# APPLICATION OF CRYOGENIC INFRARED AND ULTRAVIOLET SPECTROSCOPY FOR STRUCTURAL AND DYNAMIC STUDIES OF GAS PHASE IONS

by

**Christopher P. Harrilal**

**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

December 2019

# THE PURDUE UNIVERSITY GRADUATE SCHOOL
# STATEMENT OF COMMITTEE APPROVAL

**Dr. Scott A. McLuckey, Co-Chair**

Department of Chemistry

**Dr. Timothy S. Zwier, Co-Chair**

Department of Chemistry

**Dr. Hilkka I. Kenttamaa**

Department of Chemistry

**Dr. Libai Huang**

Department of Chemistry

**Approved by:**

Dr. Christine Hrycyna

*This thesis is dedicated to my family and friends.*

## ACKNOWLEDGMENTS

First, and foremost, I want to thank my parents. Mom, you are the hardest working person I know. I've watched you work your way up from the humblest of beginnings to the successful person that you are now. You have always been my inspiration and I strive to match your work ethic in everything I do. Dad, you helped me in the most crucial part of my life and put me on track to succeed. You opened up opportunities that transformed my life for the better and I will always be grateful. You also remind me of where I come from and pull me back in when feel like I'm lost. Mike, much of my success would not have been possible without you. Thank you for being part of our lives and providing support at every step. In the words of Dave, you the man!

To my aunts (Pam, Chan, Vid, and Naomi) and uncles (Latch, Ramesh, Ram, Pradeep, Mahen, and Jason) I cannot thank you enough for always being there whenever I need help or someone to talk to. I am blessed to have this family. Uncle Latch and Auntie Naomi, you guys were like my second parents growing up. You guys helped shape my childhood and put me on the right track from an early age. Auntie Vid and Uncle Mahen, I'm not sure I would have made it out of Florida without you guys. You guys treated me like a son since the day I came down and gave me the tools to succeed. Thank you for investing in me.

My two sisters, Sara and Christine, you guys know me better than I know myself. I can always be myself around you guys and its always a relief coming back home. Sara, thank you for helping me through all the tough times in my life and always telling me what I need to hear rather than what I want to hear. Christine, as you grow up, I can't help but to be proud of you. No matter what happens I know you guys will always have my back and that is one of the best feelings in the world.

To my cousins Jen, Lea, Corey, Anup, Matthew, and Anthony the strong family ties will always be there. You guys push me to be a better person and seeing your success motivates me. You guys helped me in ways that you may not know so thank you for every conversation along the way.

Sydney Chow, thank you for being an amazing girlfriend and my escape from reality. I got to experience so many things with you in such a short time. You constantly bring me out of my comfort zone and into some adventure. Thank you for all your support and for shaking things up, these last few years definitely would not have been the same without you.

Edward Ramirez and Robert Delanuez, thank you for all the wisdom that you passed down to me, you guys have been great mentors during my time in Florida.

You guys played a huge role in shaping me into who I am today. I have a lot of great memories thanks to you guys and the rest of the CrossFit family.

Shantell Rolle! I'm not even sure if I would have been here if it were not for you. Thank you for helping me pass organic chemistry way back in FIU. You are one of my best friends and helped me so much, in and out of school, along the way. Naomi Pierre, I'm still convinced that you tricked me into being the president of the Science Club, but it has brought me a long way. Thank you for all your support.

Dr. Delgado and Mayra Exposito, thank you for being amazing mentors and pushing me to branch out and providing me every opportunity to succeed. I'm glad you guys were as involved as you were, it has brought me a long way.

Dabargha, thanks for all the phone calls and motivation. You've always been supportive, and I can talk to you about anything in life. Thanks for being the positive person that you are, I needed that.

Anthony and Elissia, I'm glad you guys joined the lab when you did. Going through the program would have been much harder without you guys. Thanks for all the memories and always being there to talk about everything besides science. I wish you guys the best of luck in everything that comes next.

To all the previous members of the McLuckey and Zwier labs, thank you for laying the foundations of the labs and being great role models. To the current members, thank you for always performing at a high level, it has always been motivational to see everyone progressing in their research.

Andrew DeBlase, thank you for training me and teaching me so much about instrumentation and spectroscopy. I wouldn't have had the success that I have without you. You have been a great mentor in showing me how to perform scientific research at a high caliber.

Dr. McLuckey and Dr. Zwier, thank you for all the guidance and support over the last few years. I have learned so much about science and life from each of you. It was a great experience working in your labs and I am truly happy I had that opportunity.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABSTRACT

The work presented here employs cryogenic ion spectroscopy for the study of protein structure, kinetics, and dynamics. The main technique used is IR-UV double resonance spectroscopy. Here peptide ions are generated through nano electrospray ionization, guided into a mass spectrometer, mass selected, and then guided into a cryogenically held octupole ion trap. Ions are subsequently cooled to their vibrational ground state through collisions with 5 K helium allowing for high resolution IR and UV spectra to be recorded. The IR spectra are highly sensitive to an ion's conformation, and the well resolved UV spectra provides a means generate conformer specific IR spectra. With the use quantum mechanical calculations, it is possible to calculate the vibrational spectra of candidate structures for comparison with experimental spectra. Strong correlations between theory and experiment allow for unambiguous structural assignments to be made.

Structural studies are performed on β-turn motifs and well as salt-bridge geometries. Beta-turns are a commonly occurring secondary structure in peptides and proteins. It is possible to artificially encourage the formation of this secondary structural element through the incorporation of the D-proline ($^{D}$P) stereoisomer followed by a gly or ala residue. Interestingly, the L-proline ($^{L}$P) stereoisomer is seen to discourage the formation of beta turn structure. Here were probe the inherent conformational preferences of the diastereomeric peptide sequences YA$^{L}$PAA and YA$^{D}$PAA. The findings agree with solution phase studies, the $^{D}$P sequence is observed to adopt a beta turn however, the $^{L}$P sequence is found to undergo a sterically driven *trans* → *cis* isomerization about the proline amide bond. We find the energetics associated with this unfavorable interaction and show the ability to reverse it by proper substitution of $Ala_2$ for a Gly.

The studies directed towards gas phase salt bridges have been limited to single amino acids or dipeptides. Generally, these species are ionized using a metal ion or adducted with water or excess electrons in order to stabilize a zwitterionic motif. Here we take the first look at a salt bridge motif incorporated into polypeptide in order to understand how the solvation from the secondary structure can aid in stabilizing these motifs in non-polar environments. We find a unique salt bridge motif in the YGRAR sequence in which the tyrosine OH acts as a neutral bridge to form a network between the C-terminal arginine and the ion pair formed between the central arginine and C-terminal carboxylate group. This binding motif has not been discussed in literature and appears as

an important structural element in non-polar environments as all salt bridge character is lost upon substituting Tyr for Phe. We are the process of mining the PDB for these types of interactions.

To better understand how cryo-cooling impacts the resulting population distribution at 10 K we measured the distribution among the two major conformation of the YGPAA ion. This was carried out using population transfer spectroscopy. In this method conformational isomerization is induced vis single conformer infrared excitation. The change in population can be related to the final population distribution at 10 K. With this number, we were able to develop a cooling model to simulate the change in the distribution as a function of cooling. The cooling rates, were experimental established, and the isomerization rates and starting population were theoretically derived through RRKM and thermodynamic calculations. With these parameters and cooling model, we found that the room temperature population distribution is largely preserved. When isomerization events involve breaking a hydrogen bond, they become too slow to complete with the cooling time scale of the experiment, effectively freezing in the room temperature structures. These are important physical parameters to characterize when performing structural studies at 10 K.

Finally, we demonstrate a 2-Color IRMPD technique that is able to generate linear spectra at varied temperatures. This is in sharp contrast to traditional IRMPD which results in non-linear and skewed spectra. The importance of generating linear spectra when making structural assignments is highlight by comparing the performance between both techniques. Furthermore, with this technique we show the ability to record the spectra of ion prepared with high internal energies. This provides spectroscopic snapshots of the unfolding events leading to dissociation. Overall, the versatility of this technique to record ground state spectra comparable to IR-UV DR, to record linear spectra at room temperature, and to probe dynamics proves this technique to be useful in the field of ion spectroscopy.

# CHAPTER 1. INTRODUCTION

## 1.1 Motivation

The strong relationship between structure and function for biological molecules has been well-documented.[1,2] For example, proteins are able to perform various functions ranging from acting as catalysts[3], transport[4], oxygen storage[5], generating movement[6], ect. The type of function is often related to the 3-D shape of a given protein. In some cases, large structural rearrangements occur after metal binding or interactions with bio-molecules allowing signaling pathways to be turned on or off.[7] The conformational heterogeneity of protein, peptides, and other bio-molecules and their ability to respond to external forces results in a rich biology that is essential for life. Understanding structure and dynamics is thus necessary to describe, model, and mimic the various functions carried out in a living cell.

## 1.2 Condensed Phase Techniques

Many techniques have been developed towards this end including nuclear magnetic resonance[8,9] (NMR), X-Ray crystallography[10], and cryogenic electron microscopy[11] (cryo-EM) to name a few. Unsurprisingly, the native state of a molecule is sensitive to its surrounding environment. Currently, there is no technique capable of probing structure in its native environment. Nevertheless, significant insights have been made with solution and solid phase-based techniques. A complete description of these techniques is outside the scope of this text; however, some important features of each technique can be noted. NMR is capable of providing structural constraints in the form of distances between atoms in a range of solvent types.[12] However, it requires high sample purity and is hindered in its ability to resolve multiple conformations that are populated at temperatures that the measurement is performed.[13] X-ray crystallography allows for single molecules to be probed however, a limited amount of proteins with low conformational heterogeneity can be crystallized. Furthermore, the crystallization process may dramatically alter the structure.[14] Cryo-EM offers some advantages in this sense as proteins can be frozen in a thin slice of solvent. The ability to distinguish between conformations along with limited resolution however, also limit this technique. While each technique is able to excel in their own right, the

major limiting factors of structural analysis in the condensed phase is the need for high purity sample, time, and the ability to resolve between distinct conformational states.

### 1.3 Gas-Phase Approaches

While probing bio-molecules in the condensed phase is a natural approach to understanding 3-D structure, studies in the gas phase have been shown to provide many unique benefits. In particular native mass spectrometry[15] along with a range of activation techniques and methods to probe 3-D structure has given way to the field of gas phase structural biology. With native mass spectrometry it has been demonstrated that relatively large proteins can be transported into the gas phase with gentle conditions such that the solution phase structure is largely preserved.[16,17] Retention of solution phase structure in the gas phase has also been observed for small peptide systems. This is evident by the ion's gas phase mobility dependence on the composition of electrospray solution in which they are dissolved.[18] The benefits of performing structural studies by means of mass spectrometry is that ions can be manipulated based on their mass to charge. This provides a high level of specificity, purity, and analysis time frames in the millisecond range, speeds which are unmatched by solution phase techniques.

In regard to protein and peptide analysis, mass spectrometry has traditionally been used for probing primary sequence. However, techniques such as oxidative footprinting[19], hydrogen deuterium exchange[20] (HDX), crosslinking[21], surface induced dissociation[22] (SID), collision induced dissociation[23] (CID), electron transfer/capture dissociation[24,25] (EXD), ion mobility[26] (IM) and ultraviolet photodissociation[27] (UVPD) have been leveraged to gain insights on protein structure ranging from secondary to quinary. Furthermore, mobility-based techniques have been used in complementary ways to provide measures of the collisional cross section of peptides and proteins under various conditions. The inherent speed associated with mass spectrometry also provides that ability to probe dynamics and kinetics ranging from the millisecond to day timescales. This allows for slow processes such as solution phase peptide aggregation to be continuously monitored. Other techniques such as collisionally induced unfolding[28] (CIU) provide the ability to monitor large scale conformational changes and the energetics and timescales associated with those processes. The ability to couple any of these techniques provide relatively loose but informative constraints on systems that would otherwise be hard to probe with more traditional methods. Furthermore, depending on the conditions used the resulting information can be

representative of the solution phase structures or the structures populated on the gas phase potential energy surface.

## 1.4 Ion Spectroscopy

Ion spectroscopy is an additional gas phase technique that can be used for structure elucidation. This technique provides the highest level of structural detail for ions within a certain size regime. In the gas phase IR spectra are not subjected to line broadening induced by external solvent interactions and appear highly sensitive to an ion's 3-D conformation.[29,30] The vibrational spectra essentially describe the hydrogen bonding network present in the ion. This provides many structural constraints that can be used in addition to quantum molecular calculations to provide unambiguous structural assignments.[31] The level of structural detail provided by the IR spectra and calculations are unmatched by the techniques described above and can be applied to systems ranging from peptides[32], sugars[33], metal complexes[34], water complexes[35], ect.

While this technique is capable of providing unambiguous structural assignments, it is sharply limited in the size of ions that can be probed. This limitation comes largely from the dependence on quantum level calculations that are needed to model the experimental IR spectrum. As such the biological systems that are studied are in a regime that is referred to as the complexity gap. In this regime the systems are small enough (>12 amino acids) such that high level calculations can be performed however, large enough such that it is impossible to intuit the conformational preference. At these sizes peptide systems are large enough such that defined secondary structure can be observed.[32,36] As such, most studies directed towards peptides are built upon understanding the relationship between the primary sequence of an ion, and its secondary structure or how charge solvation takes place in the absence of solvent. Understanding these relationships are key to building models that can accurately predict peptide folding. The gas-phase environment provides the benefit in that the inherent conformational preference can be probed free of solvent perturbations. Furthermore, stepwise solvation with individual water molecules can be performed such that the solvent interactions can also be interrogated.[37] The high resolving power of ion spectroscopy lends itself well to a bottom approach to understanding the subtle effects of how the primary sequence affects the 3-D shape through intramolecular hydrogen bonding and steric effects. The ability to perform these studies on mass spectrometry-based platform further increases the specificity and utility of the technique as it can be coupled to other techniques providing

orthogonal information. Currently, the field is being push in a more analytical direction as vibrational spectra can be used to differentiate between isomeric/isobaric systems that appear at the same mass to charge ratio but are chemically distinct.[38] Often in cases like this the fragmentation patterns are similar preventing the ability to distinguish between positional isomers. Towards this goal libraries are being made with vibrational spectra for systems like glycans. This text focuses primary on the use of IR-UV double resonance and IRMPD in a more fundamental approach and are discussed in a much greater detail below.

### 1.4.1 Basis of Action Spectroscopy

There are three major forms of ion spectroscopy commonly employed, messenger tagging[39], infrared multiple photon dissociation[40] (IRMPD), and IR-UV double resonance[32] (IR-UV DR) spectroscopy. A detail discussion of each technique is given individually in chapter 3. Of the three, messenger tagging and IR-UV DR require ions to be cryo-cooled prior to being probed. A benefit of operating at these temperatures is the vibrational spectra appear linear and well resolved. IRMPD on the other hand can be performed at any temperature however, the resulting spectra appear non-linear and distorted in nature, in most cases the spectra differ drastically from the linear spectra. While each technique is carried out in a unique manner there are some fundamental similarities by nature that all measure the vibrational spectra of gas phase ions. A brief overview of the necessary considerations to perform gas phase spectroscopy as well as the general operating principles of each technique is discussed below.

In traditional spectroscopic measurements the attenuation of light is monitored as it passes through a sample and can be modelled using beer-lamberts law:

$$I(\nu) = I_0 e^{-\sigma(\nu)Ln}$$

where $I_0$ and $I(\nu)$ are the intensities of the light before and after passing through the sample, $\nu$ is the frequency of light, $\sigma$ is the absorption cross-section, L and n are the path length and number density. The number density of ions in a typical trapping device in a mass spectrometer is on the order of $2x10^3$ $cm^{-3}$, which is sufficiently low such that a difference between the incoming and transmitted light is negligible.[30,41] To circumvent the low number density, absorption events are monitored by detecting an action or coincidence event such as a change in the m/z ratio of the ion

being monitored. A full description of how action spectra are recorded in given in chapter 3. The basis of all ion spectroscopy experiments however, is as follows, the ion of interest is mass isolated from all other ions created during the ionization process and then exposed to laser irradiation. If the wavelength of the laser is resonant with an electronic or vibrational transition a mass shift will occur via chemical fragmentation or through the evaporation of a non-covalently bound tag molecule. The fragmentation event forms an ion with a m/z ratio distinct from the initially isolated precursor ion. The lighter fragment ions are then detected through a subsequent mass analysis step. A full action spectrum is created by monitoring the formation of fragment ions as a function of the IR or UV laser wavelength. The number of ions affected by the incoming light can be expressed as:

$$N(\nu) = N_0 e^{-\sigma(\nu)\Phi(\nu)}$$

where $\Phi(\nu)$ is the photon fluence experience by the sample. For linear depletion measurements recorded using IR-UV DR, $N(\nu)$ is on the order of about 30 % of the fraction of ions initially excited by the UV laser. Only by correlating absorption and action events is it routine to perform spectroscopy on gas phase ions.

