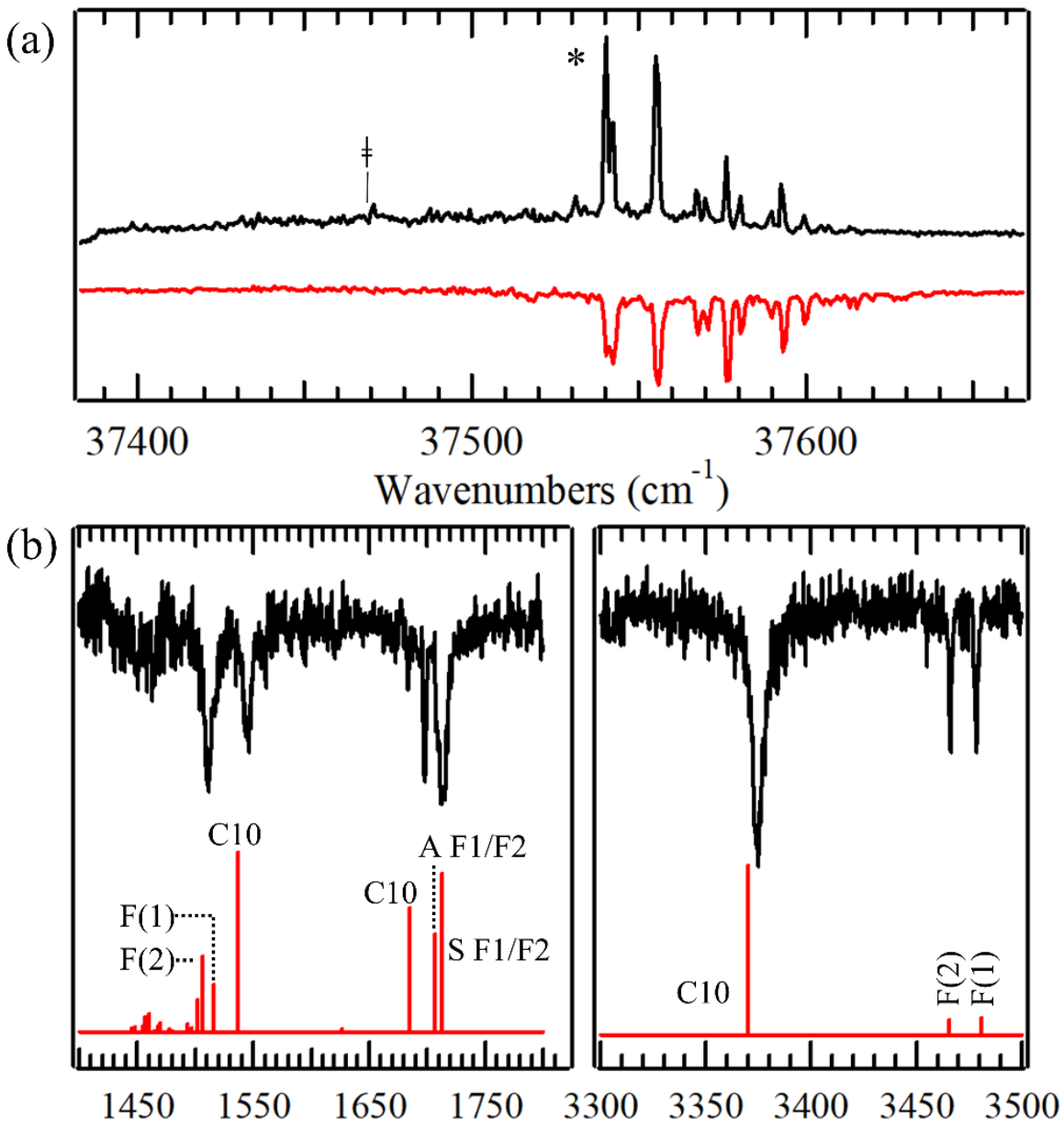


Figure 3.3 (a) R2PI and IR-UV HB spectra of Ac-$(Aib)_2$-NHBn. (b) RIDIR spectra for Ac-$(Aib)_2$-NHBn. The experimental vibrational spectra on top (black) cover the amide II region (1400–1600 cm^{-1}), the amide I region (1600–1800 cm^{-1}), and the NH stretch region (3300–3500 cm^{-1}). The predictions of DFT B3LYP/D3BJ 6-31+G*(d) calculations are shown in red below.

3.2.3 Ac-(Aib)$_2$-NHBn

The R2PI (black, top) and IR-UV HB (red, bottom) spectra are shown in Figure 3.3. Most of the transitions found in the R2PI spectrum are reproduced in the IR-UV HB spectra, indicating that there is one major conformer, with its S_0-S_1 origin transition at 37540 cm^{-1}. Close inspection of the region near 37470 cm^{-1} shows a small transition, marked with a dagger, which is part of a broad background not present in the IR-UV-HB spectra. This transition may belong to a low population second conformer; however, there was not enough signal to perform double resonance experiments on it. In addition, there is a 2 cm^{-1} splitting of the origin. However, we verified that both peaks involved in the splitting belong to the same conformer: RIDIR spectra taken on both peaks produced identical IR spectra. Additionally, due to the consistent difference in relative intensities of the doublet between the IR-UV-HB and R2PI spectra, IR-UV-HB spectra were taken using all the peaks in the NH stretch region and one peak in the amide I region (1697 cm^{-1}) as hole burn transitions, with all scans modulating the intensities of the same set of UV transitions, as shown in Figure 3.3. This shoulder is then most likely a sequence band from the same conformer.

The RIDIR spectrum of this main conformer of Ac-(Aib)$_2$-NHBn in the NH stretch region (3300-3500 cm^{-1}) is shown in Figure 3.3b (top, black). This region is highly sensitive to the intramolecular hydrogen bonding architecture of the peptide and is commonly used to make unique structural assignments of peptides when supported by sufficient theory.[20–40] The NH stretch region shows three transitions, as anticipated given the three NH groups in the molecule. Two weak, narrow transitions are found in the free amide NH stretch region, at 3466.2 and 3478.5 cm^{-1}. The intense, broad transition at 3361 cm^{-1} indicates that the third NH bond is involved in a strong hydrogen bond. The calculated stick spectrum shown below the experimental spectrum belongs to the global minimum structure and is our best fit to the experimental spectrum. Its close match with experiment across the NH stretch, amide I, and amide II regions lends confidence to the assignment.

The assigned structure has a single C10 hydrogen bond and is labeled as F/F/10 in our nomenclature, with ϕ_2/ψ_2 Ramachandran angles of -59°/115° and ϕ_3/ψ_3 of +58°/+30°, matching those characteristic of a Type II β-turn. The formation of a β-turn instead of a 3_{10}-helix is a straightforward result of the central amide group in the backbone being unconstrained by hydrogen bond formation due to the short length of the capped dipeptide. The loss of stabilization from an additional hydrogen bond on the central amide results in the unconstrained β-turn being the more stable structure.

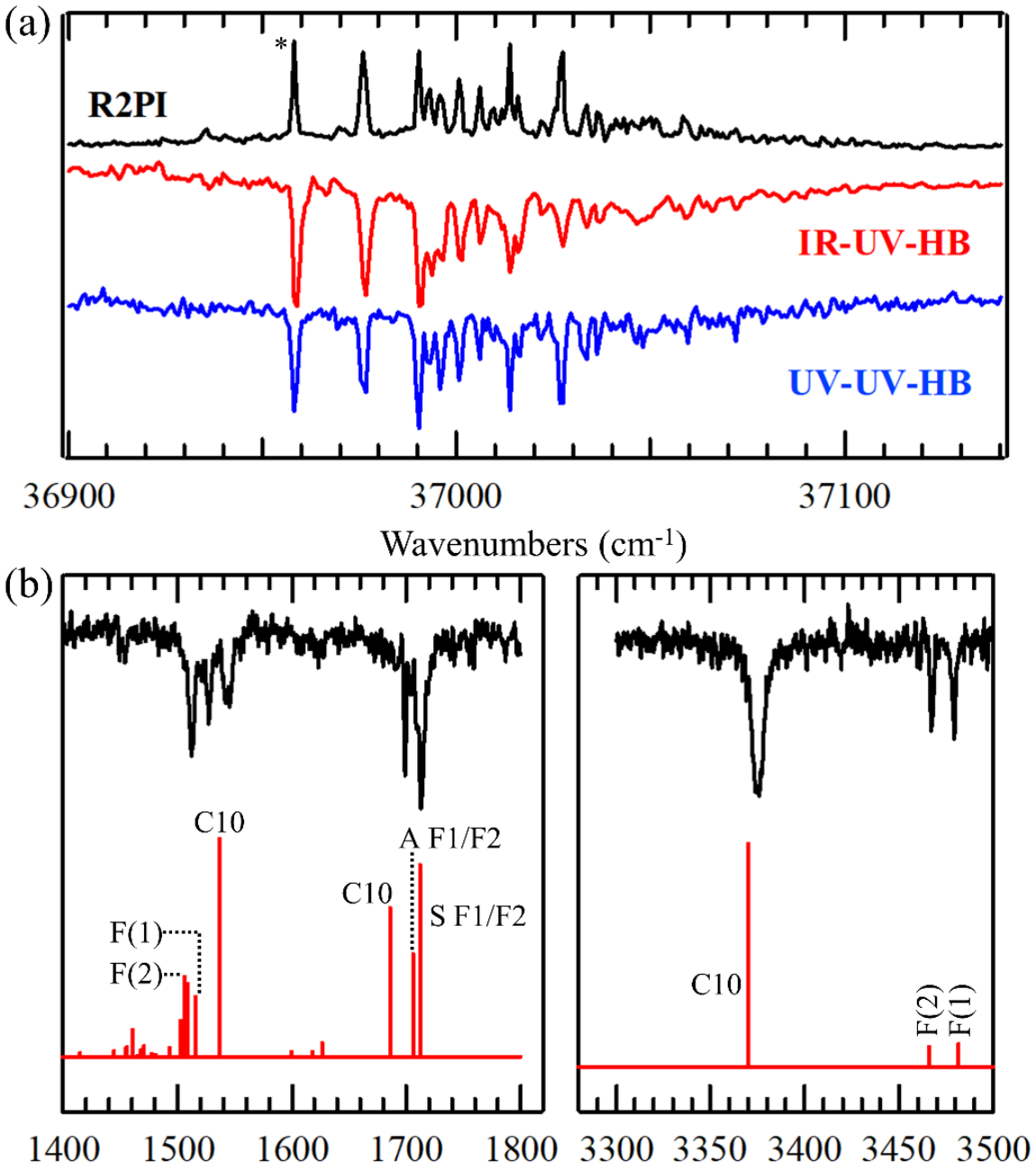


Figure 3.4 The single and double resonance spectra for Ac-Aib$_2$-NHBn-F. (a) R2PI spectrum (black), UV-UV HB (blue), and IR-UV-HB scans. The asterisk labels the peak used to as monitor transition in obtaining the RIDIR spectrum. (b) RIDIR spectra in the amide I, amide II, and NH stretch regions, compared with the predictions of theory for the assigned structure. See text for further discussion.

3.2.4 Ac-(Aib)$_2$-NHBn-F

The R2PI spectrum in the S_0-S_1 origin region for the NHBn-F capped dipeptide is shown in Figure 3.4a. It was taken with a power 75% less than in NHBn to avoid saturation effects (not seen in NHBn), demonstrating the positive benefits of the larger UV oscillator strength of the NHBn-F cap. Importantly, the increased oscillator strength of the S_0–S_1 transition allowed UV-UV HB spectra (bottom, blue) to be acquired, which is difficult with NHBn caps, and was one of the main motivations for considering this fluorinated phenyl cap. An IR-UV-HB scan is also shown for comparison (middle, red). Being able to record UV-UV HB scans is significant because UV spectra typically suffer from less overlap than IR spectra, affording better specificity when selecting conformer specific transitions for double resonance methods. Since all transitions in the R2PI spectrum are also present in the hole-burn scans, there is one major conformer in the expansion, despite the presence of more than ten transitions in the UV.

Table 3.1 List of dihedrals, hydrogen bond lengths, and hydrogen bond angle for the NHBn β-turn structures, obtained from DFT calculations. Dihedrals start from the N-terminus. Rows with typical Ramachandran values are filled gray.

	ϕ_1 (°)	ψ_1 (°)	ϕ_2 (°)	ψ_2 (°)	l_{HBond} (Å)	∠COH (°)
Type II β-turn	-60	120	80	0	--	--
NHBn	-59	115	58	30	1.96	120
NHBn-F	-59	116	59	30	1.96	120
AMBA Conf A	-60	117	58	30	1.96	121

The RIDIR spectrum in Figure 3.4b is virtually identical to the corresponding spectrum acquired for NHBn (Figure 3.3b), indicating that the structures are very similar. Two free NH transitions are identified at 3467 cm^{-1} and 3479 cm^{-1} and a third, broad transition belonging to a C10 hydrogen bond is located at 3375 cm^{-1}. Indeed, the best fit structure is another type-II β-turn that has nearly identical ϕ/ψ angles (Table 3.1). We surmise on this basis that fluorine substitution on the phenyl ring does not perturb the structure significantly. This structure is also the calculated global minimum in energy. Since the F-F-10 conformer, like its NHBn counterpart, does not contain an NH••π interaction, it is not possible to evaluate the extent to which fluorine substitution

on the aromatic ring reduces the strength of π••H hydrogen-bonds. A spectroscopic difference between the NHBn and NHBn-F caps becomes apparent in the 1400–1600 cm^{-1} region, typically ascribed to Amide II vibrations. However, in the spectrum of the NHBnF capped molecule, a new transition appears at 1527.5 cm^{-1} that is due to the C—F stretch fundamental. If the fluorine atom is involved in intramolecular or intermolecular bonding, it is possible that the C-F transition could have structural diagnostic capabilities in future studies.

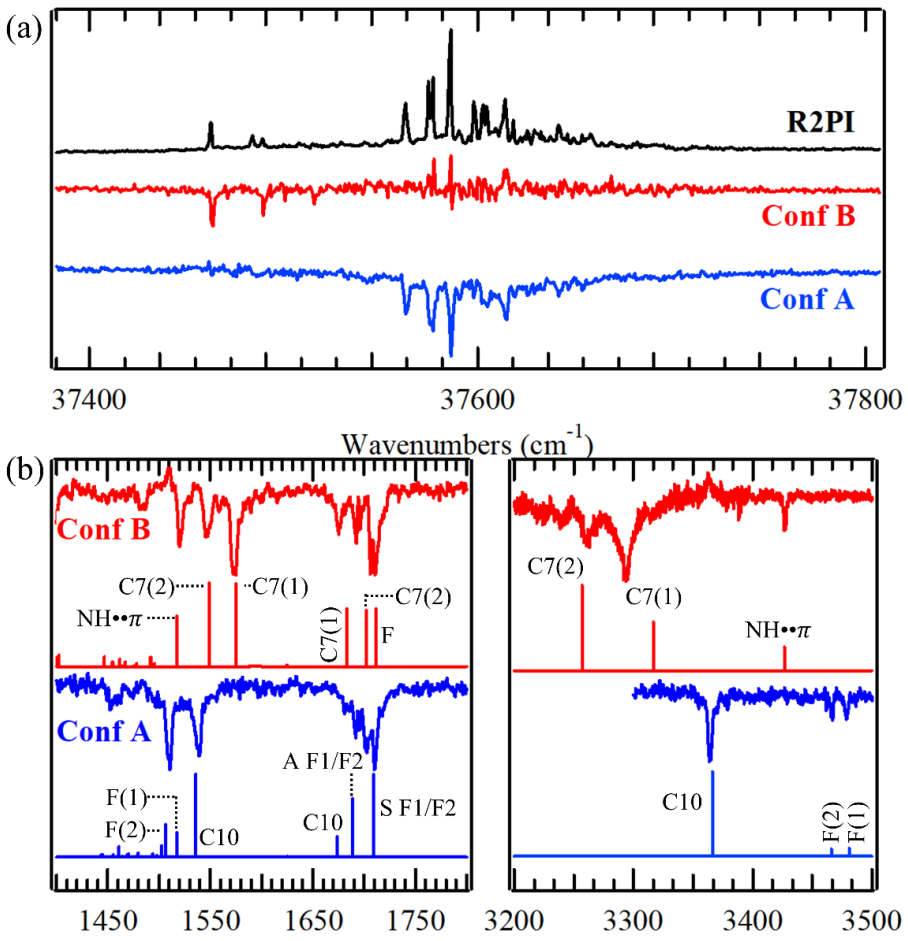


Figure 3.5 (a) R2PI (top, black) and IR-UV HB spectra for conformer A (bottom, blue) and conformer B (middle, red) of AMBA. The asterisk indicates the frequency the UV laser was fixed on to produce the RIDIR spectra (Figure 3.5b), with the color matched to the respective conformer. (b) RIDIR spectra for conformers A and B of AMBA in the NH stretch region (3200-3500 cm^{-1}) and amide I/II regions (1400-1800 cm^{-1}). Stick spectra are the predictions of scaled, harmonic vibrational frequency calculations for the two assigned structures. F indicates transitions involving free N—H bonds, C*n* indicates hydrides stretches closing rings containing *n* atoms, A indicates antisymmetric, S indicates symmetric.

3.2.5 Ac-(Aib)$_2$-AMBA

The R2PI spectrum of the AMBA-capped dipeptide in the S_0-S_1 origin region is shown in Figure 3.5a (top, black). IR-UV HB scans are shown below the R2PI spectrum, demonstrating the presence of two conformers with S_0-S_1 origin transitions at 37562 cm^1 (conformer A, blue) and 37462 cm^{-1} (conformer B, red). The small band +22 cm^{-1} above the B origin does not appear in the IR-UV HB spectrum, and thus must be due to a third minor conformer. Given the small size of this transition, we did not pursue its characterization further.

RIDIR spectra were collected while monitoring both origin peaks, with results shown in Figure 3.5b. Incorporation of the AMBA cap was intended to create two diastereomers out of otherwise indistinguishable R- and L-handed structural pairs. However, inspection of the NH stretch RIDIR spectra (Fig. 5b) indicates that conformer A (blue, bottom) and conformer B (red, top) have different hydrogen bonding networks. Indeed, cursory inspection shows one (conformer B) and two (conformer A) free NH stretch fundamentals, eliminating them as members of a diastereomeric pair. Conformer A of AMBA shows the same hydrogen bonding pattern as the sole conformers of NHBn, with two free NH transitions at 3467 and 3479 cm^{-1} and a broad transition belonging to a C10 hydrogen bond at 3585 cm^{-1}. Not surprisingly, the structure is assigned to the same Type II β-turn as in NHBn and NHBn-F. By contrast, conformer B of AMBA shows only one peak above 3400 cm^{-1}, assigned to an amide NH••π stretch at 3426.4 cm^{-1}. The experimental spectrum of conformer B shows two broadened transitions at 3293.5 and 3260.5 cm^{-1}, with the latter particularly broadened and split, complicating the match-up between experiment and theory. However, the excellent match in the amide I/II regions strengthens the assignment of Conformer B to a structure with a sequential C7-C7 double ring, with two γ-turns of first left and then right-handed orientation. This structure is similar to a β-ribbon, and we refer to it as a 2_7-ribbon.[8] It is also worth noting that C7 hydrogen bonds often exhibit broadened NH stretch fundamentals, as seen in previous work with Aib peptides.[31]

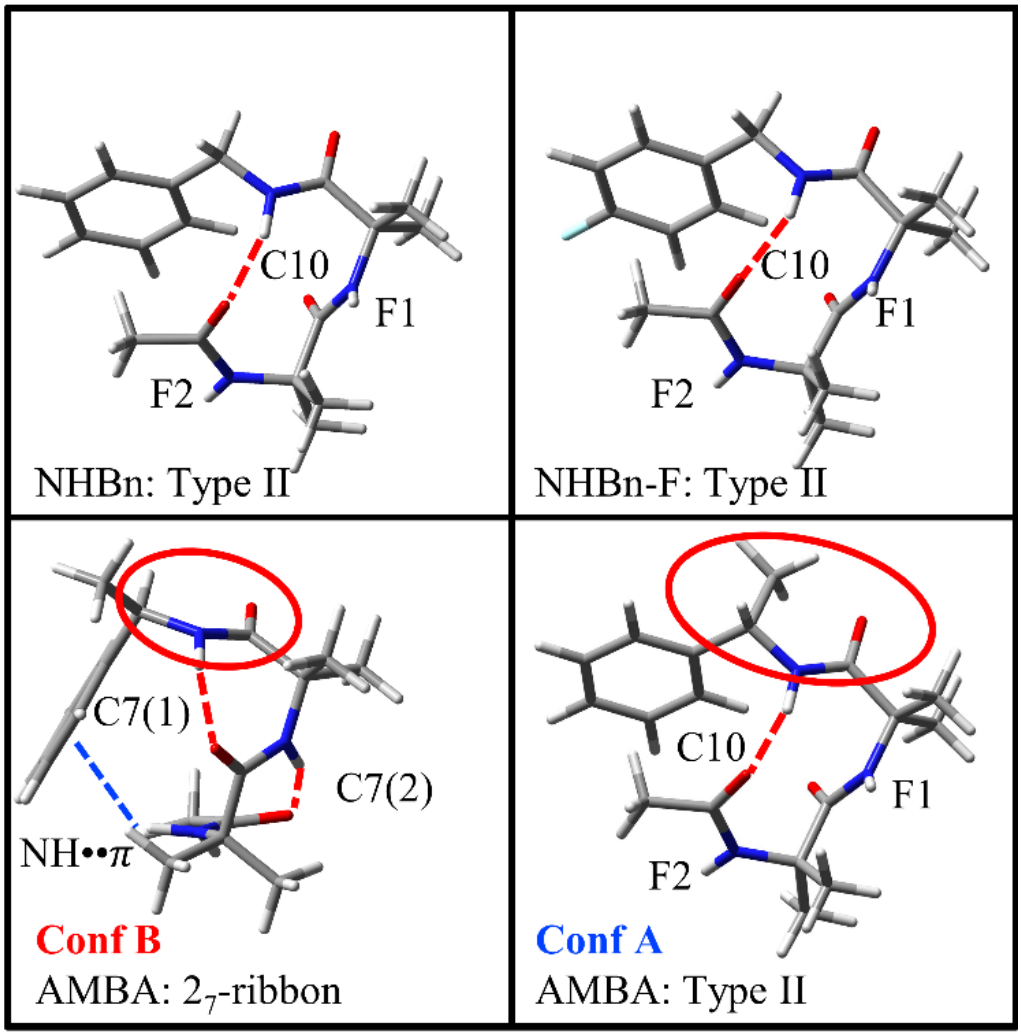


Figure 3.6 Comparison of the structures assigned for NHBn (top, left), NHBn-F (top, right) and AMBA (bottom).

The assigned structures for all observed conformers of the three capped peptides are presented in Figure 3.6. Despite the different nature of the caps, the same β-turn structure is observed in all three molecules, with NHBn and NHBn-F having it as the only observed conformer. The striking similarity is also evident in the ϕ/ψ Ramachandran angles, hydrogen bond lengths and NHO bond angles presented in Table 3.1. When the chiral AMBA cap is present, a second important conformer is present, the 2_7-ribbon. The increase in the population of this conformer based on the addition of a methyl group to C(α) is somewhat surprising, as adding a methyl group outside of the folding region of the peptide would not be expected to perturb the intramolecular interactions and the methyl group can only provide stabilization through weak dispersive

interactions. We believe that preferential stabilization of the 2_7-ribbon by the methyl group seems less likely than a destabilizing effect of the methyl group due to steric effects in the β-turn structure.

Table 3.2 List of dihedrals, hydrogen bond length, and hydrogen bond angle for the NHBn β-turn and 3_{10}-helix-like structures, obtained from DFT calculations. Dihedrals start from the N-terminus. Type refers to the hydrogen bonding present at each amide NH, with F representing free and numerals n representing Cn hydrogen bonds, starting with the N-terminus and progressing towards the C-terminus. The S represents a Schellman motif, NS represents a non-Schellman motif. Rows with typical Ramachandran values are filled gray. C.S. is short for crystal structure.

	Assigned 2° structure	H-bonds	ϕ_1	ψ_1	ϕ_2	ψ_2	ϕ_3	ψ_4
Type II/II' β-turn			$\mp$60	$\pm$120	$\pm$80	$\pm$0		
3_{10}-helix (right)			-57	-30	-57	-30		
NHBn	Type II	F-F-10	-59	115	58	30		
NHBn-F	Type II	F-F-10	-59	116	59	30		
AMBA (A)	2_7-ribbon	π-7-7	-75	60	68	-71		
AMBA (B)	Type II	F-F-10	-60	117	59	30		
AMBA-F (C.S.)	Type III	F-F-10	-55	-35	-62	-23		
AMBA (Theory)	Type II'	F-F-10	57	-122	-58	-34		
3_{10}-helix-like		F-F-10	-67	-14	-53	-35		
Z-Aib$_2$-OMe (A)		F-5	-57	-38	-178	-179		
Z-Aib$_2$-OMe (B)		F-7/5	-72	70	180	179		
Z-Aib$_2$-OMe (C)		5-5	180	-180	180	180		
Z-Aib$_4$-OMe (S)	3_{10}-helix	F-F-10-7	-61	-26	-55	-31	51	40
Z-Aib$_4$-OMe (NS)	3_{10}-helix	F-F-10-7	-61	-27	-57	-22	-48	-40

3.3 Discussion

3.3.1 Comparison to Z-Aib$_2$-OMe and Z-Aib$_4$-OMe

Previous work on Aib dipeptides[31] utilized a N-terminal benzyloxycarbonyl (Z) cap to sensitize Aib$_2$-OMe for UV and IR-UV double resonance experiments. Gord *et al.* showed that Z-Aib$_2$-OMe adopts three different conformers, an F-5 conformer, an F-7/5 structure with a unique C7/C5 bifurcated hydrogen bond, and an extended 5-5 structure. The dihedral angles for these conformers are included in Table 3.2. An important distinction between the capping groups used by Gord *et al.* (N-terminal Z cap and C-terminal OMe cap) and the caps used in this study (Ac N-terminal and NHBn C-terminal cap) is how the caps alter the number of hydrogen bond donor/acceptor groups. Z-Aib$_2$-OMe has only two amide NH groups whereas Ac-Aib$_2$-R caps studied here all incorporate three. To that end, Z-Aib$_2$-OMe cannot form a C10 hydrogen bond,

precluding the formation of a β-turn. Nevertheless, conformers were identified that are pre-configured for 3_{10}-helix formation, at least in part. This is particularly true of the F-5 structure, which was the global minimum energy conformer (Table 3.2).

In the capped Ac-Aib$_2$-R dipeptides studied here, the global minimum is a structure with a single C10 hydrogen bond. However, while the dihedral angles associated with the Aib(2) residue are those of a 3_{10}-helix, Aib(1) takes on dihedrals (ϕ_1 = -59°, ψ_1 = +116°) that enable formation of a Type II β-turn. We will return to this point when considering energetic grounds shortly. The steric influence of the α-dimethyl residue of Aib amino acids leads to the formation of a 3_{10}-helix in the gas-phase once there are two or more C10 hydrogen bonds possible, which re-orient the central amide group(s) to form these additional hydrogen bonds.[24, 31] We will provide a more in-depth exploration in section C of what forces may contribute to the unexpected formation of the Type II β-turn over the expected 3_{10}-helix..

The IR spectroscopy itself provides some evidence that the C10 hydrogen bond in the Type II β–turn is somewhat stronger than the Type III C10 hydrogen bond of the 3_{10}-helix. Recall that the vibrational transitions in the NH stretch region of NHBn, NHBn-F, and AMBA Conf B contain two free NH transitions at ~3466 and ~3479 cm^{-1} and a strong, broad transition at ~3361 cm^{-1}. Z-Aib$_4$-OMe also has two free NH stretch fundamentals at 3467 cm^{-1} and 3478 cm^{-1}, and two hydrogen bonded NH stretches at 3383 cm^{-1} and 3406 cm^{-1}. While the intensity patterns between the spectra of Z-Aib$_4$OMe and the Ac-Aib$_2$-R series are strikingly similar, the hydrogen bonded NH stretch fundamental in the Ac-Aib$_2$-R peptides is shifted 20 cm^{-1} lower in frequency than the lowest energy transition found in Z-Aib$_4$-OMe. This suggests the presence of a stronger C10 hydrogen bond in the Ac-Aib$_2$-R peptide β-turns than in the 3_{10}-helix of Z-Aib$_4$-OMe. The strength of this hydrogen bond likely contributes to the formation of the β-turn over the formation of the 3_{10}-helix, as is discussed further below.

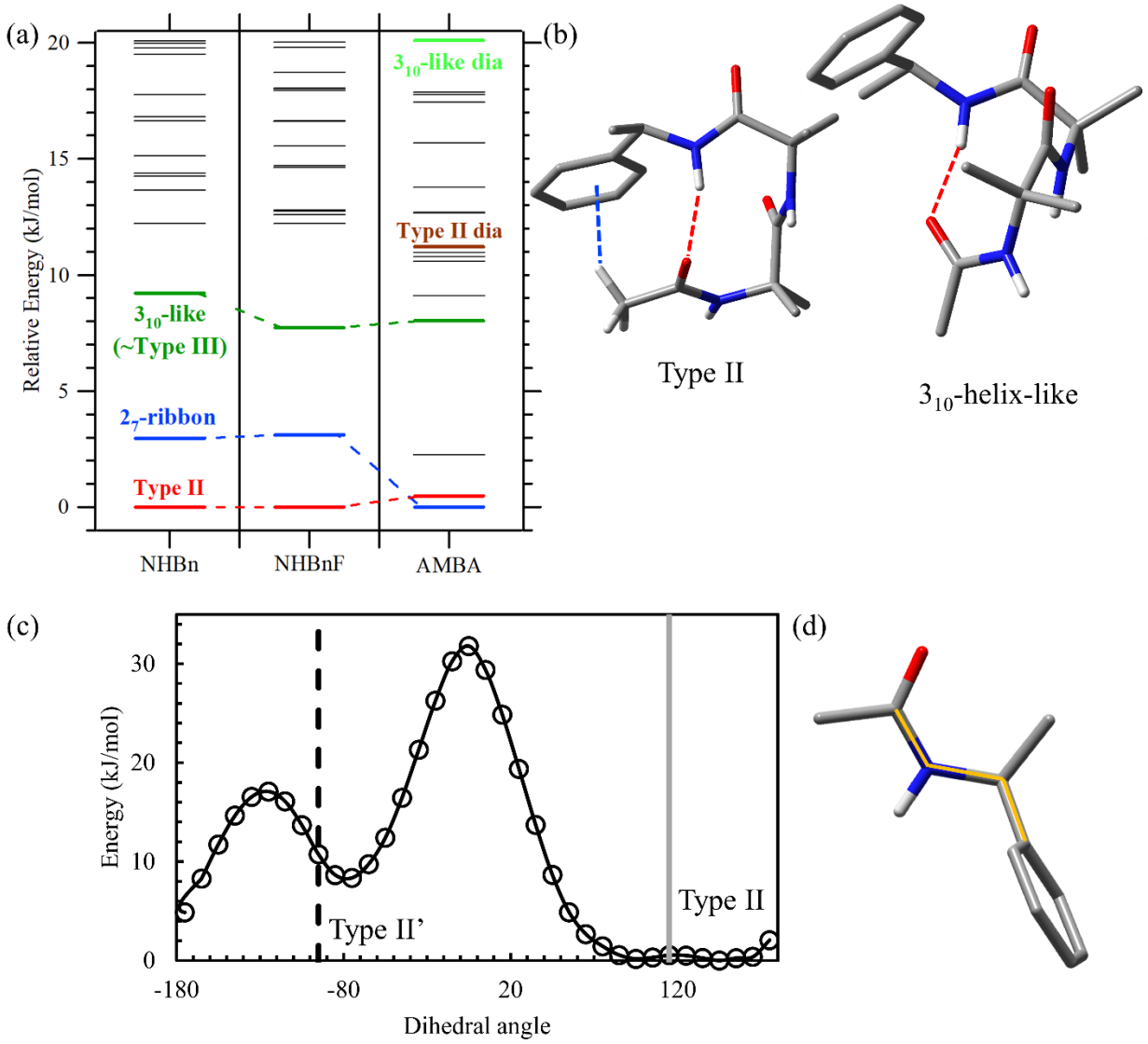


Figure 3.7 (a) Potential energy diagram for Ac-Aib$_2$-R, where the R label is given on the x-axis. Zero-point corrected relative energies at the DFT B3LYP/6-31(g)+d level of theory with the D3BJ dispersion correction. (b) Comparison of β-turn (left) to 3$_{10}$-helix-like conformation (right), with non-polar hydrogens removed to improve clarity. Red lines indicate hydrogen bonds, blue lines indicate π-interactions. (c) Relative energy versus dihedral angle along the coordinate indicated in (d) by an orange line. The same level of theory was used as in (a). (d) Image to illustrate which dihedral coordinate was scanned.

3.3.2 Energetic Analysis

Figure 3.7a presents energy level diagrams for each member of our series, depicting the relative zero-point corrected energies of all conformers within 20 kJ/mol of the global minimum for each member of our series. The assigned structures for the NHBn and NHBn-F caps are both Type II β-turns, while the chiral AMBA cap had population in both a Type II β-turn and a 2_7-ribbon conformation. For NHBn and NHBn-F, the DFT calculations predict that the β-turn is the global minimum structure. The next lowest energy structure in both conformers, ~ 3 kJ/mol higher in energy, is a 2_7-ribbon nearly identical to the assigned structure for AMBA Conformer B. Notably, in AMBA, the 2_7-ribbon structure drops in energy relative to the Type II β-turn, becoming more stable than the Type II β-turn by 0.5 kJ/mol. This stabilization is consistent with the experimental detection of 2_7 ribbon in the presence of the AMBA cap. In the 2_7-ribbon structure of AMBA, the α-methyl group that renders the benzylic carbon chiral points away from the rest of the molecule and likely has little interaction, suggesting that its effect on the 2_7-turn energetics is small. It seems more likely, then, to ascribe the energy difference in AMBA relative to NHBn to the destabilizing steric effect of the methyl group on the Type II β-turn, a change in relative energy of 3.4 kJ/mol.

A structure with Ramachandran angles similar to a 3_{10}-helix (in green, Fig. 7a) was located in each of the capped Aib dipeptides, as might have been anticipated based on the 3_{10}-helix folding propensity of oligo-Aib peptides. This structure is compared with the Type II β–turn in Figure 3.7b. Note that the principle difference between the two is the orientation of the central amide group. In the 3_{10}-helix-like structure (hereafter referred to as 3_{10}-like), the central C=O group points in the same direction along the helix as those adjacent to it, as it must if it is to form a hydrogen bond in a longer 3_{10}-helix.

In all three capped dipeptides, the 3_{10}-helix-like conformer is 8–10 kJ/mol above the global minimum. Typical ϕ/ψ angles for a 3_{10}-helix with several complete turns are loosely defined due to strain from direct overlap between the *i* and *i+3* residues of the peptides, preventing the formation of perfect (e.g., idealized) 3_{10}-helices. The ϕ/ψ angles are usually around -57°/-30° for ϕ_1/ψ_1 and ϕ_2/ψ_2. The $\phi + \psi$ sums typically range between -90° and -120°, which is close to the sum of first set of ϕ/ψ angles of a Type II β-turn, explaining the similarity of the structures. In the Type II β-turn, the ϕ_2/ψ_2 angles are -59°/115°, while in the 3_{10}-helix-like structure the

ϕ/ψ angles are -53°/-35°. As discussed above, the Type II β-turn may be more likely to form in the short peptides studied here because the unconstrained central amide group does not need to be involved in hydrogen bonding present in the 3_{10}-helix. It may also be inferred based on the results from Z-Aib$_4$-OMe that the 3_{10}-helix forms a weaker C10 hydrogen bond than the β-turn.[31]

Depending on the barriers to conformational isomerization and the timescales for cooling, conformer populations can sometimes be dictated more by their equilibrium populations (via ΔG(T)) prior to cooling rather than by their zero-point corrected relative energies. To investigate the populations at these temperatures, we calculated ΔG(300 K) for all conformers of the NHBn and AMBA capped dipeptides with zero-point energies below 10 kJ/mol. The free energy diagram is shown in Supplemental Figure 3.2, with an emphasis on the ΔE in the diastereomer of the Type II β-turn, which is potentially more stable according to ΔG values. Note that, since laser desorption was used to bring the samples into the gas phase, we have no direct measure of the appropriate temperature to use; hence, the choice of $T_{conformer}$ = 300 K is used to track how the conformer populations respond to increasing temperature.

The structure with lowest free energy did not change in either molecule, nor did the ordering of all structures below 5 kJ/mol, indicating that our focus on zero-point corrected internal energies is justified. At the same time, all conformers seem to have a relative free energy closer to that of the global minimum than was the case in the ZPE-corrected internal energies. In particular, the 3_{10}-like structure are lower in relative free energies by ~4 kJ/mol in NHBn and ~6 kJ/mol in AMBA compared to the type II β-turn. Furthermore, the diastereomer of the Type II β-turn lowers from 11 kJ/mol in ΔE to 5 kJ/mol in ΔG. While none of these conformers had population trapped in them by the jet-cooling process, it is interesting that the free energies suggest that both the 3_{10}-helix and the type II' β-turn diastereomer are close enough in free energy to the type II β-turn that further modification of the position/nature of the cap or elongation of the Aib oligomer could bring them into competition.

Indeed, in considering how the caps might affect the energetics, we have also considered the CH_3-π interaction present in the Type II β-turn (Figure 3.7b). To investigate the influence of the aromatic ring cap computationally, we re-optimized all structures below 15 kJ/mol after substituting a methyl group for the phenyl ring, so that NHBn becomes NHEt. The relative energies of the ethyl-capped structures can be found in Supplemental Figure 3.1. The Type II β-turn remains the global minimum energy structure with the NHEt cap, indicating the primary

forces driving the formation of the β-turn are not linked to the aromatic ring. The 2_7-ribbon, which has an NH-π interaction, increases its relative energy from 3.0 to 4.6 kJ/mol in the absence of the phenyl ring.

One of the most striking energetic changes between the NHBn and NHEt caps is that the 3_{10}-like helix collapses from 9.2 kJ/mol in relative energy to 1.1 kJ/mol, adopting a different structure that becomes the second lowest energy conformer. The new structure to has ψ_1/ϕ_1 angles of -61° and -28°, and ψ_2/ϕ_2 angles of -62° and -20°, which fall within range of the ϕ/ψ angles of the first turn of a 3_{10}-helix. Additionally, two other structures, found at 13.6 kJ/mol and 14.2 kJ/mol, which were phenyl rotors of the 3_{10}-like helix (Supplemental Figure 3.1), optimized to the same structure. These results suggest that the phenyl ring of the NHBn cap creates unfavorable interactions that hinder the formation of the 3_{10}-helix. Though this result is interesting, the Type II β-turn remains the calculated global minimum both in the presence and absence of the phenyl ring cap.

A principal reason for considering the chiral AMBA cap is that it would enable us to distinguish right- from left-handed structures in the otherwise achiral Aib dipeptide. The energy level diagram in Figure 3.7a shows the presence of a significant energy difference between the left-handed and right-handed diastereomeric structures created by the S-chiral AMBA cap. Surprisingly, the left-handed diastereomer of the lowest energy right-handed β-turn conformer, which we label as Type II', is found 11.2 kJ/mol above the global minimum, indicating that the steric differences between left- and right-handed β–turns are significant. To explain this difference, we consider the chiral center associated with the AMBA cap and how it interacts with the rest of the molecule. In the Type II' diastereomeric structure, the methyl group of the AMBA cap is *cis* to the nearest acyl group, while it is *trans* in the Type II structure, indicating the presence of an unfavorable steric interaction between the methyl and acyl groups (Figure 3.6). It is initially surprising that this interaction would result in an 11 kJ/mol difference in energy.

To investigate the energetic cost of eclipsing the acyl group with the methyl group of the AMBA cap, we performed a relaxed dihedral scan of an AMBA fragment, N-1(phenylethyl)acetamide (shown in Figure 3.7d), along a dihedral between the acyl and methyl group (acyl/methyl dihedral), highlighted in orange on the model. All other initial dihedral angles match the lowest energy Type II β-turn, shown in Figure 3.7b. The energy difference of the fragment when the dihedral angle matches the Type II β-turn diastereomers is 10.2 kJ/mol, a value

in reasonable agreement with the fully optimized structures. While the dihedral scan of N-1(phenylethyl)acetamide does not provide a perfect account for the energetic cost of eclipsing the acyl and methyl group of the AMBA cap, it does point towards this interaction as the major determinant in destabilizing the left-handed Type II' β-turn relative to the right-handed Type II turn. A similar comparison between the diastereomers of the 3_{10}-like structures shows a similar difference in energy (~10 kJ/mol), also shown in Figure 3.7a, consistent with the energetic cost of the aforementioned *cis* acyl/methyl group.

3.3.3 Comparison to X-ray Data

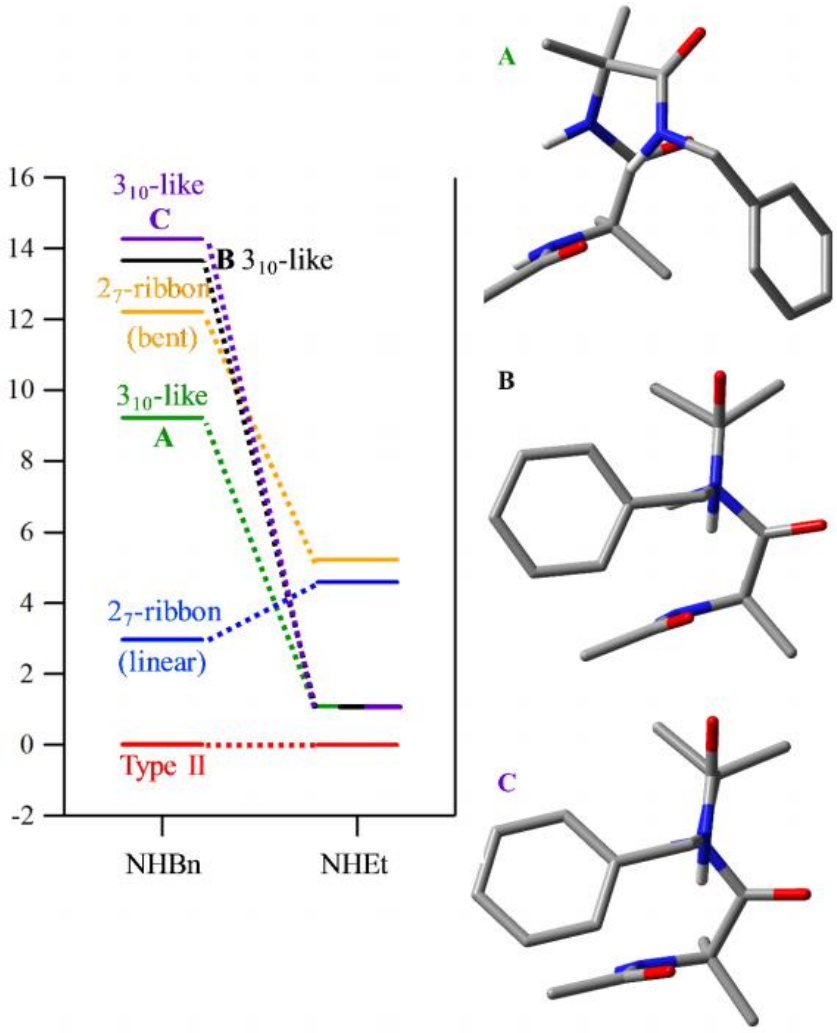
As a further point of comparison with the gas phase conformational results for Ac-Aib_2-AMBA, we obtained single crystal X-ray diffraction results for its para-fluorinated analogue, Ac-Aib_2-AMBA-F. This molecule crystallized as a monohydrate with four peptide molecules per unit cell (Fig. S3). Interestingly, the peptide crystallized in a right-handed β-turn, with a single C10 intramolecular hydrogen bond (See supplemental info for .cif file), as in the gas phase. However, peptide molecules in the crystalline solid exhibit a Type III β –turn rather than the Type II turn found as the global minimum in the gas phase. The preference for right-handed Type III turn over left-handed Type III' must presumably be imposed by the S-configuration of the C-terminal AMBA-F group, with the intramolecular methyl/acyl interaction likely contributing to the preference for right-handed over left-handed β–turn.

The peptides and water molecules of the crystalline structure exhibit significant *intermolecular* hydrogen bonding, further stabilizing the observed Type III configurations. The positions of hydrogen atoms suggest more than twice as many hydrogen bonding interactions between co-crystallized water and peptide molecules than C10-type intramolecular hydrogen bonds within the peptide molecules themselves. Therefore, we attribute the discrepancy between the crystal structure and our gas phase structural determination to the significant structural influence of hydrogen bonding interactions with co-crystallized water and with neighboring molecules in the crystal lattice.

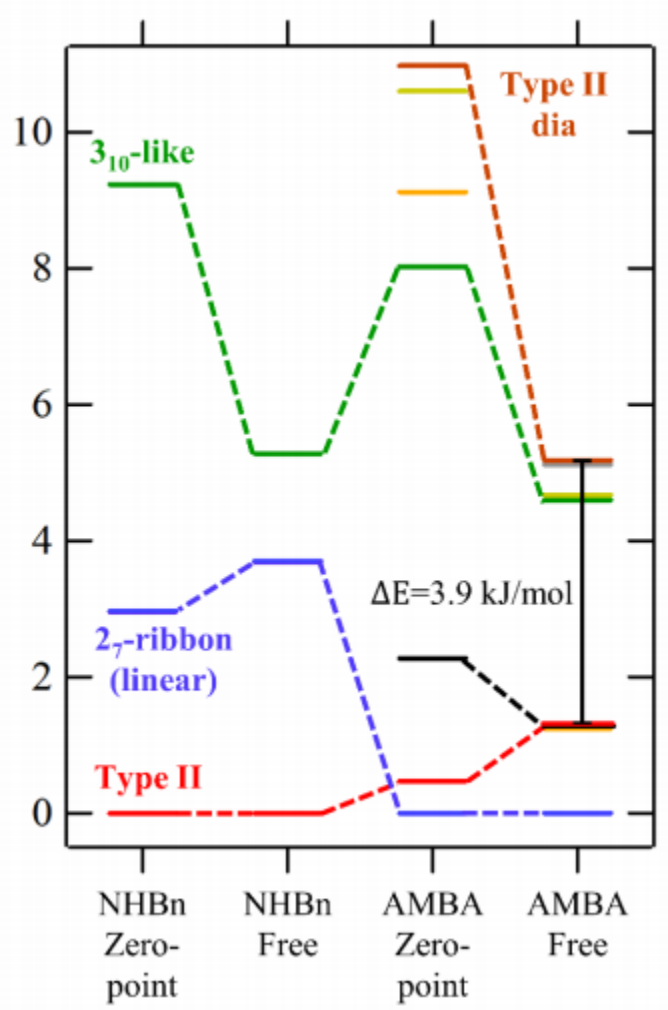
3.4 Conclusion

Three chromophore caps were tested on a model Aib dipeptide system. The addition of fluorine to the NHBn fluorophore produced no significant structural perturbation while affording increased signal and UV-UV HB capabilities. The addition of a chiral methyl group to the benzyl carbon in AMBA induces a strong, destabilizing steric interaction between the methyl substituent and the nearest amide carbonyl. Inspection of the energy level diagram for AMBA indicates that the oppositely handed β-turn is ~11 kJ/mol above the global minimum. Finally, due to the improved signal and hole-burning capabilities arising using the NHBn-F cap, we plan in future studies to carry out single-conformation spectroscopy on Ac-Aib$_3$-AMBA-F with the goal of determining whether the large energy gap calculated to exist in Ac-Aib$_2$-AMBA-F is mitigated by increasing the length of the peptide, and therefore the number of C10 hydrogen bonds present. The search for a chiral cap that interacts less strongly with the adjacent amide group will also be pursued. The extension to a longer Aib oligomer is crucial because Aib2 is too short to form a single turn of a helix, while Aib3 can do so. This will provide the first opportunity to see and characterize both handed helices in the gas-phase.

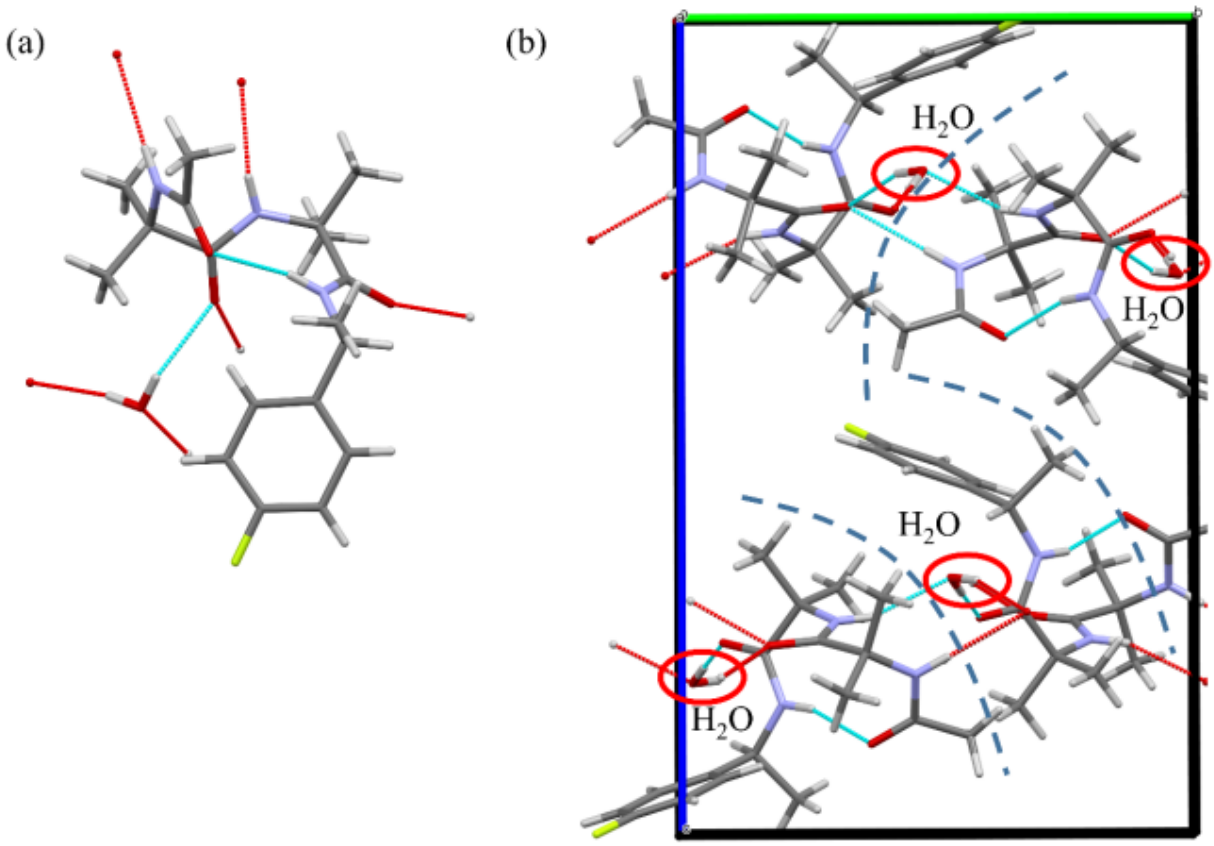
3.5 Supplemental Figures



Supplemental Figure 3.1. *Left*: Potential energy diagram comparing the lowest energy structures found in the NHBn capped dipeptide with the same structures after substituting the NHBn cap with an NHEt cap. *Right*: Structures of the three conformers that collapsed to an identical structure resembling the first turn of a 310-helix. These structures, labeled A, B, and C, are phenyl rotamers of the 310-like helix structure; thus, removal of the phenyl group results in the 310-like structures relaxing to adopt the first turn of a 310-helix. By contrast, in the 27-ribbon linear structure, an NH-π interaction is lost by removal of the phenyl ring, destabilizing it relative to the others.



Supplemental Figure 3.2. Energy level diagram comparing the relative energies of the calculated lowest energy conformers, zero-point corrected internal energies with ΔG (300K). Dashed lines connect structures with similar dihedral angles.



Supplemental Figure 3.3. (a) Crystal structure of a single molecule from the unit cell of Ac-Aib2-AMBA-F. Relative dihedral angles can be found in the main text (Table 3.2). Intramolecular hydrogen bonds are indicated in turquoise, while red lines denote intermolecular bonding to other molecules in the unit cell. (b) Unit cell of Ac-Aib2-AMBA-F. Red circles used to illustrate the location of water molecules in the crystal. Blue dashed lines are used to help the reader separate molecules in different unit cells.

	AMBA-F
Crystal data	
Chemical formula	$C_{18}H_{26}FN_3O_3 \cdot 0.94(H_2O)$
M_r	368.38
Crystal system, space group	Orthorhombic, $P2_12_12_1$
Temperature (K)	150
a, b, c (Å)	8.5837 (3), 12.3394 (4), 19.5795 (6)
V (Å^3)	2073.82 (12)
Z	4
$F(000)$	789.6
D_x (Mg m^{-3})	1.180
Radiation type	Cu K□
No. of reflections for cell measurement	9958
□ range (°) for cell measurement	4.2–79.9
□ (mm^{-1})	0.74
Crystal shape	Block
Colour	Colourless
Crystal size (mm)	0.23 × 0.22 × 0.14
Data collection	
Diffractometer	Bruker AXS D8 Quest CMOS diffractometer
Radiation source	I-μ-S microsource X-ray tube
Monochromator	Laterally graded multilayer (Goebel) mirror
Scan method	□ and phi scans
Absorption correction	Multi-scan, *SADABS* 2016/2: Krause, L., Herbst-Irmer, R., Sheldrick G.M. & Stalke D., J. Appl. Cryst. 48 (2015) 3-10
T_{min}, T_{max}	0.651, 0.754
No. of measured, independent and observed [$I > 2$□(I)] reflections	15856, 4402, 4131
R_{int}	0.029
□ values (°)	□$_{max}$ = 80.8, □$_{min}$ = 4.2
(sin □/□)$_{max}$ (Å^{-1})	0.640
Range of h, k, l	h = -9□10, k = -15□15, l = -24□23
Refinement	
Refinement on	F^2
$R[F^2 > 2$□$(F^2)]$, $wR(F^2)$, S	0.035, 0.098, 1.07
No. of reflections	4402

Supplemental Figure 3.4. Details of experimental parameters for X-ray crystallography. Computer programs: Apex3 v2017.3-0 (Bruker, 2017), SAINT V8.38A (Bruker, 2016), SHELXS97 (Sheldrick, 2008), SHELXL2018/3 (Sheldrick, 2015, 2018), SHELXLE Rev937 (Hübschle et al., 2011).

Supplemental Figure 3.4. continued.

No. of parameters	257
No. of restraints	5
H-atom treatment	H atoms treated by a mixture of independent and constrained refinement
Weighting scheme	$w = 1/[\sigma^2(F_o^2) + (0.0499P)^2 + 0.2091P]$ where $P = (F_o^2 + 2F_c^2)/3$
$(\Delta/\sigma)_{max}$	0.001
$\Delta\rho_{max}$, $\Delta\rho_{min}$ (e Å^{-3})	0.14, -0.16
Absolute structure	Flack x determined using 1688 quotients [(I+)-(I-)]/[(I+)+(I-)] (Parsons, Flack and Wagner, Acta Cryst. B69 (2013) 249-259).
Absolute structure parameter	0.08 (6)

3.6 References

1. Dill, K. A., Dominant forces in protein folding. *Biochem.* **1990,** *29*, 7133-55.

2. Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R., The protein folding problem. *Annu. Rev. Biophys.* **2008,** *37*, 289-316.

3. Nick Pace, C.; Scholtz, J. M.; Grimsley, G. R., Forces stabilizing proteins. *FEBS Lett.* **2014,** *588*, 2177-2184.

4. Eisenberg, D., The discovery of the α-helix and *β*-sheet, the principal structural features of proteins. *Proc. Natl. Acad. Sci.* **2003,** *100*, 11207-11210.

5. Donohue, J., Hydrogen bonded helical configurations of the polypeptide chain. *Proc. Nat. Acad. Sci.* **1953,** *39*, 470-478.

6. Venkatachalam, C. M., Stereochemical criteria for polypeptides and proteins. V. Conformation of a system of three linked peptide units. *Biopolymers* **1968,** *6*, 1425-36.

7. O'Neil, K. T.; DeGrado, W. F., A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* **1990,** *250*, 646-51.

8. Pauling, L.; Corey, R. B.; Branson, H. R., The structure of proteins: two hydrogen-bonded helical configurations of the polypeptide chain. *Proc. Nat. Acad. Sci.* **1951,** *37*, 205-211.

9. Anderson, C. A. F.; Rost, B., Structural Bioinformatics. In *Structural Bioinformatics*, 2 ed.; Gu, J.; Weissig, H., Eds. John Wiley & Sons, Inc.: Hoboken, 2003; pp 339-363.

10. Toniolo, C.; Benedetti, E., The polypeptide 310-helix. *Trends Biochem. Sci.* **1991,** *16*, 350-353.

11. Beck, D. A. C.; Alonso, D. O. V.; Inoyama, D.; Daggett, V., The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. *Proc. Natl. Acad. Sci.* **2008,** *105*, 12259-12264.

12. Atilgan, C.; Gerek, Z. N.; Ozkan, S. B.; Atilgan, A. R., Manipulation of Conformational Change in Proteins by Single-Residue Perturbations. *Biophys. J.* **2010,** *99*, 933-943.

13. Park, S. H.; Opella, S. J., Conformational changes induced by a single amino acid substitution in the trans-membrane domain of Vpu: Implications for HIV-1 susceptibility to channel blocking drugs. *Prot. Sci.* **2007,** *16*, 2205-2215.

14. Toniolo, C.; Bonora, G. M.; Barone, V.; Bavoso, A.; Benedetti, E.; Di Blasio, B.; Grimaldi, P.; Lelj, F.; Pavone, V.; Pedone, C., Conformation of pleionomers of .alpha.-aminoisobutyric acid. *Macromolecules* **1985,** *18*, 895-902.

15. Marshall, G. R.; Hodgkin, E. E.; Langs, D. A.; Smith, G. D.; Zabrocki, J.; Leplawy, M. T., Factors governing helical preference of peptides containing multiple alpha,alpha-dialkyl amino acids. *Proc. Nat. Acad. Sci.* **1990,** *87*, 487-491.