### 1.4.2 Mass Analyzer - Quadrupole

Quadrupole mass filters, as shown in Fig. 1.1, are one of the most common mass analyzers in mass spectrometry. In the experimental setup described below all the mass isolation and well as mass analysis takes place within quadrupoles. The basis of quadrupole operation is that ions can be trapped within the inscribed radius formed by the 4 rods. Trapping occur laterally through collisions with a buffer gas that effectively removes the translational energy of ions as the enter the quadrupole trap. Repulsive voltages can be applied to lenses on either side of the trap such that ions are encouraged to stay within the length of the quadrupole. Radial confinement is achieved by applying radio frequency (RF) of about 1 MHz, referred to as the drive frequency, to each rod that makes up the quadrupole. Adjacent rods are set such that they have opposite phases of RF, which effectively creates a trapping quadrupole field. Ions within the rod set experience attractive forces towards one set of rods for half a cycle of RF followed by repulsive force for the second half. The motion occurs in 2 dimensions keeping the ions with the boundaries of the rod set. The rate which ions move towards and away from the rods is referred to as their secular frequency.

This value is directly related to the ions mass and charge. Lighter ions tend oscillate at higher frequencies, while heavy ones at lower frequencies.

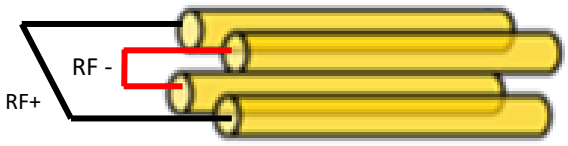


Figure 1.1: Schematic of quadrupole mass analyzer

The fact that ions of different masses oscillate at different frequencies allows ions to be manipulated in a mass dependent manner. This provides the ability to either eject all masses out of the trapping region beside while retaining a particular mass, eject a particular mass from the trapping region while retaining all other masses, or to eject all masses out in a mass dependent manner towards a detector. In the latter case, this is how a mass spectrum can be generated. In the case of ejecting masses out of the trap a supplemental AC wave in required, this waveform is superimposed on one set of rods. Effectively, by tuning the frequency of the AC waveform to match that of the ion's secular frequency, it is possible to selectively excite or accelerate ions of a particular mass.

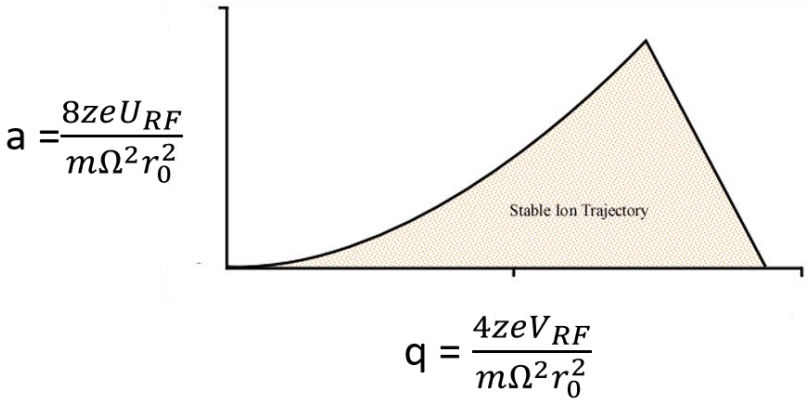


Figure 1.2: Mathieu stability diagram with Mathieu equations

The ion motion within the trapping volume is governed by the Mathieu stability equations. One set of solution to those equations is shown on the above graph. The motion of the ions can be

related to the dimensionless parameters (a) and (q) where z is the charge or the ions, e is the elementary charge, $V_{RF}$ is the voltage of the RF applied, $U_{DC}$ is the voltage of the DC applied, $\Omega$ is the frequency of the RF, and $r_0$ is the inscribed radius of the trap. DC potential are also applied to each rod set, again adjacent rods have opposite signs, the DC applied in this manner is referred to as quadrupolar DC.

A full description of the operation of a quadrupole will not be given here. The basis is that ions can be radially trapped inside the volume of a quadrupole ion trap and then be manipulated in a mass dependent manner. Manipulations can be performed by applying supplemental AC on to the rod, adjusting the voltage of the drive RF, and by adjusting the DC voltage. The stabilities of the ions are governed by the solutions to the Mathieu equations. By adjusting the parameters mentioned before ions can be move in and out of regions of stability. A basic understanding of this theory is necessary to perform ion spectroscopy as it is a tandem mass spectrometry experiment. That is, fragment ions are created from the isolated precursor ions and subsequently measured.

### 1.4.3 Outline of Thesis

In the following chapter a brief introduction to the instrumentation used for these studies is given. The overall concept of how the instrument functions is provided along with a description of significant modification that have led to an over increase in performance.

In chapter 3 a detail description of all the ion spectroscopy techniques is given. The types of scans that are possible using IR-UV DR are explained and illustrated using real data. New triple resonance laser schemes are explained and well as the best manner in which to align the three beams. Finally, a detailed description of the IRMPD mechanism is given followed the mechanism of 2-Color IRMMPD.

The final four chapters go into the application of these techniques to solve various problems. The goal of this work is to demonstrate the many ways in which ion spectroscopy can be used. To this end 3 distinct types of experiments will be discussed. In the first type the static structure of gas phase ions is interrogated in order to better understand the relationship between secondary and primary sequence. The two projects discussed here take advantage of high structural resolving power of IR-UV double resonance to shed light on particular folding and hydrogen binding modes. In the first project the folding propensity of L ($^{L}$P) and D ($^{D}$P) proline stereoisomers are investigated when incorporated into short peptide (YA$^{L/D}$PAA) sequences. Solution phase studies

have demonstrated the ($^{D}$P) isomer to encourage the formation of the β-turn motif, while the L-proline ($^{L}$P) often discourages its formation. Using IR-UV double resonance we make structural assignments for each diastereomer to better understand the difference in folding propensity. The results of the gas phase structure for the $^{D}$P sequence agree well with the solution phase structure in that it adopts a β-hairpin turn. Similar to solution, we find that the $^{L}$P sequence indeed does not adopt the β-hairpin turn. Rather we find that this sequence undergoes a *trans* → *cis* isomerization of the proline amide bond. By use of quantum mechanical calculations we find that this isomerization is driven largely by unfavored steric interactions between the proline and N-terminal alanine side chain. We experimental confirm this hypothesis by making the appropriate substitution of Ala for Gly, forming a sequence YGPAA. The substitution for Gly results in the reformation of the β-hairpin turn. These results highlight the structural resolving power to investigate the intimate relationships between primary and secondary structure.

Along similar lines we take advantage of the high structural resolving power provided by IR-UV DR spectroscopy to study the relationship between salt-bridge formation and primary sequence. Salt bridges are the strongest electrostatic interactions in peptide and protein systems. They are formed through an interaction between oppositely charges residues. In the field of gas phase structural biology, they are not easily detected yet are important structural factors. Using IR spectroscopy can readily detect the presence of salt-bridge species as the give rise to unique spectral features that can easily be discerned. By using model systems (YGRAR and FGRAR), it has been possible to characterize these interactions in systems large enough that the solvation provided by the backbone can play a role in salt-bridge formation. In the YGRAR sequence we find that a salt bridge is formed between the central arginine, $ArgH^{+}$(3) and the deprotonated C-terminus. Salt-bridges involving the C-terminus are not often observed and presents a new binding motif. Even more interesting however, is the role that the tyrosine OH group plays in stabilizing the salt bridge. The polar OH is able to bind to the carboxylate and form a neutral bridge between the C-terminal arginine, $ArgH^{+}$(5) and the salt bridging partners. Such a motif is not discussed in literature and may be a common occurrence. We plan to mine the protein data bank to investigate the importance of this interaction. Within the confines of our study we find that is highly important as the substitution of Tyr for Phe (effectively removes the OH group) destabilizes the salt bridge interaction altogether. The presence of this polar group may also be of importance as it can aid offsetting the dehydration penalty of the COOH group in the non-polar gas environment.

Both structural studies briefly described here take show case the structural resolving power attainable with IR-UV DR. The ability to attain unambiguous structural assignments along with quantum calculation lead to a deeper understanding of the relationship between primary and secondary structure in both cases. These types of studies are often the most performed in the field of ion spectroscopy.

In addition to studies pertaining to structural elucidation it is also possible to study dynamics using IR-UV double resonance as well. A study along this line has been performed with the model peptide ion, YGPAA. This system was found to adopt two major structurally distinct conformations. The energetics associated with isomerization as well relative abundance between the two structures were measured using population transfer spectroscopy. In this scheme it is possible to excite a single conformation such that it can isomerize to the available wells on the potential energy surface. The results of this study lead to a much better understanding of the relationship between the conformational distribution at room temperature to that at 10 K. This was realized by comparing the measured and distribution after cooling to the calculated distribution at room temperature. With these values we were able to build cooling model to simulate the kinetics of cooling compared to the kinetics of isomerizing to the most preferred structure at 10 K. These results show that for barrier heights above 3000 $cm^{-1}$ isomerization rates are two slow to complete with cooling, preserving the room temperature structures. Furthermore, isomerization processes with barrier heights less than 3000 $cm^{-1}$ were found to lead to very minor and almost insignificant changes in geometry. By studying the dynamics between conformations in a well-defined system physical insight between isomerization rates, barrier heights, and the origin of the observed structures were realized

In the last project we demonstrate a 2-color IRMPD technique is adopted and applied for the first time to flexible bio-molecules. This technique was characterized and further developed as it portrays several favorable characteristics for structural and dynamical studies of gas phase peptide ions. IRMPD is the most broadly applicable technique in the field of ion spectroscopy however, as will be explained in chapter 3, results in spectra that do not fully describe the ion structure. With this technique we show the ability to attain linear spectra with similar integrity to the spectra attainable with IR-UV DR. We show the versatility of this technique in it ability to record conformer specific spectra under favorable conditions and the ability to record linear spectra at any trap temperature between 10 and 300 K. The latter is of great importance as there is not

technique that can currently record linear spectra at room temperature. The importance of this is highlighted by the large discrepancies between the linear and non-linear IR spectra recorded fro YGGFL. The latter of which will result in erroneous assignments. Furthermore, we demonstrate the unique ability to record spectra of ions prepared with internal energies near the dissociation limit. This opens the doorway to study the dynamic processes that occur as the ion begins to dissociate. Much of this work is new and the correct model systems must be chosen carefully to successfully monitor dynamic processes such as proton migration.

## 1.5 References


(1) Horisberger, J.; Lemas, V.; Kraehenbuhl, J.; Rossier, B. C. Structure-Function Relationship of Na,K-ATPase. *Annual Review of Physiology* **1991**, *53* (1), 565–584. https://doi.org/10.1146/annurev.ph.53.030191.003025.

(2) Strosberg, A. D. Structure/Function Relationship of Proteins Belonging to the Family of Receptors Coupled to GTP-Binding Proteins. In *EJB Reviews 1991*; Christen, P., Hofmann, E., Eds.; EJB Reviews 1991; Springer: Berlin, Heidelberg, 1992; pp 45–54. https://doi.org/10.1007/978-3-642-77200-9_4.

(3) Cooper, G. M. The Central Role of Enzymes as Biological Catalysts. *The Cell: A Molecular Approach. 2nd edition* **2000**.

(4) Prediction of Membrane Transport Proteins and Their Substrate Specificities Using Primary Sequence Information https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0100278

(5) Hayashi, T. Iron: Heme Proteins & Dioxygen Transport & Storage Based in Part on the Article Iron: Heme Proteins & Dioxygen Transport by Harold M. Goff Which Appeared in the Encyclopedia of Inorganic Chemistry, First Edition. In *Encyclopedia of Inorganic and Bioinorganic Chemistry*; American Cancer Society, 2011. https://doi.org/10.1002/9781119951438.eibc0097.

(6) Lodish, H.; Berk, A.; Zipursky, S. L.; Matsudaira, P.; Baltimore, D.; Darnell, J. Myosin: The Actin Motor Protein. *Molecular Cell Biology. 4th edition* **2000**.

(7) McCudden, C. R.; Hains, M. D.; Kimple, R. J.; Siderovski, D. P.; Willard, F. S. G-Protein Signaling: Back to the Future. *Cell Mol Life Sci* **2005**, *62* (5), 551–577. https://doi.org/10.1007/s00018-004-4462-3.

(8) Rehm, T.; Huber, R.; Holak, T. A. Application of NMR in Structural Proteomics: Screening for Proteins Amenable to Structural Analysis. *Structure* **2002**, *10* (12), 1613–1618. https://doi.org/10.1016/S0969-2126(02)00894-8.

(9) Shortle, D. R. Structural Analysis of Non-Native States of Proteins by NMR Methods. *Current Opinion in Structural Biology* **1996**, *6* (1), 24–30. https://doi.org/10.1016/S0959-440X(96)80091-1.

(10) Flack, H. D.; Bernardinelli, G. The Use of X-Ray Crystallography to Determine Absolute Configuration. *Chirality* **2008**, *20* (5), 681–690. https://doi.org/10.1002/chir.20473.

(11) Danev, R.; Yanagisawa, H.; Kikkawa, M. Cryo-Electron Microscopy Methodology: Current Aspects and Future Directions. *Trends in Biochemical Sciences* **2019**, *44* (10), 837–848. https://doi.org/10.1016/j.tibs.2019.04.008.

(12) Marion, D. An Introduction to Biological NMR Spectroscopy. *Mol. Cell Proteomics* **2013**, *12* (11), 3006–3025. https://doi.org/10.1074/mcp.O113.030239.

(13) Konrat, R. NMR Contributions to Structural Dynamics Studies of Intrinsically Disordered Proteins. *Journal of Magnetic Resonance* **2014**, *241*, 74–85. https://doi.org/10.1016/j.jmr.2013.11.011.

(14) Nielsen, B. G.; Jensen, M. Ø.; Bohr, H. G. The Probability Distribution of Side-Chain Conformations in [Leu] and [Met]Enkephalin Determines the Potency and Selectivity to μ and δ Opiate Receptors. *Peptide Science* **2003**, *71* (5), 577–592. https://doi.org/10.1002/bip.10539.

(15) Heck, A. J. R. Native Mass Spectrometry: A Bridge between Interactomics and Structural Biology. *Nat. Methods* **2008**, *5* (11), 927–933. https://doi.org/10.1038/nmeth.1265.

(16) Leney, A. C.; Heck, A. J. R. Native Mass Spectrometry: What Is in the Name? *J Am Soc Mass Spectrom* **2017**, *28* (1), 5–13. https://doi.org/10.1007/s13361-016-1545-3.

(17) Mikhailov, V. A.; Mize, T. H.; Benesch, J. L. P.; Robinson, C. V. Mass-Selective Soft-Landing of Protein Assemblies with Controlled Landing Energies. *Anal. Chem.* **2014**, *86* (16), 8321–8328. https://doi.org/10.1021/ac5018327.

(18) Shi, L.; Holliday, A. E.; Bohrer, B. C.; Kim, D.; Servage, K. A.; Russell, D. H.; Clemmer, D. E. "Wet" versus "Dry" Folding of Polyproline. *J Am Soc Mass Spectrom* **2016**, *27* (6), 1037–1047. https://doi.org/10.1007/s13361-016-1372-6.

(19) Guan, J.-Q.; Chance, M. R. Structural Proteomics of Macromolecular Assemblies Using Oxidative Footprinting and Mass Spectrometry. *Trends in Biochemical Sciences* **2005**, *30* (10), 583–592. https://doi.org/10.1016/j.tibs.2005.08.007.

(20) Konermann, L.; Pan, J.; Liu, Y.-H. Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem Soc Rev* **2011**, *40* (3), 1224–1234. https://doi.org/10.1039/c0cs00113a.

(21) Tang, X.; Bruce, J. E. Chemical Cross-Linking for Protein-Protein Interaction Studies. *Methods Mol. Biol.* **2009**, *492*, 283–293. https://doi.org/10.1007/978-1-59745-493-3_17.

(22) Wysocki, V. H.; Joyce, K. E.; Jones, C. M.; Beardsley, R. L. Surface-Induced Dissociation of Small Molecules, Peptides, and Non-Covalent Protein Complexes. *J Am Soc Mass Spectrom* **2008**, *19* (2), 190–208. https://doi.org/10.1016/j.jasms.2007.11.005.

(23) Wells, J. M.; McLuckey, S. A. Collision-Induced Dissociation (CID) of Peptides and Proteins. *Meth. Enzymol.* **2005**, *402*, 148–185. https://doi.org/10.1016/S0076-6879(05)02005-7.

(24) Kim, M.-S.; Pandey, A. Electron Transfer Dissociation Mass Spectrometry in Proteomics. *Proteomics* **2012**, *12* (0), 530–542. https://doi.org/10.1002/pmic.201100517.

(25) Zubarev, R. A. Electron-Capture Dissociation Tandem Mass Spectrometry. *Curr. Opin. Biotechnol.* **2004**, *15* (1), 12–16. https://doi.org/10.1016/j.copbio.2003.12.002.