16. Bavoso, A.; Benedetti, E.; Di Blasio, B.; Pavone, V.; Pedone, C.; Toniolo, C.; Bonora, G. M., Long polypeptide 3₁₀-helices at atomic resolution. *Proc. Nat. Acad. Sci.* **1986,** *83*, 1988-1992.

17. Jones, J. E.; Diemer, V.; Adam, C.; Raftery, J.; Ruscoe, R. E.; Sengel, J. T.; Wallace, M. I.; Bader, A.; Cockroft, S. L.; Clayden, J.; Webb, S. J., Length-Dependent Formation of Transmembrane Pores by 310-Helical α-Aminoisobutyric Acid Foldamers. *J. Am. Chem. Soc.* **2016,** *138*, 688-695.

18. Maekawa, H.; Formaggio, F.; Toniolo, C.; Ge, N.-H., Onset of 310-Helical Secondary Structure in Aib Oligopeptides Probed by Coherent 2D IR Spectroscopy. *J. Am. Chem. Soc.* **2008,** *130*, 6556-6566.

19. Zeko, T.; Hannigan, S. F.; Jacisin, T.; Guberman-Pfeffer, M. J.; Falcone, E. R.; Guildford, M. J.; Szabo, C.; Cole, K. E.; Placido, J.; Daly, E.; Kubasik, M. A., FT-IR Spectroscopy and Density Functional Theory Calculations of 13C Isotopologues of the Helical Peptide Z-Aib6-OtBu. *J. Phys. Chem. B* **2014,** *118*, 58-68.

20. Cable, J.; Tubergen, M. J.; Levy, D. H., Electronic spectroscopy of small tryptophan peptides in supersonic molecular beams. *J. Am. Chem. Soc.* **1988,** *110*, 7349-7355.

21. Cable, J.; Tubergen, M. J.; Levy, D. H., Fluorescence spectroscopy of jet-cooled tryptophan peptides. *J. Am. Chem. Soc.* **1989,** *111*, 9032-9039.

22. Rijs, A. M.; Oomens, J., *Gas-phase IR spectroscopy and structure of biological molecules*. Springer: 2015; Vol. 364.

23. Alauddin, M.; Biswal, H.; Gloaguen, E.; Mons, M., *Intra-residue interactions in proteins: interplay between Serine or Cysteine side chains and backbone conformations, revealed by laser spectroscopy of isolated model peptides*. 2014; Vol. 17.

24. Chin, W.; Piuzzi, F.; Dognon, J.-P.; Dimicoli, I.; Tardivel, B.; Mons, M., Gas Phase Formation of a 310-Helix in a Three-Residue Peptide Chain: Role of Side Chain-Backbone Interactions as Evidenced by IR−UV Double Resonance Experiments. *Journal of the American Chemical Society* **2005,** *127*, 11900-11901.

25. Dian, B. C.; Longarte, A.; Zwier, T. S., Conformational dynamics in a dipeptide after single-mode vibrational excitation. *Science* **2002,** *296*, 2369-2373.

26. Dian, B. C.; Longarte, A.; Winter, P. R.; Zwier, T. S., The dynamics of conformational isomerization in flexible biomolecules. I. Hole-filling spectroscopy of N-acetyl tryptophan methyl amide and N-acetyl tryptophan amide. *J. Chem. Phys.* **2004,** *120*, 133-147.

27. James, W. H.; Baquero, E. E.; Choi, S. H.; Gellman, S. H.; Zwier, T. S., Laser Spectroscopy of Conformationally Constrained α/*β*-Peptides: Ac-ACPC-Phe-NHMe and Ac-Phe-ACPC-NHMe. *J. Phys. Chem. A* **2010,** *114*, 1581-1591.

28. Dean, J. C.; Buchanan, E. G.; Zwier, T. S., Mixed 14/16 Helices in the Gas Phase: Conformation-Specific Spectroscopy of Z-(Gly)n, n = 1, 3, 5. *J. Am. Chem. Soc.* **2012,** *134*, 17186-17201.

29. Kusaka, R.; Zhang, D.; Walsh, P. S.; Gord, J. R.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S., Role of Ring-Constrained γ-Amino Acid Residues in α/γ-Peptide Folding: Single-Conformation UV and IR Spectroscopy. *J. Phys. Chem. A* **2013,** *117*, 10847-10862.

30. Gord, J. R.; Walsh, P. S.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S., Mimicking the First Turn of an α-Helix with an Unnatural Backbone: Conformation-Specific IR and UV Spectroscopy of Cyclically Constrained *β*/γ-Peptides. *J. Phys. Chem. B* **2014,** *118*, 8246-8256.

31. Gord, J. R.; Hewett, D. M.; Hernandez-Castillo, A. O.; Blodgett, K. N.; Rotondaro, M. C.; Varuolo, A.; Kubasik, M. A.; Zwier, T. S., Conformation-specific spectroscopy of capped, gas-phase Aib oligomers: tests of the Aib residue as a 310-helix former. *Phys. Chem. Chem. Phys.* **2016,** *18*, 25512-25527.

32. Walsh, P. S.; Dean, J. C.; McBurney, C.; Kang, H.; Gellman, S. H.; Zwier, T. S., Conformation-specific spectroscopy of capped glutamine-containing peptides: role of a single glutamine residue on peptide backbone preferences. *Phys. Chem. Chem. Phys.* **2016,** *18*, 11306-11322.

33. Blodgett, K. N.; Zhu, X.; Walsh, P. S.; Sun, D.; Lee, J.; Choi, S. H.; Zwier, T. S., Conformer-Specific and Diastereomer-Specific Spectroscopy of α*β*α Synthetic Foldamers: Ac–Ala− *β*ACHC–Ala–NHBn. *J. Phys. Chem. A* **2018,** *122*, 3697-3710.

34. Fricke, H.; Funk, A.; Schrader, T.; Gerhards, M., Investigation of Secondary Structure Elements by IR/UV Double Resonance Spectroscopy: Analysis of an Isolated *β*-Sheet Model System. *J. Am. Chem. Soc.* **2008,** *130*, 4692-4698.

35. Lee, J. J.; Albrecht, M.; Rice, C. A.; Suhm, M. A.; Stamm, A.; Zimmer, M.; Gerhards, M., Adaptive aggregation of peptide model systems. *J. Phys. Chem. A* **2013,** *117*, 7050-7063.

36. Schwing, K.; Reyheller, C.; Schaly, A.; Kubik, S.; Gerhards, M., Structural analysis of an isolated cyclic tetrapeptide and its monohydrate by combined IR/UV spectroscopy. *Chem. Phys. Phys. Chem.* **2011,** *12*, 1981-1988.

37. Řeha, D.; Valdes, H.; Vondrášek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; de Vries, M. S., Structure and IR Spectrum of Phenylalanyl–Glycyl–Glycine Tripetide in the Gas-Phase: IR/UV Experiments, Ab Initio Quantum Chemical Calculations, and Molecular Dynamic Simulations. *Chem. Eur. J.* **2005,** *11*, 6803-6817.

38. Rijs, A. M.; Kabeláč, M.; Abo-Riziq, A.; Hobza, P.; de Vries, M. S., Isolated gramicidin peptides probed by IR spectroscopy. *Chem. Phys. Phys. Chem.* **2011,** *12*, 1816-1821.

39. Barry, C. S.; Cocinero, E. J.; Çarçabal, P.; Gamblin, D. P.; Stanca-Kaposta, E. C.; Remmert, S. M.; Fernández-Alonso, M. C.; Rudić, S.; Simons, J. P.; Davis, B. G., 'Naked' and Hydrated Conformers of the Conserved Core Pentasaccharide of N-linked Glycoproteins and Its Building Blocks. *J. Am. Chem. Soc.* **2013,** *135*, 16895-16903.

40. Stanca-Kaposta, E. C.; Çarçabal, P.; Cocinero, E. J.; Hurtado, P.; Simons, J. P., Carbohydrate–Aromatic Interactions: Vibrational Spectroscopy and Structural Assignment of Isolated Monosaccharide Complexes with p-Hydroxy Toluene and N-Acetyl l-Tyrosine Methylamide. *J. Phys. Chem. B* **2013,** *117*, 8135-8142.

41. Levene, P. A.; Steiger, R. E., The action of acetic anhydride on tertiary amino acids and dipeptides. On catalytic effects. The hydrolysis of acetyldipeptides. *J. Biol. Chem.* **1931,** *93*, 581-604.

42. Paterson, Y.; Stimson, E. R.; Evans, D. J.; Leach, S. J.; Scheraga, H. A., Solution conformations of oligomers of α-aminoisobutyric acid. *Int. J. Pept. Protein Res.* **1982,** *20*, 468-80.

43. Weiner, P. K.; Kollman, P. A., AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions. *J. Comp. Chem.* **1981,** *2*, 287-303.

44. Frisch,M.J.; Trucks,G.W.; Schlegel,H.B.;Scuseria,G.E.; Robb,M.A.; Cheeseman,J.R.; Scalmani,G.; Barone, V.; Mennucci, B.; Petersson, G. A., et al. *Gaussian 16 Rev. B.01*, Wallingford, CT, 2016.

45. Walsh, P. S.; Dean, J. C.; McBurney, C.; Kang, H.; Gellman, S. H.; Zwier, T., Conformation-specific spectroscopy of capped Glutamine-containing peptides: Role of a single glutamine residue on peptide backbone preferences. *Phys. Chem. Chem. Phys.* **2016**. *18*, 11306-22.

46. Bruker, **2016**.

47. Sheldrick, G. M., A short history of SHELX. *Acta Crystallogr. A* **2008,** *64*, 112-122.

48. Sheldrick, G. M. Crystal rerinement with ShelXL. *Acta Crystallogr.* **2015,** *71*,3-8.

49. Hübschle, C. B.; Sheldrick, G. M.; Dittrich, B., ShelXle: a Qt graphical user interface for SHELXL. *J. Appl. Crystogr.* **2011,** *44*, 1281-1284.

CHAPTER 4. CONFORMATIONAL ANALYSIS OF AC-γ^4-PHENYLALANINE-NHME MONOMER AND SINGLE WATER HYDRATION: DOMINANT PHENYL ROTOR INTERACTIONS AND THE EFFECT OF SOLAVTION ON AMIDE STACKING

4.1 Introduction

In most peptides and proteins, hydrogen bonding plays an integral role in directing the molecule's secondary and tertiary structure, and therefore potential bioactivity, creating an imperative demand to understand the propensities driving these interactions.[1–5] Synthetic foldamers are oligomers designed to display discrete folding propensities determined by the chemist, ideally allowing control over bioactivity. To that end, synthetic chemists have endeavored to create designer peptides that allow control over the formation of hydrogen bonds to create tailor-made tertiary structures. Having more flexibility in both the peptides substituents and substitution location allows more control over folding propensities, creating an interest in peptides with extended backbones, such as in β- and γ-peptides, which extend the peptide backbone between the N- and C- terminus by one and two methylene groups, respectively.

Interrogating the conformations adopted by synthetic peptides is traditionally done in the solid and liquid phase using X-ray crystallography and NMR spectroscopy. However, gas-phase spectroscopic methods supported with the latest density functional theory *ab initio* methods provide a unique opportunity to study the conformation of molecules at a high resolution (debatably the highest resolution) free from intermolecular interactions, allowing the molecule to fold under its own propensities while having ready control over the sample to form adducted species.[6–26] Spectroscopy also allows the identification of spectroscopic signatures that are affected by the inter and intramolecular forces responsible for the molecules adopted conformation, providing insights into the structural propensities that drive conformer preference, as well as obtaining conformer specific information using double resonance methods.[27-29]

The role of dispersive interactions in directing molecular conformation, particularly in amide groups,[30-38] is more difficult to identify in comparison to hydrogen bonds. This is due to a difficulty in detection shared between all the detection methods. Dispersive interactions do not have a strong dipole moment relative to hydrogen bonds, which have strong signatures in IR spectroscopy, nuclear magnetic resonance, and X-ray crystallography. However, unique and interesting

dispersive interactions, such as amide stacking, have been documented by our group before, in the synthetic gamma peptides Ac-γ^2-Phe-NHMe (AG2P)[39] and Ac-γ^2-Phe-NMe_2.[40]

The original amide stacking work by James was a cursory inspection of AG2P, where he showed the presence of amide stacking alongside the existence of two phenyl rotamers of structures that form a hydrogen bond between the N-terminal C=O and the H from the protecting NHMe. Additionally, the relative populations of the conformers in the experimental conditions used were measured, showing that the amide stacking was present at roughly 20%, a surprisingly high abundance considering the amide stack has no actual hydrogen bonds. Later, this work was expanded upon by trying to restrict formation of the other conformers by removing the hydrogen used in the hydrogen bond through methylation (Ac-γ^2-Phe-NMe2) and steric hindrance by bulking up the same space on the molecule (Ac-γ^2-Phe-NH*i*Pr). Removing competing structures through methylation did indeed drive the formation of only amide stackings structures, but the introduction of the *i*Pr group lead to reduction of amide stack formation, with the bulky *i*Pr group preventing proper alignment of the amide groups for stacking. James also looked at extended γ^2-peptides to see if amide stacking survived in longer peptides, but found no evidence to support that stacked structures persisted.

Later, Buchanan et al furthered this work by hydrating AG2P with a single water.[41] He found that the binding energy of water to the amide stacked structures was ~15 kJ/mol stronger than to the other observed conformers. He furthered this study further by performing irradiating the water cluster and showing the strong binding of the water to the amide stack structure resulted in isomerization to different water clusters, while in the non-amide stacked structure, competition between isomerization and water-loss occurred.

This series of work was comprehensive, but all based on a γ^2-backbone. Here, we begin investigating the effect of moving the phenyl group from the 2 position to the 4 position, which inverts the orientation of the C=O and amide NH in relation to the phenyl ring. We are interested in the effect this change in orientation (a hydrogen bond donor for an acceptor) has on the stability of the amide stacking, as well as how hydration is affected.

4.2 Methods

4.2.1 Experimental

Sample Handling

Ac-γ^4-Phe-NHMe (AG4P) was dried using a lyophilizer to remove water from the solid, allowing spectra of only the AG4P monomer to be taken. Solid dry AG4F was wrapped in glass wool and placed inside of a glass insert and inserted into a stainless-steel sample holder located immediately before a pulsed general valve (Parker General Valve, Series 9) filled with neon carrier gas (5 bar). The dried sample was then heated to 180–185 °C by wrapping the pulsed valve with a heating rope powered by a Variac, then the resulting neon/AG4F mixture was pulsed into a vacuum chamber, resulting in a supersonic expansion, generating a cold molecular beam that provided collisional cooling of the sample to its zero-point vibrational levels. The resulting expansion was skimmed through a ~2 mm skimmer into the extraction region of a Wiley-McLaren time-of-flight (TOF) mass spectrometer. Water was introduced through a 3-way manifold which allowed regulation of the flow of carrier gas through the water sample holder using a needle valve, such that 0.0–0.8% of the backing gas was water.

Spectroscopic Methods

The output of a Nd:YAG (Continuum, Surelite II, 20 Hz) pumped dye laser (NarrowScan, Coumarin 540A) was frequency doubled and introduced into the chamber orthogonally to the molecular beam as it passed through the extraction region of the TOF. Ions are generated by resonant two-photon ionization (R2PI), and the extracted into the TOF tube and detected using an MCP detector. The ion intensity at an arrival time specific to the mass of the target analyte was monitored as a function of laser frequency to generate UV spectra.

The output of a Nd:YAG (Continuum, Surelite III, 10 Hz) was used to pump an optical parametric oscillator/amplifier stage (LaserVision) which output IR radiation in the NH stretch region. The IR radiation was introduced antiparallel to the molecular beam such that the IR radiation passed through the skimmer and intersected the pulse valve at the orifice, ensuring the maximum number of molecules in the cold region of the expansion are exposed to IR radiation.

Double resonance methods were used to generate conformer specific spectra. These methods utilize the 20 Hz/10 Hz firing rate of the UV and IR lasers, respectively, with the IR laser spatial

overlapping the UV laser. The timing of the IR laser allows control of how far the molecular beam travels from the pulse valve towards the extraction region of the ToF extraction region before interacting with the IR radiation. In resonant infrared ion depletion spectroscopy (RIDIRS), the UV laser is fixed on the frequency of a conformer specific transition and the frequency of the IR laser is scanned and introduced temporally preceding the UV laser by ~60 μs (which is approximately the time it takes for the molecules to travel from the skimmer, where they meet the IR laser, to the extraction region of the ToF). When the IR laser frequency is on a vibrational transition, population is removed from the zero-point vibrational level of every other shot of the UV laser, creating a signal with the IR-on and IR-off. Because the UV laser generates signal only from one conformer, ion signal will only be depleted when the IR and UV laser are on transitions that belong to the same conformer, and a conformer specific IR spectrum is generated by taking the difference of the IR-on/IR-off ion currents.

IR ion gain spectroscopy (IRIGS) uses the same timings and alignment of lasers as RIDIRS, with the exception being that the UV laser is fixed where there are no resonant transitions. When the IR laser is scanned over the transition of *any* conformer, the zero-point population that is removed is spread over numerous states through redistribution energy absorbed from the IR photon, resulting in spectral broadening of the UV transition. This non-conformer specific broadening will cause a gain in on current, resulting in an IR spectra that is not conformer specific, but contains IR transitions for all of the conformers present in the expansion.

IRUV hole burning (IRUVHB) follows the same laser alignment and timings, but instead the IR laser is fixed on a conformer specific frequency and the UV laser is scanned. Differences in ion current between IR-on/IR-off shots is again recorded as a function of UV frequency, and because the IR laser is only removing the zero-point population of a specific conformer, modulation of intensity is only observed when the UV frequency and IR frequency belong to the same conformer, generating conformer specific UV spectra.

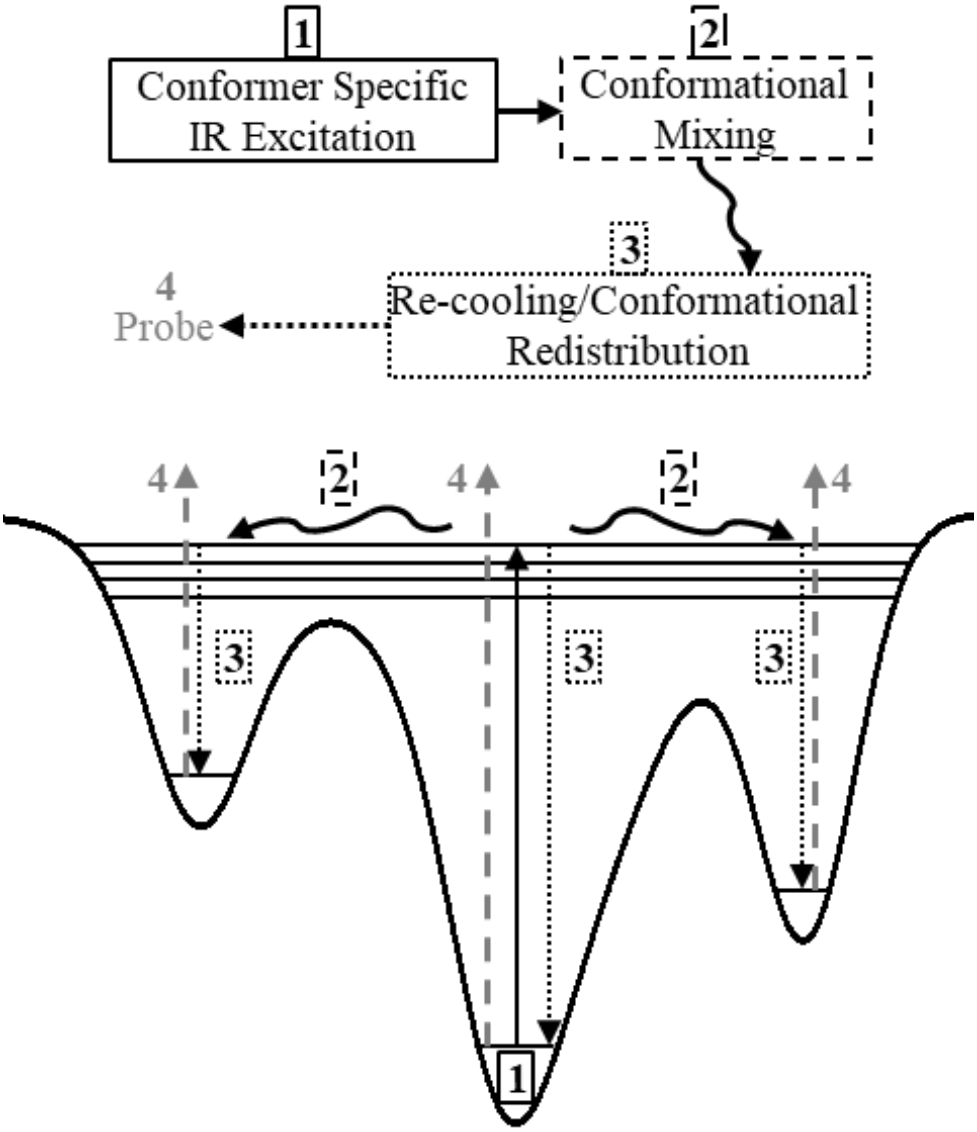


Figure 4.1. *top*) Flow diagram of IRPT. *bottom*) Cartoon depicting IRPT steps after initial cooling to the molecules zero-point vibrational level. The outline of the boxes is consistent between the top and bottom of the figure to help the reader track which arrow belongs to which step.

IRPT uses the same laser alignment as RIDIRS but varies the timing to so that the laser intersects the molecular beam immediately after the molecules have cooled to their zero-point vibrational levels. The IR laser frequency is set to a conformer specific transition, vibrationally exciting molecules early in the expansion such that they can then be recooled if the density of the expansion is still high enough to facilitate collisions. Assuming the IR photon provides enough energy to overcome isomerization barriers and the cooling rate allows sufficient time for

isomerization to occur, the population of vibrationally excited conformers will be depleted as they isomerize, and the population of conformers that were not vibrationally excited will grow, all relative to the populations initially generated in the super-sonic expansion. If no new conformers are populated, it can be shown that the following equation describes the populations:[12]

$$0 = \sum_i F_i * I_i^{PT}(\tilde{\nu}) \qquad \text{Equation 4.1}$$

where F_i is the fractional abundance of conformer i and $I_i^{PT}(\tilde{\nu})$ is the intensity of the IRPT scan of conformer i as a function of wavenumber, $\tilde{\nu}$. By solving Eq. 1 for all $\tilde{\nu}$, a stringent, overexpressed criterion is met to determine the fractional abundances of all the conformers observed. An in-house script that systematically stepped each scaling factor from 0 to 100 and then took the weighted sum of all IRPT spectra. The spectra with the smallest sum of the absolute value of each data point it shown as the sum, and the scaling factors were renormalized to sum to 1.

UV spectra showing the gain and depletion of population can also be acquired. These spectra are called IRUV hole-filling (IRUVHV) spectra and are acquired similarly to IRPT spectra, but the IR laser is fixed on a conformer specific transition and the UV laser is scanned. However, IRUVHB spectra cannot be used to determine relative populations due to a reliance on the intensity on Frank Condon factors. These spectra, however, can be used to provide an idea of how conformers redistribute and sharp gain peaks provide an indication of successful recooling after IR radiation.

Computational Methods

A conformational search of the potential energy surface using the OPLS3 force field was carried out using Monte Carlo methods to vary the dihedral angles of the peptide with Macromodel, included in the Maestro software suite, allowing a population of low energy conformers to be identified. [42] The lowest 100 energy structures were further optimized with density functional theory (DFT) using either the B3LYP with Grimme's D3BJ (B3LYP-GD3BJ) dispersion correction or M05-2X functional, both with the 6-31(g)+d basis set using Gaussian 16. [43] It was found that B3LYP handled relative energies relative to experiment well, but struggled predicting frequencies of the dispersive interactions. Accordingly, when comparing energies to relative

populations, B3LYP-GD3BJ is used, otherwise M05-2x is used. Dihedral scans were also carried out using Gaussian 16 with the B3LYP-GD3BJ. An anharmonic correction factor was used to in predicting the NH spectra as follows. M05-2x: NH stretches were scaled by 0.94; OH stretches were scaled by 0.95. B3LYP: NH stretches were scaled by 0.957, OH stretches were scaled by 0.97. For frequencies, M05-2x functional was used and for energies B3LYP with Grimme's D3BJ dispersion correction was used.

Binding energies were calculated by taking the difference in energy between the complex and each component of the complex, then correcting for basis set superposition error using the counter-poise method:

$$BE_{BSSE}(AB) = E_{AB}^{AB}(AB) - E_{A}^{AB}(A) - E_{B}^{AB}(B) \qquad \text{Equation 4.2}$$

where $BE_{BSSE}(AB)$ is the binding energy of a complex consisting of molecules A and B. For the following components, E is the energy of each component where the super script denotes the basis used, the subscripts denote the geometry under which the energy was calculated, and the symbol in the parenthesis denotes the component considered.

4.3 Results and Analysis

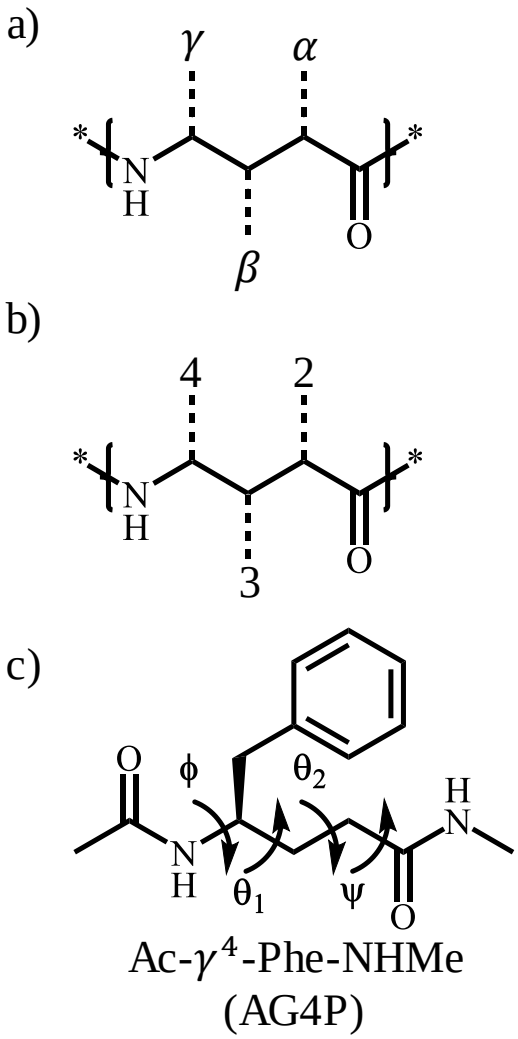
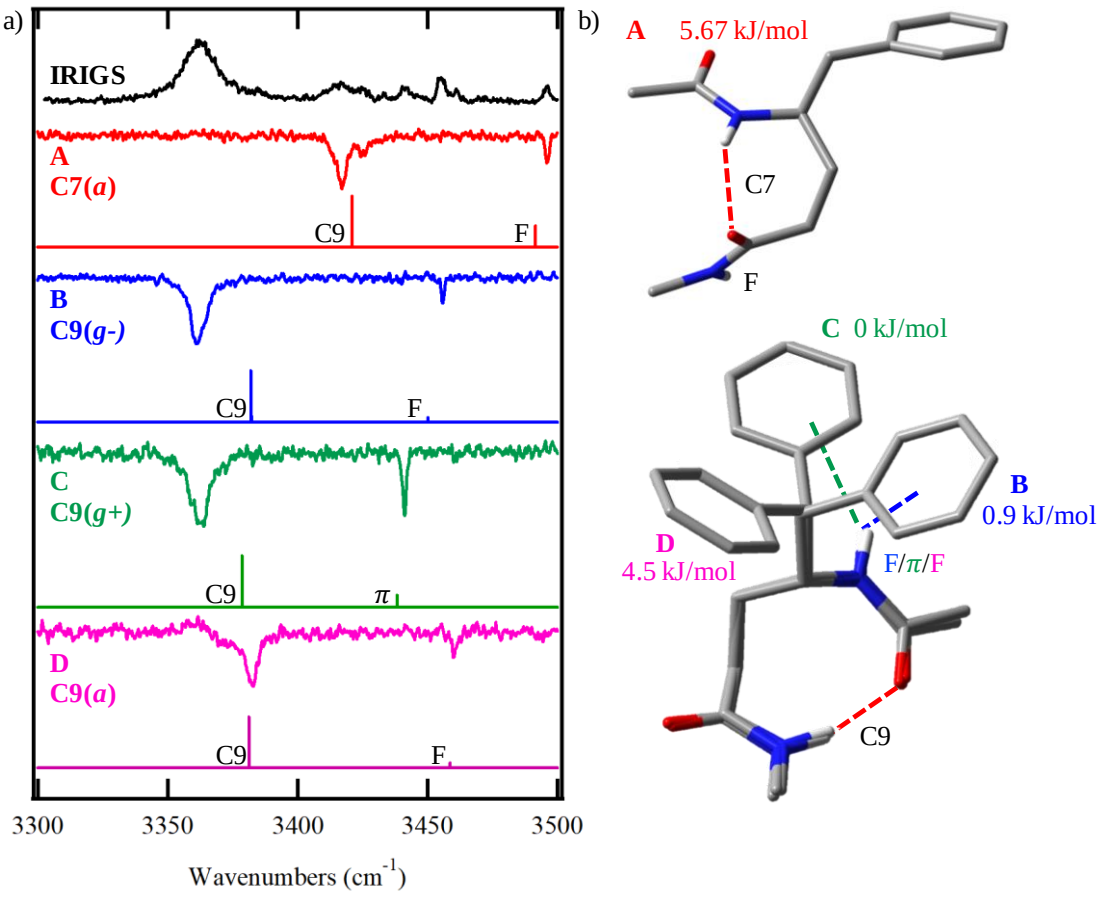


Figure 4.2. Scheme depicting a) carbons responsible for determining if a peptide is $\alpha/\beta/\gamma$, b) the numbering assigned to substitution along the peptide backbone, and c) the AG4P structure used in this study.

4.3.1 Nomenclature

Figure 4.2 shows how the synthetic peptides are classified based on their backbone length and where along that backbone substitution occurs. Figure 4.2c shows AG4P, studied in this work, as well as how the dihedral angles afforded to γ-peptides in addition to the ϕ/ψ angles found in naturally occurring α-amino acids are named. The first dihedral following the ϕ angle is denoted θ_1, and the dihedral angle following θ_1 is θ_2, then ψ dihedral angle preceding the c-terminus of the peptide. Intramolecular hydrogen bonds inherently close a ring and will be defined by the number of members in the ring closed by the hydrogen bond. A ring closed by a hydrogen bond that has *n* members will be called a C*n* hydrogen bond.

The orientation of the phenyl ring relative to the backbone also plays an important role in distinguishing conformers and accordingly the orientation will be defined as either gauche ($g\pm$) or anti (a), with a + or – defining whether or not the gauche angle is approximately +60° or -60°, respectively or the *a* corresponding to a dihedral angle of ~180°. Dihedral angles are defined by starting with the carbon closest to the N-terminus. The water adducted conformers will be handled in a similar manner, using subscripts to help the reader track the location of the water and will be discussed in more detail closer to discussion of the AG4P•H_2O clusters.



p

Figure 4.3. a) The IRIGS scan of AG4P is plotted above the conformer specific RIDIRS scans of AG4P. The corresponding best-fit calculated structure is plotted below each colored RIDIRS scan. b) Assigned conformers of AG4P. Note that 3 conrformers share nearly identical dihedral anagles, excluding the phenyl rotomermer, and have been stacked to illustrate how similar the structures are. The color coded letter next to each phenyl ring assigns the conformer to the colored spectrum. The color also is used to assign the non-C9 hydrogen bond. *We refer the reader to the onine version if coloring is not available.*

4.3.2 Spectroscopy of AG4P Monomer and Conformational Assignments

The electronic spectrum initially taken of AG4P (Supplemental Figure 4.1) showed an unexpectedly high peak density. Upon taking vibrational spectra of higher energy electronic transitions, it became apparent that enough water was present in the sample to generate AG4P•H_2O clusters in an abundance rivaling the monomer. This suggests that either upon electronic excitation, or after ionization in the R2PI process, the AG4P•H_2O complex undergoes fragmentation through water loss, indicated by the presence of AG4P•H_2O transitions in the parent mass channel. While we eventually do study the spectroscopy of said clusters, their presence interfered with our ability

to do IRPT of the bare monomer, specifically if vibrational transitions of the cluster overlapped with electronic transitions of the parent or the electronic excitation resulted in fragmentation of water from the cluster, increasing the population of one of the AG4P parent conformers. Upon drying the sample, the spectral density lowered considerably (Supplemental Figure 4.1).

We also found that electronic exitation of AG4P provided a high fragmentation yield (Supplemental Figure 4.2). The ion current genreated by the loss of the benzyl group (m/z 157.1) was greater than the ion current generated by the parent AG4P (m/z 248.2) by nearly a factor of two. Additinally, the R2PI spectrum of the parent and fragment show a difference in the intesities of UV transitions, suggesting that certain conformers may fragment more readily than others. Particullarly, the ratio of intensities of the most abundant transition at 37505 cm^{-1} and the second most abundant transition at 37544 cm^{-1} differ from between the spectra taken monitoring the ion current of the parent and fragment ions (Figure 2.1). RIDIRS spectroscpy complimented by IRUVHB spectroscpy (Figure 4.3, Supplemental Figure 4.3) show that, indeed, these transitions belong to different conformers, suggesting that certain conformers are more likely to fragment than others. The high fragmentation yeild as well as transistion specific fragmentation distributes the parent absortption over two mass channels, so the sum of both channels is used to generate the UV spectrum for the monomer.

The IR spectra of the four unique conformers identified through RIDIRS/IRUVHB spectroscopy are shown in Figure 4.3a, where the RIDIRS spectra reveal two transitions in the NH stretch region. Additionaly, and IRIG's spectrum (top,black), shows the transitions of all of the conformers of AG4P present in our expansion and confirms that we have spectra for all conformers in significant relative abundance. Below the IRIGS spectrum are the conformation specific RIDIRS spectra, as well as the best fit stick spectra of the assigned conformers, calculated using the M05-2x functinoal. While there is substantial spectral congestion around 3360 cm^{-1}, preventing confident matching of IRIGS peaks to RIDIRS peaks, comparison of the other transitions of the IRIGS spectrum against the four RIDIRS spectra does account for all major peaks.

For conformer A, shown in red, a broad, intense, transition is found at 3416 cm^{-1} with a shoulder at 3425 cm^{-1}, belonging to a C7 hydrogen bond between the C-terminal amide hydrogen (added the NHMe group) and the N-terminal carbonyl oxygen. The best fit structure does not account for the shouldering we observe at 3425 cm^{-1}, but instead falls between the intense peak and the shoulder. We verified the shouldering was not a low abundance conformer that shares the

same UV transition by taking conformer specific IRUVHB spectra with the IR laser parked on both the intense transition at 3416 cm^{-1} and the local maximum created by the shoulder at 3425 cm^{-1} as well as the shoulders of the peaks. Identical IRUVHB spectra were produced, suggesting that both transitions belong to the same conformer and, due to the number of NH oscillators are likely the same NH stretch. Conformer A has a sharp, low intensity transition at 3496 cm^{-1}, indicating the C-terminal amide hydrogen is free and not participating in any hydrogen bond. Considering the shoulder and the main peak belong to the same NH stretch, the best fit stick spectra provides excellent agreement to experimental results.

Conformer A is assigned to a C7 structure, with the phenyl ring *anti* to the peptide backbone starting from the N-terminus and is thus refered to as C7(*a*) structure. The structure can be found at the top of Figure 4.2b. It is noteworthy that the C7 hydrgogen bond formed is strained, having a ∠COH angle of 88°. This strained hydrogen bonding angle aligns the amide groups into a nearly stacked arrangement, though there does appear to be an approach to a hydrogen bond. Based on the AG2P amide stacking RIDIRS spectra, the amide stacked transitions are expected to occur around 3470 cm^{-1} and 3480 cm^{-1}. The shifting of the hydrogen bond to lower energy clearly indicates that this is not true amide stacking. The zero-point corrected energy of conformer A, however, is 6.3 kJ/mol above the global minumum (the sixth lowest energy structure), which is unexpected for a conformer that appears to be in relatively high abundance based on the peak intensities of the electronic spectrum. We do acknowledge that other influences, such as Frank-Condon factors, determine the intensities of electronic transitions, but also acknowledge that in the benezene based chromophore systems, the the relative intensities would not normaly be expected to vary significantly enough to explain why conformer A has such a large relative intensity, especially considering the lack of Frank-Condon activity in all conformers, suggesting no significantly large geometry changes between the ground and excited electronic states. The energetics that lead to this aparent discrepancy will be discussed in detail in the Energetic Analysis section of the Discussion, where we will explain our confidence in assigning this surprsingly high energy structre.

The RIDIRS spectra in the NH stretch region for conformers B, C, and D are all shown above their corresponding best fit line spectra in Figure 4.3a. The best fit line spectra all show reasonable agreement with the experimental spectra, lending confident sturctural assignemnts. The three spectra are noticeably similar, suggesting they all belong to similar conformers. All three

conformers show a broad peak in the ~3360–3380 cm^{-1} region, suggesting they could have a similar hydrogen bond. The remaining NH stretch transition is found from ~3440–3460 cm^{-1} in all three conformers. Indeed, all three assigned structures are stacked in Figure 4.3b, showing all three conformers are phenyl rotomers of the same C9 hydrogen bond. The hydrogen bonded transition for Conformer B is centered at ~3361 cm^{-1}, for conformer C at 3363 cm^{-1}, and conformer D at 3382 cm^{-1}, and belongs to the C-terminal amide NH being hydrogen bound to the N-terminal carbonyl oxygen. The remaning NH stertch transition for conformer B and conformer D are found at 3455 cm^{-1} and 3460 cm^{-1}, respecively, both belonging to a free NH stretch. Conformer C presents a transition at 3441 cm^{-1} with increased relative intensity, suggesting an NH-π interaction. Conformer C is found to be the energetic zero-point corrected minimum, followed by conformer B at 0.9 kJ/mol, and conformer D at 4.5 kJ/mol. The other two low energy conformers predicted by our conformational search are phenyl rotomers of conformer A and will also be discussed in greater detail in the Energetic Analysis section of the Discussion.

Considering the distinguishing characteristic of conformers B, C, and D, is the phenyl rotomer, we will use that to distinguish conformers in a more intuitive way hereafter, instead of using the A,B,C,D notataion used thusfar. Conformer B has phenyl rotor dihedral angle of ~-60° and conformer C having a phenyl rotomer dihedral of ~+50°. Conformer B is hereafter refered to as C9(g-) and conformer C as C9(g_+). The phenyl rotomer dihedral angle assigned to conformer D is -172° and is hereafter reffered to as C9(a).

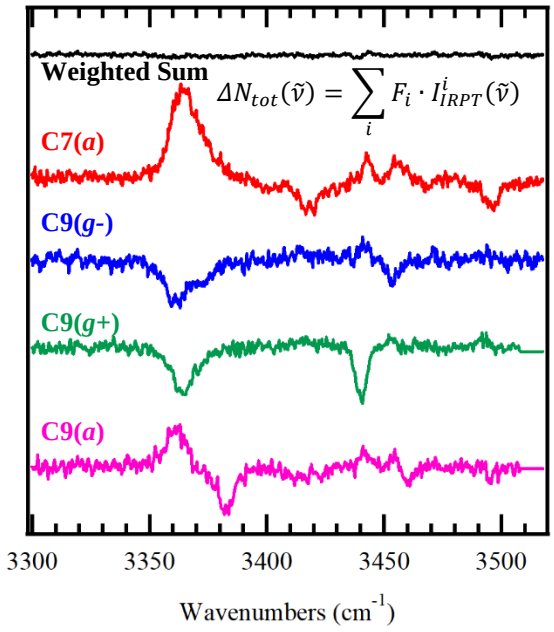


Figure 4.4. (*top*)The sum of all fractionally weighted IRPT spectra is shown above the IRPT spectrum of each labeled conformer before weighting. Each spectra has been normalized o the gain or depletion relative to the ion current when there is no IR laser, with signal above the baseline corresponding to gain and signal below the baseline corresponding to depletion of population.

Table 4.1. Relative populations for Ac-γ^2-Phe-NHMe[39] and Ac-γ^4-Phe-NHMe determined using IRPT.

	γ^2-Phe	γ^4-Phe
C7(*a*)	n/a	27±5
C9 (*g*+)	n/a	13±2
C9 (*g*-)	41±2	53±5
C9 (*a*)	38±1	7±1
Amide Stack	21±1	n/a

4.3.3 Relative Population Distribution of AG4P

IRPT will provide us information about the relative population of each conformer as it exists under the expansion conditions. This will be useful in evaluating the C7(*a*) population relative the nearly equal intensity C9(*g*-) conformer. The black trace shown at the top of Figure 4.4 is the weighted sum of all the colored IRPT spectra below it, where the weighting factors for each IRPT spectra used in the sum are the relative populations of each conformer, determined by satisfying

Eq (1). The IRPT spectra shown below are normalized to ion current generated when only the UV laser is fired. Accordingly, in the IRPT spectra, values above the baseline signify a gain population and values below the baseline correspond to depletion of the parent population. IRUVHF spectra can be found in Supplemental Figure 4.5, showing effective recooling after undergoing vibrational transitions, indicating population depleted from the zero-point vibrational level is only redistributed to low-energy conformers and not lost to a vibrationally excited population. Supplemental Figure 4.5 also shows all population is redistributed to only conformers identified in the R2PI spectrum, meting all conditions required to use Eq (1). The weighted sum pleasingly shows little deviation from zero. The average of the absolute value of the sum of the intensity over all frequencies provides a useful metric to evaluate how well the weighted sum approaches an ideal zero value. The sum of the absolute value of the intensity over all frequencies of the weighted sum spectrum is $2.65 \times 10^{-3} \pm 1.60 \times 10^{-3}$, which is impressive considering the relative maximum intensity of the weighted components was 4.86×10^{-1}, providing confidence in our population assignments. The relative populations found were 53±5% C9(*g*-), 27±5% C7(*a*), 13±5% C9(*a*), and 7±1 C9(*a*). The values are listed in comparison to the population analysis determined for Ac-γ^2-Phe-NHMe in Table 4.2, which is relevant due to the similarities in conformers observed. Indeed, the C9(*g*-) conformer is the most abundant in both the γ^2 and the γ^4 species. Additionally, the assigned populations show no agreement with the zero-point corrected energies, but instead agree more with the 'eyeball' assignment based on relative intensities of electronic transitions. An explanation for this is also found in the Energetic Analysis section of the Discussion.

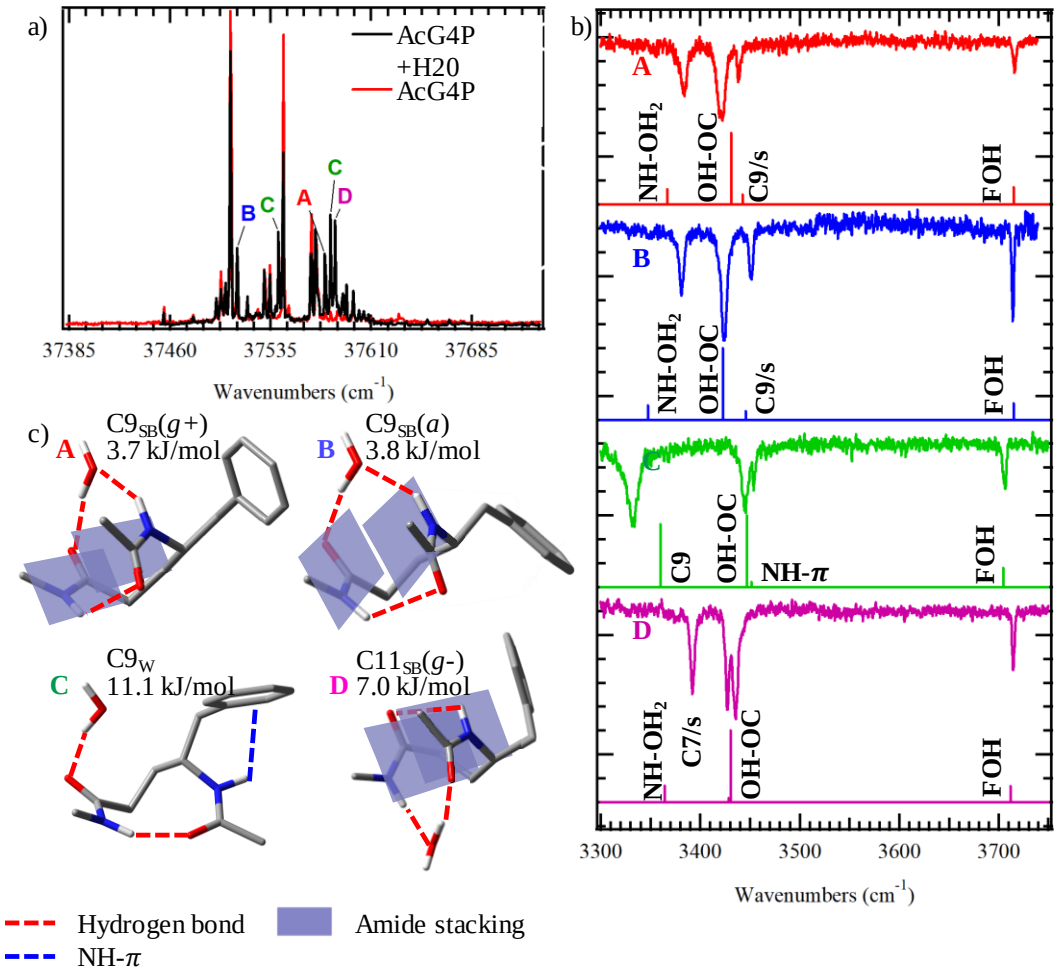


Figure 4.5. a) R2PI spectra of AG4P (red) and the AG4P•H_2O complex (black). b) RIDIRS spectra of the four identified conformers of AG4P•H_2O complex. c) Assigned conformers of AG4P•H_2O conformers. Conformers assigned to letters and colors are consistent across all aspects of the figure. Readers are referred to the online text if a colored version of the text is not available.

4.3.4 Spectroscopy of AG4P•H_2O Complex and Conformational Assignment

Figure 4.5a shows the R2PI spectrum of the AG4P monomer (black) overlaid with the R2PI of the AG4P•H_2O complex. It was found that upon electronic excitation, the AG4P•H2O complex dissociated almost entirely to the AG4P monomer. Thus, we identified transitions to perform RIDIRS on by overlaying the R2PI spectra before and after the addition of water. The use of conventional IRUVHB spectroscopy to match electronic transitions to conformers was not possible due to spectral congestion, particularly in the 3300–3400 cm^{-1} range, leaving a process of

elimination as the most available method to determine which transitions belonged to the AG4P•H_2O cluster or AG4P monomer. Overlaying the electronic spectra (Figure 4.5a) provides a means of identifying which electronic transitions belong to the AG4P•H_2O dimer. It should be noted that the R2PI spectrum of only the unfragmented parent is used for comparison, as only that mass channel was used to record the R2PI spectrum of the AG4P•H_2O complex.

Comparing the R2PI spectra of AG4P and AG4P•H_2O, the relative intensity of the monomer peaks with and without the presence of water have changed. This is a result of taking the AG4P•H_2O electronic spectrum in only the parent mass channel while the comparison spectra of the AG4P monomer is the sum of the parent and the fragment resulting from the loss of the phenyl ring. A very similar relative intensity pattern for just the parent can be seen in Supplemental Figure 4.3. Four conformers were found in a significant enough abundance to perform double resonance spectroscopic studies on. Unique electronic origins were found at 37575 cm^{-1} (Conformer A), 37510 cm^{-1} (Conformer B), 37541 cm^{-1} (Conformer C), and 37583 cm^{-1} (Conformer D), all with limited Frank Condon activity. Inferring from the spectra of the AG4P monomer, the origin is dominantly the most abundant peak, and due to the inherently lower intensity of the water complex signal, some of the Frank Condon activity may not even be visible within the limit of detection of the instrument, though some activity is apparent at ~37590 cm^{-1} and higher frequencies.

The vibrational RIDIRS spectra of the identified four conformers is shown in Figure 4.5b, with each best fit calculated stick spectrum presented beneath the respective RIDIRS spectrum. Each spectrum contains four peaks, one for each of the two amide NH oscillators, as well as one for each of the two water OH oscillators. The assigned conformers calculated from the best fit structures are presented in Figure 4.5c. The high degree of similarity and spectral overlap between the conformers complicates assignment, leading us to tentative assignments. However, we do feel the assignments are within the families of conformers identified through the conformational search and provide suitable grounds for discussion. See Supplemental Figure 4.6 for a comparison of the experimental spectra to the families identified.

Our conformational search found that there were six major families for AG4P•H_2O to populate, which we will discuss briefly now to try to aid with the clarity and flow of the text. The water adduct can simply add itself to the existing conformers present in the expansion once the internal energy of AG4P has been cooled away, inducing minimal conformational changes, which will continue to be referred to as $C7_W$ or $C9_W$ structures, where the W stands for water adducted. The

water can bridge hydrogen bonding sites, as is common in water clusters, but interestingly the bridging water can leave the structure nearly unperturbed or induce amide stacking. That is, bridging of the water in the AG4P C9 structure can result in a bridged C11 structure ($C11_B$), or generate an amide stacked structure bridged by the water with a C11 hydrogen bond ($C11_{SB}$). Note the $_{SB}$ stands for stacked bridged and $_B$ stands just for bridged. The same process can happen for the AG4P C7 structure, making a stacked bridged C9 ($C9_{SB}$) or bridged C9 ($C9_B$) structure. We will continue with the same *g+*/*g-*/*a* to distinguish phenyl rotor dihedrals. With this nomenclature in place, we can look to the RIDIRS spectra and conformational assignments.

The RIDIRS spectra for conformers A, B, C, and D can be found in Figure 4.5b, each with the best fit stick spectra shown below the spectra it is assigned to. Conformer A, shown in red, is in good agreement with the best fit spectrum presented below it and assigned to a $C9_{SB}$(*g+*) structure, where the water bridges what would be the C7 hydrogen bond in the AG4P monomer, but induces a conformational change to an amide stacked structure. The lowest intensity transition of conformer A is a sharp transition found at 3715 cm^{-1}, in the free OH stretch region of the vibrational spectrum. The remaining peaks can be found in the 3350-3450 cm^{-1} region, where hydrogen bound NH and OH bonds are expected to have transitions. A sharp, low intensity peak can be found at 3437 cm^{-1}, nearly residing on the shoulder of the most intense peak in the spectrum, a broad peak centered at 3421 cm^{-1}. The transition at 3437 cm^{-1} belongs to an OH stretch where the water hydrogen is bound to the C-terminal C=O, forming a C11 hydrogen bond and stapling the amide stack into its geometry. The transition at 3421 cm^{-1} belongs to the hydrogen bound NH oscillator, where the N-terminal hydrogen is bound to the oxygen of the water. The lowest frequency peak is broad and of the second most intensity, found at 3384 cm^{-1}, and is assigned to a frustrated C9 hydrogen bond. The best fit spectrum presented below the vibrational spectrum of conformer A and shows good agreement.

Note that we have assigned the structure to both an amide stack and a C9 hydrogen bond. This is due to the planar amide group needing to pucker in order to form the C9 hydrogen bond, resulting in an unfavorable hydrogen bonding angle and a weaker hydrogen bond. The dihedral angle of the amide C–N bond is -165°, showing a 15° pucker from the expected value of 180° in order to satisfy the frustrated C9 hydrogen bond. The COH bond angle is 90.9°, which is nearly ideal for stacking but less so for hydrogen bonding conditions, which prefer bond angles of 120° or 180°.

The NH/OH stretch spectrum of conformer B, shown in blue below conformer A in Figure 4.5b, has a very similar peak intensity and spacing pattern, suggesting a similar hydrogen bonding network. The best fit spectrum, shown below the experimental spectra, is in good agreement and we assign his conformer to a $C9_{SB}(a)$ structure, only differing significantly from the $C9_{SB}(g+)$ structure assigned to conformer A in the phenyl rotor dihedral angle. A sharp, strong transition can be found at 3714 cm^{-1}, corresponding to a free OH stretch, with three transitions between 3380 cm^{-1} and 3450 cm^{-1} following the same intensity pattern as conformer A. These peaks are located at 3380 cm^{-1}, 3422 cm^{-1}, and 3450 cm^{-1}. The peak at 3380 cm^{-1} corresponds to a C-terminal amide NH hydrogen bound to the water oxygen. The most intense, broad peak at 3422 cm^{-1} is assigned to the hydrogen bond water OH stretch while bound to the N-terminal acyl group. The peak found at 3450 cm^{-1} belongs to a frustrated C9 hydrogen bond. The best fit predicted stick spectrum is shown below, again showing reasonable agreement and being assigned to a similar structure where the water bridges an amide stacking structure. This generates a similar frustrated amide stacking/hydrogen bonding effect as in conformer A, resulting in a similar spectrum.

The spectrum for conformer C is unique in comparison to conformers A and B. Most notably, the lowest frequency peak (found at 3331 cm^{-1}) has shifted to a lower frequency by about 50 cm^{-1} and broadened in intensity relative to all the other conformers, suggesting stronger hydrogen bonding. The best fit spectrum presented below the RIDIRS spectrum is in good agreement, and being the only unique spectrum, is our most confident assignment. The structure is a $C9_W(g-)$ structure with the water adducted to the C-terminal carbonyl oxygen. The stronger hydrogen bond relative to the $C9_{SB}$ structures assigned results from the water not having to bridge between the two amide groups, leaving the hydrogen bond unperturbed. Interestingly, the water appears to have a cooperative effect, strengthening the backbone C9 hydrogen bond, indicated by the shift from 3361 cm^{-1} in the monomer to 3331 cm^{-1} in the AG4P•H_2O complex. The highest frequency peak, found at 3706 cm^{-1} exhibits a modest shift to lower frequency in comparison to the other three conformers, though the OH oscillator involved is entirely free, oscillating into the empty space above the C-terminal amine, away from the phenyl ring. This is likely a consequence of the water molecule in the $C9_W$ being involved in only one hydrogen bond instead of two as in the $C9_{SB}$ structures. This allows more electron density to remain in the OH bond, resulting in a decrease in the frequency.

Conformer D presents interesting spectroscopy, but first we will discuss the best fit spectra as well as the conformational assignment. The best fit spectrum for conformer D is presented below the experimental spectrum in Figure 4.5b, with the assigned structure being a $C11_{SB}$(*g*-). Upon first inspection, conformer D has a unique spectrum, suggesting the adopted structure will belong to a unique family. However, upon closer inspection, it is found that the frequencies of the transitions are very close the that of the $C9_{SB}$ structures already assigned, with the significant difference being in the intensity pattern of the two peaks that make a doublet, with peaks of nearly equal intensity at 3427 cm^{-1} and 3435 cm^{-1} standing out in particular. The intensity pattern could result from energetic coupling in which the lower intensity transition at 3427 cm^{-1} is nearly isoenergetic with the transition at 3435 cm^{-1} and borrows intensity from the much stronger transition. In this case, the relative intensity suggests the peak at 3435 cm^{-1} belongs to the hydrogen bound OH of the water, bound to the C-terminal carbonyl of the AG4P monomer. The peak at 3427 cm^{-1} is still at a similar frequency to the frustrated C9 hydrogen bond found in conformers A and B and is assigned to the NH stretch of a frustrated C11 hydrogen bond between the C-terminal amide N and N-terminal C=O. The frequency of this transition being so like the frustrated C9 hydrogen bond observed in conformers A and B suggests that this transition may be characteristic of water bridging a hydrogen bond in such a way that a stacking motif is adopted. The transition at 3714 cm^{-1} belongs to the free water OH stretch, unsurprisingly.

4.4 Discussion.

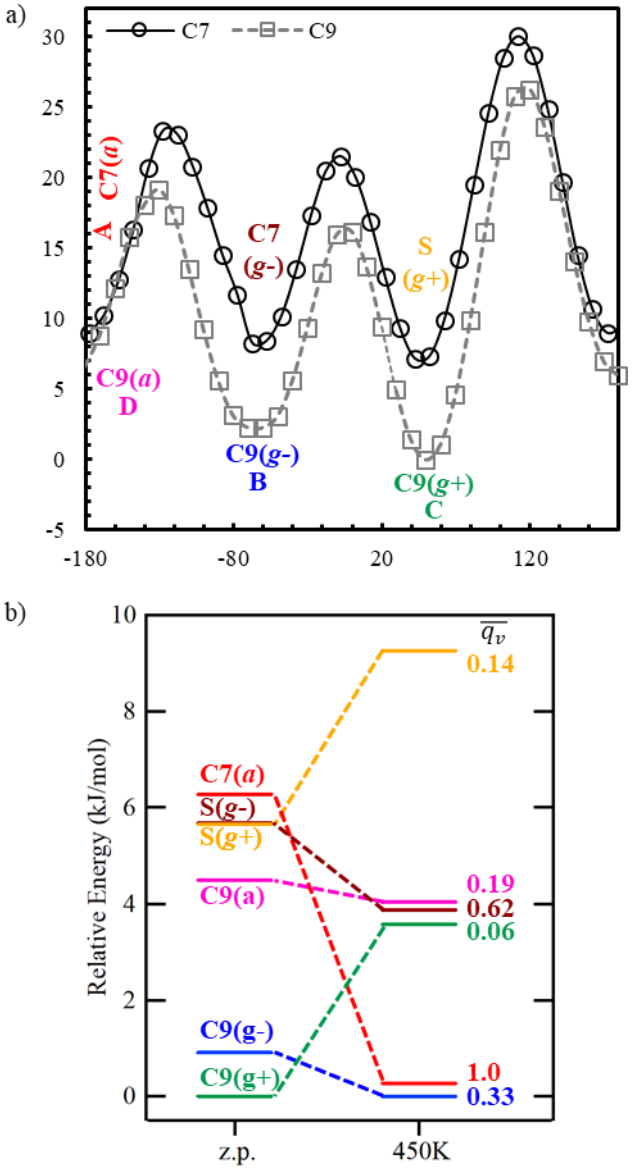


Figure 4.6. a) Relative energies of conformers generated by scanning along the phenyl rotor dihedral of AG4P and b) the relative energies of relevant conformers observed experimentally and only computationally (S(*g*-) and S(*g*+)). All energetic calculations were performed using DFT with the B3LYP-GD3BJ using the 6-31+g(d) basis set.