(26) Eiceman, G. A.; Karpas, Z.; Jr, H. H. H.; Karpas, Z.; Jr, H. H. H. *Ion Mobility Spectrometry*; CRC Press, 2013. https://doi.org/10.1201/b16109.

(27) Brodbelt, J. S. Photodissociation Mass Spectrometry: New Tools for Characterization of Biological Molecules. *Chem Soc Rev* **2014**, *43* (8), 2757–2783. https://doi.org/10.1039/c3cs60444f.

(28) Dixit, S. M.; Polasky, D. A.; Ruotolo, B. T. Collision Induced Unfolding of Isolated Proteins in the Gas Phase: Past, Present, and Future. *Curr Opin Chem Biol* **2018**, *42*, 93–100. https://doi.org/10.1016/j.cbpa.2017.11.010.

(29) Boyarkin, O. V. Cold Ion Spectroscopy for Structural Identifications of Biomolecules. *International Reviews in Physical Chemistry* **2018**, *37* (3–4), 559–606. https://doi.org/10.1080/0144235X.2018.1547453.

(30) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. Spectroscopic Studies of Cold, Gas-Phase Biomolecular Ions. *International Reviews in Physical Chemistry* **2009**, *28* (3), 481–515. https://doi.org/10.1080/01442350903069931.

(31) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V. Highly Resolved Spectra of Gas-Phase Gramicidin S: A Benchmark for Peptide Structure Calculations. *J. Am. Chem. Soc.* **2010**, *132* (12), 4040–4041. https://doi.org/10.1021/ja910118j.

(32) Burke, N. L.; Redwine, J. G.; Dean, J. C.; McLuckey, S. A.; Zwier, T. S. UV and IR Spectroscopy of Cold Protonated Leucine Enkephalin. *International Journal of Mass Spectrometry* **2015**, *378*, 196–205. https://doi.org/10.1016/j.ijms.2014.08.012.

(33) Mucha, E.; Marianski, M.; Xu, F.-F.; Thomas, D. A.; Meijer, G.; Helden, G. von; Seeberger, P. H.; Pagel, K. Unravelling the Structure of Glycosyl Cations via Cold-Ion Infrared Spectroscopy. *Nat Commun* **2018**, *9* (1), 1–5. https://doi.org/10.1038/s41467-018-06764-3.

(34) C. Polfer, N.; Oomens, J.; C. Dunbar, R. IRMPD Spectroscopy of Metal-Ion/Tryptophan Complexes. *Physical Chemistry Chemical Physics* **2006**, *8* (23), 2744–2751. https://doi.org/10.1039/B603665A.

(35) Douberly, G. E.; Walters, R. S.; Cui, J.; Jordan, K. D.; Duncan, M. A. Infrared Spectroscopy of Small Protonated Water Clusters, H+(H2O)n (n = 2−5): Isomers, Argon Tagging, and Deuteration. *J. Phys. Chem. A* **2010**, *114* (13), 4570–4579. https://doi.org/10.1021/jp100778s.

(36) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. Spectroscopy and Conformational Preferences of Gas-Phase Helices. *Phys. Chem. Chem. Phys.* **2008**, *11* (1), 125–132. https://doi.org/10.1039/B814143F.

(37) Fischer, K. C.; Sherman, S. L.; Voss, J. M.; Zhou, J.; Garand, E. Microsolvation Structures of Protonated Glycine and L-Alanine. *J. Phys. Chem. A* **2019**, *123* (15), 3355–3366. https://doi.org/10.1021/acs.jpca.9b01578.

(38) Ben Faleh, A.; Warnke, S.; Rizzo, T. R. Combining Ultrahigh-Resolution Ion-Mobility Spectrometry with Cryogenic Infrared Spectroscopy for the Analysis of Glycan Mixtures. *Anal. Chem.* **2019**, *91* (7), 4876–4882. https://doi.org/10.1021/acs.analchem.9b00659.

(39) Goebbert, D. J.; Wende, T.; Bergmann, R.; Meijer, G.; Asmis, K. R. Messenger-Tagging Electrosprayed Ions: Vibrational Spectroscopy of Suberate Dianions. *J. Phys. Chem. A* **2009**, *113* (20), 5874–5880. https://doi.org/10.1021/jp809390x.

(40) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. Gas-Phase Infrared Multiple Photon Dissociation Spectroscopy of Mass-Selected Molecular Ions. *International Journal of Mass Spectrometry* **2006**, *254* (1), 1–19. https://doi.org/10.1016/j.ijms.2006.05.009.

(41) *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*; Rijs, A. M., Oomens, J., Eds.; Topics in Current Chemistry; Springer International Publishing: Cham, 2015; Vol. 364. https://doi.org/10.1007/978-3-319-19204-8.

# CHAPTER 2. EXPERIMENTAL APPARATUS

## 2.1 Instrument for Cryogenic Ion Spectroscopy

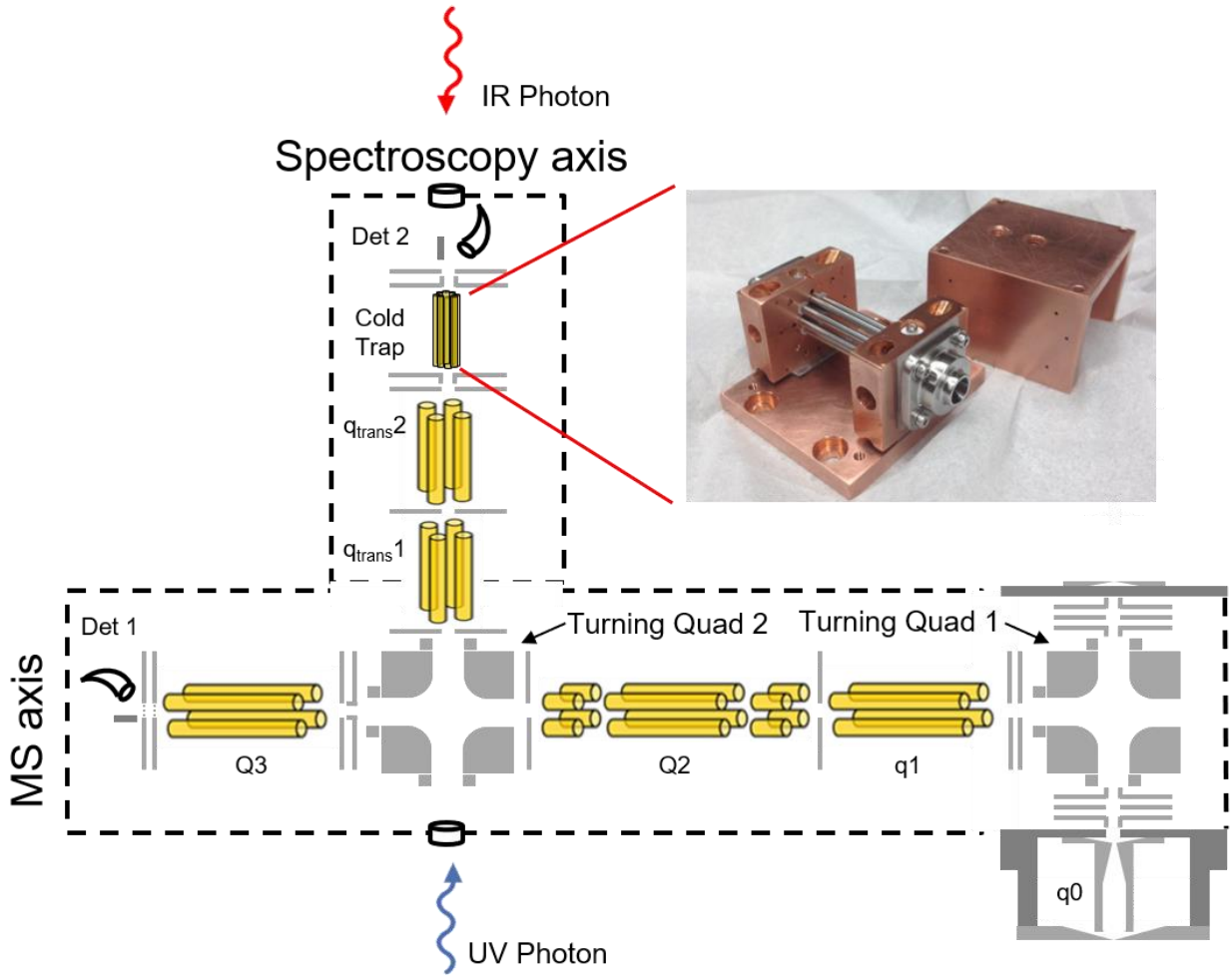


Figure 2.1: Schematic of cryogenic ion spectroscopy instrument

All of the work described in this thesis has been collected using a home built mass spectrometer that has been modified for spectroscopy. The instrument is comprised of two axes, a triple quadrupole mass spectrometer, the mass spectrometry axis, and an orthogonal spectroscopy axis. The schematic for this instrument is shown in Fig. 2.1. A simple description of experiment is as follows, ions are generated using nano-electrospray ionization and subsequently trapped in q0, a quadrupole ion trap. After the q0 ion trap is filled for the specified amount of time the ions are

transferred to Q2, another quadrupole ion trap. Here the ion of interest is mass isolated using RF DC apex isolation. The mass selected ion is turned down the spectroscopy axis via turning quadrupole 2 and guided in the cold trap. The cold trap is an octupole ion trap that is mounted in a copper housing. The trap is held at a temperature of 4 K via a closed cycle helium cryostat (Sumitomo Heavy Industries, Tokyo Japan). Prior to the ions entering the trap, a pulse of helium fills the ion trap housing and acts as a buffer gas. As the helium collides with the copper housing it is cooled to the trap temperature. As the ions enter the octupole ion trap, they undergo a series of collisions with the helium buffer gas and are in turn cooled to a final temperature of 10 K. At this temperature the ion exists in their vibrational zero-point level, allowing them to be probed using IR-UV DR techniques. This technique will be explained in detail in a following section but requires the cryo-cooled ion to be UV excited. Excitation at a resonant electronic transition imparts sufficient energy into the ion such that it fragments. The fragments created via UV excitation are referred to as photofragments. After photofragmentation has been induced the newly formed fragment ions along with residual precursor ions are extracted back down the spectroscopy axis and turned into Q3 via turning quadrupole 2. Q3 is another quadrupole ion trap and can be operated in two modes, a mass spectrometry mode, and a spectroscopy mode. In spectroscopy mode, the precursor ions are ejected out of the trap using a supplemental AC waveform. The remaining photofragments are then extracted, in a non-mass selective manner, onto the channeltron detector. In spectroscopy mode this sequence of events is repeated as a function of the UV laser's wavelength at a rep rate of 10 Hz. By monitoring the photofragment signal as a function of the UV wavelength it is possible to record the 10 K electronic spectrum of the ion that was mass selected in Q2. An example of a 10 K electronic spectrum for the peptide system $[YGPAA+H]^+$ is shown in Fig. 2.2. This spectrum was recorded by stepping the laser wavelength by 0.005 nm (UV wavelength) at a rep rate of 10 Hz across the entire spectral region. At each wavelength the photofragmentation signal was recorded. In mass spectrometry mode mass selective axial ejection is used to record the mass spectrum of the photofragment ions such that they can be identified. As previously stated, this process is repeated at a rep rate of 10 Hz as determined by the firing rate of the UV and IR lasers employed. In order for the entire process to be completed in a 100 ms time frame an overlap scan function is used. In the true sequence of events there are multiple ion packets

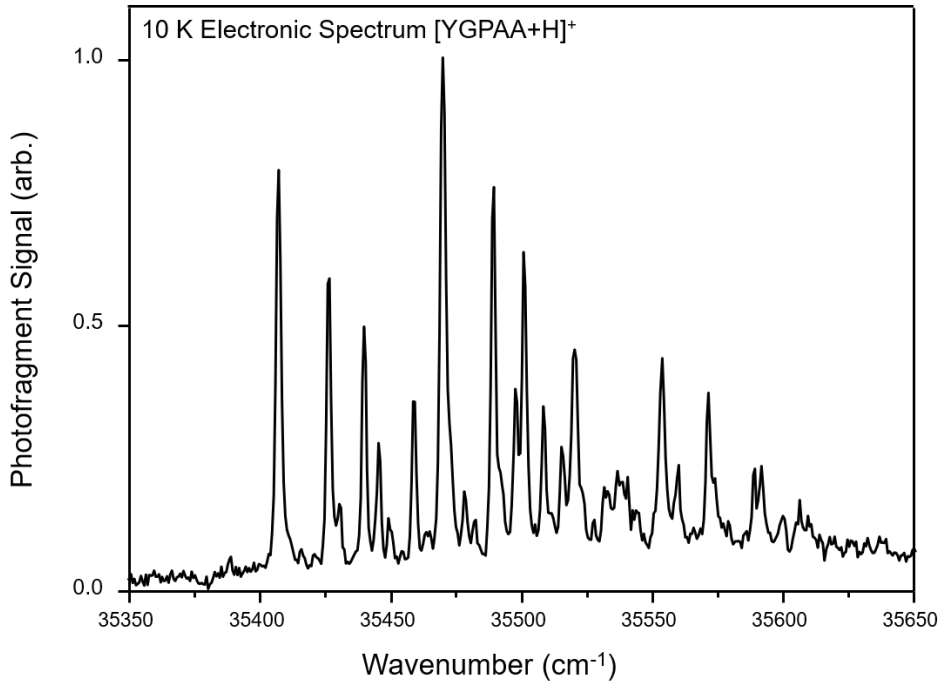


Figure 2.2: Example of 10 K electronic spectrum for $[YGPAA+H]^+$ system

being shuffled around the instrument at a single time. As ions are being ejected and extracted out of Q3, a new ion packet is being sent from Q2 to the cold trap to be photo-dissociated. As ions as ions are leaving Q2, new ions are being transfer to Q2 from q0, and in a similar manner as ions leave q0 new ions are being loaded. Thus, in a single scan function there are 4 individual ion packets at different stages of the experiment. A pictorial representation of the scan function shown in Fig. 2.3. The different colored lines represent the various digital triggers used during the course of the experiment. The triggers initiate different parts of the experiment by switching the voltage, between two predefined values, applied to the various ion optics.

This text assumes the reader is well acquainted with quadrupole theory. The following is a brief description. Quadrupoles are one of the most common mass analyzers in mass spectrometry. They consist of four rods to which an RF voltage is applied to each. Adjacent rods have the RF 180° out of phase with one another such that two rods have a positive phase, and the other two

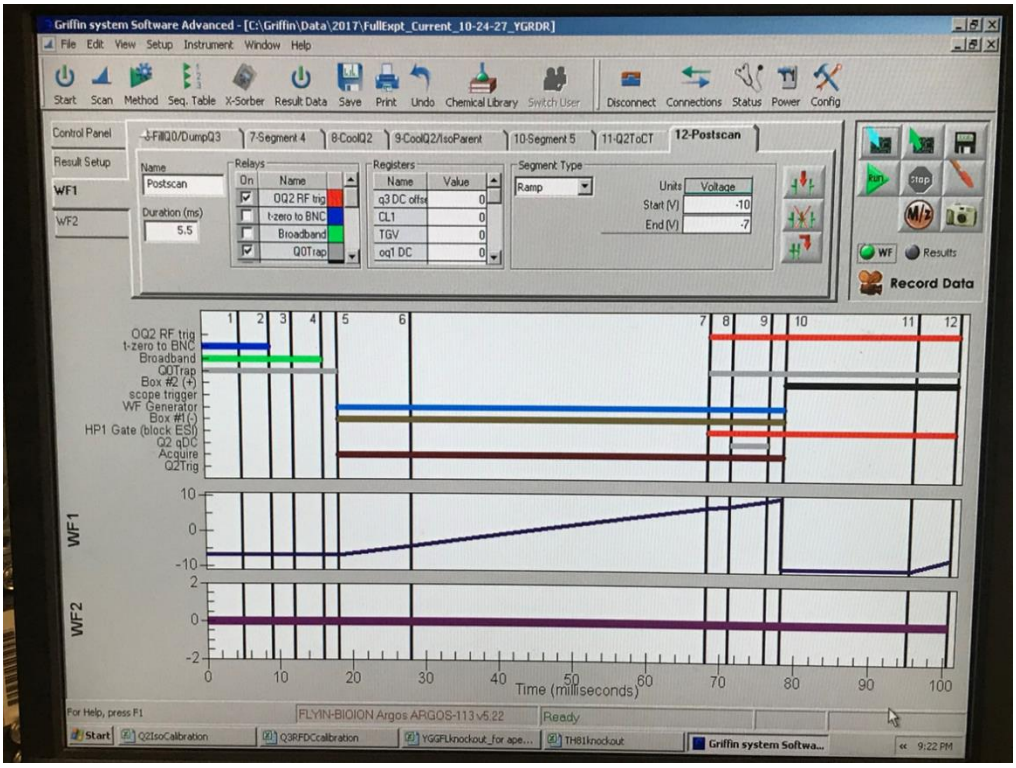


Figure 2.3: Pictorial representation of the overlapping scan function used in the 10 Hz experiment

have a negative phase. In addition to RF, DC is also applied to each rod, the adjacent rods have opposite signs of DC applied to them. The frequency of the RF in the quadrupoles that are used as a mass analyzer (Q2 and Q3) is 816 kHz and are operated in a voltage range of a few 100 volts to a max of 3000 volts peak to peak. As ions enter a quadrupole their translational kinetic energy can be dampened via collisions with buffer gas. This prevents incoming ions from leaving through the other side of the quadrupole in addition to the trapping potentials that are applied to the lenses on either side of the quadrupole. The RF voltage applied to the rods create a quadrupolar field that radially confines the ions. When in this field ions begin to oscillate at a frequency that is related to their mass to charge (m/z) ratio. At a particular RF amplitude, a range of m/z values are stable giving rise to a region of stability.

Since ions of various mass to charges oscillate at different frequencies it is possible to manipulate them such that they become unstable. This is how the precursor ions in Q3 are ejected from the trap when in spectroscopy mode. By applying a supplemental AC frequency to the Q3 rod set it is possible to increase the kinetic energy of ions that oscillate only at that frequency. In this case a waveform that contains the secular frequency of the precursor ion is applied to the Q3

rod set such that ions of a single mass are ejected from the trap, while the remaining fragment ions retain stable trajectories. In Q2 it is necessary to isolate a single mass to charge while ejecting all other masses. For this the voltage of the RF and DC amplitudes can be adjusted such that only a single mass has a stable trajectory.

### 2.1.1 Instrument Upgrade – New Source

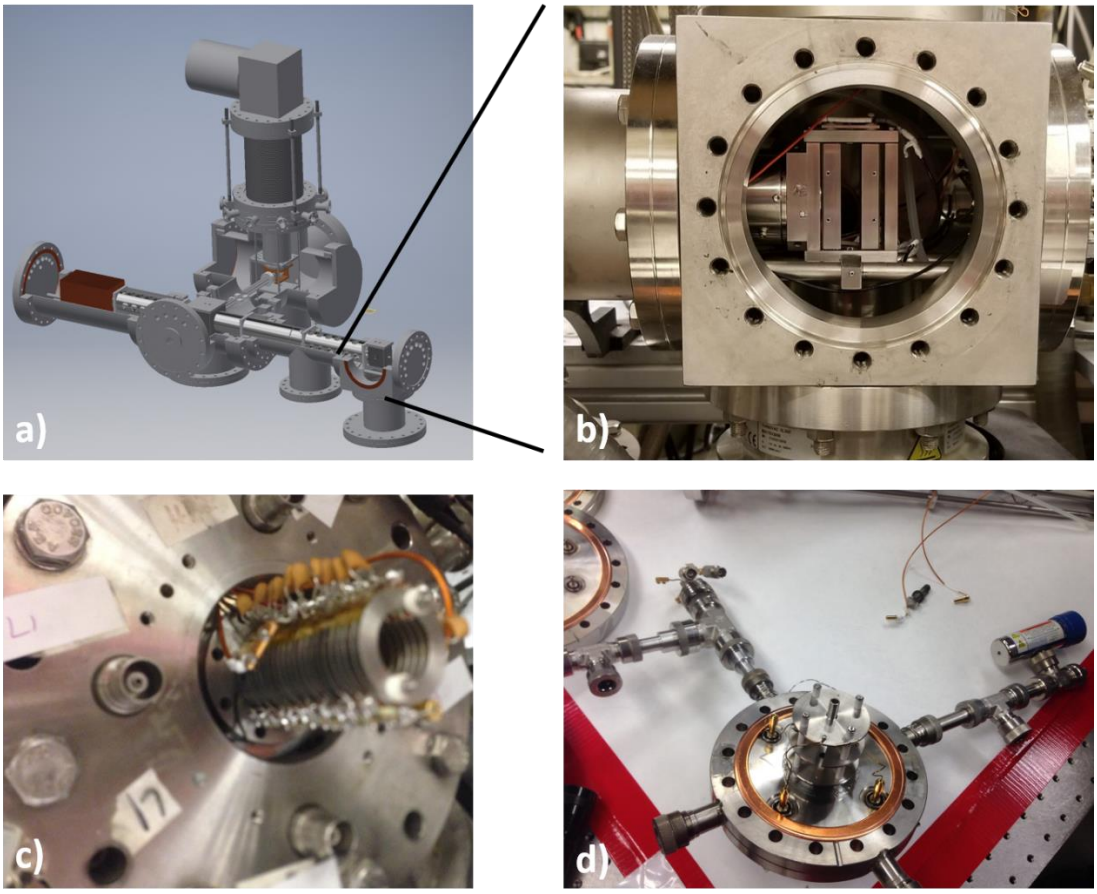


Figure 2.4: AutoCad drawing of the instrument (a), photograph of the turning first turning quadrupole (b), old funnel design (c), einzel (d)

Several major upgrades to the instrument have been made between 2015 and 2019. The first is the upgrade to the source. The AutoCAD drawing of the instrument is shown in Fig 2.4a. The source region is the set of ion optics that guides ions as they transition from atmosphere into the vacuum chamber. As ions make their way through the source, they are guided into the first turning

quadrupole, shown in Fig. 2.4b. The initial source was an ion funnel, Fig. 2.4c followed by an einzel stack, Fig. 2.4d. In this geometry the einzel stack was mounted on one side of the flange and led up to the turning quadrupole. The ion funnel was mounted on the other side of the same flange. This funnel was a miniaturized version of the funnel described by Jarrold. While it was possible to transmit ions through the ion funnel, the signal was often quite low and unstable. This made recording data quite difficult as a stable ion signal is needed in order to generate a UV and IR spectra.

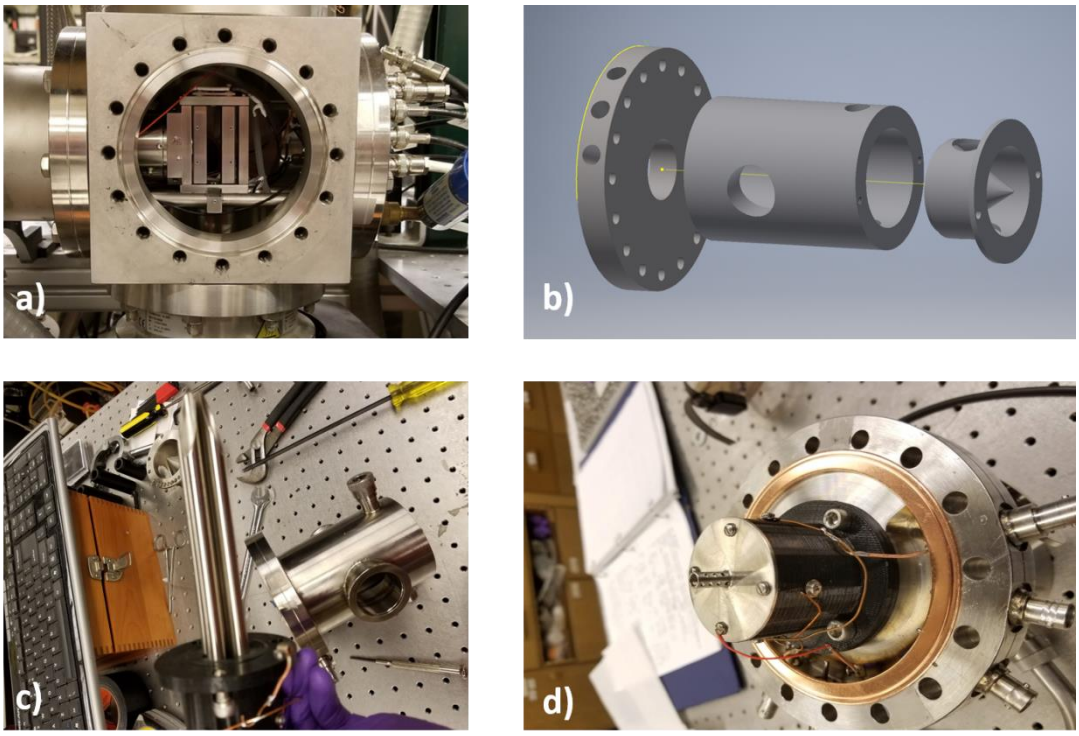


Figure 2.5: Photograph of turning quadrupole 1 (a), q0 rod set and plastic mount (c), partially hollowed flange, tube lens, and electrical connections (d), AutoCAD drawing of q0 housing and skimmer

This source geometry was replaced with a tapered quadrupole rod set taken from an old Sciex instrument. Fig. 2.5c shows the tapper rod set, fastened in a 3-D printed mount. The mount was printed using plastic that was compatible was vacuum compatible such that it does not collapse on itself when under vacuum. Tapped holes in the plastic mount allows the individual rods to be fastened sturdily into the mount with minimal motion. The rod set already had holes tapped from the manufacturer. The screws that fastened the rod set to the mount were also used to electrically

connect the opposing rods as shown in Fig. 2.5d. A tube lens that leads directly up to turning quadrupole 1 was also attached to the end of the mount, Fig. 2.5d, to act as a trapping lens. A hole was made through a standard 6-inch flange such that the mount could be screwed into the backside of the flange as shown in Fig. 2.5d. The AutoCAD drawing of the flange, the metal sleeve, and the skimmer is shown in Fig. 2.5b. The feedthroughs for the electrical connections were made through the edge of the flange, as shown in Fig. 2.5b and d. The side of the flange that is on the inside of the vacuum chamber was partially hollowed out such that inline connectors could easily fit over the ends of the feedthroughs, Fig. 2.5d. A custom sleeve was made to fit over the rod set and was welded to the flange. The skimmer that came with the rod set fit into this sleeve and contacted a lip on the inside of the sleeve. An O-ring placed between the lip and the skimmer allowed the sleeve to be differentially pumped. Two KF-40 connections were made to pump on the part of the sleeve behind the skimmer, while two KF-25 holes were made to pump on the skimmer. The entire sleeve assembly welded to the flange can be seen in the background of Fig. 2.5c. A plate with a small orifice is placed on top of the skimmer to complete the new source design.

Signal stability in addition to total ion count improved tremendously with this new design. Switching between samples, maintaining cleanliness of the ion optics, and tuning for signal became routine with the new source. Data sets that would normally take 2 weeks could be taken in 2 days. Much of the overall success of this geometry compared to the funnel design may be from the fact that the ions can be trapped and thermalized prior to being turned by turning quadrupole 1. This allows the ions to fully desolvate and be released from q0 with well-defined kinetic energies. In the funnel design, ions would pass through the orifice and get focused down through the RF ring electrodes and be turned immediately. It is likely that not only a small portion of the ion were desolvated by the time they reached the end of the funnel as a counter-current drying gas or heated capillary is not used in the process of creating ions. This, in addition to a spread in kinetic energies may have led to poor focusing conditions with the funnel design. With the q0 in place it is possible to saturate the trap in about 60 ms with most samples. This ensures that a steady ion signal is maintained through the course of the experiment.

### 2.1.2 Instrument Upgrade – Commercial RF Power Supplies

Commercial RF power supplies from an old Sciex instrument was also adapted to the instrument in addition to the q0 rod set. The commercial power supplies provide the ability to

perform apex isolation in Q2 and Q3, the ability to obtain resolving powers comparable to a commercial triple quadrupole mass spectrometer, and the ability to perform monoisotopic isolation. Additionally, it is also now possible to study larger systems as the RF amplitudes in Q2 and Q3 can be extended to 3000 V peak-to-peak. Much of this work was performed by Dr. Eric Dziekonski from Dr. McLuckey's group and the details of how to adapt the commercial power supplies onto the home-built instrument can be found in his thesis.

The ability to perform apex isolation in Q2 makes the overall experiment progress much faster. Previously, a SWIFT type waveform was used to isolate the precursor ion of interest. The software used for this process was quite slow at building the waveform and the correct placement of the notch was found by trial and error. With the apex isolation the isolation parameters are still found mostly through trial and error however, the parameters are easily found. The RF and DC amplitudes can be scanned in real time using reference voltages in from two independent waveform generators allowing for quick isolations to be performed.

### 2.1.3 Instrument Upgrade – Transport Quadrupoles

Once the ions leave Q2 they are tuned down the spectroscopy axis and focusing using the two transport quadrupole labelled $q_{trans}1$ and 2. These rods sets were home built and only serve as focusing optics to guide ions. The copper wires used to electrically connect the rods to the homebuilt RF power supplies were a thin gauge wire, limiting the max voltage that can be applied to the rods. It was found that there was a mass dependent focusing issue when ions leave the cold trap and are guided back down the spectroscopy axis to be turned into Q3. As ions leave the cold trap the different masses are ejected from the cold trap with roughly the same kinetic energy. This causes ions of different masses to have different velocities. As different mass ions travel at different velocities, they spend different amounts of time in the transport quadrupole and in turn get focused slightly differently. This would not be a problem if the ions were trapped on the other side of the transport quadrupoles, as they are going into the cold trap, however when they are leaving the spectroscopy axis, they are turned by turning quadrupole 2. Stable trajectories through the turning quadrupole is sensitive to how the ions are focusing as they enter the turning quadrupole. The different focusing conditions ultimately leads to a mass discrimination in the ions that make it to Q3.

A remedy to this problem would be to transport ions with smaller kinetic energies, leaving from the cold trap. Much effort has gone into changing the float voltage of the cold trap to release the ions with smaller kinetic energies however, a large decrease in signal is observed when the cold trap voltage is less than 7 volts. Another remedy to this problem is to increase the RF amplitude on the transport rod sets, however because of the wiring there is a limit to voltage that can be applied. The wiring of transport quadrupole 1 was changed to a thicker gauge wire because it is accessible with minimal dismantling of the instrument. A homebuilt switch box was also built in JAFCI that allows the voltage applied to this rod set to be switched between two different reference voltages. This ultimately allows the precursor ions the be focused going into the cold trap and the lighter fragment ions to be focused efficiently when leaving the spectroscopy axis. No adjustments have been made to transport quadrupole 2 since it is not easily accessible. Under normal operation transport quadrupole 2 is set at a single voltage while transport quadrupole 1 is changed between two voltages. The ability to switch between RF voltages applied to a single rod set seems to be fine in terms of allowing stable trajectories for the lighter fragment ions.

# CHAPTER 3. TECHNIQUES AND METHODS

## 3.1 Messenger Tagging

While messenger tagging is not used at all in this work it will be described for completeness as the other two spectroscopic methods will described in detail. Briefly, for messenger tagging ions are trapped in a cryogenically held 3-D quadrupole, attaining final temperatures anywhere between 10-30 K. At these temperatures it is possible to form non-covalent complexes between the ion and neutral inert gases such as helium, nitrogen, or argon. The complexes are held together through Van Der Waals forces which are too weak to keep the complex together at thermal energies. Once the ion is tagged, the mass of the complex is $[M+tag]^+$, where M is the mass of the ion. Through the use of a reflectron time of flight is it possible to monitor the complexes' flight time which is proportional to its mass. Once this is recorded, the complex can be irradiated with an IR laser in transit to the reflectron. If the laser is resonant with a vibrational transition the ion is able to absorb a photon. This imparts sufficient energy into the ion such that the tag is evaporated. The change in mass affects the total flight time which is measured as the bare ion hits the detector. Measuring the appearance of the lighter, bare ions as a function of the IR wavelength results in a linear action spectrum. In this method the action event is caused directly by the IR laser, and is operable under single photon conditions, which is highly favorable for maintaining the integrity of the resulting IR spectrum.

Since spectra are only recorded using an IR laser, tagging is a necessity. Typical onset dissociation energies for peptide bonds are about ~20,000 $cm^{-1}$ for pentapeptide systems. Energies at which the fragmentation rate is on the ms timescale (relevant for mass spectrometry) is about ~40,000 $cm^{-1}$. IR spectroscopy takes place in a region much lower in energy, between 1400-1800 $cm^{-1}$ and 2600-3700 $cm^{-1}$. Thus, to cause an action event with single IR absorption events it is necessary for ions to be cryo-cooled and tagged with inert gases. The binding energies of the tagged complexes are much lower than the energy of the incoming photons, allowing for mass changes to be readily observed.

## 3.2 Description of IR-UV Double Resonance Spectroscopy

The majority of the work in this thesis was recorded using IR-UV double resonance spectroscopy. This technique has long been employed to study neutral molecules in isolated gas phase environments and through the pioneering work of Rizzo and co-workers has been adopted to study ions. In this technique bare ions, without tags, are interrogated. As previously mentioned, the energy imparted by an IR photon is insufficient to break a chemical bond and without a tag does not produce a mass shift. Instead, in this technique a packet of trapped ions is irradiated with UV laser. A resonant UV photon imparts about 37,000 $cm^{-1}$ of internal energy into the absorbing ion, readily inducing chemical fragmentation. Since a UV laser is used to induce dissociation the samples to be analyzed must contain a UV chromophore. The wavelength of light used to induce dissociation is near 280 nm for tyrosine and 270 nm for phenylalanine. As stated previously, the formation of the fragment ion is detected in subsequent mass analysis step after the irradiation event, Q3, Fig. 1.1. By monitoring the appearance of the photofragment ions as a function of the UV wavelength a complete electronic spectrum can be recorded. As will be explained in more detail below, the relative UV fragmentation signal can then be modulated by irradiating the ion packet with an IR photon prior the UV irradiation. Monitoring the relative changes in the fragmentation yield as a function of the IR wavelength results in a vibrational spectrum.