4.4.1 Energetic Analysis of AG4P Monomer and Population Analysis

A cursory inspection of the assigned conformers and relative energies of the AG4P monomer (Figure 4.6) presented earlier raises an immediate question: Why do we see phenyl rotamers of the C9 structure but only a single phenyl rotamer of the C7 structure, none the less the highest energy phenyl rotamer of the C7 structure? To address the former question, we performed a relaxed dihedral scan of the phenyl rotamer coordinate of both the C7 and C9 structure, shown in Figure 4.6a to estimate the relative energies of the C7 phenyl rotamers and barriers between isomerization. These values are not zero-point corrected but will suffice to supplement discussion.

We explored to see if there were larger barriers separating the C9 structures, potentially trapping population in those conformers while the C7 rotamers would be free to isomerize. The trace represented by a dashed, grey line was produced by seeding the dihedral scan with the assigned C9(*g*-) conformer and shows similar 10–15 kJ/mol barriers between conformers with relative energies in rough agreement with what we reported with the assigned structures. The solid black trace results from seeding the dihedral scan with the assigned C7(*a*) structure and rotating the phenyl dihedral, and shows two structures that are lower in predicted energy than the experimentally observed conformer, separated by barriers of roughly 10–15 kJ/mol. Interestingly, in the C7 scan, the hydrogen bond becomes increasingly frustrated as you move away from the *a* conformer and start approaching the *g*+ or *g*- rotamers, until an amide stacking configuration is produced as the local minima for the C7(*g*+) and C7(*g*-) conformers (See Figure 4.6b). This stacking is must put an unusual strain on the amide plane to adopt a stacking configuration. In order to adopt these stacked configurations, the amide plane is greatly perturbed to a dihedral angle of 162–163° to facilitate what appears to be a weak, frustrated hydrogen bond. Comparison of the predicted line spectra of these two new conformers do not match our experimental spectra, with predicted peaks at around 3444 and 3454 cm^{-1} for both conformers (recall the experimental peaks were found at 3416 and 3496 cm^{-1}, separated by 80 $cm\text{-}^{1}$ and not 10 cm^{-1}). These two stacked structures found in the dihedral scan will be referred to as S(*g*+) and S(*g*-) in convention with the nomenclature we have been using, where the S stands for stacked.

Unfortunately, the barriers between the phenyl rotamers of the C7 or C9 structures are consistent, suggesting even more so that we should observe the S(*g*+) and S(*g*-) phenyl rotamers of the C7(*a*) conformer. However, the barriers between the rotamers in both the C7 and C9 structures are substantial enough that we consider kinetic trapping induced by fast cooling of

conformers rather than thermodynamic redistribution of populations. Indeed, upon calculating Gibb's free energies for the molecules at 450 K (the temperature required to volatilize the sample inside of the pulse valve), we see the energetic ordering of each conformer changes dramatically, with conformer B becoming the lowest in energy, followed by A, C, and D, respectively (Figure 4.6b). Additionally, the two phenyl rotamers of the C7 structure, the sought after amide stacking conformers that were not observed become considerably higher in energy than the assigned C7(a) structure.

Further, considering we have experimental relative populations, we can predict and compare the populations based on the calculated relative Free energies using a partition function,

$$P_n = \frac{e^{-\frac{\Delta G_n}{kT}}}{\sum_n e^{-\frac{\Delta G_n}{kT}}} \qquad \text{Equation 4.3}$$

where P_n is the population of conformer n, $\Delta G_n$is the relative energy of conformer n from the global minimum, k is the Boltzmann constant, and T is the temperature. Given the level of theory used to calculate the energies as well as some population possibly undergoing isomerization while cooling, this provides only a rough estimate at best. The partition function using the relative Free energies predicts populations of 30% C7(a), 32% C9(g-), 13% C9(a), 11% C7(g-), 3% C7(g+). Given the potential error in theoretical energy and the fact that cooling could result in isomerization or kinetic trapping, these populations are in a respectable agreement with our experimental populations, but importantly show that if the populations of the C7(g-) and C7(g+) conformers are below 11%. There is a high probability that we may not have been able to detect them in our experiment, given the low signal abundance of the C9(a) conformer (refer to Supplemental Figure 4.4 for the to see the relatively low R2PI signal intensity of conformer D).We can also consider the Free energies of the S(g+) and S(g-). The large drop in free energy of the C7(a) conformer likely dominates between the phenyl rotors, the S(g+) and S(g-). The S(g-) conformer soars to over 9 kJ/mol in relative energy and a predicted population of 3%, so it likely isn't observed, but the S(g-) conformer could still be observable. It is possible that with higher sensitivity, the S(g+) conformer could have been observed but was missed in this study.

The large change in relative energy between the conformers when considering the Free energy or the zero-point corrected energy is no intriguing, especially when considering that there are really two conformational families composed of three phenyl rotors each. The entropic contribution to Free energy is typically more temperature sensitive than the enthalpic contribution. Qualitatively, we propose that this change can be accounted for by the rigidity of the molecule: as the temperature increases a more rigid molecule will have less possible vibrational states available for population and therefore a lower entropy and higher Free energy. Moving the phenyl ring shifts the energy of the low frequency modes to lower energy, equivalent of being more flexible, such that these conformers entropically favored at higher temperatures. Looking to the energy of the vibrations below ~300 cm^{-1}, roughly the temperature AG4P was heated to and therefore the vibrational modes most accessible and able to affect free energy, we find that each conformer has the same number of vibrational modes below 300 cm^{-1}, suggesting a more subtle effect may be determining the energetic rearrangement. However, we do note that the lowest frequency vibration of each conformer, which contributes significantly to the possible states as temperature increases, agrees with our observations. That is, we find the ordering of the lowest energy vibration to be C7(*a*)<C9(*g*+)<C9(*g*-)<C9(*a*). Exploring this further, we use the vibrational partition function to track the number of states accessible at 450K using

$$q_v = \prod_N \frac{1}{1-e^{\frac{\tilde{\nu}_N}{k_B T}}} \qquad \text{Equation 4.4}$$

where q_v is the vibrational partition function and $\tilde{\nu}_N$ is the frequency of the vibrational mode N, using all calculated 108 vibrational modes. The output of q_v is on the order of $1x10^{22}$ and therefore has been normalized (represented $\bar{q}_v$) to simplify interpretation. The larger the normalized value of $\bar{q}_v$, the more we expect the conformer to have a lower Free energy at higher temperatures due to the flexibility of the conformer. The normalized values are presented alongside the free energies in Figure 4.6b. Clearly, we see that the C7(*a*) conformer ($\bar{q}_v$=1.0) has the largest number of populatable states at 450K, followed by the C9(*g*-) ($\bar{q}_v$=0.33), C9(*a*) ($\bar{q}_v$=0.19), and C9(*g*+) ($\bar{q}_v$=0.06), respectively, which is agreement with the calculated Free energies.

Table 4.2. Table of dihedral angles of indicated AG4P/AG2P conformers.[39]

Structure	ϕ	θ_1	θ_2	ψ
AG4P C9	101	-69	-74	104
AG4P C7(a)	-154	63	-73	147
AG2P S(a)	-102	55	-76	136
AG2P C9	99	-70	-73	105

4.4.2 Comparison of Stacked AG2P Structures and Populations to AG4P

First, we would like to direct the reader's attention to the fact that there are very similar structures adopted by both the AG4P and AG2P structural isomers. To that end, we will introduce AG4P or AG2P before the assigned structure to help guide the reader and prevent confusion. In the AG2P studies, two C9 structures, the AG2P C9(*a*) and AG2P C9(*g*-) structures, were assigned. The C9 structures have nearly identical dihedral angles and hydrogen bonding characteristics, suggesting that the location of the phenyl group along the γ-backbone does not significantly influence the ability of the AG4P to adopt the C9 structure. This result is unsurprising, as in the C9 structures, the phenyl ring is oriented such that it is pointing away from the backbone and not able to interact much with the hydrogen bonding network, though we do spectroscopically observe weak NH-π interactions in one of the three C9 structures that differ from AG2P, likely resulting in differences in observed populations and relative energies. That is, in the AG2P molecule, the AG2P C7(*g*-) and AG2P C7(*a*) structure are found within 0.3 kJ/mol of each other, while the AG4P analogs are found within 0.9 kJ/mol of each other. This difference in energy is slight but worth noting and is likely a result of the phenyl ring interacting both at a shorter distance and at a different angle to the amide NH and CO. While this interaction does not have significant impact on the C9 structures due to C9 hydrogen bond naturally orienting the phenyl ring away from the amide groups it will be important when considering the C7/stacked conformer, and discuss that further in the C9 structures now.

When we consider the relative populations of the conformers in AG4P and AG2P, the C9(*g*+) conformer becomes interesting to consider. While in low abundance, the AG4P C9(*g*+) conformer is experimentally visible, it is not found in experimentally as a conformer of AG2P. The phenyl ring has dictated the majority differences between conformers and structural isomers determined

so far, so we compared how the phenyl ring is interacting in the C9($g+$) structural isomers. What we find in AG4P is that the $g+$ phenyl rotor dihedral places the benzene ring above the π-cloud of the nearest acyl group, which is confirmed by observing a shift to lower frequency of the non-hydrogen bound amide NH. In AG2P, noting the minimal change in the dihedral angles, the benzene ring is similarly over the nearest amide group, though this time interacting directly with the acyl oxygen, likely destabilizing the conformer enough that the population would not be experimentally observable. That is, the phenyl ring interacts with the amide from the opposite side, switching the interaction of the phenyl π-cloud from an NH group in AG4P to an acyl C=O (AG2P), which seem s to be highly unfavorable, as shown in other similar studies with phenyl rotors.[44] Though the structural details of the AG2P C9(g_+), the authors of the previous work do find the lowest energy AG2P C9($g+$) structure at least 7 kJ/mol above the global minimum,[39] in agreement with our interpretation of the structure.

The remaining AG4P C7(a) conformer has no direct analog to compare to AG2P, but it does appear to have be a close analog of the AG2P S(a) conformer at a cursory glance; the molecule folds tightly in the middle rather than adopting a more ring-like structure as seen in the C9 structures. This is verified by examining the dihedral angles, of which three of the four are nearly identical, shown in Table 2. The only dihedral angle with significant difference is the ψ dihedral, which indicates that only that one dihedral must change in order to convert the between amide stacked or C7 structures, as found by James in his study of Ac-γ^2-Phe-NMe_2. Considering that in both AG4P and AG2P, the molecule is folded, almost pinched in the center, maintaining a phenyl rotor dihedral angle of 180° places the benzene ring pointing out and away from the backbone in AG4P and over the amide stack, in the case of AG2P. The fold occurs at the 3 position, so this is a natural consequence in comparing the γ^2 peptide to the γ^4 peptide. The AG2P phenyl ring being oriented above the amide π-cloud could provide stability to the existing stacking motif, which promotes formation of amide stack. The contrast in this observation against the instability of the AG2P C9($g+$) conformer results from which part of the amide acyl group is interacting with the ring, as mentioned above. In the AG2P C9($g+$) conformer, rotation of the ring 120° forces the benzene ring to be in close proximity, but not stacked with the amide oxygen, and therefore is electrostatically destabilized.

Now that all the structural, conformational, and energetic differences between the two structural isomers have been elaborated upon, the difference in relative populations we see

between AG2P and AG4P can be accounted for, assuming that the AG4P C7(*a*) and AG2P S(*a*) structures are substitutable (which seems reasonable give their close dihedral angles). The C9(*g*-) conformer is the dominant conformer in both cases, accounting for 53% of the population in AG4P and 41% of the population in AG2P. However, the position of the phenyl ring reorders all the remaining populations, with the second most abundant conformer in AG4P, the C7(*a*), accounting for 27% of the remaining population and the AG2P C9(*a*) structure accounting for 38% of the remaining population in AG2P. Though we do not have access to the Free energies at the temperature AG2P was heated in order to volatilize the sample, we surmise that in AG4P, C7(*a*), the phenyl ring is oriented away from the backbone in AG4P, allowing the molecule to be more flexible and therefore the AG4P C7(*a*) structure is more entropically favored at experimental temperatures, relatively, than both the AG2P S(*a*) analog and the AG2P C9(*a*) structure. Interestingly, the C9(*g*+) is the next most abundant conformer of AG4P, accounting for 13% of the population, while not being observed in the AG2P structural isomer. Recall the AG2P C9(*g*+) conformer is likely unobserved due to the orientation of the phenyl ring in AG2P C9(*g*+) relative to the nearest amide oxygen. Lastly, the AG4P C9(*a*) structure accounts for the remaining 7% of the population. The fact that in AG2P the C9(*a*) conformer is the second most abundant, but the AG4P C9(*a*) conformer is the least abundant can be accounted for by considering the C9(*g*+) conformer is not present in AG2P, and population funnels into that conformer.

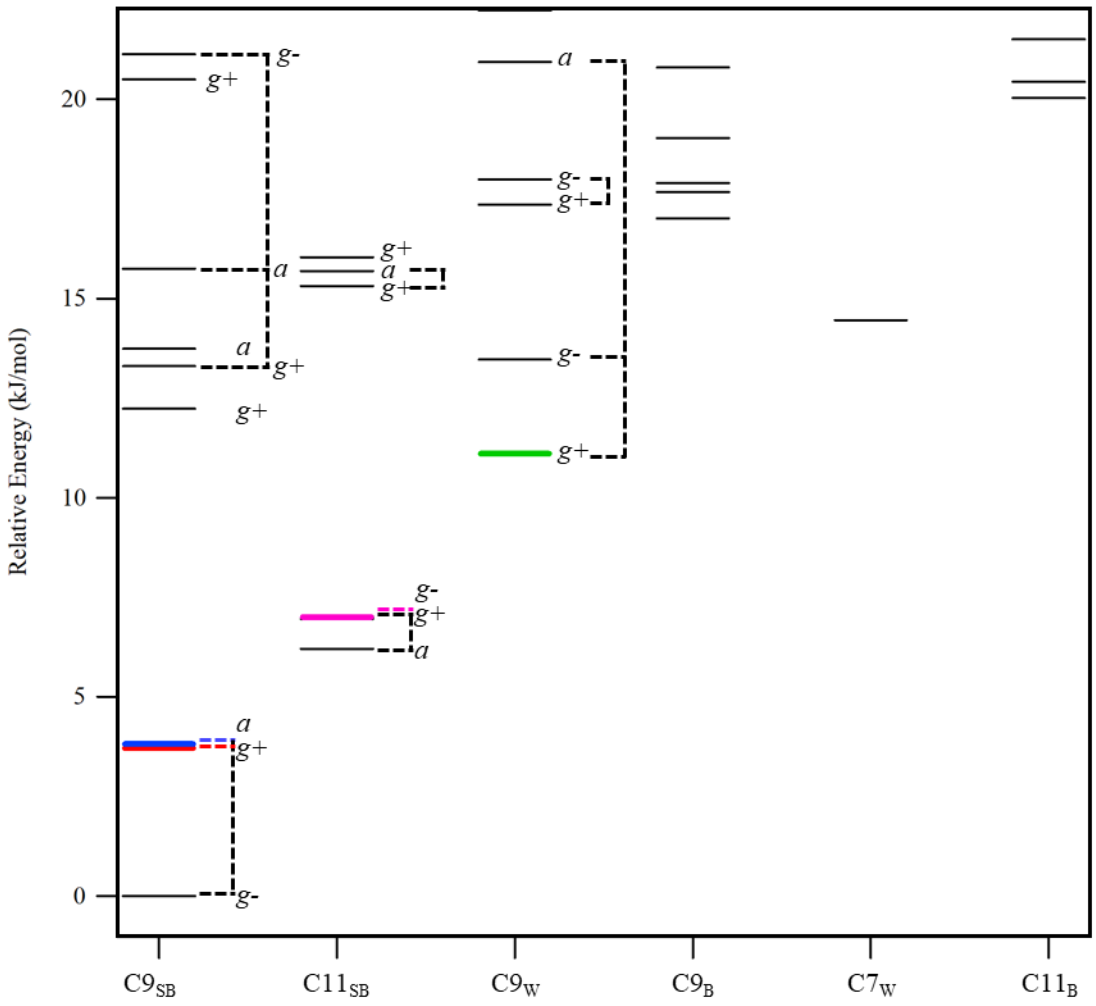


Figure 4.7. Potential energy diagram for the potential AG4P•H_2O conformers. Columns are families determined by the hydrogen bonding patterns observed. Dashed lines are used to connect phenyl rotamers of conformers with all other dihedral angles similar. The colored lines/dashed lines indicate where the assigned conformers fall on the potential energy landscape, matching the colors presented in Figure 4.6 as well as Supplemental Figure 4.6. Collection of all predicted stick spectra grouped by hydrogen bonding family, as discussed in the main text. The left side of each stick spectrum has both the relative energy to the global minimum as well as the orientation of the phenyl rotor, indicated by either a *g+*, *g-*, or *a*. The bolded, colored stick spectra are assigned to experimental spectra, with the color being consistent with the main text. The actual structure can be determined from the family and the orientation of the phenyl ring presented in the figure.

and Table S3.

4.4.3 AG4P•H_2O Cluster

As stated above, the large degree of spectral overlap in the infrared spectrum of AG4P•H_2O cluster spectra complicates interpretation and we have made only tentative assignments. We do think we have enough data to instead assign all of the experimentally observed conformers into families confidently. The supplemental information contains a complete library of calculated spectra grouped by family, Supplemental Figure 4.6, if the reader is interested. That being said, our calculations found that the lowest energy families were structures where the water bridges across the amide groups terminating the molecule, forming either C11 or C9 structures in amide stacking configurations. This is particularly of interest because the studies on the AG4P monomer showed no signs of amide stacking, showing the adduction of water generates new conformations as a preferred low-energy motif rather than just binding to one of the low-energy monomer conformations present in the expansion.

Looking to the potential energy diagram presented in Figure 4.7, we see that the lowest energy structure, the $C9_{SB}(g\text{-})$ is not assigned to an experimentally observed conformer. The $C9_{SB}(g\text{-})$ conformer has one water OH bridging the amide stack and the other water OH pointed into the center of the phenyl ring, forming an OH–π bond. The OH–π bond should shift the frequency of the oscillator to a lower frequency around 3660 cm^{-1} according to the calculated stick spectrum for the $C9_{SB}(g\text{-})$ structure, but clearly there is no experimental evidence that such a transition exists in any spectra. This is puzzling, as positioning the water in such a way such that both OH oscillators are participating in hydrogen bonding is the intuitive (and calculated) way to generate a low energy structure, but we do not observe it experimentally. We cautiously suggest that this conformer must be difficult to form in the cooling/adduction process, as there is a conformational change induced by the adduction of water, but have no evidence support what kind of mechanisms may be responsible. The second and third lowest energy structures assigned to $C9_{SB}(g+)$ and $C9_{SB}(a)$, found 3.7 and 3.8 kJ, respectively, which is reassuring given the lack of the lowest energy $C9_{SB}(g\text{-})$ conformer.

The next lowest energy family is the $C11_{SB}$ family, where the water bridges an amide stack and the strained hydrogen bond occurs through puckering of the preterminal amide NH and C=O. The conformational assignment of the $C11_{SB}$ structure is particularly difficult due to what we

interpret as the mixing and splitting of two nearly isoenergetic transitions. This leads us to look for transitions that have nearly isoenergetic energies to explain the intensity pattern of the two peaks. The best fit spectrum belongs to the third highest energy $C11_{SB}$ conformer in the phenyl rotamer family identified in Figure 4.7 (with the predicted peaks separated by only 2 cm^{-1}). However, the best fit stick spectrum for the lowest energy conformer, the $C11_{SB}(a)$ structure does provide a reasonable match in terms of frequency, but not intensity, with the coupled transitions separated by $17cm^{-1}$. The conflict of intensity, frequency an energy presents a dilemma for making an absolute assignment, but we have decided to favor the assigned structure, $C11_{SB}(g-)$, over the lowest energy structure in the family, $C11_{SB}(a)$, due to the *g*- conformer having a better match of intensity and frequency coupled with the fact that all three structures are within 0.8 kJ/mol in relative energy, a modest difference at our current level of theory.

The last experimentally observed conformer is the $C9_W(g+)$ conformer, which is the lowest energy conformer of the $C9_W$ family, and the next lowest energy family. This structure is nearly identical to the C9 conformer assigned to the AG4P monomer, with the water orienting adducting at the C-terminal amide C=O. Additionally, the water has arranged itself so that a hydrogen bond could occur between the water O and one of the phenyl hydrogens, though this should be a weakly stabilizing interaction, particularly at the distance of 2.6 Å.

The prevalence of multiple stacked, bridged structures instead of adducts to the naturally existing C7 and C9 structures found in the monomer is interesting, but has been explored in the previous work on AG2P.[41] We can reinforce and further the understanding reached previously with our current observations. In the previous work on AG2P•H_2O, it was found that the binding energy of water in the amide stacked conformer was 15 kJ/mol greater than the binding energy of the $C9_W$ adducted conformer.[41] We observe a very similar trend (the BSSE corrected binding energies can be found in Table S4), showing the binding energy is roughly 11 kJ/mol higher for all amide stacking conformer than for the $C9_W$ conformer. Buchanan proposed that this high degree of stabilization of the water-bridged amide stack structures meant that the presence of the $C9_W$ structure was a result of preferential addition of H_2O to the highly abundant C9 structure rather than the stability of the final product. In our experiment, where little to no stacked structures are present, the majority of the amide stacked structures must be generated from a conformational change upon water adduction, validating Buchannan's proposal as that is the only possible route for the formation of the amide stacked structures. Additionally, the $C9_W$ structure is predicted to

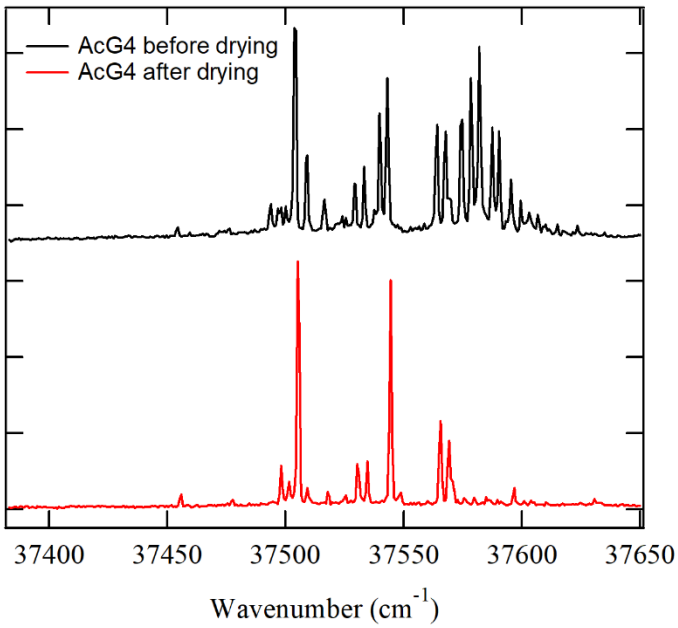
be 11 kJ/mol higher in energy than he global minimum. Seeing such a high energy structure provides even more support that conformers survive from the adduction to the population.

4.5 Conclusion

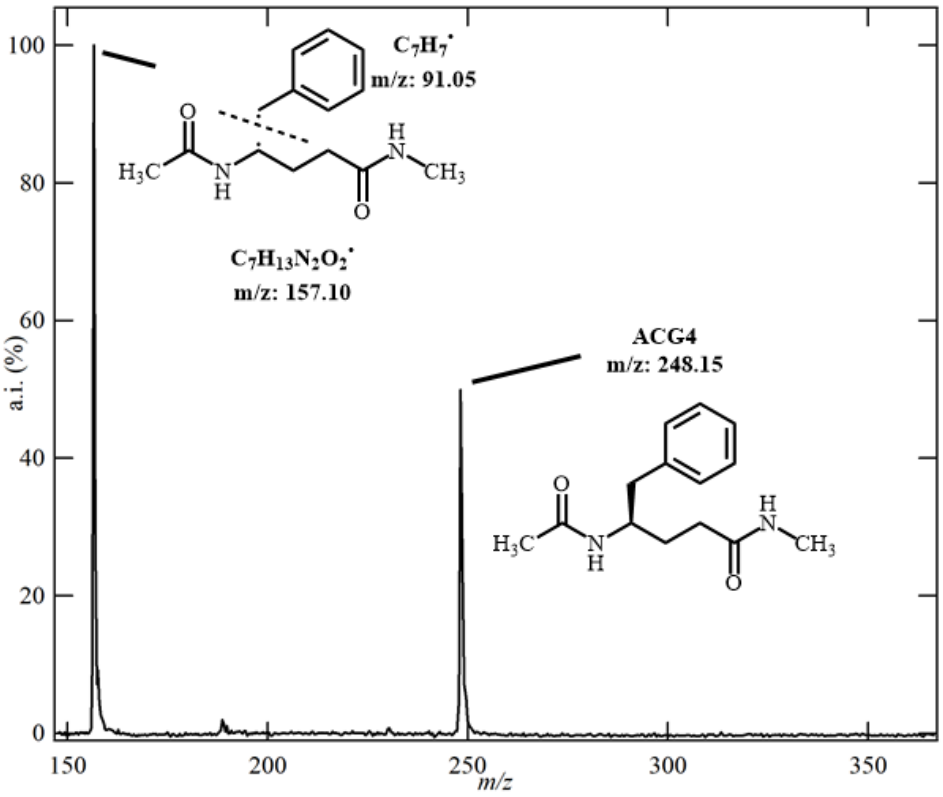
AG2P was shown, and extensively studied, in the gas-phase previously, showing not only the existence of a unique amide stacking motif, but also that various factors can determine the relative stability of amide stacking to more traditional hydrogen bonding motifs. Here, we explored the effect of moving the phenyl ring along the gamma peptides backbone, from the 2 position, to the 4 position and showed that AG4P behaves differently than AG2P through single-conformer spectroscopy and IRPT spectroscopy. Namely, a unique C7 structure was found not observed in AG2P, but no amide stacking was observed. Additionally, all three phenyl rotamers of a C9 structure were observed. Interestingly, no phenyl rotamers of the C7 structure were observed. This was attributed to high isomerization barriers between the rotamers in conjunction as well as there being a low population of the unobserved C7 rotamers at the experimental temperature before cooling. Cooling resulted in kinetic trapping and therefore no C7 rotamers. This claim is qualitatively supported by matching the relative abundances of population determined experimentally to populations predicted by a Boltzmann distribution. The large difference in energies between the zero-point corrected energies commonly used to evaluate conformers and the free energies is explained by considering the effect a more flexible molecule (lower force constants, lower frequency vibrations) has on the vibrational partition function and therefore Free energy.

Additionally, the adduction of a single water to AG4P was performed, bringing amide stacking structures in two different families seemingly large relative abundances, though IRPT was unable to be performed within the dynamic range of our instrument. A high degree of spectral overlap hinders absolute assignments of conformers, but families of conformers were tentatively assigned, including two separate amide stacking families as well as water adducted to the C9 structure. The presence of amide stacked structures are the product of isomerization after water adducts to one of the C7 or C9 structures of the AG4P monomer present in the expansion, and show a surprising 15 kJ/mol more stabilization through binding energy than the C7 or C9 water clusters that do not undergo isomerization.