### 3.2.1 IR-UV Double Resonance Spectroscopy – Electronic Spectra

Similar to messenger tagging it is essential that the ions are cryo-cooled prior to being probed. In this case cryo-cooling is necessary in order to suppress thermal broadening in the electronic spectrum. Fig. 3.1 compares the UV spectrum of $[YGPAA+H]^+$ taken at 300 K (a) and 10 K (b). Generally, UV excitation takes place from the ground state ($S_0$) of a UV active mode to the excited state ($S_1$) within the same normal mode. At room temperature, several vibrational levels of various low frequency modes can be populated, including the UV active modes. Under these conditions UV transitions can occur from all the populated vibrational levels in the ground state UV active modes to various vibrational levels in the excited state, Fig. 3.2a. This leads to significant thermal broadening as a UV transition is possible at nearly any given wavelength near the origin transition, Fig. 3.1a. By cryocooling ions to 10 K all of the vibrational population is brought down to its ground state (zero-point level). Under these conditions UV transitions are only possible from the

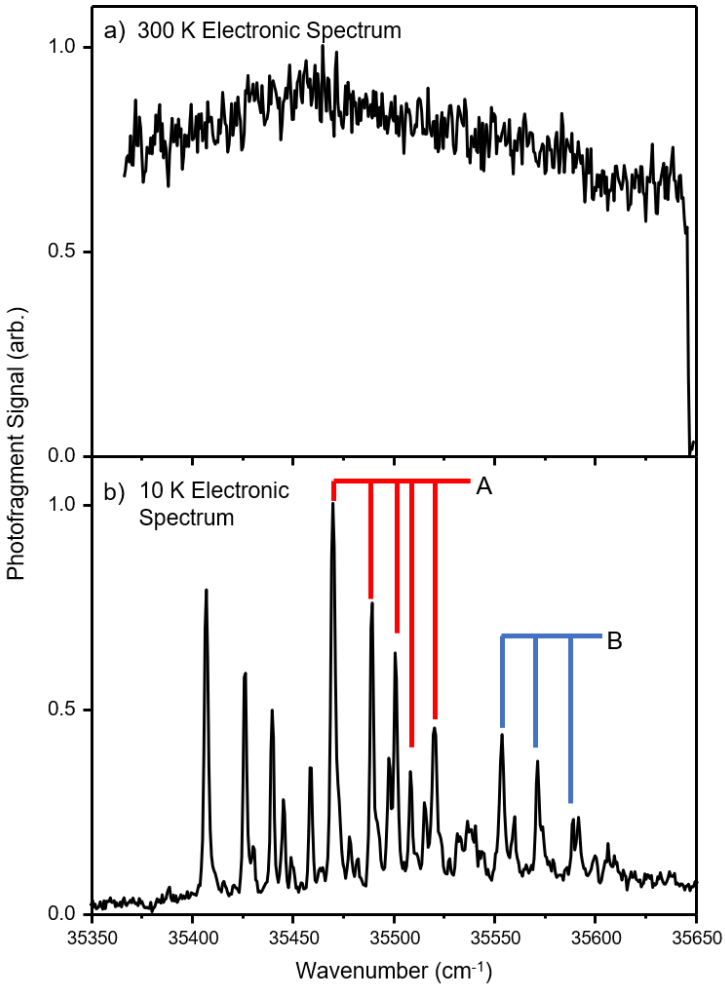


Figure 3.1: Electronic spectrum of $[YGPAA+H]^+$ taken at 300 K (a) and 10 K (b).

ground state of the UV active mode to the vibrational levels in the excited state, Fig. 3.2b. This results in a much simpler spectrum free from thermal congestion, as shown in Fig. 3.1b. Furthermore, under cryogenic conditions clear Franck Condon progressions can be seen in the experimental spectrum as the laser energy is scanned to higher wavelengths. This is illustrated in Fig 3.2b and shown experimentally with the red and blue drop lines in Fig. 3.1b. In this example the progressions are about 20 $cm^{-1}$. For perspective, at room temperature this mode is expected to be populated up to v=10. Another important feature in the 10 K electronic spectrum shown in Fig. 3.1b is that the origin transitions of different conformations are resolved from one another. Two specific progressions are highlighted with the red and blue lines and labeled A and B, respectively. Not all the conformations are labeled in this spectrum, but several different progressions are clearly observable and are attributed to different conformations. The following sections show the different

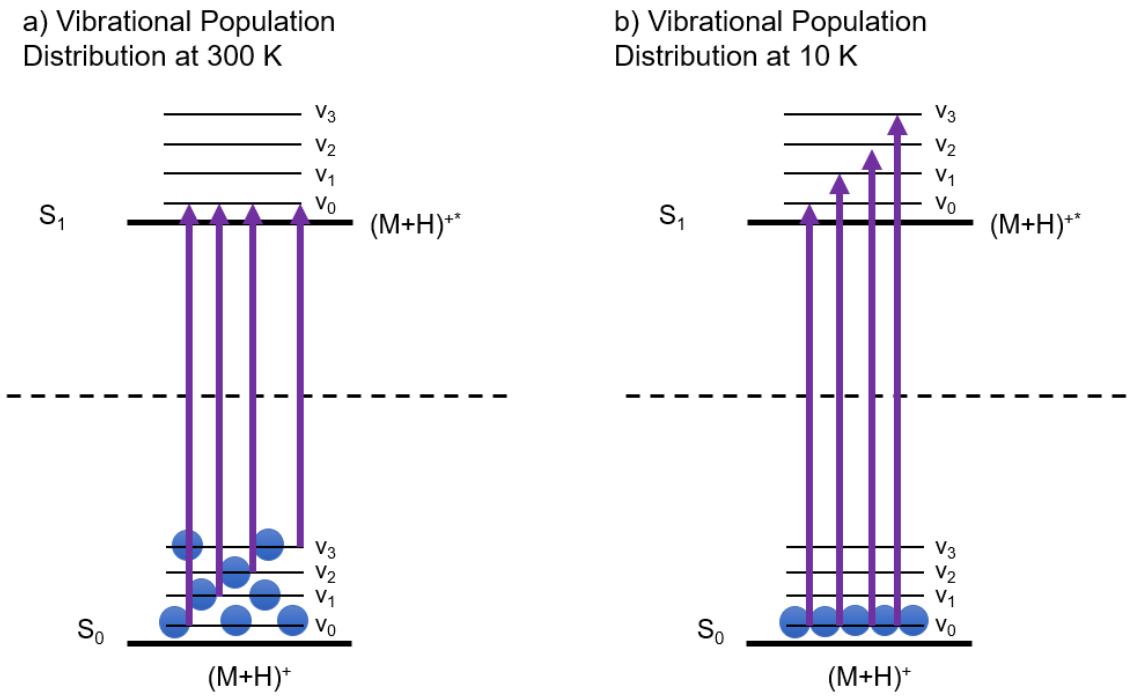


Figure 3.2: Illustration of the ground state population at 300 and 10 K (a) and (b)

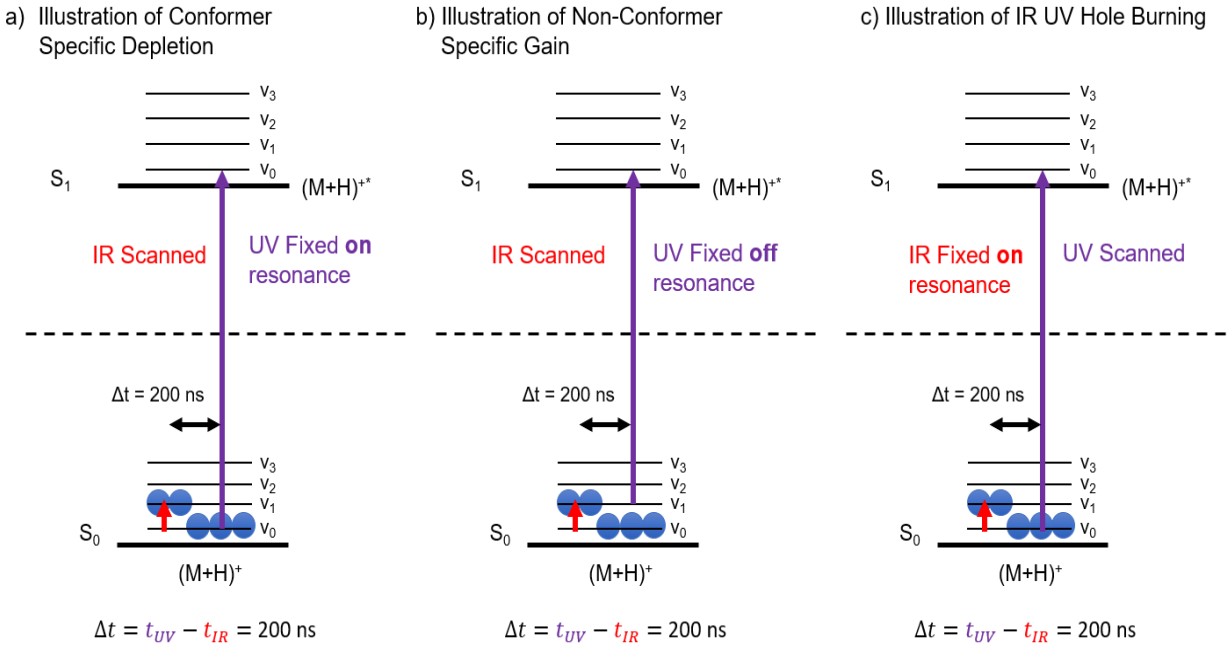


Figure 3.3: Schematic of common IR-UV DR techniques

types of scans that can be recorded using IR-UV double resonance spectroscopy. In most applications the IR laser first irradiates the ion packet followed by the UV laser. One laser is always fix at a predetermined frequency while the other is scanned. Depending on the which laser is scanned and which is fixed, and where it is fixed, several types of scans can be recorded. A schematic of all the scans are given in Fig. 3.3 and will be referred to as needed.

### 3.2.2 IR-UV Double Resonance – Conformer Specific IR Spectra

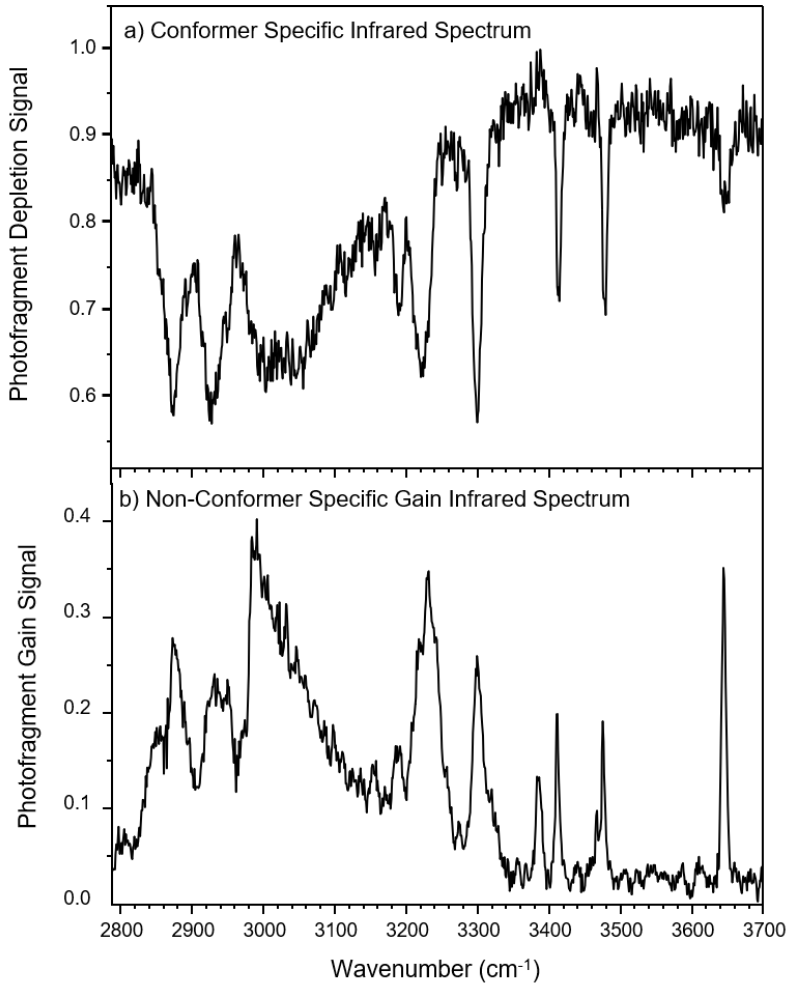


Figure 3.4: Conformer specific depletion and non-conformer specific gain spectra (a) and (b)

The conformer resolved 10 K electronic spectrum provides a means to record conformer specific IR spectra. As stated previously, the relative photofragment signal can be modulated via IR absorption prior to UV absorption. By fixing the UV laser on a specific electronic transition it

is possible to selectively monitor the fragmentation of a single conformation. If that conformation absorbs an IR photon prior to the UV photon it will gain internal energy. This removes some of the population from the vibrational zero-point level, as shown in Fig. 3.3a and results in an overall depletion in photofragment signal. By monitoring the depletion as a function of the IR wavelength it is possible to record a conformer specific spectrum, such as the one shown in Fig. 3.4a. This is the IR spectrum of the conformation responsible for the progression labeled A in the 10 K electronic spectrum, Fig. 3.1b. By fixing the UV wavelength on other electronic transitions present in the electronic spectrum it is possible to independently record the IR spectra of all the conformations present under the experimental conditions. These spectra are referred to as depletion spectra.

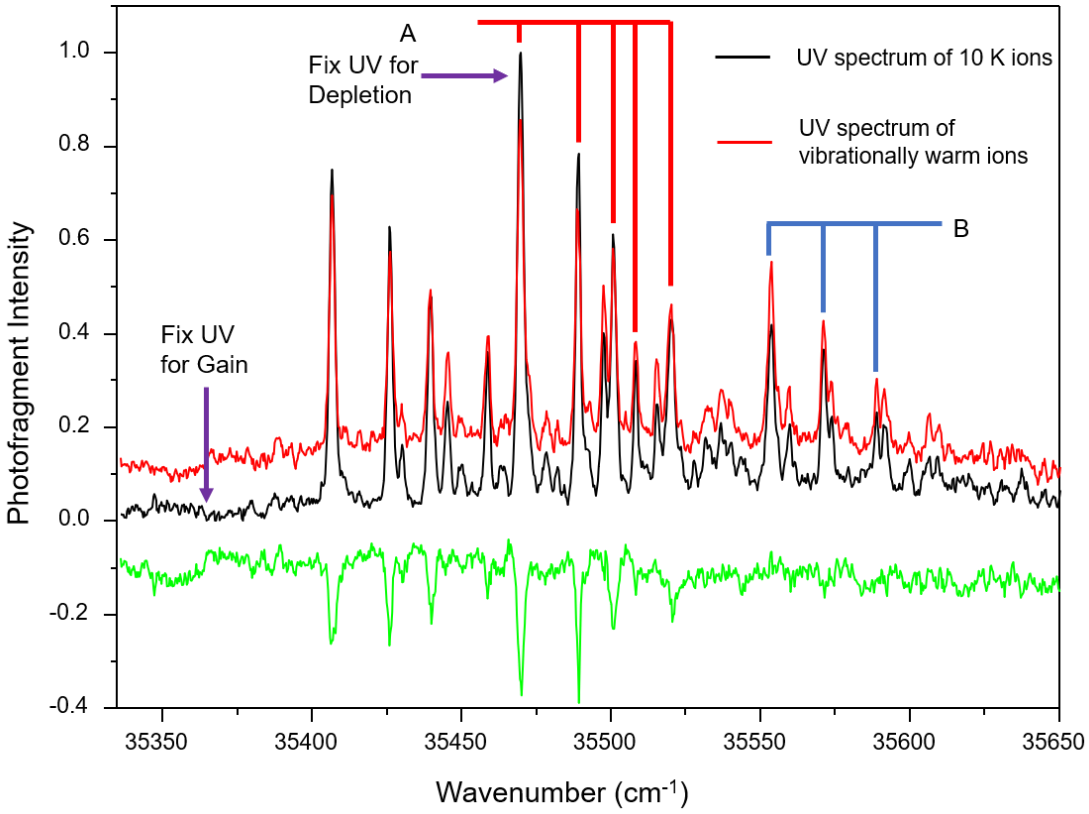


Figure 3.5: 10 K electronic spectrum of $[YAPAA+H]^+$ (black trace). Electronic spectrum response to absorbing IR photon (red trace). Hole burn spectrum obtained by subtracting black and red trace (green) after IR absorption (red trace), difference between the black and (green trace)

For perspective, the red trace in Fig. 3.5 shows how the UV spectrum responds to 10 K ions becoming vibrationally heated. In this figure the IR photon was fix on a resonant vibrational transition of conformation A. The UV laser was then scanned across the spectra region at some time delay after IR photon irradiation event, red trace Fig. 3.5. For reference the 10 K spectrum is superimposed in the background, black trace. There are two important features associated with the UV spectrum of the vibrationally heated ions. The first is the transitions that belong to conformation A experience an overall depletion. This is the depletion that is monitored when recording a conformer specific spectrum. Secondly, the UV spectrum appears broadened with increased intensity off resonance of where the 10 K ions absorb. In the following section it will be shown how to exploit this to generate a non-conformer specific spectrum.