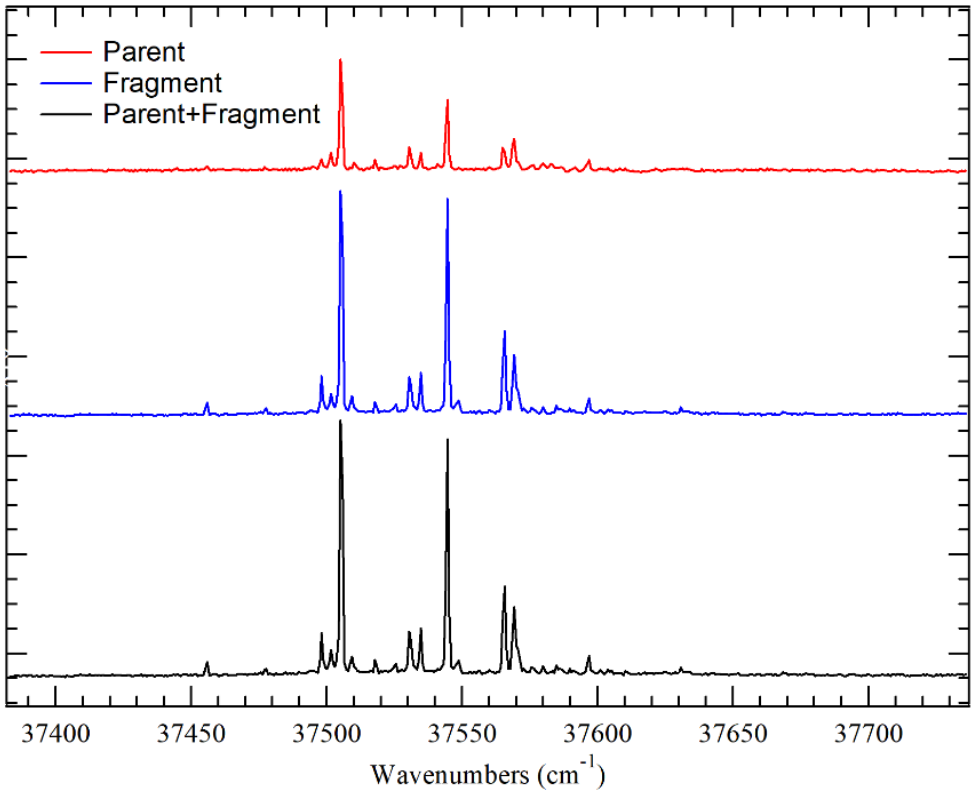
4.6 Supplemental Figures



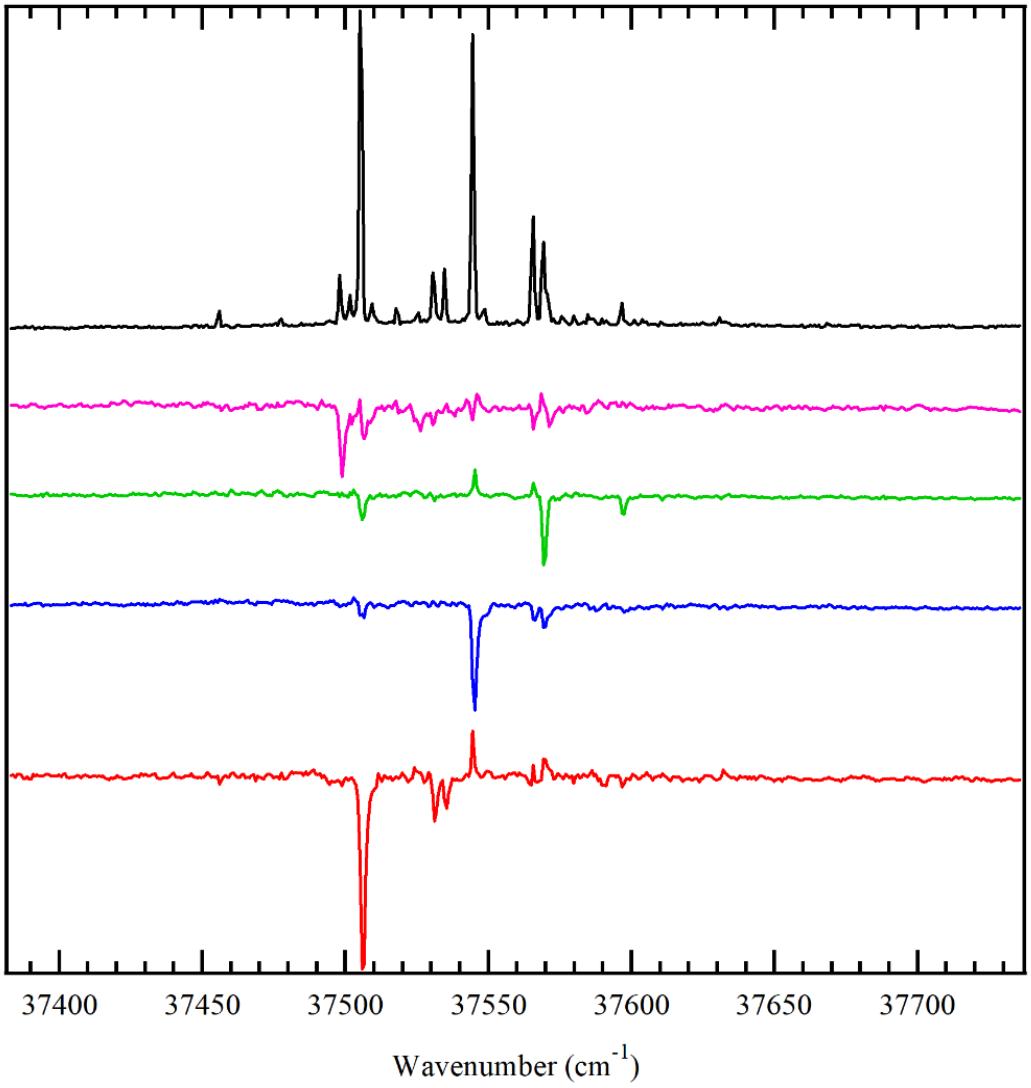
Supplemental Figure 4.1. Shown is black is the R2PI spectrum of AG4P before drying in a lyophilizer. Shown in red is the R2PI spectrum of AG4P after lyophilizing.



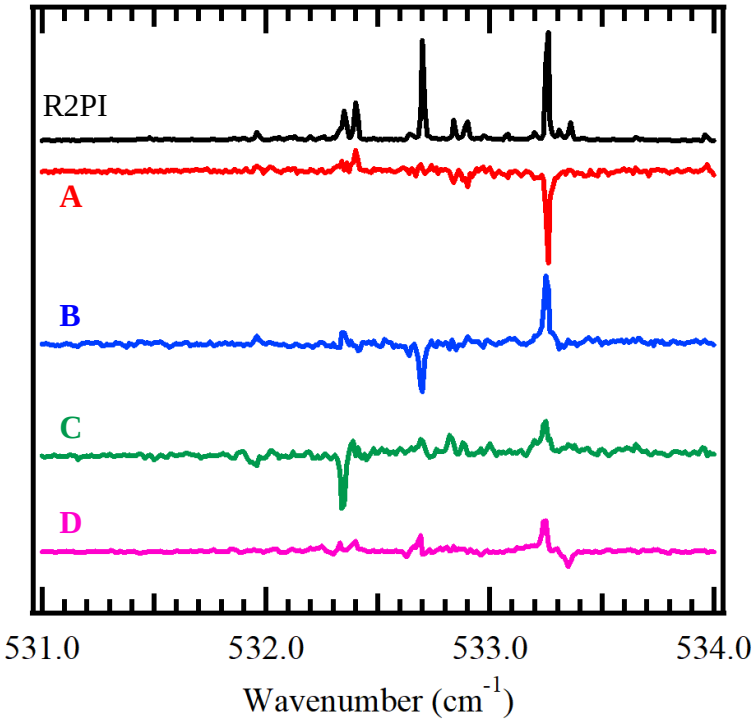
Supplemental Figure 4.2. Mass spectrum produced with the UV laser frequency set to 37505 cm^{-1}, showing the presence of the parent at m/z 248.2 and the photodissociated fragment resulting from the loss of the benzyl group at 157.1 m/z.



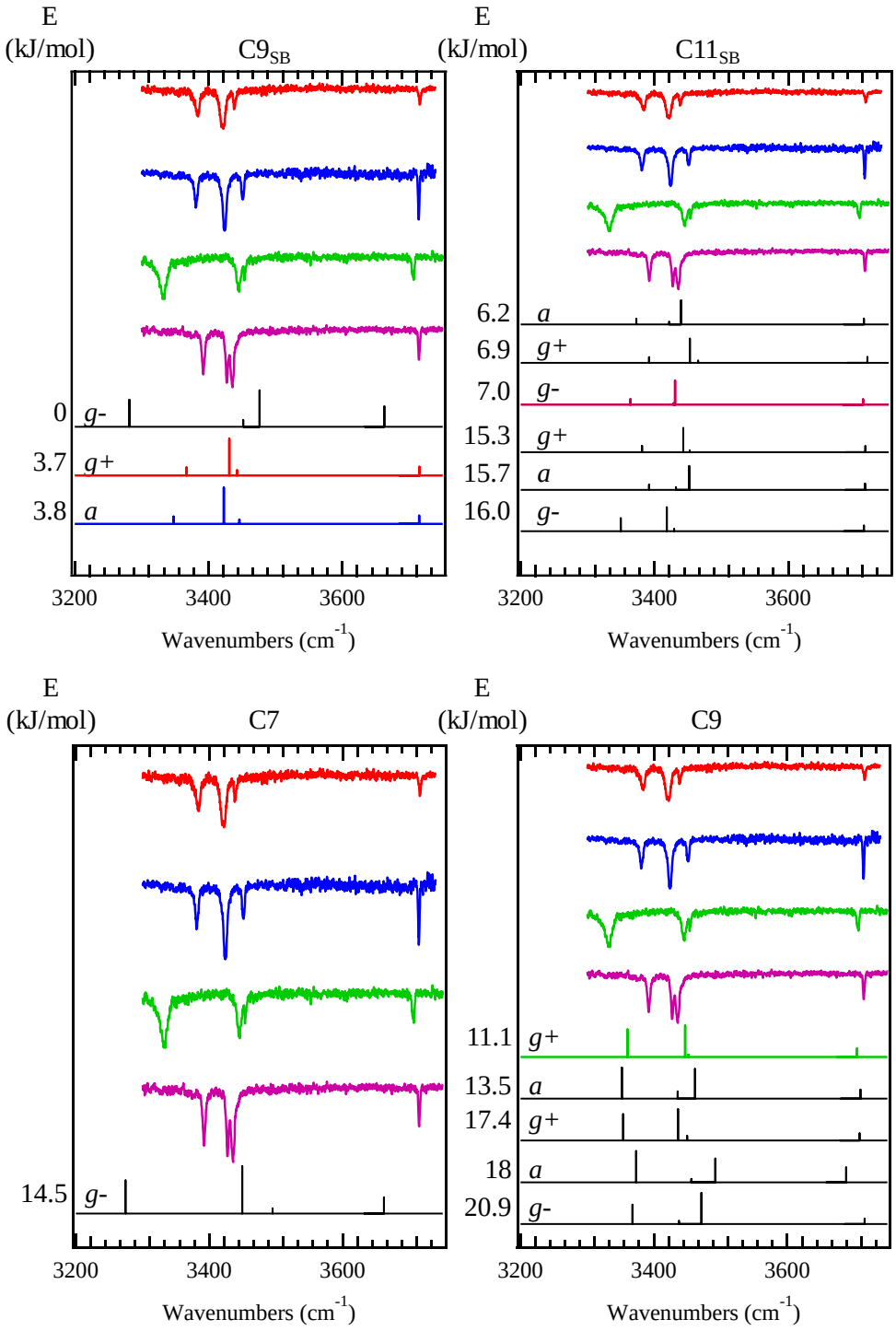
Supplemental Figure 4.3. R2PI spectra of *top, red*) AG4P parent, *middle, blue*) AG4P benzyl loss fragment, and *bottom, black*) the sum of the two spectra.



Supplemental Figure 4.4. *Top, black*) R2PI and *bottom, colored*) IRUVHF spectra of AG4P. Each holeburn is labeled with which conformer it belongs to on the left side of the spectra. Readers are referred to Figure 4.3b of the main text to see conformers.

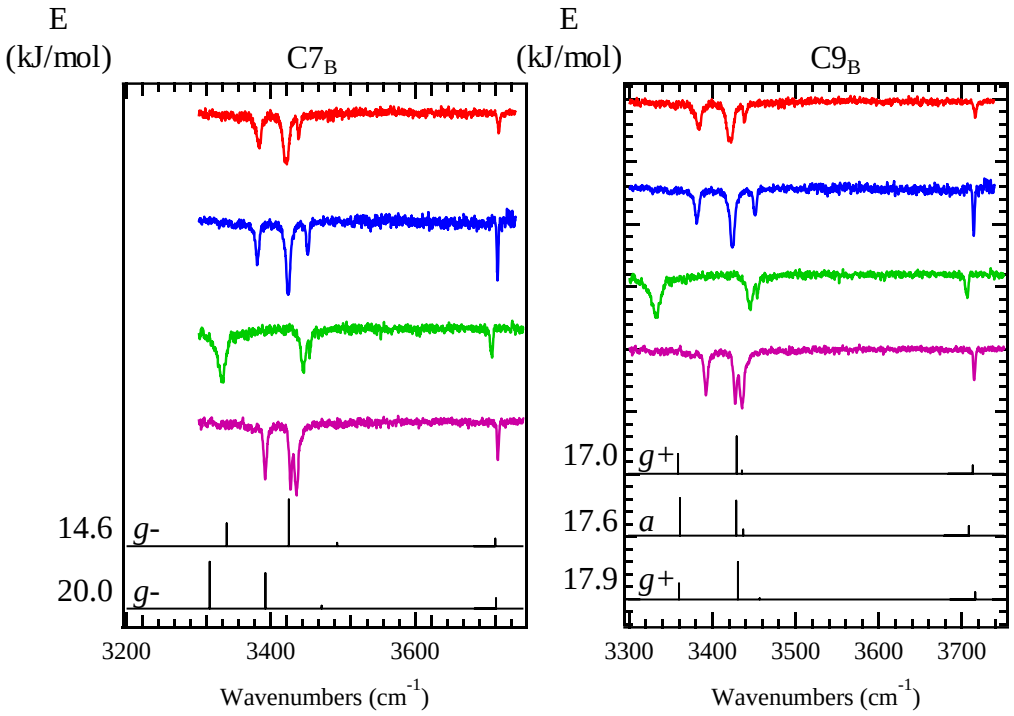


Supplemental Figure 4.5. *Top, black*) R2PI spectrum and *colored, bottom*) IRUVHF spectra of AG4P. Each hole fill is labelled with the corresponding conformer on the left side of the spectra. Readers are referred to Figure 4.4 for corresponding IRPT spectra.

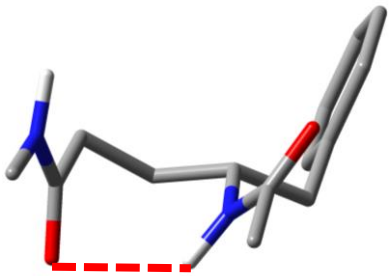


Supplemental Figure 4.6. Collection of all predicted stick spectra grouped by hydrogen bonding family, as discussed in the main text. The left side of each stick spectrum has both the relative energy to the global minimum as well as the orientation of the phenyl rotor, indicated by either a *g*+, *g*-, or *a*. The bolded, colored stick spectra are assigned to experimental spectra, with the color being consistent with the main text. The actual structure can be determined from the family and the orientation of the phenyl ring presented in the figure.

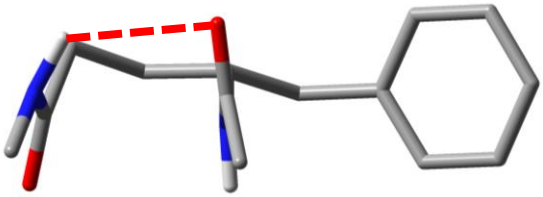
Supplemental Figure 4.6. Continued on next page



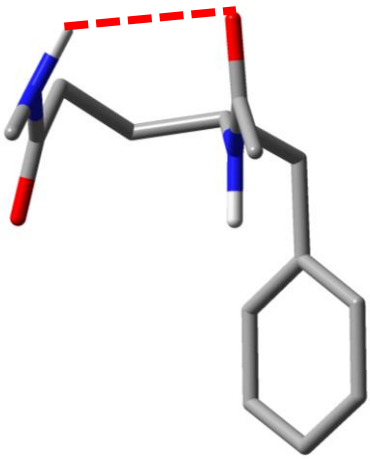
C7(a)



S(*g*+)



S(*g*-)



Supplemental Figure 4.7. Potential energy minimum found in phenyl rotor dihedral scan of the C7(*a*) conformer of AG4P.

4.7 References

1. Dill, K. A., *Biochem.* **1990,** *29*, 7133-55.

2. Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R., *Annu. Rev. Biophys.* **2008,** *37*, 289-316.

3. Nick Pace, C.; Scholtz, J. M.; Grimsley, G. R., *FEBS Lett.* **2014,** *588*, 2177-2184.

4. Eisenberg, D., *Proc. Natl. Acad. Sci.* **2003,** *100*, 11207-11210.

5. Anderson, C. A. F.; Rost, B., Structural Bioinformatics. In *Structural Bioinformatics*, 2 ed.; Gu, J.; Weissig, H., Eds. John Wiley & Sons, Inc.: Hoboken, 2003; pp 339-363.

6. Cable, J.; Tubergen, M. J.; Levy, D. H., *J. Am. Chem. Soc.* **1988,** *110*, 7349-7355.

7. Cable, J.; Tubergen, M. J.; Levy, D. H., *J. Am. Chem. Soc.* **1989,** *111*, 9032-9039.

8. Rijs, A. M.; Oomens, J., *Gas-phase IR spectroscopy and structure of biological molecules*. Springer: 2015; Vol. 364.

9. Alauddin, M.; Biswal, H.; Gloaguen, E.; Mons, M., *Intra-residue interactions in proteins: interplay between Serine or Cysteine side chains and backbone conformations, revealed by laser spectroscopy of isolated model peptides*. 2014; Vol. 17.

10. Chin, W.; Piuzzi, F.; Dognon, J.-P.; Dimicoli, I.; Tardivel, B.; Mons, M., *Journal of the American Chemical Society* **2005,** *127*, 11900-11901.

11. Dian, B. C.; Longarte, A.; Zwier, T. S., *Science* **2002,** *296*, 2369-2373.

12. Dian, B. C.; Longarte, A.; Winter, P. R.; Zwier, T. S., *J. Chem. Phys.* **2004,** *120*, 133-147.

13. James, W. H.; Baquero, E. E.; Choi, S. H.; Gellman, S. H.; Zwier, T. S., *J. Phys. Chem. A* **2010,** *114*, 1581-1591.

14. Dean, J. C.; Buchanan, E. G.; Zwier, T. S., *J. Am. Chem. Soc.* **2012,** *134*, 17186-17201.

15. Kusaka, R.; Zhang, D.; Walsh, P. S.; Gord, J. R.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S., *J. Phys. Chem. A* **2013,** *117*, 10847-10862.

16. Gord, J. R.; Walsh, P. S.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S., *J. Phys. Chem. B* **2014,** *118*, 8246-8256.

17. Gord, J. R.; Hewett, D. M.; Hernandez-Castillo, A. O.; Blodgett, K. N.; Rotondaro, M. C.; Varuolo, A.; Kubasik, M. A.; Zwier, T. S., *Phys. Chem. Chem. Phys.* **2016,** *18*, 25512-25527.

18. Walsh, P. S.; Dean, J. C.; McBurney, C.; Kang, H.; Gellman, S. H.; Zwier, T. S., *Physical Chemistry Chemical Physics* **2016,** *18*, 11306-11322.

19. Blodgett, K. N.; Zhu, X.; Walsh, P. S.; Sun, D.; Lee, J.; Choi, S. H.; Zwier, T. S., *J. Phys. Chem. A* **2018,** *122*, 3697-3710.

20. Fricke, H.; Funk, A.; Schrader, T.; Gerhards, M., *J. Am. Chem. Soc.* **2008,** *130*, 4692-4698.

21. Lee, J. J.; Albrecht, M.; Rice, C. A.; Suhm, M. A.; Stamm, A.; Zimmer, M.; Gerhards, M., *J. Phys. Chem. A* **2013,** *117*, 7050-7063.

22. Schwing, K.; Reyheller, C.; Schaly, A.; Kubik, S.; Gerhards, M., *Chem. Phys. Phys. Chem.* **2011,** *12*, 1981-1988.

23. Řeha, D.; Valdes, H.; Vondrášek, J.; Hobza, P.; Abu-Riziq, A.; Crews, B.; de Vries, M. S., *Chem. Eur. J.* **2005,** *11*, 6803-6817.

24. Rijs, A. M.; Kabeláč, M.; Abo-Riziq, A.; Hobza, P.; de Vries, M. S., *Chem. Phys. Phys. Chem.* **2011,** *12*, 1816-1821.

25. Barry, C. S.; Cocinero, E. J.; Çarçabal, P.; Gamblin, D. P.; Stanca-Kaposta, E. C.; Remmert, S. M.; Fernández-Alonso, M. C.; Rudić, S.; Simons, J. P.; Davis, B. G., *J. Am. Chem. Soc.* **2013,** *135*, 16895-16903.

26. Stanca-Kaposta, E. C.; Çarçabal, P.; Cocinero, E. J.; Hurtado, P.; Simons, J. P., *J. Phys. Chem. B* **2013,** *117*, 8135-8142.

27. Gellman, S. H., *Acc. Chem. Res.* **1998,** *31*, 173-180.

28. Horne, W. S.; Gellman, S. H., *Acc. Chem. Res.* **2008,** *41*, 1399-1408.

29. Seebach, D.; Gardiner, J., *Acc. Chem. Res.* **2008,** *41*, 1366-1375.

30. Choudhary, A.; Gandla, D.; Krow, G. R.; Raines, R. T., *J. Am. Chem. Soc.* **2009,** *131*, 7244-7246.

31. Yonezawa, T.; Morishim, I. B., *Chem. Soc. Jpn.* **1966,** *39*, 2346.

32. Rabinovitz, M.; Pines, A., *Journal of the Chemical Society B: Physical Organic* **1968**, 1110-1111.

33. Betson, M. S.; Clayden, J.; Lam, H. K.; Helliwell, M., *Angew. Chem.* **2005,** *44*, 1241-1244.

34. Clayden, J., *Chem. Soc. Rev.* **2009,** *38*, 817.

35. Maccallum, P. H.; Poet, R.; Milnerwhite, E. J., *J. Mol. Biol.* **1995,** *248*, 361.

36. Maccallum, P. H.; Poet, R.; Milnerwhite, E. J., *J. Mol. Biol.* **1995,** *248*, 374.

37. Hinderaker, M. P.; Raines, R. T., *Protein Sci.* **2003,** *12*, 1188.

38. Fischer, F. R.; Wood, P. A.; Allen, F. H.; Diederich, F. P., *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105*, 17290.

39. James III, W. H.; Müller, C. W.; Buchanan, E. G.; Nix, M. G.; Guo, L.; Roskop, L.; Gordon, M. S.; Slipchenko, L. V.; Gellman, S. H.; Zwier, T. S., *J. Am. Chem. Soc.* **2009,** *131*, 14243-14245.

40. James III, W. H.; Buchanan, E. G.; Guo, L.; Gellman, S. H.; Zwier, T. S., *J. Phys. Chem. A* **2011,** 115, 11960-11970.

41. Buchanan, E. G.; James III, W. H.; Gutberlet, A.; Dean, J. C.; Guo, L.; Gellman, S. H.; Zwier, T. S., *Faraday Discuss.* **2011,** *150*, 209-226.

42. Weiner, P. K.; Kollman, P. A., *J. Comp. Chem.* **1981,** *2*, 287-303.

43. Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16 Rev. B.01*, Wallingford, CT, 2016.

44. Fischer, J. L.; Elvir, B. R.; DeLucia, S.-A.; Blodgett, K. N.; Zeller, M.; Kubasik, M. A.; Zwier, T. S., *The Journal of Physical Chemistry A* **2019,** *123*, 4178-4187.

CHAPTER 5. KINETIC STUDY OF LEAVING GROUPS IN GAS-PHASE COVALENT MODIFICATION CLICK REACTIONS USING ISOMERS OF NHS TRIAZOLE ESTERS

5.1 Introduction

Tandem mass spectrometry is often used to determine the sequence of biologically relevant molecules, such as proteins and lipids, through fragmentation of a precursor ion into its substituent parts. Although complete structural coverage is rarely obtained through a single method and often complimentary methods are required to generate adequate sequence coverage.[1] Different fragmentation spectra are often obtained from different ion types, making ion/ion reactions a convenient method to change ion types and gather complimentary fragmentation spectra for sequencing.[2] Common examples of ion/ion reactions include proton transfer, which can be used to change the charge state of an ion,[3-7] metal cation transfer reactions, [8-11] and conversion of even-electron ions to odd-electron ions.[12, 13] Another commonly applied ion/ion chemistry is selective covalent modification, affording chemical modification to ions in the gas-phase.[14-18] Multiple gas-phase covalent modification reactions have been studied previously, including Schiff base reactions between amines and aldehydes,[19] Click chemistry between azides and amines or alkenes,[16] and oxidation reactions,[20] among others.

A commonly used solution-state reaction used to modify peptides and proteins is the nucleophilic substitution reaction between *N*-hydroxysuccinimide (NHS) esters and various nucleophiles.[21] The reaction mechanism proceeds through the nucleophilic attack of the NHS acyl group, displacing the NHS leaving group and forming a new amide bond between the tag and the nucleophile. This reaction has been applied successfully to gas-phase systems to selectively tag neutral amines[16] and guanidines[18] by our group as well as others.[22, 23] These studies were extended to the triazole-esters 1-hydroxy-7-aza-benzotriazole (HOAt) and 1-hydroxy-benzotriazole (HOBt), extending the reactivity to include neutral histadines.[17]

Covalent bond-forming gas-phase reactions are thought to proceed through a long-lived complex resulting from the condensation of opposite polarity ions. The electrostatic potential energy is transferred to the internal energy of the resulting ion/ion electrostatic complex. The lifetime of these complexes is sufficiently long that collisional cooling in the relatively high-pressure environment stabilizes the complex,[24] though if during the cooling process the

complexation energy is significantly greater than either the covalent bond-forming critical energy or the proton transfer critical energy, the respective products can be produced before cooling can occur.[17, 18] Surviving complex is then subjected to collisional activation that produces either covalent bond formation or proton transfer product, depending on the potential energy surface. Bu et al used DFT calculations to explore the potential energy surface of both the proton transfer and covalent modification pathways using model systems. These model systems were doubly protonated *tris*(3-aminopropyl)methane with deprotonated ethyl sulfonate (for proton transfer) and methyl amine reacted with NHS-acetate for covalent modification. The results showed that the relative critical energy of both pathways determines the relative yield.[18] Naturally, having a lower covalent modification critical energy relative to the proton transfer critical energy will improve the yield of covalent modification product.

When attempting to gain additional sequence information from gas-phase covalent modification reactions, converting the complex into proton transfer product is undesirable, especially if the peptide or protein of interest is of low abundance, which is often the case. To that end, Bu et al additionally showed that the covalent modification product yield using the triazole esterified tags instead of NHS esters is improved,[18] and further used DFT calculations to explore the potential energy surface and show that the covalent modification critical energy is decreased by 3 kcal/mol relative to NHS esters. Here, we attempt to further that achievement by using isomers of the 3-sulfobenzoic acid (3-SBA) tags used by Bu et al and acquiring direct dipolar current (DDC) activation kinetics to evaluate the kinetics of the isomers. By moving the sulfate group from *meta* position to the *para* position (e.g., 4-SBA), the electronic structure of the leaving group is altered, changing the critical energy of the covalent modification pathway. The applied DDC voltages are converted into effective temperatures using the Tolmachev approach,[25] allowing thermodynamic constants to be measured and inform the interpretation of the role of the different leaving group esterified tags.

One of the benefits of performing peptide tagging reactions of ions in the gas-phase is the control offered by mass isolation. However, the initial formation of the complex inherently leads to a large conversion of potential energy into internal energy (Figure 5.1), and depending on the potential energy surface and cooling rates, can drive the ion/ion reaction to either covalent modification (CMR), proton transfer (PT), or simply complex formation. The gas-phase ion/ion reaction can proceed through various competing channels, limiting the yield of the targeted

covalent modification product. Therefore, when performing gas-phase covalent modification ion/ion reactions, the competing process of proton transfer (Figure 5.1) must not only be considered, but ideally mitigated to maximize product signal. Inspired by previous studies from Bu et al,[17] we intend to modify the tagging reagents to explore their effects on the ion/ion reaction kinetics to inform future tagging choices. With regards to the SBA tag, which has been studied previously in detail, moving the location of the sulfate group from the *meta* position (3-SBA) to the *para* position (4-SBA) will change the electronic structure of the ester leaving group (either NHS or HOBt), potentially improving or hindering the production of the covalent modification product relative to the proton transfer product. Here, we perform reactions of both primary amines and arginines with these reactive leaving groups. By performing kinetic studies, we aim to use the obtained thermodynamic values to gain insight into the forces guiding these reactions to not only improve understanding of the reactions, but to guide future choices in tag selection. In this work, both the 3- and 4-isomer of SBA substituted with either NHS or HOBt will be reacted with the peptides $[KGAILKGAILR+2H]^{2+}$, $[RARARAA+2H]^{2+}$, and $[RKRARAA+2H]^{2+}$.

5.2 Experimental

5.2.1 Materials

Peptides KGAILKGAILR, RARARAA, and RKRARAA were obtained as standards from stock solutions of peptides were prepared in MeOH:H_2O (v/v, 50/50) in a concentration of 0.1 mg/mL or 10 μmol. Isomers of n-sulfobenzoic acid (n=3,4) were obtained from Sigma Aldrich as sodium salts and prepared as 10 mM solutions in dimethyl formamide. *N*-hydroxysuccinimide (NHS) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) were obtained from Thermo Scientific (Rockford, IL, USA) and prepared as 10 mM solutions. 1-hydroxybenzotriazole hydrate (HOBt•xH_2O) was purchased from Sigma Aldrich (St. Louis, MO, USA) and prepared as a 10 mM solution in DMF. To generate the reactive NHS/HOBt ester tag, a 1:1:1 mixture of n-SBA, NHS or HOBt, and EDC was prepared fresh before each experiment.

5.2.2 Mass Spectrometry

For simplicity, the tag will be referred to by the isomer, followed by the leaving group. For example, 3-SBA reacted with NHS or HOBt will simply be referred to as 3-SBANHS or 3-SBAOBt.

All experiments were performed on a modified TripleTOF 5600 quadrupole time-of-flight mass spectrometer (SCIEX, Concord, ON, Canada). Modification details are available in previous literature.[26] Briefly, the instrument was modified to facilitate a dual nano-electrospray (nESI) source, allowing for the sequential injections of cations and anions. Additional modifications to the pre-TOF quadrupole include the ability to apply a DDC potential to a pair of opposite rods, as well as an RF voltage to the trapping lenses to facilitate trapping of ions of both polarities in the axial dimension. Doubly protonated peptide cations were generated via nESI and mass selected in Q1, operated in transmission mode, as they traveled into q2 for mutual storage. While holding the positive ions in q2, singly charged deprotonated anions of n-SBANHS or n-SBAOBt were generated by nESI, mass selected as they travelled through Q1 into q2 and stored with the cations for 30 ms. The complex of the resulting IIRXN resulting from mutual storage of the cations and anions was then isolated by sequential resonance ejection ramps.[27] The isolated ions were then subjected to DDC activation (V_{RF}=1086V) by applying a DC voltage to a pair of opposing rods from 40 to 120 ms. The resulting ions were then injected into the TOF region to be mass analyzed.

5.2.3 Rate Measurements

Rate measurements of the IIRXN resulting from DDC activation have been discussed previously.[17, 18] In review, the dissociation of the precursor ion follows pseudo-first-order reaction kinetics, [28] as indicated in:

$$[M]_t = e^{-k_{diss}t} \qquad \text{Equation 5.1}$$

where $[M]_t$ is the molar fraction of the ion at time t, and k_{diss} is the dissociation rate constant. The molar fraction, $[M]_t$, is determined by dividing the intensity of the precursor signal by the

sum of the intensities of the precursor signal and all product ions, effectively making $[M]_t$ the survival yield.

By applying the Tolmachev's Effective Ion Temperature model, the effective temperature (T_{eff}) can be approximated, allowing a series of DDC voltages to be treated as a series of effective temperatures. Expressing rate constants as a function of temperature instead of voltage allows Arrhenius plots to be generated, and can provide insight into thermodynamic values afforded by the Arrhenius equation,

$$k(T) = Ae^{-E_a/RT} \qquad \text{Equation 5.2}$$

namely the critical energy, E_a, and the collision factor, A. It is worth noting that the physical interpretation of the A factor is ambiguous, but is often thought of as a scalar for the conversion of a collision between the reactants into products.

The Tolmachev model assumes that after an appropriate time, ions accelerated by a voltage will acquire an average internal energy proportional to their collision kinetics following a Boltzmann distribution. This internal energy will effectively be identical to ions stored to thermal equilibrium at a certain temperature. When considering fragmentation kinetics, the fragmentation rates must be slow relative to the energy exchange, otherwise the Boltzmann distribution necessary for considering thermal equilibrium will become truncated as the hotter ions fragment and leave the population, perturbing the Tolmachev interpretation. Rates where these conditions are met are considered to be within the rapid energy exchange limit.[25] Rapid energy exchange conditions are assumed (and required) to apply the Tolmachev model here. The Tolmachev model, after making the aforementioned assumptions, is fairly straightforward:

$$T_{eff} \approx T + \Delta T_k \qquad \text{Equation 5.3}$$

where T is the ambient temperature of the bath gas when the experiment was performed and ΔT_k is the increase in temperature resulting from DDC activation. The value of ΔT_k can be obtained using the following relationship:[25]

$$k\Delta T_k = \frac{m_g \omega^2 r_0^2}{24}\left(\frac{V_{DDC}}{V_{RF}}\right)^2 \qquad \text{Equation 5.4}$$

where k is the Boltzman constant, m_g is the mass of the bath gas, ω is the drive frequency of RF voltage applied to the quadrupole, r_0 is 5 half the distance between opposing rods of the quadrupole, V_{DDC} is the amplitude of the DDC voltage applied to opposing rods of the quadrupole, and V_{RF} is the 0-peak amplitude of the RF drive voltage.

5.3 Results and Discussion

5.3.1 $[KGAILKGAILR+2H]^{2+}$: Kinetics of Covalent Modification with $[n\text{-}SBAOBt\text{-}H]^-$ and $[n\text{-}SBANHS\text{-}H]^-$ -

Figure 5.3a shows the resulting ion/ion reaction between $[KGAILKGAILR+2H]^{2+}$ and $[3\text{-}SBAOBt\text{-}H]^-$. Upon simply converting potential energy into internal energy upon complex formation, 77% of total ion abundance belongs to the proton transfer product and 14% to covalently modified product, with the remaining 9% left existing in a complex, illustrating the need to mitigate proton transfer product production. The majority of potential signal is lost before the experiment can even begin. Figure 5.3c shows the same reaction and conditions as Figure 5.3a, with the exception that 4-SBAOBt has been substituted for 3-SBAOBt. An immediate difference is that no covalently modified product has been generated, with over 70% of the complex still being converted into proton transfer product, illustrating that simply changing the location of the sulfate group on the SBA tag has significantly altered the reaction trajectory. This is further illustrated by subjecting the resulting ion/ion complexes from each reaction to DDC activation. The products of 19 V DDC activation for 90 ms are shown for the complex of $[KGAILKGAILR+2H]^{2+}$ with $[3\text{-}SBAOBt\text{-}H]^-$ and $[4\text{-}SBAOBt\text{-}H]^-$ in Figure 5.3b and Figure 5.3d, respectively. Both reactions show the formation of covalently modified product exclusively, which is expected based on the 'stewing' conditions used.[17] However, the reaction rate with the two different anions is remarkably different, with 82% of the parent complex being converted into covalently modified product in the 3-SBAOBt reaction, and only 9% of the product being converted in the 4-SBAOBt reaction.

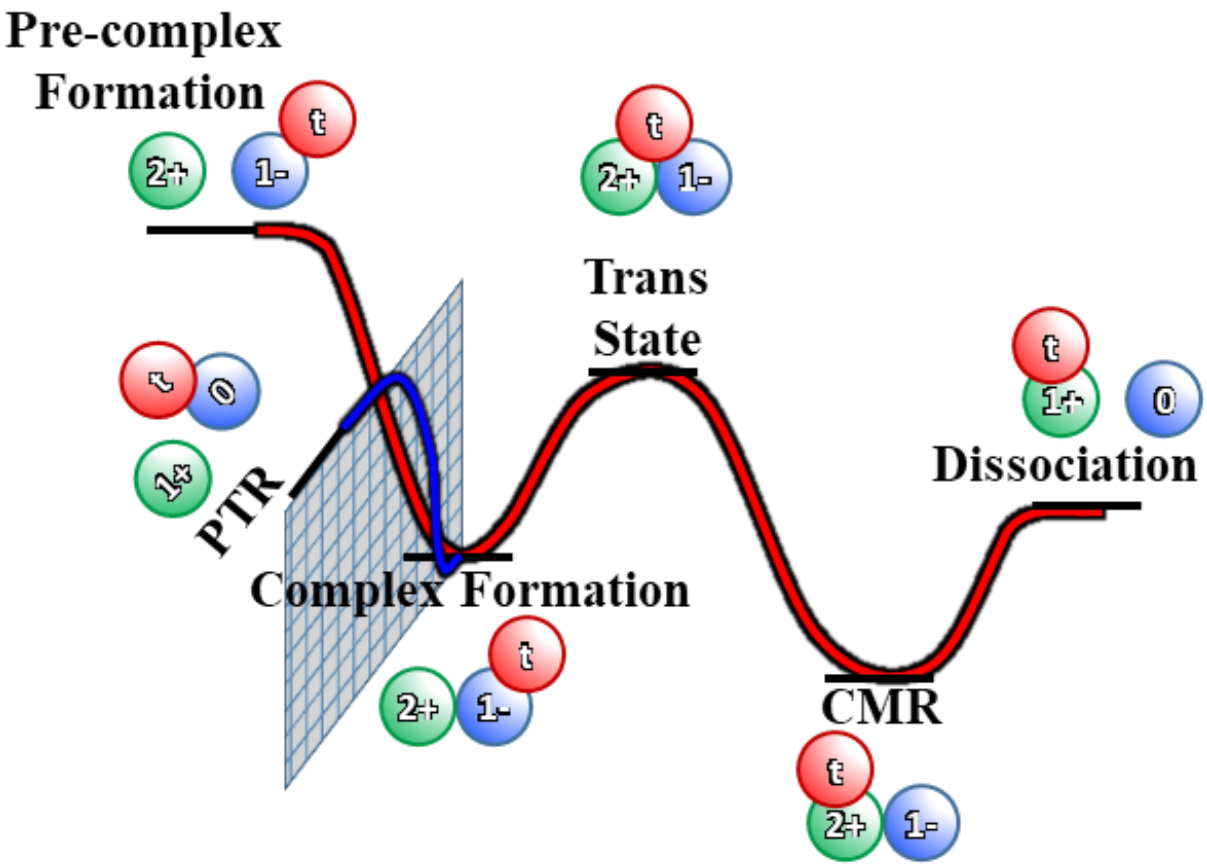


Figure 5.1. A generalized potential energy diagram for the reaction between a singly charged anion with a reactive electrophile and a doubly charged cation with a reactive nucleophile. The reactant is shown with a green circle, the tag with a red circle, and the anion with a blue circle. Relative energies are arbitrary. The gridded plane represents an 'out-of-plane' pathway.

Observing similar rates of conversion to proton transfer product upon complexation but dramatically differ rates of formation for covalently modified product upon DDC activation suggests that the subtle change of moving the location of the sulfate substitution induces a significant change on the potential energy surface of the reaction. This was further investigated by obtaining rate constants for the reactions at various voltages, shown in Figure 5.4a and Figure 5.4b. In agreement with the differing rate constants observed in Figure 5.3, the rate constants are drastically different. Activating of the complex with 16V DDC results in rate constant of 10.2 s^{-1} when using 3-SBAOBt, and a rate constant of 0.6 s^{-1} when using 4-SBAOBt. Insight into thermodynamic values such as activation energy provides insight into the potential energy landscape of the different reactions. Accordingly, a variety of rate constants were obtained using a variety of DDC activation energies, shown in Figure 5.3a. The nearly two-fold difference in

rate constants required higher activation energies to be used when acquiring survival yields from the 4-SBAOBt reaction, as at 19V DDC (compare Figure 5.3b and Figure 5.3d), the 3-SBAOBt complex had almost completely reacted while the majority of the 4-SBAOBt complex had not.

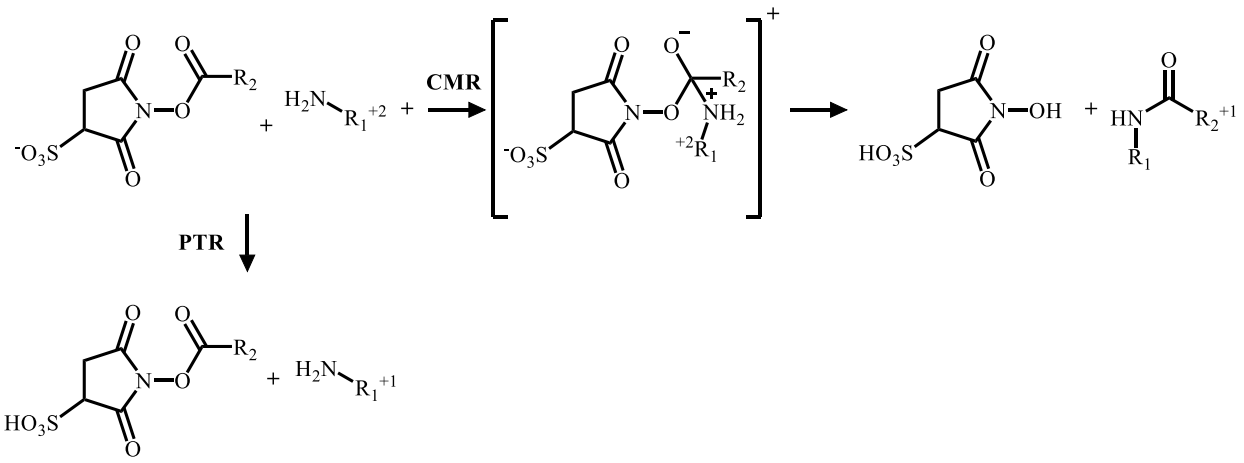


Figure 5.2. General mechanism of the possible pathways for the ion/ion reaction to proceed. From left to right is the CMR, from top to bottom is the proton transfer pathway (PTR).

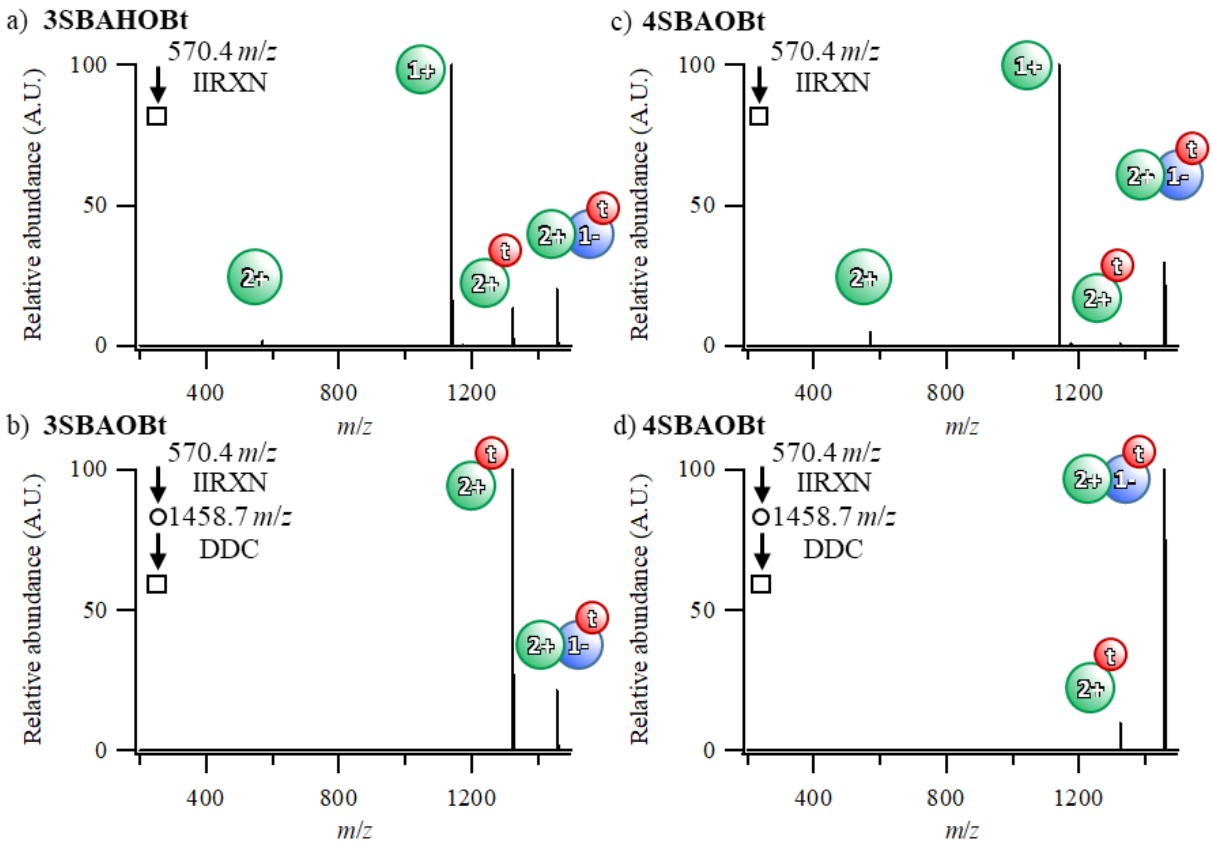


Figure 5.3. a) Ion/ion reaction resulting from mutual storage of the doubly charged ion of KGAILKGAILR and the singly charged anion of 3-SBAOBt. b) Products produced after 19 V DDC activation for 90 ms of isolation ion/ion complex shown in a). c) Ion/ion reaction resulting from mutual storage of the doubly charged ion of KGAILKGAILR and the singly charged anion of 3-SBAOBt. d) Products produced after 19 V DDC activation of isolation for 90 ms of ion/ion complex shown in a).

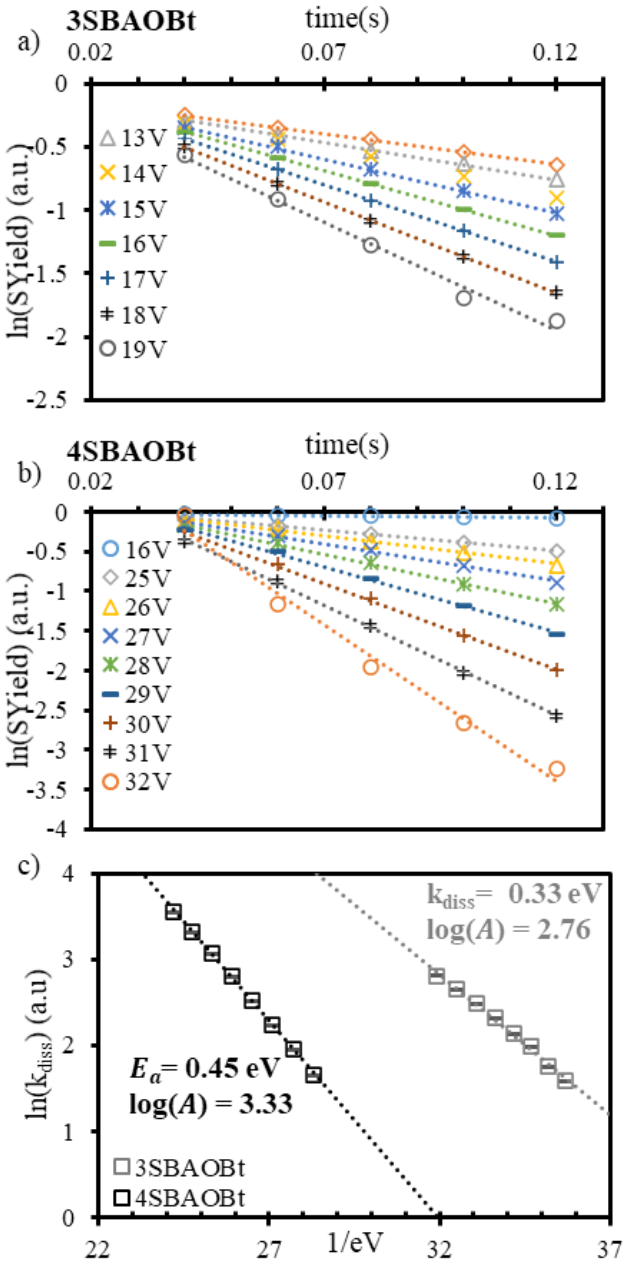


Figure 5.4. Plot of the natural logarithm of survival yield against the reaction time of the covalent modification reaction of a) [KGAILGAILR+3-SBAOBt+H]$^+$ and b) [KGAILKGAILR+4-SBAOBt+H]$^+$ at the various DDC activation voltages, indicated on the respective plot. c) The rate constants obtained from the slope of the survival yield plots were plotted against T_{eff}, obtained from the Tolmachev model.

The natural logarithm of the rate constants is plotted against the inverse of the effective temperature obtained from the Tolmachev model in Figure 5.4c, showing good linearity. Tolmachev's model requires the assumption of REX conditions to maintain a Boltzmann distribution of internal energies. Conversion of the product to covalently modified product before energy exchange is achieved will truncate the Boltzmann distribution, resulting in the Tolmachev model predicting higher effective temperatures than the ions are actually experiencing, resulting in a deviation from linearity in the Arrhenius plot.[29] Given the Tolmachev model has been used to study leucine enkephalin[29] and determine reasonable Arrhenius parameters compared to other studies,[30] we assume, at worst, the Tolmachev model will provide a fair estimation of the Arrhenius parameters. However, we do acknowledge further study of the accuracy of the Tolmachev model is warranted.The values of E_a and the $\log(A)$ factor extracted from the Arrhenius plot of the DDC activation of the [KGAILKGAILR+3-SBAOBt+H]$^+$ complex are 0.33±0.085 eV and 2.76±0.15, respectively. Comparatively, the value of E_a and the $\log(A)$ factor for the DDC activation of the [KGAILKGAILR+4-SBAOBt+H]$^+$ complex are 0.45±0.051 eV and 3.33±0.070, respectively. The lower E_a value observed for the reaction with 3-SBAOBt is expected, given a larger rate constant was determined. However, it is worth noting that the A factor of the 3-SBAOBt reaction is smaller than the A factor of the 4-SBAOBt reaction, which would result in a lowering of the rate constant. Considering the similarity between 3-SBAOBt and 4-SBAOBt, seeing any significant deviation in the Arrhenius parameters is noteworthy and merits further investigation.

To further explore the differences imparted by modifying the location of the sulfate group on the SBA tag, we studied the system with the NHS leaving group using the same methods (spectra can be found in Supplemental Figure 5.1. DDC dissociation spectra for the reactions of [RARARAA-SBA+H]$^+$, where the tagged peptide used the indicated tagging reagent.). The use of an NHS ester for similar reactions has been shown to be more selective for primary amines and guanidines, but have slower reaction rates relative to HOBt.[17] Rate constants at various DDC amplitudes were obtained for the activation of [KGAILKGAILR+3-SBANHS+H]$^+$ and [KGAILKGAILR+4-SBANHS+H]$^+$ and plotted against T_{eff}, obtained from the Tolmachev model, show in Figure 5.5a, as well as the previously reported Arrhenius plots for the same reaction with the HOBt esterified SBA tag. The extracted values of E_a and $\log(A)$ for the activation of [KGAILKGAILR+3-SBANHS+H]$^+$ are 0.48±0.012 eV and 6.63±0.017, respectively. For the activation of

[KGAILKGAILR+4-SBANHS+H]$^{+}$, the extracted values of E_a and $\log(A)$ are 0.46±0.012 eV and 6.23±0.060, respectively. Comparing the rate constants of using 3-SBANHS and 4-SBANHS, reactions using 3-SBANHS appear to proceed marginally faster, though the difference is small relative to the difference observed between the rate constants of 3-SBAOBt and 4-SBAOBt. Interestingly, the critical energy of 3-SBANHS is statistically indistinguishable to the critical energy of 4-SBANHS, which is the differs from the trend observed relative to HOBt leaving groups. The difference is offset by the pre-exponential factor, A, in this case, illustrating the interplay between the two parameters. Further, the $\log(A)$ factor being greater in the reaction of 3-SBANHS than 4-SBANHS is the opposite trend observed when using HOBt leaving groups.

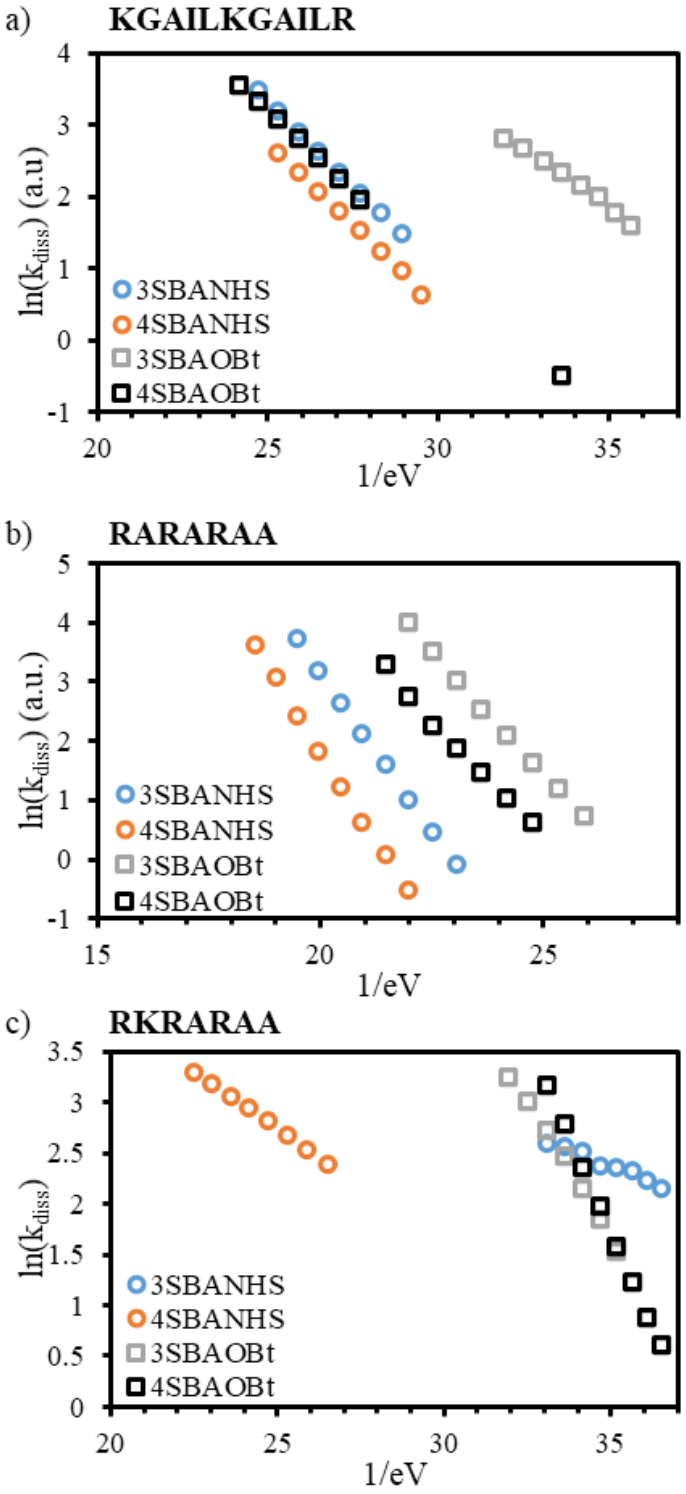


Figure 5.5. Arrhenius plots for the activation of complex of $[n\text{-SBAOBt-H}]^-$ and $[n\text{-SBANHS-H}]^-$ (where n=3 or 4) with a) $[KGAILKGAILR+2H]^{2+}$, b) $[RARARAA+2H]^{2+}$, and c) $[RKRARAA+2H]^{2+}$.

5.3.2 [RARARAA+2H]$^{2+}$ and [RKRARAA+2H]$^{2+}$: Kinetics of Covalent Modification with [n-SBAOBt-H]$^{-}$ and [n-SBANHS-H]$^{-}$

The interplay between the impact of the Arrhenius parameters when reacting the primary amine of KGAILKGAILR with the SBA tags prompted study of the reaction with other functional groups. To study the reactivity of guanidine groups, which have also been studied previously,[17] the model peptides RARARAA and RKRARAA were studied with the same combination of esterified SBAs. Arrhenius plots were generated following the same methods used above.

The results for the same reactions studied above, but with the model peptide RARARAA, are shown in Figure 5.5b. Comparing the Arrhenius plots from the reactions with KGAILKGAILR with the reaction with RARARAA, some immediate similarities can be seen. Most notably, the reactions with the HOBt esterified SBA tags proceed faster than the reactions of NHS esterified tags, in agreement with the results reported by Bu et al. Additionally, reactions with 3-SBA proceed faster than reactions with 4-SBA. Dissimilarly to the reaction with KGAILKGAILR, the difference in rate constants between 3-SBAOBt and 4-SBAOBt are smaller, suggesting that the difference in rate constants observed in the reaction with [KGAILKGAILR+2H]$^{2+}$ could be a special case rather than a result of changing the substitution of the sulfate group on the tag.

Of note is the similarity in slopes between the isomers of SBA when reacting with both peptides. Similar slopes are obtained for both isomers of the SBA tag with both leaving groups, suggesting that the pre-exponential A factor is dictating the differences in the observed reaction rates. Considering the A factor is thought of as a collisional efficiency term, similar to a cross section, this would indicate that the change in electronic structure induced by isomerism does not contribute as significantly to the rate constants as much as accessibility to the reactive site by the nucleophile upon complexation. Considering the charge site of the n-SBA tags is on the sulfate group, the angle that the ring group, and consequently the reactive site of the leaving group, approaches the peptide changes by roughly 60°, which would indeed be expected to have a significant impact on the accessibility of the reactive site to nucleophilic attack.

When comparing the reaction with either of the previously discussed peptides to RKRARAA, however, large deviations from the observed trends occur. Immediately, the change in the relative slopes of the plots, the critical energy E_a, is apparent. In both previously studied peptides, all the critical energies were similar, with differences between HOBt and NHS leaving groups appearing to be dictated by the A factor. Here, the difference is apparent in the critical

energies (indicated by the change in slope between NHS and HOBt leaving groups), as well as pre-exponential factors (indicated by the shifts in the x-intercept, particularly with the NHS leaving groups). Particularly of interest is the low critical energies of the NHS esterified tags, with extracted critical energies of only 0.13 and 0.23 eV for 3-SBANHS and 4-SBANHS, respectively. These are the lowest critical energies evaluated in this study. A compilation of these values can be found in. Some similarities are still observable, notably that when using NHS leaving groups, 3-SBANHS reacts faster than 4-SBANHS. Unexpectedly, the reaction of the SBA tags with RKRARAA appears to produce similar rate constants at similar temperatures to KGAILKGAILR, and not RARARAA. In the reaction with KGAILKGAILR, a primary amine is reacting with the tag, whereas with RARARAA the guanidine functional group of an arginine is reacting with the tag. Using only the trends observed to this point, this suggests that the reaction of RKRARAA with our tags is occurring with a primary amine instead of a guanidine group. To verify this, DDC was used generate fragment spectra of the products of the covalent modification reaction (Supplemental Figure 5.1 and Supplemental Figure 5.2), which revealed indeed the tag is substituted onto the lysine in RKRARAA with all tags tested, and not on an arginine. Regardless, the reaction with RKRARAA and n-SBA tags with NHS or HOBt leaving groups appears to be a special case and further study would be needed to identify what forces are driving the thermodynamic values obtained.