### 3.2.3 IR-UV Double Resonance – Non-Conformer Specific Gain Spectrum

In addition to being able to generate conformer specific spectra it is also possible to record a non-conformer specific composite gain spectrum. The laser timings and overall physical process is the same as recording a depletion spectrum, however in this arrangement the UV laser is fixed at a wavelength slightly lower than the lowest energy origin transition in the electronic spectrum, as indicated by a purple arrow in Fig. 3.5. This process is graphically illustrated in Fig. 3.3b. At this frequency there is only a UV resonance for the vibrationally warm ions as the cold ions do not absorb. Fixing the UV laser at this frequency, whenever any conformation absorbs an IR photon a resonance will be created for the UV laser leading to the observation of photofragment signal. In this manner by scanning the IR laser and monitoring the appearance of photofragment signal as a function of the IR wavelength a composite IR spectrum can be recorded. An example of this type of spectrum is shown in Fig. 3.4b. A comparison of these two spectra show that the vibrational transitions of conformation A appears as well as some new resonances which belong to an addition conformer. In this $[YGPAA+H]^+$ example it is clear that there are multiple conformations based on the Franck Condon progressions that appear in Fig. 3.1b. However, in many cases interpretation of the UV spectra is not as easy. Generally, a composite spectrum is recorded to ensure that there are no conformations that go unaccounted.

### 3.2.4 IR-UV Double Resonance Hole Burning

As demonstrated in the last two sections fixing the UV either on or off resonance and scanning the IR laser will result in the either conformer specific or non-conformer specific spectra. In an alternative method referred to as IR-UV Hole Burning, the IR is fixed, and the UV is scanned. This method is illustrated in Fig. 3.3 C and has already been demonstrated in Fig. 3.5. As stated previously, by fixing the IR on a vibrational transition and scanning the UV laser at some short time delay afterwards it is possible to record how the UV spectrum responds to a subset of conformations becoming vibrationally excited. The red trace in Fig. 3.5 represents irradiating conformation A using a 3418 $cm^{-1}$ photon. The UV transitions of all the conformations that contain this vibrational resonance will broaden and partially be depleted. Again, this occurs because the IR photon removes some of the vibrational population from the vibrational zero-point level of these conformations. By subtracting the 10 K electronic spectrum, black trace, Fig. 3.5, from the spectrum of the vibrationally heated ions, red trace, Fig. 3.5, it is possible to obtain a hole burn spectrum, green trace Fig. 3.5. The hole burn spectrum indicates all the electronic transitions that share the particular vibrational frequency used to perform the hole burn. In this case all the transitions labeled A appear as a negative signal along with some other transitions lower in energy. The second set of transitions that burn out, as will be discussed in a later chapter, are representative of a very similar conformer. The transitions that are labelled B do not burn out as this conformation did not become vibrational heated. This technique is often applied to distinguish between electronic transitions belonging to different conformations in hard to interpret or congested electronic spectra.

In sections 3.3.1-3.3.3 it is demonstrated how removing population from the ground state prior to UV absorption can be used to record informative IR and UV spectra. The following sections will describe the types of modulation that can occur once the vibrational population is in the excited state (after UV excitation), and the types of spectra and information that can be obtained. However, the process of using IR and UV spectra to make structural assignments will first be discussed.

## 3.3 Sensitivity of IR Spectrum to Conformational Shape

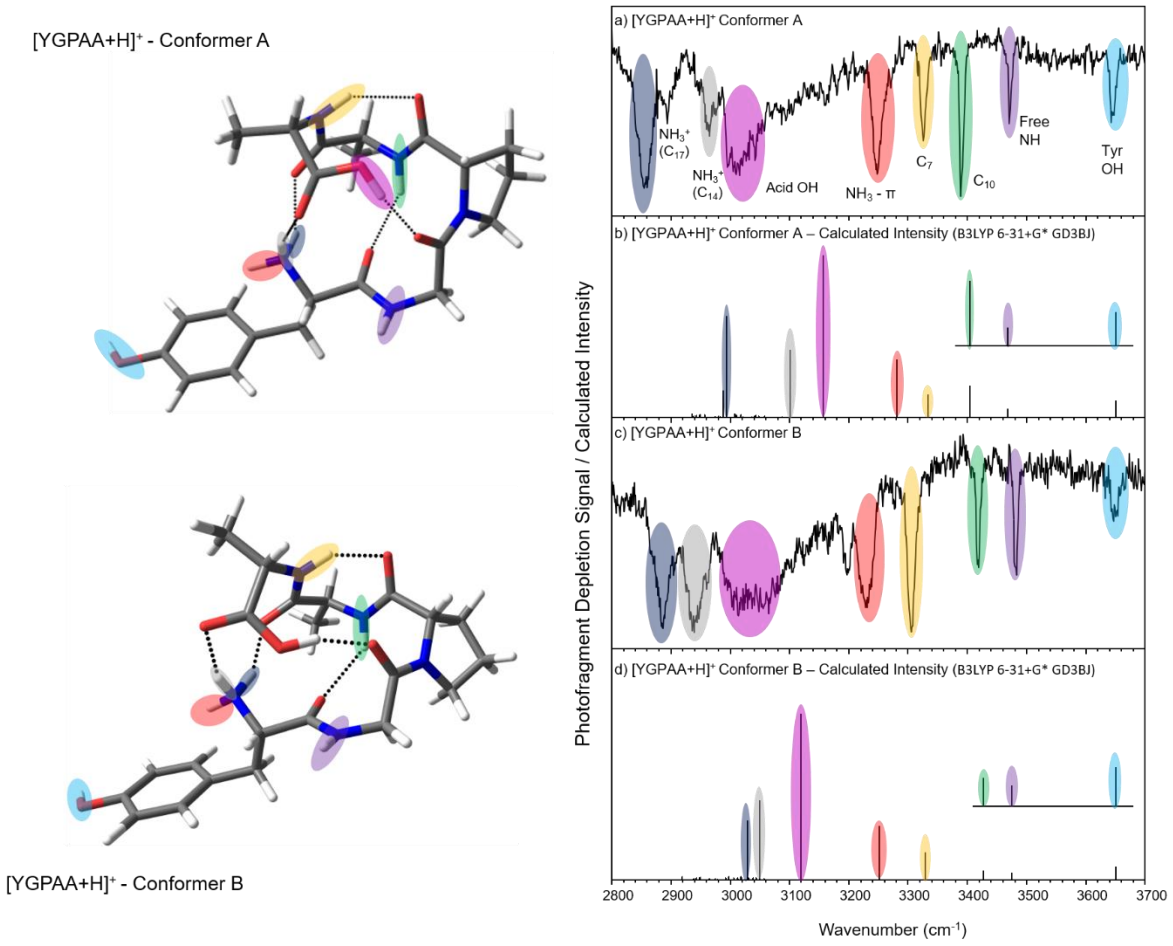


Figure 3.6: Assigned structures of $[YGPAA+H]^+$ (left). IR spectra of conformer A and B (a) and (c), respectively. Scaled calculated IR spectra of conformers A and B (b) and (d), respectively.

The IR spectra in the hydride stretch region of the two major conformations adopted by the $[YGPAA+H]^+$ system is displayed in Fig. 3.6a and c, along with the assigned structures shown on the left and their calculated vibrational frequencies, displayed in panels b and d. The frequency range shown in Fig. 3.6 reports on NH and OH fundamentals. Each fundamental is shaded with a different color in the structure and the corresponding vibrational bands are shaded with the same color in the experimental and calculated IR spectra.

The resulting IR spectra are sensitive to an ion's conformation because the vibrational frequency for each normal mode is dependent on its local hydrogen bonding environment. For example, the 3 NH oscillators shaded with purple, green, and yellow circles appear experimental

at different frequencies even though they each have the same fundamental frequency. The purple shaded NH is the least perturbed and appears near its fundamental frequency while the yellow and green NH's are shifted to a lower frequency. The magnitude of the shift is dependent on the relative strength of the hydrogen bond. The same is true for the OH and carbonyl oscillators making the IR spectrum highly sensitive to the ion's 3-D shape.

The sensitivity of the IR spectrum in distinguishing between conformations is highlighted by being able to differentiate between the two assigned structures shown in Fig. 3.6. Both structures are beta turns and possess the same C10, C7, C14, and C17 hydrogen bonds. The nomenclature for the hydrogen bonds is derived from the number of atoms that make up the cycle formed by hydrogen bonding partners. For example, the C10 hydrogen bond includes 10 atoms starting from the hydrogen of the green shaded NH going clockwise around to the carbonyl. The major difference between the two structures is the *cis* vs *trans* geometry of the carboxylic acid. While the hydrogen bonding pattern is preserved between the two structures the relative distance between some of the hydrogen bonding partners are slightly different. The most notable is the distance between the hydrogen bonding partners that make up the C10. In conformer B this hydrogen bond is slightly longer, making it weaker compared the C10 in conformer A. This slight difference is reflected experimentally as the C10 band in conformer B appears 30 $cm^{-1}$ higher in energy compared to the C10 band in conformer A. The ability for the experimental IR spectra to capture these subtle changes between structures along with the ability to differentiate between the structures using DFT makes IR-UV DR highly sensitive to the differences between structures.

## 3.4 Computational Details

While the IR spectrum is representative of an ion's conformation is it near impossible to make a structural assignment without the aid of molecular modeling and DFT. Assignments are generally made by first generating a list of candidate structures through a series of Monte Carlo searches performed using the MacroModel software. The search is performed using the multiple minimum method with either the Amber Star or OPL3 forcefields. The forcefield energies are often unreliable however, often times the assigned structure is found within the lowest 100 energy structures. The candidate structures are optimized using DFT with a specified basis set and level of theory. Calculation performed by Dr. DeBlase revealed that the most robust and inexpensive

method and basis set that would yield reliable results was B3LYP-GD3BJ/6-31+G*. The details of this calculations can be found in the supplemental materials of the cited work.

Once the geometries are optimized the vibrational frequencies of the structures are calculated. For each structure the calculated vibrational spectrum is scaled to account for anharmonicity and then plotted against the experimental spectrum. Different modes require different scaling factors, generally the NH's are scaled by 0.958, free carboxylic acid OH's by 0.973, tyrosine OH's by 0.975, carbonyl stretches and NH bends by 0.981, and arginine NH stretches by 0.96. Once the calculated vibrational spectra are scaled, they can be plotted against the experimental spectra to determine if the structure accurately recreates the experimental spectrum.

The calculated and scaled vibrational spectra for each conformer of YGPAA are plotted under their respective experimental IR spectrum in Fig. 3.6. Both calculated spectra shown a good agreement to the experimental spectrum. In both cases the modes that appear between 3200 and 3700 $cm^{-1}$ line up very closely to the experimental vibrational bands. The bands below 3200 $cm^{-1}$ however, consistently appear higher in frequency compared to the experimental bands. These bands are associated with the hydrogen bonds formed by the ammonium NH's and the carboxylic acid OH. The hydrogen bonds formed by these modes are generally very strong and are not accurately represented by the scale factor used for the amide NH stretches or the free acid OH stretch. Because of this the accuracy of the fit is judged not only by how closely the calculated bands are to the experimental frequencies but also on the overall pattern of the spectrum. Furthermore, the relative energies of the candidate structures are also considered when making the structural assignments. In this case both conformer A and B are the two lowest energy structures. Generally, if the assigned structure is not within 5 kJ/mol of the lowest energy structure additional Monte Carlo searches are performed.

One particular thing worth noting is that the vibrational spectra recorded at 10 K are not affected by thermal broadening and result in well resolved vibrational bands. This aids in making structural assignments as bands that are close in frequency can often be resolved from one another. Furthermore, recording spectra at 10 K also allows for conformer specific spectra to be recorded which further aids in distinguishing between conformations.

## 3.5 Fate of Electronically Excited Ions

The previous sections demonstrated some of the traditional IR-UV DR techniques and how to interpret and assign resulting IR spectra. In this section the fate of the electronically excited ions will be discussed and how manipulation of the excited state population can lead to large signal increases and provide a means to record excited state IR spectra. IR-UV-IR triple resonance techniques will also be discussed.

After UV excitation there are multiple modes of relaxation. The major pathways that occur on the time scale of the experiment include, fluorescence, intersystem crossing to a triple state, internal conversion back to the ground state, and coupling to a dissociative state from the excited state. Action spectroscopy uses the formation of fragment ions to detect an absorption event. In this case it is important understand the factors which govern photofragmentation. UV excitation

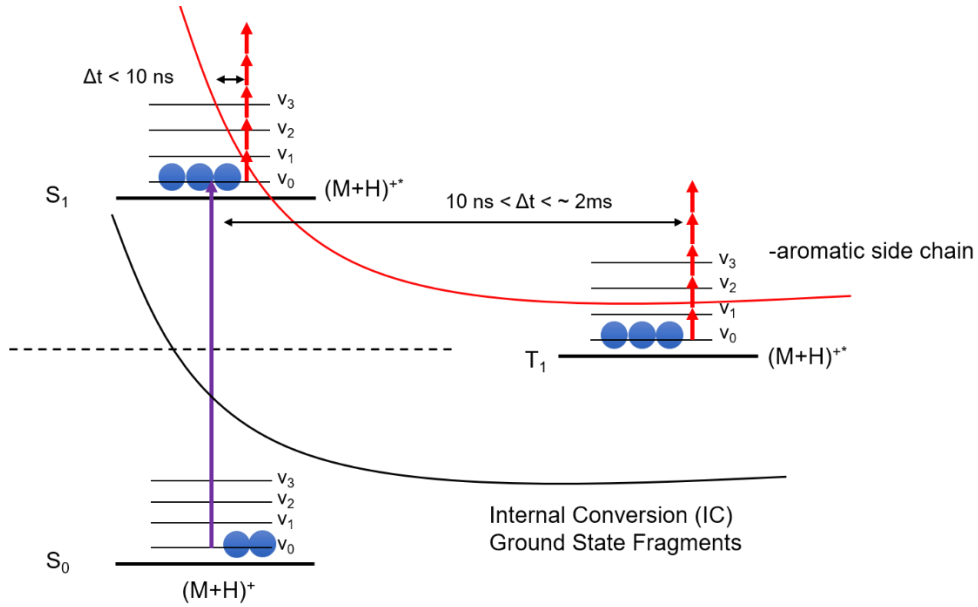


Figure 3.7: Relaxation pathways out of the excited state resulting in fragmentation

of aromatic chromophores are known to result in a selective cleavage of the $C_\alpha$-$C_\beta$ bond resulting in a neutral loss of -107 and -91 for tyrosine and phenylalanine, respectively. The abundance of these fragment ions depends on the rate of fragmentation relative to other relaxation pathways. Fig. 3.7 shows a simplified illustration of the relaxation pathways after UV excitation. The commonly observed side chain loss that occurs after UV excitation of aromatic side chains is non-

statistical and is known to occur specifically out of the excited state. The dissociative curve that the cleavage takes place from is known to intersect both the $S_1$ and $T_1$ states, as shown in Fig. 3.7. The fragmentation rate for this specific cleavage is thus dependent on the relative energies between the $S_1$ state and dissociative curve. Additional fragmentation can occur as the excited peptide internal converts resulting in a hot ground state. In this case the dissociative curve intersects the ground state resulting in thermal like fragments. Generally, the pentapeptide systems that have been studied in this thesis show the neutral loss associated with the aromatic side chain as the most dominate loss after UV excitation.

Often, the relative fragmentation is quite low, especially under single UV photon conditions. Under the normal operating conditions fragmentation efficiencies are less than 1 %, requiring additional gain in the preamplifier to detect the fragment ions. In most systems that we have studied the small photofragmentation quantities have not generally been a limiting factor. However, systems that contain arginine or have phenylalanine as the main chromophore result in especially low photofragment signal and in some cases have prevented data from being recorded. Singly protonated arginine systems are known to have relatively high dissociation thresholds due to the increased proton affinity of the guanidium side chain. This hinders fragmentation after internal conversion and also limits fragmentation from the excited state. For phenylalanine systems the fragmentation is quite limited due to the lower cross-section for UV absorption and the overall lower laser power near the origin region of phenylalanine.