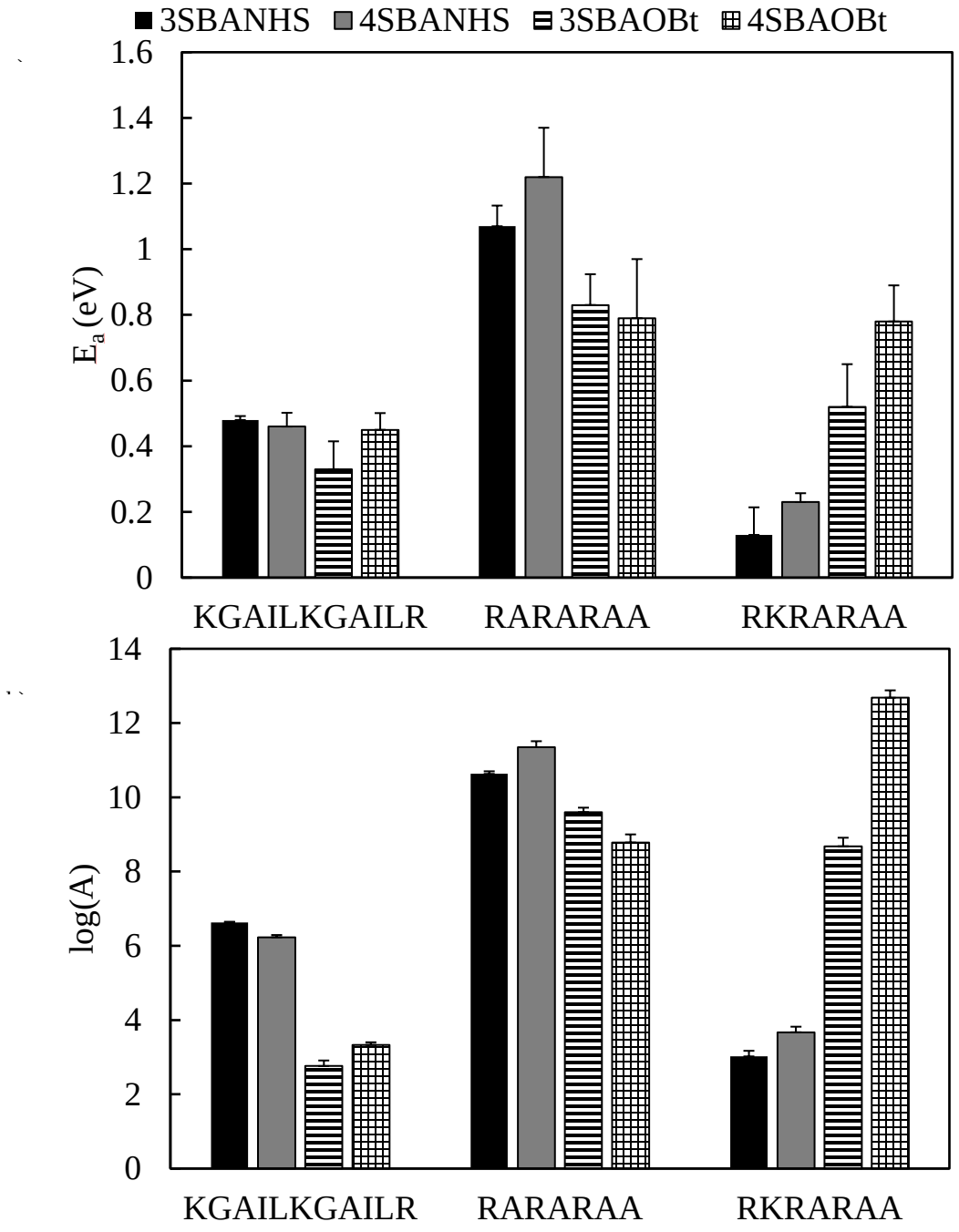


Figure 5.6. Plot of the a) critical energy, E_a, and b) logarithm of the pre-exponential factor, $\log(A)$ for the indicated peptides.

Table 5.1 Compilation of Arrhenius parameters for reactions of n-SBA-(OBt/NHS) with KGAILKGAILR, RARARAA, and RKRARAA.

		Ea (eV)		log(A)	
KGAILKGAILR	3SBANHS	0.48	(0.012)	6.63	(0.017)
	4SBANHS	0.46	(0.042)	6.23	(0.060)
	3SBAOBt	0.33	(0.085)	2.76	(0.15)
	4SBAOBt	0.45	(0.051)	3.33	(0.070)
RARARAA	3SBANHS	1.07	(0.063)	10.63	(0.070)
	4SBANHS	1.22	(0.15)	11.35	(0.16)
	3SBAOBt	0.83	(0.094)	9.60	(0.12)
	4SBAOBt	0.79	(0.18)	8.78	(0.22)
RKRARAA	3SBANHS	0.13	(0.084)	3.02	(0.15)
	4SBANHS	0.23	(0.027)	3.67	(0.034)
	3SBAOBt	0.52	(0.13)	8.68	(0.23)
	4SBAOBt	0.78	(0.11)	12.68	(0.20)

When selecting peptide tags and the leaving group used in the ion/ion reaction, having a general-purpose tag that works consistently in many cases would be ideal. The thermodynamic values obtained from the Arrhenius plots provide insight to not only which tags to use but can also provide foresight for future tag selection. The critical energies, E_a, and pre-exponential factors, A, have been compiled for comparison in Figure 5.6. The work by Bu et al using a model system estimated that using an HOBt leaving group instead of an NHS leaving group effectively lowers the transition state energy by about 0.3 eV,[17] which agrees with what we see here, and led us to believe this would be the determining variable dictating the difference in kinetics. The E_a and log(A) values are paired depending on whether the tag has been esterified with NHS or HOBt, rather than which isomer is used. This is indicated by a similar change in the E_a and A values depending on the leaving group. This is unsurprising, but still of note now that thermodynamic values have been provided to support this observation. A point of interest is that using either the 3- or 4-SBA isomer does still impact the reaction kinetics, though not as significantly as the leaving group. Unfortunately, no clear general trends when using 3-SBA or 4-SBA can be identified. For

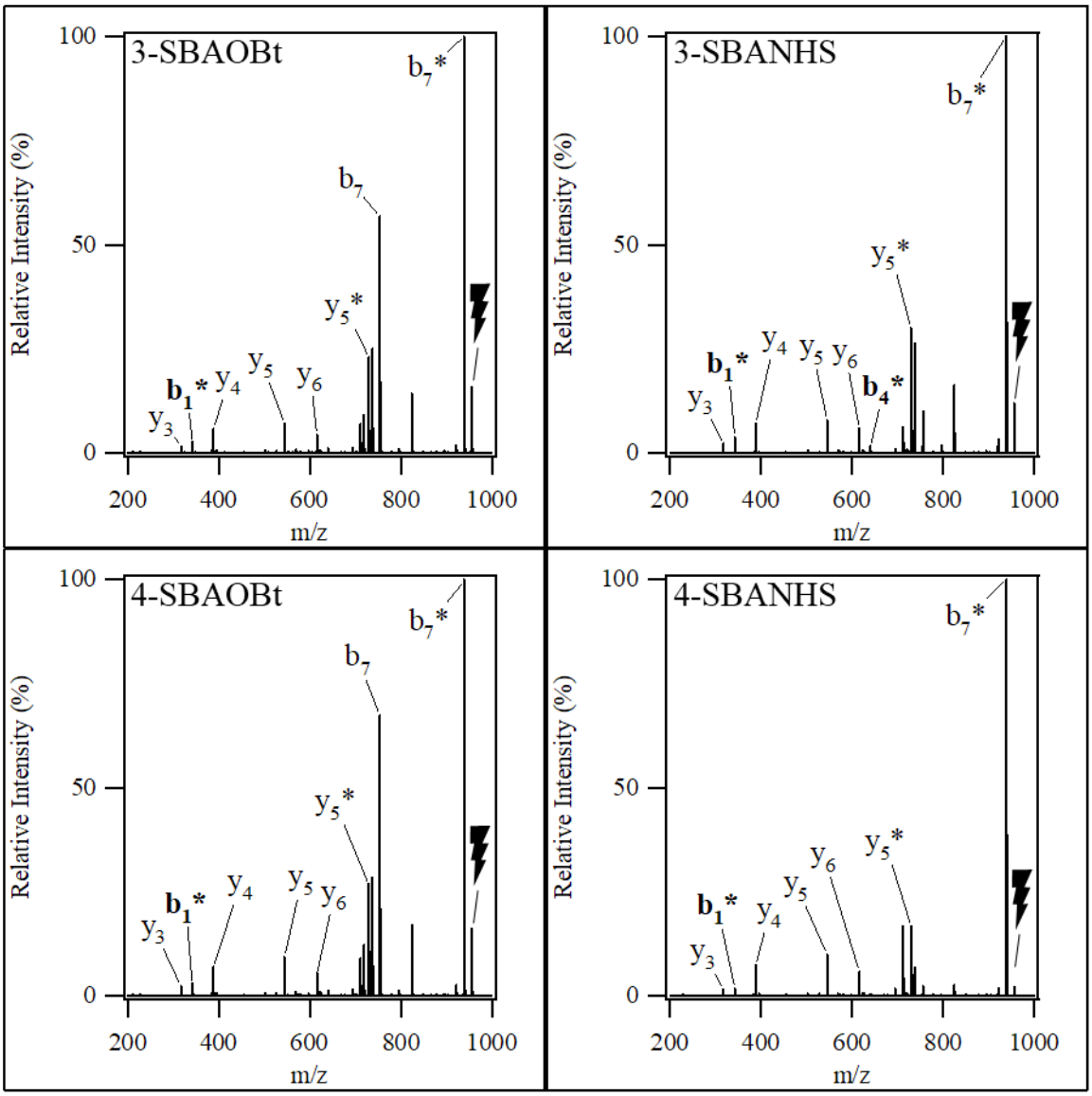
example, when reacting NHS esterified tags with KGAILKGAILR, a higher critical energy is observed with 3-SBANHS than 4-SBANHS, but observing the reaction between RARARAA and RKARARAA, a lower critical energy is observed using 3-SBANHS. However, when observing 3-SBAOBt esterified tags, a lower critical energy is observed with KGAILKGAILR and RKRARAA when reacting with 4-SBAOBt rather than 3-SBAOBt. Considering the larger variation in the pre-exponential *A* factors, we tentatively attribute these differences to be conformational differences. As discussed above, the geometry in which 3- and 4-SBA tags can complex with the peptide is altered by approximately 60°, which could influence other interaction between the tag and the peptide and likely perturbs the conformer.

Identifying that the pre-exponential factors are playing as significant of a role in the ion/ion reaction as the critical energies, if not more, motivates future studies using leaving groups of different tag geometries. Potentially, using smaller tags such as acyl groups could facilitate faster covalent modification reactions, as we are currently stewing the ions for hundreds of milliseconds. Further, using longer tags that don't place the reactive site close to the cation/anion pairing site of the complex and instead on a longer, flexible chain could facilitate more consistent and universal results, leading to the generation of a more general-purpose tag.

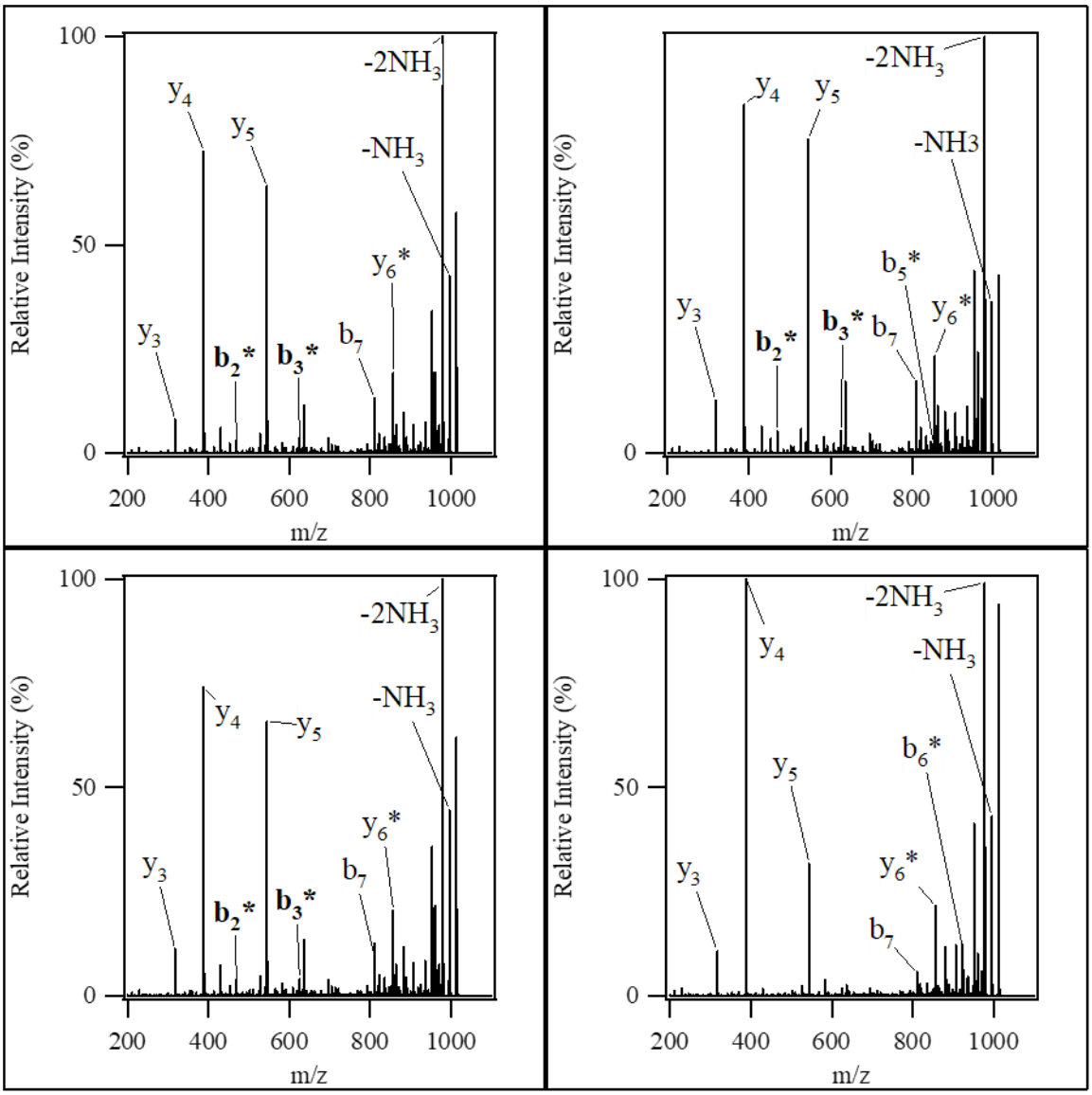
5.4 Conclusion

Here, the kinetic Arrhenius parameters of gas-phase nucleophilic ion/ion reactions between triazole-ester and NHS-ester reagents is quantitated by converting DDC voltages to effective temperatures using the Tolmachev model and applied to the Arrhenius equation. In agreement with previous studies by Bu et al,[16, 17] HOBt esters were found to react with faster rate constants than NHS esters. However, with the kinetic parameters obtained from the resulting Arrhenius plots, we observed that while the HOBt reactions indeed have lower critical energies, which was observed by Bu et al through DFT calculations, the reaction efficiency pre-exponential factor also plays an important role in the kinetics.

5.5 Supplemental Figures



Supplemental Figure 5.1. DDC dissociation spectra for the reactions of $[RARARAA\text{-}SBA+H]^+$, where the tagged peptide used the indicated tagging reagent.



Supplemental Figure 5.2. DDC dissociation spectra for the reactions of $[RKRARAA\text{-}SBA+H]^+$, where the tagged peptide used the indicated tagging reagent.

5.6 References

1. Reid, G. E.; McLuckey, S. A., 'Top down'protein characterization via tandem mass spectrometry. *Journal of mass spectrometry* **2002,** *37* (7), 663-675.

2. McLuckey, S. A.; Mentinova, M., Ion/neutral, ion/electron, ion/photon, and ion/ion interactions in tandem mass spectrometry: do we need them all? Are they enough? *Journal of the American Society for Mass Spectrometry* **2011,** *22* (1), 3-12.

3. Abdillahi, A. M.; Lee, K. W.; McLuckey, S. A., Mass Analysis of Macro-molecular Analytes via Multiply-Charged Ion Attachment. *Analytical Chemistry* **2020,** *92* (24), 16301-16306.

4. Loo, R. R. O.; Winger, B. E.; Smith, R. D., Proton transfer reaction studies of multiply charged proteins in a high mass-to-charge ratio quadrupole mass spectrometer. *Journal of the American Society for Mass Spectrometry* **1994,** *5* (12), 1064-1071.

5. Scalf, M.; Westphall, M. S.; Krause, J.; Kaufman, S. L.; Smith, L. M., Controlling Charge States of Large Ions. *Science* **1999,** *283*, 194.

6. Robb, D. B.; Brown, J. M.; Morris, M.; Blades, M. W., Method of atmospheric pressure charge stripping for electrospray ionization mass spectrometry and its application for the analysis of large poly (ethylene glycol) s. *Analytical chemistry* **2014,** *86* (19), 9644-9652.

7. He, M.; McLuckey, S. A., Two Ion/Ion Charge Inversion Steps To Form a Doubly Protonated Peptide from a Singly Protonated Peptide in the Gas Phase. *J. Am. Chem. Soc.* **2003,** *125*, 7756.

8. Randolph, C. E.; Foreman, D. J.; Betancourt, S. K.; Blanksby, S. J.; McLuckey, S. A., Gas-Phase Ion/Ion Reactions Involving Tris-Phenanthroline Alkaline Earth Metal Complexes as Charge Inversion Reagents for the Identification of Fatty Acids. *Anal. Chem.* **2018,** *90*, 12861.

9. Newton, K. A.; McLuckey, S. A., Gas-Phase Peptide/Protein Cationizing Agent Switching via Ion/Ion Reactions. *J. Am. Chem. Soc.* **2003,** *125*, 12404.

10. Luongo, C. A.; Bu, J.; Burke, N. L.; Gilbert, J. D.; Prentice, B. M.; Cummings, S.; Reed, C. A.; McLuckey, S. A., Selective Removal of Alkali Metal Cations from Multiply-Charged Ions via Gas-Phase Ion/Ion Reactions Using Weakly Coordinating Anions. *J. Am. Soc. Mass Spectrom.* **2015,** *26*, 404.

11. Chao, H.-C.; McLuckey, S. A., Differentiation and Quantification of Diastereomeric Pairs of Glycosphingolipids Using Gas-Phase Ion Chemistry. *Analytical Chemistry* **2020,** *92* (19), 13387-13395.

12. Coon, J. J., Collisions or electrons? Protein sequence analysis in the 21st century. ACS Publications: 2009.

13. Syka, J. E. P.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proceedings of the National Academy of Sciences* **2004,** *101* (26), 9528-9533.

14. Mentinova, M.; McLuckey, S. A., Covalent modification of gaseous peptide ions with N-hydroxysuccinimide ester reagent ions. *Journal of the American Chemical Society* **2010,** *132* (51), 18248-18257.

15. Cheung See Kit, M.; Shepherd, S. O.; Prell, J. S.; Webb, I. K., Experimental Determination of Activation Energies for Covalent Bond Formation via Ion/Ion Reactions and Competing Processes. *Journal of the American Society for Mass Spectrometry* **2021**.

16. Bu, J.; Pilo, A. L.; McLuckey, S. A., Gas phase click chemistry via ion/ion reactions. *Int. J. Mass Spectrom.* **2015,** *390*, 118.

17. Bu, J.; Fisher, C. M.; Gilbert, J. D.; Prentice, B. M.; McLuckey, S. A., Selective Covalent Chemistry via Gas-Phase Ion/ion Reactions: An Exploration of the Energy Surfaces Associated with N-Hydroxysuccinimide Ester Reagents and Primary Amines and Guanidine Groups. *J. Am. Soc. Mass Spectrom.* **2016,** *27*, 1089.

18. Bu, J.; Peng, Z.; Zhao, F.; McLuckey, S. A., Enhanced Reactivity in Nucleophilic Acyl Substitution Ion/Ion Reactions Using Triazole-Ester Reagents. *J. Am. Soc. Mass Spectrom.* **2017,** *28*, 1254.

19. Hassell, K. M.; Stutzman, J. R.; McLuckey, S. A., Gas-phase bioconjugation of peptides via ion/ion charge inversion: Schiff base formation on the conversion of cations to anions. *Analytical chemistry* **2010,** *82* (5), 1594-1597.

20. Pilo, A. L.; McLuckey, S. A., Oxidation of Methionine Residues in Polypeptide Ions Via Gas-Phase Ion/Ion Chemistry. *J. Am. Soc. Mass Spectrom.* **2014,** *25*, 1049.

21. Wang, P.; Hadjar, O.; Gassman, P. L.; Laskin, J., Reactive landing of peptide ions on self-assembled monolayer surfaces: an alternative approach for covalent immobilization of peptides on surfaces. *Physical Chemistry Chemical Physics* **2008,** *10* (11), 1512-1522.

22. Peng, Z.; McGee, W. M.; Bu, J.; Barefoot, N. Z.; McLuckey, S. A., Gas phase reactivity of carboxylates with N-hydroxysuccinimide esters. *Journal of the American Society for Mass Spectrometry* **2014,** *26* (1), 174-180.

23. Webb, I. K.; Morrison, L. J.; Brown, J., Dueling electrospray implemented on a traveling-wave ion mobility/time-of-flight mass spectrometer: Towards a gas-phase workbench for structural biology. *Int. J. Mass Spectrom.* **2019,** *444*, 116177.

24. Foreman, D. J.; McLuckey, S. A., Recent Developments in Gas-Phase Ion/Ion Reactions for Analytical Mass Spectrometry. *Analytical Chemistry* **2020,** *92* (1), 252-266.

25. Tolmachev, A. V.; Vilkov, A. N.; Bogdanov, B.; PĂsa-Tolić, L.; Masselon, C. D.; Smith, R. D., Collisional activation of ions in RF ion traps and ion guides: The effective ion temperature treatment. *Journal of the American Society for Mass Spectrometry* **2004,** *15* (11), 1616-1628.

26. Xia, Y.; Chrisman, P. A.; Erickson, D. E.; Liu, J.; Liang, X.; Londry, F. A.; Yang, M. J.; McLuckey, S. A., Implementation of ion/ion reactions in a quadrupole/time-of-flight tandem mass spectrometer. *Analytical chemistry* **2006,** *78* (12), 4146-4154.

27. Liang, X.; Han, H.; Xia, Y.; McLuckey, S. A., A Pulsed Triple Ionization Source for Sequential Ion/Ion Reactions in an Electrodynamic Ion Trap. *J. Am. Soc. Mass. Spectrom.* **2007,** *18* (3), 369-376.

28. Vékey, K., Internal energy effects in mass spectrometry. *Journal of mass spectrometry* **1996,** *31* (5), 445-463.

29. Webb, I. K.; Londry, F. A.; McLuckey, S. A., Implementation of dipolar direct current (DDC) collision-induced dissociation in storage and transmission modes on a quadrupole/time-of-flight tandem mass spectrometer. *Rapid Communications in Mass Spectrometry* **2011,** *25* (17), 2500-2510.

30. Sztáray, J.; Memboeuf, A.; Drahos, L.; Vékey, K., Leucine enkephalin—a mass spectrometry standard. *Mass spectrometry reviews* **2011,** *30* (2), 298-320.

VITA

Joshua Lee Fischer was born on August 24th, 1988 in Ann Arbor, MI and is the oldest of Mark and Dawn Scheu's three children. Joshua graduated from Stockbridge High School in 2006 and initially pursued a career at McDonald's to fund his upper education, becoming a salaried manager at the age of 18. While working at McDonalds, he pursued an associate degree in chemistry at Lansing Community College, but was unable to maintain the workload and grades, and eventually stopped attending college. Joshua worked at McDonald's for six more years before deciding to return to school to pursue his original passion, science. He enrolled at Wayne State University in 2011 and completed his bachelor's in chemistry (Cum Laude) in 2016. Joshua began laboratory research in the summer of his freshman year with Prof. Matthew Allen and went on to receive numerous accolades, including an undergraduate research grant in 2013 and the ACD Division of Analytical Chemistry Award in 2014, to name a few. Later, in his Junior year, Joshua would begin studying with Prof. Sarah Trimpin, where he would continue his academic success, winning the American Society for Mass Spectrometry Undergraduate Student Award in 2015, as well as the Best Paper Award for Advances in Academic Research at the ANACHEM symposium, also in 2015. This would culminate into his acceptance to the analytical chemistry graduate program at Purdue University in 2016, in which he would receive the Ross Fellowship, and later the Arthur Kelly Teaching Award in 2018. While at Purdue, Joshua studied the conformational preferences of gas-phase peptide molecules with Dr. Timothy Zwier, as well as the ion/ion reaction kinetics of gas-phase covalent modification reactions with Dr. Scott McLuckey.

PUBLICATIONS

1. Fischer, J.L; Merhnart, S. A.; McLuckey, S.A., Kinetic Study of Leaving Groups in Gas-Phase Covalent Modification Click Reactions. *In writing.*

2. Trimpin, S.; Marshall, D. D.; Karki, S.; Madarshahian, S.; Hoang, K.; Meher, A. K.; Pophristic, M.; Richards, A. L.; Lietz, C. B.; Fischer, J. L.; Elia, E. A.; Wang, B.; Pagnotti, V. S.; Lutomski, C. A.; El-Baba, T. J.; Lu, I-C.; Wager-Miller, J.; Macki, K.; McEwen, C. N.; Inutan, E. D., An Overview of Biological Applications and Fundamentals of New *Inlet* and *Vacuum* ionization technologies. *Rapid Commun. Mass Spectrom.* **2020,** DOI: 10.1002/rcm.8829.

3. Inutan, E. D.; Jarois, D. R.; Meher, A. K.; Karki, S.; Fischer, J. L.; Imperial, L. F.; Foley, C. D.; El-Baba, T. J.; Lutomski, C. L.; Trimpin, S., New MS Concepts for Synthetic Polymers and Additives Characterization. *Rapid Commun. Mass Spectrom.* **2020,** *34*, e8768.

4. Austin, C. A.; Inutan, E. D.; Bohrer, B. C.; Li, J.; Fischer, J. L.; Wijerathane, K.; Foley, C. D.; Lietz, C. B.; Woodall, D. W.; Imperial, L. F., Resolving Isomers of Star-Branched Poly (Ethylene Glycols) by IMS-MS using Multiply-Charged Ions. *J. Am. Soc. Mass. Spectrom.* **2020**, *32,* 21-32.

5. Blodgett, K. N.; Fischer, J. L.; Zwier, T. S.; Sibert, E. L., The missing NH stretch fundamental in S 1 methyl anthranilate: IR-UV double resonance experiments and local mode theory. *Physical Chemistry Chemical Physics* **2020,** *22*, 14077-14087.

6. Fischer, J.L.; Blodett, K. N.; Harrilal, C. P.; Walsh, P. S.; Choi, S. H.; Zwier, T. S., Conformational Analysis of Ac-γ^4-Phenylalanine-NHMe Monomer and Single Water Hydration: Dominant Phenyl Rotor Interactions and the Effect of Solvation on Amide Stacking. *In writing*.

7. Blodgett, K. N.; Sun, D.; Fischer, J. L.; Sibert, E. L.; Zwier, T. S., Vibronic Spectroscopy of Methyl Anthranilate and its Water Complex: Hydrogen Atom Dislocation in the Excited State. *Phys. Chem. Chem. Phys.* **2019**, *21,* 21355–21369.

8. Fischer, J. L.; Elvir, B. R.; DeLucia, S.-A.; Blodgett, K. N.; Zeller, M.; Kubasik, M. A.; Zwier, T. S., Single-Conformation Spectroscopy of Capped Aminoisobutyric Acid Dipeptides: The Effect of C-Terminal Cap Chromophores on Conformation. *J. Phys. Chem. A* **2019**, *123,* 4178–4187.

9. Blodgett, K. N.; Fischer, J. L.; Lee, J.; Choi, S. H.; Zwier, T.S., Conformation-Specific Spectroscopy of Asparagine-Containing Dipeptides: Influence of Single or Adjacent Asn Residues on the Inherent Conformational Preferences. *J. Phys. Chem. A,* **2018**, 8762–8775.

10. Harrilal, C. P.; DeBlase, A. F.; Fischer, J. L.; Lawler, J. T.; McLuckey, S. A.; Zwier, T. S., Infrared Population Transfer Spectroscopy of Cryo-Cooled Ions: Quantitative Tests of the Effects of Collisional Cooling on the Room Temperature Conformer Populations. *J.Phys. Chem. A* **2018**, *122*, 2096–2107.

11. Hewett, D. M.; Tabor, D. P.; Fischer, J. L.; Sibert, E. L.; Zwier, T. S., Infrared-Enhanced Fluorescence-Gain Spectroscopy: Conformation-Specific Excited-State Infrared Spectra of Alkylbenzenes. *J. Phys. Chem. Let.* **2017**, *8*, 5296–5300.

12. Devereaux, Z. J.; Reynolds, C. A.; Fischer, J. L.; Foley, C. D.; DeLeeuw, J. L.; Wager-Miller, J.; Naraya, S. B.; Mackie, K; Trimpin, S., "Matrix-Assisted Ionization (MAI) on a Portable Mass Spectrometer: Analysis Directly from Biological and Synthetic Materials." *Anal. Chem.* **2016**, *88*, 10831–10836.

13. Fischer, J. L.; Lutomski, C. A.; El-Baba, T. J.; Sirwardena-Mahanama, B. N.; Weidner, S. M.; Falkenhagen, J.; Allen, M. J.; Trimpin, S., "Selective Ionization and Characterization of a Europium-PEG Complex From a Crude Sample Using Matrix Assisted Ionization-Ion Mobility Spectrometry." *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 2086–2095.