The more ions exposed to the laser irradiation will ultimately lead in an increase in the amount of photofragment signal. There is however, a limit as to how many ions can be held in the octupole ion trap and under normal conditions it is often filled to capacity. Thus, increases in the amount of isolated precursor will only lead to a limited increase in photofragment signal if the trap is not already saturated with ions. A rather large increase in the amount of photofragments formed however, can be achieved by modulating the population once in the excited state. This can be done through IR irradiation subsequent to UV irradiation. As illustrated in Fig. 3.7, IR irradiation can lead to large increases in the internal energy of the ions that have been UV excited. Driving the population further up the vibrational manifold in the excited state allows for a larger fraction of ions to undergo the excited state specific neutral loss. A stronger coupling to the dissociative state is evidence in the UVPD mass spectrum through an increase in the aromatic side chain loss.

This process was initially demonstrated by the Rizzo group using a line tuned 10.6 μm $CO_2$ laser. The increases in fragmentation was found to be non-conformer selective, allowing the fragmentation signal across all the conformations to be equally enhanced. In a similar manner the Rizzo group also demonstrated the ability to record conformer specific IR spectra from the excited state by using a tunable IR light source as the enhancement laser. The results from this study have shown that depending on the time delay between the UV and IR laser pulses the resulting IR spectrum could be representative of the $S_1$ or the $T_1$ states. At short time delays, less than 10 ns, the IR spectrum is most similar to the $S_1$ state, while longer time delays result in the $T_1$ spectrum. Furthermore, the spectral changes between the excited and ground state spectrum can be used to assign certain bands. NH oscillators that are near the excited UV chromophore are seen to shift about 10 $cm^{-1}$ lower in frequency in the excited state. This observation can be used to confirm conformational assignments as will be discussed below.

### 3.5.1 UV-IR Double Resonance – Photofragmentation Enhancements for UV Spectra

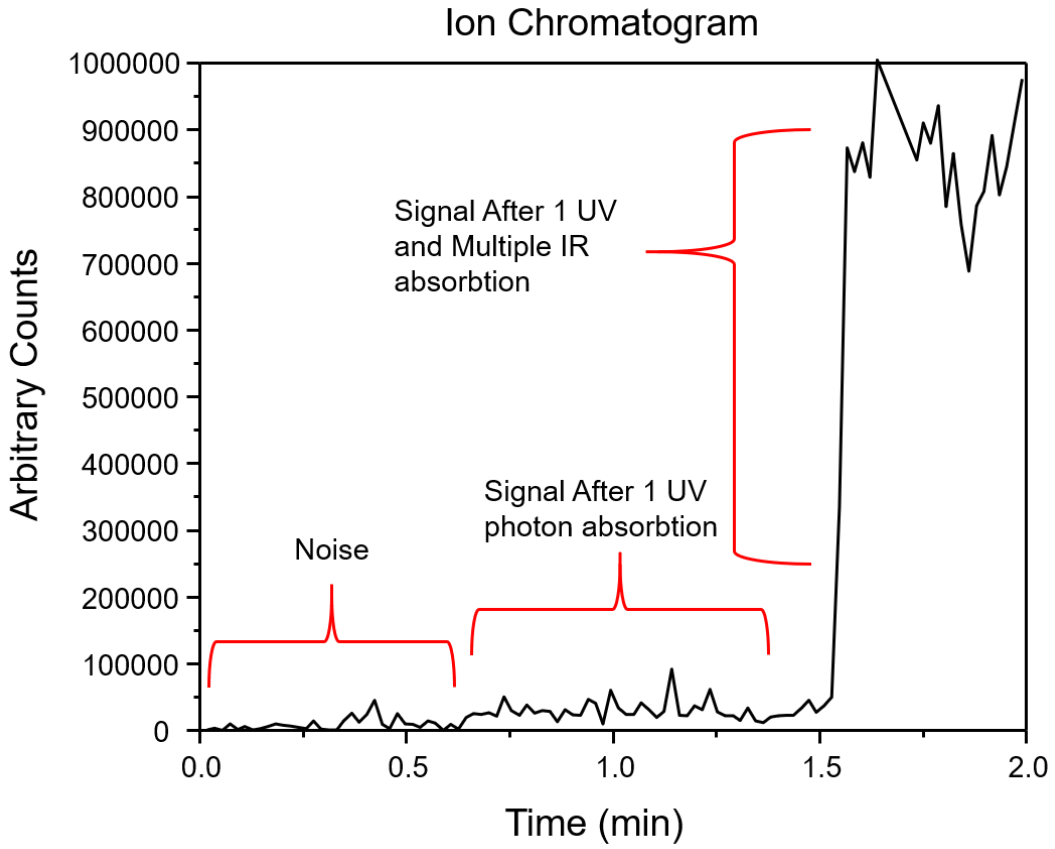


Figure 3.8: Ion chromatogram of YGRAR with no UV irradiation, UV irradiation, and UV followed by IR irradiation.

In our experimental setup we have employed similar double resonance and additional triple resonance schemes to enhance photofragmentation signal. Generally, by using the same IR-UV double resonance scheme described in section 2.2.1 it is possible to observe an order of magnitude increase in the UV photofragmentation signal by simply changing the timing of the IR laser to irradiate the ion packet after UV excitation. An example of this increase in fragmentation signal is shown in Fig. 3.8, for the $[YGRAR+H]^+$ system. With this particular system, very low fragmentation levels were observed after a single UV photon absorption event. The photofragmentation levels were hardly discernible from the background noise in the experiment, Fig. 3.8. By using the IR laser to irradiate the UV excited ions large increases in

photofragmentation signal was observed, allowing the UV spectrum to be recorded with a high signal to noise ratio.

Generally, the UV-IR setup is employed to collect a high SNR electronic spectrum due the low photofragmentation levels, or to record the IR spectrum of the excited state. In the case of the first scenario, it is most likely the case that the IR spectrum has not been recorded. This prevents the user from knowing what IR frequency to use for signal enhancement. In this case, for peptide systems, it is best to place the IR wavelength at a time delay of greater than +50 ns from the UV laser and set at a frequency between 3300 - 3400 $cm^{-1}$. Based on experimental data the IR spectrum of most peptides in the triple state have an absorption near this frequency. In doing this it is possible to scan the UV laser such that the electronic transitions become observable. Once an electronic transition is found, the UV laser can be set to that wavelength and IR laser can be scanned in real time such that the max absorption in the excited state can be found. Again, by allowing the delay time between the UV and IR laser to be ~50 ns allows for non-conformer specific fragmentation enhancements. Generally, with this time delay the conformers in the excited state have undergone IVR and are vibrationally hot causing them to possess similar IR spectra.

As shown in Fig 3.7 the signal enhancement is caused by driving the excited state population further up the vibrational ladder. This is done via multiple IR photon absorption from the excited state and the more IR photons absorbed in the excited state ultimately leads to greater signal increases. The power of the IR laser however, should not be so high such that it can induce fragmentation in the absence of the UV laser. This leads to unwanted background signal as fragments are being created when the UV laser is not resonant with an electronic transition. While using the IR to enhance the photofragmentation after UV absorption the linearity of the UV spectrum will be maintained as long as the UV power is kept in a linear regime.

### 3.5.2 UV-IR Double Resonance – Excited State IR Spectra

Using the UV-IR timings it is also possible to record excited state conformer specific IR spectra. To do this the UV laser is fixed on an electronic transition of interest and the IR laser is scanned across the spectral region. In this case, since the UV laser is resonant with an electronic transition there is a background signal due to the fragmentation induced by the UV laser. As the

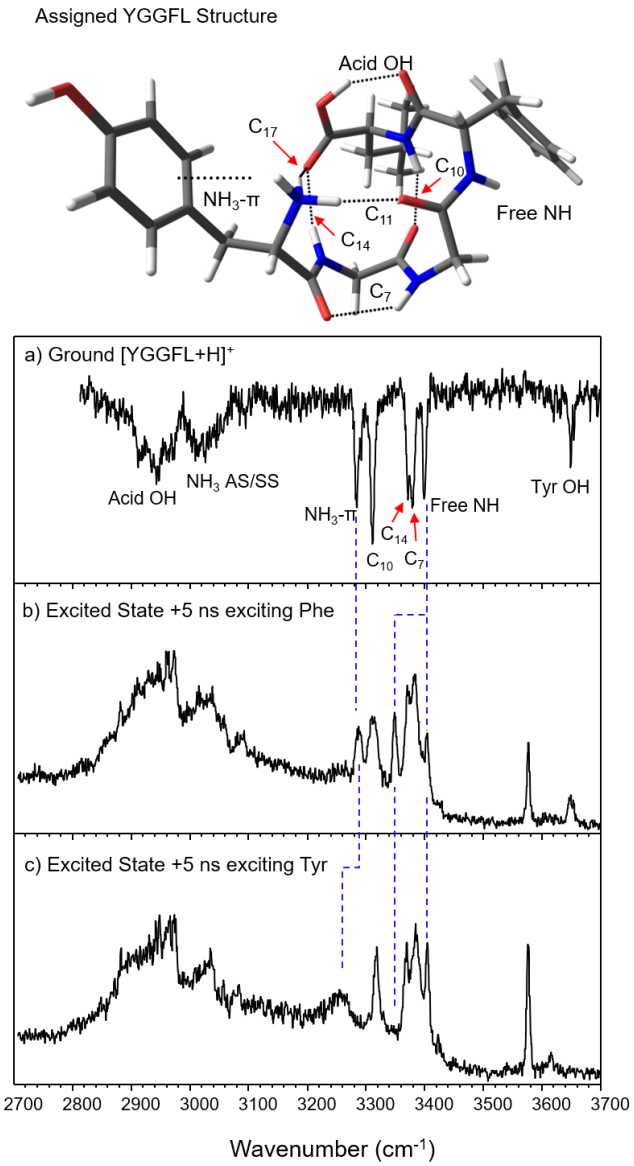


Figure 3.9: Assigned strucutre of [YGGFL +H]$^+$ (top). IR spectrum of [YGGFL+H]$^+$ in the ground state (a), phenylalanine excited state (b), tyrosine excited state (c).

IR laser is scanned across the spectral region at some delay time increases in the fragmentation signal will appear when the IR laser is resonant with a vibrational transition in the excited state. Thus, all the excited state spectra are taken on a background signal and appear as a positive going signal.

In some cases, recording the excited state spectra can help with structural assignments as vibrational bands near the UV excited chromophore are known to undergo distinct shifts. This phenomenon can be exemplified best using the [YGGFL+H]$^+$ system as it adopts a single major

conformation and has two chromophores. The assigned structure is shown in Fig. 3.9 along with its ground state vibrational spectrum in panel (a). The assigned structure shows the two chromophores are on opposite sides of the 3-D structure and interact with different NH functional groups. An ammonium NH is pointed towards the tyrosine ring, this band is labelled as $NH_3$-π in the ground state spectrum and has a frequency of 32xx $cm^{-1}$. The backbone NH of the phenylalanine is pointed towards the phenylalanine ring and is labelled as a free NH in the ground state spectrum with a frequency of 3400 $cm^{-1}$. The other modes that make up the C10, C14, C7, C11, C14, and Acid OH hydrogen bonds are sufficiently distanced from the aromatic chromophore that they are not expected to under any shifts unless a large-scale isomerization occurs after UV excitation.

Under the cryogenic conditions it is possible to uniquely excite a single chromophore such that the excited state spectrum can be taken in the $S_1$ state of either tyrosine or phenylalanine. Fig. 3.9b and c show the excited state spectrum 5 ns after UV excitation at the phenylalanine and tyrosine chromophore, respectively. At this short time delay both spectra are mostly $S_1$ character. In the excited state of both phenylalanine and tyrosine, the hydrogen bonded modes between 3200 and 3420 $cm^{-1}$ are slightly broadened due to the increase in internal energy but are all still resolvable. In the excited state of phenylalanine, Fig. 3.9b, all these modes appear in the same location as in the ground state. The band associated with the Free NH however, appears with a reduced intensity relative to the C14 and C7 modes and a new band appears near 3350 $cm^{-1}$. The reduced intensity of this band in addition to the appearance of the new band at 3350 $cm^{-1}$ is attributed to the Free NH oscillator responding to the local excitation on the phenylalanine chromophore. UV excitation at the tyrosine chromophore, Fig. 3.9c, on the other hand, results in the $NH_3$-π mode broadening and shifting to a lower frequency by about 20 $cm^{-1}$. In this spectrum the Free NH appears at its ground state frequency and has a similar intensity to the C14 and C7 bands. The band at 3350 $cm^{-1}$ is also missing in the excited state of tyrosine further supporting that the appearance of this band is related to the Free NH responding to the excitation on the phenylalanine ring.

The characterization of these types of interactions can be very helpful when making structural assignments. In fact, this procedure was used to confirm the assignments made for conformation A of the YAPAA system and the conformation of YGRAR, as will be discussed in later chapters. As stated in the previous section the goodness of fit for an assigned structure is not

analytically calculated, meaning there are no numerical criteria that define a good or bad match. Each vibrational mode in the ion must be accounted for when making a suitable match. This provides very stringent criteria in itself when making matches, but ultimately relies heavily on the theory to accurately calculate the vibrational frequencies in addition to the scaling factors sufficiently accounting for second and third order parameters such as anharmonicity and coupling between modes that also define the vibrational frequency of a given mode. Thus, it becomes useful to have other experimental techniques that can be used to confirm or dismiss candidate structures based on other measurable values.

Another interesting feature in both excited state spectra is the appearance of a band at ~3570 $cm^{-1}$. This position of this band appears where a free carboxylic acid OH stretch would appear and is attributed as such. In the ground state conformation this functional group is hydrogen bound however, UV excitation provides enough energy such that this bond can be broken. The dynamics of structural rearrangement will be discussed in detail in a separate chapter. Also, it is worth noting that the tyrosine OH band that appears at ~3640 $cm^{-1}$ in the ground state shifts to ~3620 $cm^{-1}$ in the excited state. This shift is notable in the IR spectrum from the $S_1$ state of tyrosine, Fig. 3.9c, and is known to occur in neutral phenol as well. In the $S_1$ state of phenylalanine this band does not shift at the short 5 ns time delay. A noticeable shift in this band however, can be seen at slightly longer time delays after energy transfer from phenylalanine to tyrosine has occurred, ~ >10 ns (data not shown).

### 3.5.3 IR-UV-IR Triple Resonance Schemes for Ground State IR Spectra

IR irradiation subsequent to UV absorption has been shown to significantly increase the photofragmentation signal by at least an order of magnitude, the largest signal increase has been observed to be a factor of 25. In a UV-IR double resonance scheme this signal enhancement can only be taken advantage of to record a UV or excited state IR spectrum. While these spectra can be quite useful, often the ground state spectrum is needed to make structural assignments. By using a triple resonance $IR_1$-UV-$IR_2$ scheme it is possible take advantage of the large signal increases provided by the second IR laser and record a ground state spectrum. This scheme is graphically illustrated 3.10. Essentially, the UV laser is fixed on an electronic transition of interest. The $IR_2$ laser is fixed on a vibrational resonance in the excited state and delayed by +50 ns from the UV laser, producing the signal enhancement, while $IR_1$ is timed to irradiate the ion packet 200 ns prior

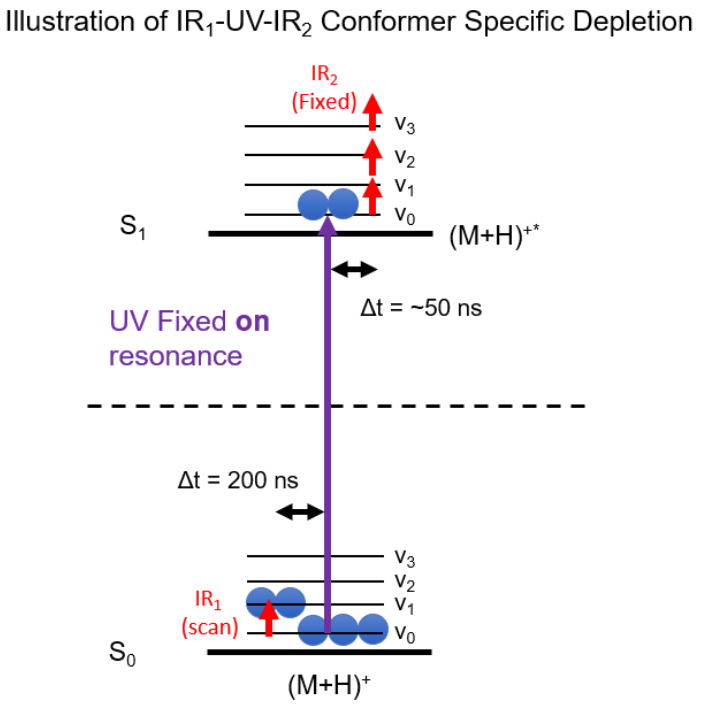


Figure 3.10: Schematic of IR-UV-IR triple resonance scheme

to the UV absorption. In this scheme as $IR_1$ is scanned across the spectra region it can deplete population from the vibrational ground state in the same manner as the traditional IR-UV DR approach. However, rather than monitoring the depletion of the UV photofragment signal, the depletion of the IR enhanced fragmentation signal is recorded. This leads to a significant increase in the signal to noise ratio of resulting IR spectrum. Furthermore, it also allows the IR spectrum of weakly fragmenting systems to be recorded. Fig. 3.11 compares the resulting amide I/II spectra recorded using the triple resonance scheme (a) and the traditional two-color scheme (b) for the YGGFL system. Both traces are comprised of a single scan. The SNR in the triple resonance scan is high enough that at most 3 scans need to be taken to get a high-quality spectrum. This cuts the data acquisition time by at least a factor of 2. The triple resonance scheme can also be taken in the hydride stretch region. The laser alignments are quite different however, depending on what part of the IR spectrum need to be recorded. The IR data for both the FGRAR and FGRAA systems were recorded with the triple resonance system. This data set would have otherwise not been able to be collected.

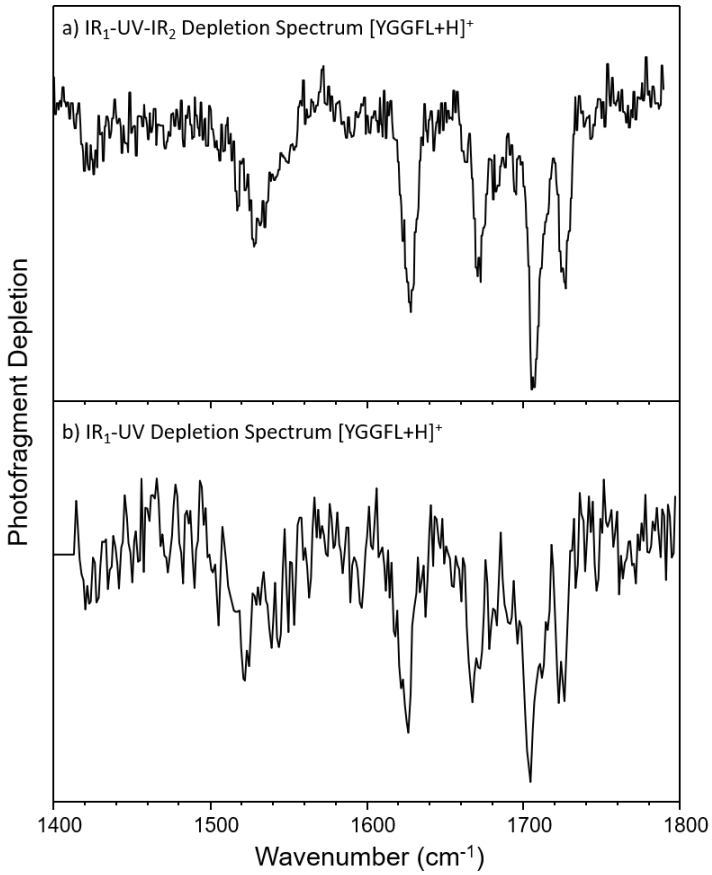


Figure 3.11: Comparison of $[YGGFL+H]^+$ amide I/II spectrum recorded with IR-UV-IR triple resonance (a) vs double resonance (b), both are a single scan.

## 3.6 IRMPD

### 3.6.1 IRMPD Mechanism

IRMPD is the most common method employed in the ion spectroscopy community as it requires minimal instrument modifications and a single high-power laser. In sharp contrast to messenger tagging and IR-UV DR this technique relies on inducing photodissociation through the absorption of multiple IR photons. The IRMPD mechanism occurs though via non-coherent absorption and is thus referred to as *multiple photon* rather than *multiphoton* absorption which would suggest a coherent absorption process. The absorption process proceeds through a non-coherent manner due to the anharmonicity associated with each normal mode. That is, the energy between adjacent bound states of a single mode decreases with increasing vibrational quantum number. This prevents the mode from being continually excited up the vibrational ladder and is

commonly referred to as the *anharmonicity bottleneck*. Rather, as population moves from v = 0 to v =1 in the initial excitation event the energy is dissipated to other modes that are coupled through anharmonic terms. These modes consist of other fundamentals but are largely combination modes. Once the excited vibrational level is deactivated it can once again be populated. This process occurs repeatedly until the internal energy is high enough for dissociation.

The mode of relation from the excited v =1 state is commonly referred to as intramolecular vibrational redistribution (IVR). The rate of this process ultimately governs the overall efficiency of the IRMPD process. Generally, IVR rates for large polyatomic molecules are picoseconds or faster at room temperature. The rate of the process is loosely related to the density of states (ρ) but can be more accurately be described by the presence of low-order coupled doorway sates and the magnitude of their coupling to the initially excited state. The states that strongly couple to the excited fundamental is often referred to as a quasi-continuum of vibrational states. The quasi-continuum, while mis-leading in name, can be though of as semi-resonant zones of absorption near the fundamental mode. The spectral width (Γ) of these zones is determined by Fermi's Golden Rule:

$$\Gamma = 2\pi\langle W^2\rangle\rho$$

where $\langle W^2\rangle$ is the mean squared coupling matrix element of the first tier of coupled states. From this equation it can be seen there is a dependence on the density of state of the molecule but how strongly the background states couple to excited fundamental will ultimately determine the IVR lifetimes (τ). Additionally, these semi-resonant zones may also borrow intensity from the bight state are couple more strongly as the internal energy of the ion is raised. While this aids with the IRMPD process it also cause the fundamental to appear shifted to lower energy (red shifted) in the resulting spectra. The criterion of stochasticity defines the onset of the quasi-continuum as the internal energy where:

$$2\pi c\tau_{IVR} = \rho(E)$$

For large pentapeptide systems IVR rates can be estimated to be a picosecond or faster. At these rates the quasi-continuum would begin at under 50 $cm^{-1}$ of internal energy. For reference the internal energies of these sized systems at room temperature is approximately 6700 $cm^{-1}$. This

indicates that all the absorption occurs through the quasi-continuum however, in line with Fermi's Golden Rule strong absorption occur only near fundamental mode.

For ridged systems the effects due to the multiple photon absorption process has been modelled and shown to correlate well with experimental spectra. For these models the anharmonicity for each mode as well the cross-anharmonicity for each mode to all other modes is needed. These values are difficult to calculate a priori however, can be established by recording the shifts of the vibrational bands as a function of temperature. This has been done for several neutral PAH's and the values were applied to model the IRMPD spectrum of the cationic coronene PAH. The resulting model displayed reasonable agreement with the observed IRMPD spectrum and display interesting features. In this ion, the internal energy distribution after excitation via three distinct modes lead to similar average internal energy distributions however, the true distribution after IR excitation was quite distinct in each mode. One mode displayed a bimodal internal energy distribution. This was attributed to a neighboring higher frequency mode being red shifted into the laser irradiation as the internal energy increased past a certain value. This ultimately leads to a second boost in fragmentation signal. The results of the model signify that the coupling between modes can significantly alter the resulting internal energy distribution in a mode specific manner, the results of which manifest themselves in appearance of the IRMPD spectra.

### 3.6.2 2-Color IRMPD

While the IRMPD process is well understood, most of the modelling was performed on ridged PAHs. The overall mechanism is not expected to change from system to system however, other parameters such as isomerization need to be taken into account when this technique is applied to much more flexible systems such as peptide ions. As stated and demonstrated in Fig. X.X the IR spectrum is highly sensitive to an ion's hydrogen bonded network. In ridged PAH's there is significantly less flexibility thus, a little change in structure is expected as the ions approach the dissociation threshold. In protonated peptides however, proton migration often proceeds fragmentation as described by the mobile proton model. This can induce structural rearrangements as the internal energy approaches the fragmentation thresholds. Furthermore, several others isomerization events that don't significantly alter the geometry can be expected to occur prior to proton migration. Small geometry changes that are substantial enough to change the hydrogen bonding pattern of the ion have barriers of about 0.5 ev while fragmentation barriers are about 1

ev. The analysis of the rates of these two processes (see chapter 5 for more detail) suggest that structural rearrangements are expected to occur much more readily compared to fragmentation. In a multiple photon absorption scheme this becomes important as the vibrational resonance used to excite the ion can fall out of resonance with the laser wavelength at internal energies below the fragmentation threshold. This would have profound effects on the resulting spectrum as bands that are associated with structural changes would appear missing from the spectra. Geometrical changes such as this were not accounted for in the early models as isomerization was not anticipated for the ridged systems that were studied. Structural changes along with the inherent band shifting associated with the IRMPD process itself can significantly distort the resulting spectra.

To minimize theses effects a 2-Color IRMPD scheme can be used. This technique is described in chapter 5 and will be outline only briefly here. The basis of this technique is to divide the IRMPD process between two lasers. One laser is set to a low fluence ($\lambda_{LF}$) such that only a limited number of photons can be absorbed. The other is set to a much higher fluence ($\lambda_{HF}$) such that dissociation can be induced This process is illustrated in Fig. 3.12. The linear spectrum recorded using IR-UV DR of $[YGGFL+H]^+$ is shown in Fig. 3.12a, and the linear 2-Color spectrum recorded at the same temperature is shown in Fig. 3.12c. For comparison the 1-Color spectrum is shown over a limited range with a blue trace in Fig. 3.12b. The discrepancies between the 1-Color IRMPD spectrum and the linear spectra are dues to the factors discussed above and thermal effects which are discussed in chapter 5.

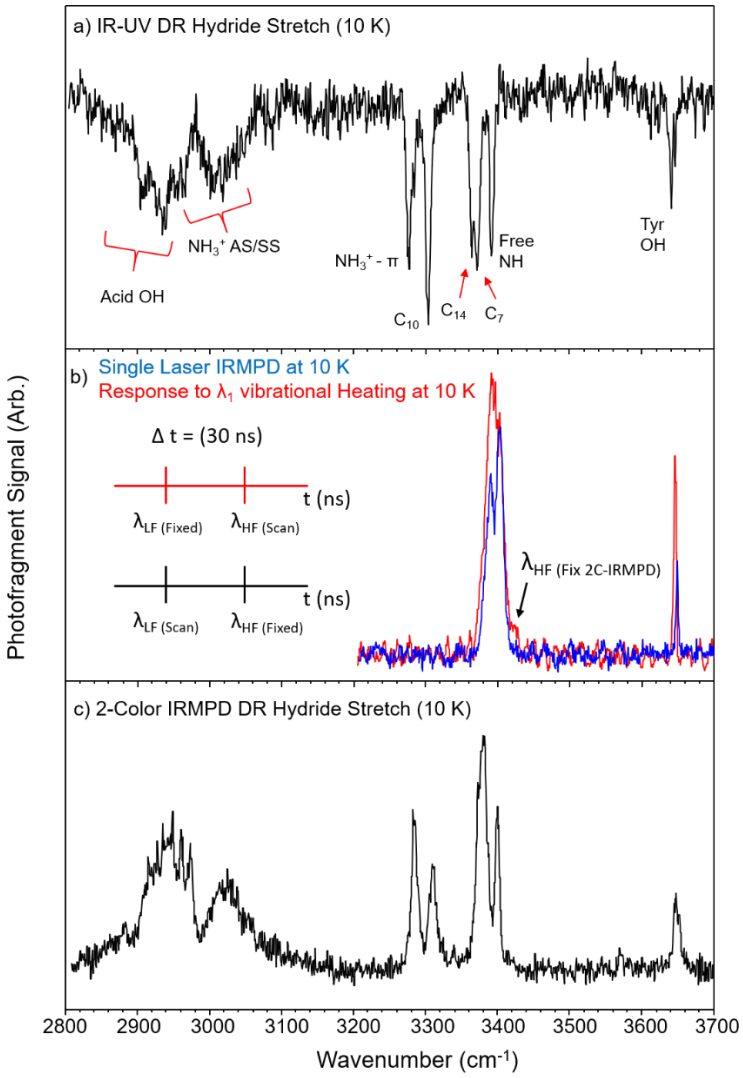


Figure 3.12: Linear spectrum of the YGGFL protonated ion recorded using IR-UV DR and 2-Color IRMPD (a and c). 1-Color IRMPD spectrum (blue trace) and the 1-Color spectra response to vibrational preheating (red trace) (b).

To generate a 2-Color spectrum a series of scans must first be taken. The first is to generate a 1-Color spectrum. This is done by recording the photofragment signal as a function of the high fluence laser's wavelength, blue trace, Fig 3.12b. The purpose of this trace is to reveal the frequency of the strongly absorbing bands. Next, the 2-Color scheme outline with the red timeline in Fig. 3.12b is employed to monitor how the absorption profile of the 1-Color spectrum changes as a function of internal energy. In this scheme the low fluence laser is fixed on a transition revealed by the 1-Color scan (3395 $cm^{-1}$) and timed to irradiate the ion packet 30 ns prior to the high fluence laser. The photofragmentation signal is again recorded as a function of the high fluence laser's

wavelength. This scan is displayed as the red trace in Fig. 3.12b. This trace reveals that as the ions internal energy is raised by the low fluence laser a new resonance appears at frequencies slightly higher that the transition at ~3400 $cm^{-1}$. This new transition is unique to the ions that have been preheated with the low fluence laser. The laser scheme that is used to record a linear 2-Color IRMPD spectrum is shown with the black timeline in Fig. 2b. In this scheme the high fluence laser is now fixed and set to be resonant with the transition that is unique to the preheated ions. The low fluence laser is now timed to irradiate the ion packet 30 ns prior to the high fluence laser, and the photofragment signal is recorded as a function of the low fluence laser's wavelength. In this manner whenever the ion gets heated with the low fluence laser a resonance is created for the high fluence laser and generates photofragment signal. Thus, fragmentation is only seen when the low fluence laser is resonant and as a result produces a linear spectrum. As stated above, the fast IVR rates produces a resonant transition for $\lambda_{HF}$ regardless of which mode is initially excited with $\lambda_{LF.}$ The ability to record linear spectra using only IR sources is highly useful and the full power of this technique will be further explained in chapter 5.

### 3.7 Laser Radiation Sources

Laser light in the UV range is generated by doubling the light produced by a Nd:YAG pumped dye laser. The dye laser (ScanMatePro Lambda Physik) is pumped using either the second (532 nm, 200 mJ/pulse) or third (355 nm, 90 mJ/pulse) harmonic of a Q-smart (450 mJ) Nd:YAG laser, depending on the dye being used (532 nm for Rhodamine 6G or 355 nm for Coumarin 540 A). The fluoresce efficiency differs between the two dyes. Pumping Rhodamine 6G with 200 mJ of 532 nm light results in upwards of 50 mJ/pulse at 552 nm which is the peak of the dye. This light is then double using an Inrad Autotracker III, producing ~3mJ/pulse of UV light after passing through a UG5 filter. Pumping Coumarin 540 A with 90 mJ/pulse of 355 nm light results in upwards of 6 mJ/pulse at 540nm, the peak of the dye. Doubling this light results in ~ 500 μJ/pulse of UV light. When working with the Coumarin 540 dye it is important to not use the UG5 filter to separate the doubled light from the non-doubled light as this reduces the power of the UV light by ~10x. Rather a Pellin-broca is used to separate the two colors of light. In the case of Rhodamine 6G it is best to use the UG5 filter to reduce the UV power into a range that is usable.

The IR light is generated by pumping a LaserVision optical parametric oscillator/optical parametric amplifier (OPO/OPA) system using the first harmonic of a Nd:YAG (Continuum

Surelite III ~780 mJ/pulse). The incoming 1064 nm light is sum and difference-frequency mixed to generate laser light in the hydride stretch (2600-3700 $cm^{-1}$) region (~20 mJ/pulse). Briefly, the incoming 1064 nm light s split using a 70/30 beam splitter. 70% of the light is passed to the OPA stage and the remaining 30% is directed through a KTP doubling crystal to generate 532 nm light which is then passed into the OPO stage. The OPO stage is comprised of two phase-matched KTP crystals that splits the 532 nm (18794 $cm^{-1}$) light into a signal ( ~12500 $cm^{-1}$) and idler (~6200 $cm^{-1}$) beam. The energy of each beam can be adjusted by angle tuning the OPO crystals, however the sum of the two beams is always 18794 $cm^{-1}$. The idler beam of the OPO stage seeds the OPA stage, which contains four KTA crystals. The incoming idler beam from the OPO stage is difference frequency mixed with the 70% of the 1064 nm light. The output idler beam of the OPA stage is then selected, using a silicon polarizer, as the output beam.

To generate light in the amide I/II region (1400 – 1800 $cm^{-1}$) the signal and idler of the OPA stage are difference frequency mixed 6 ft down field using an $AgGaSe_2$ crystal (3 - 0.75 mJ/pulse).

## 3.8 Laser Focusing Conditions and Beam Paths

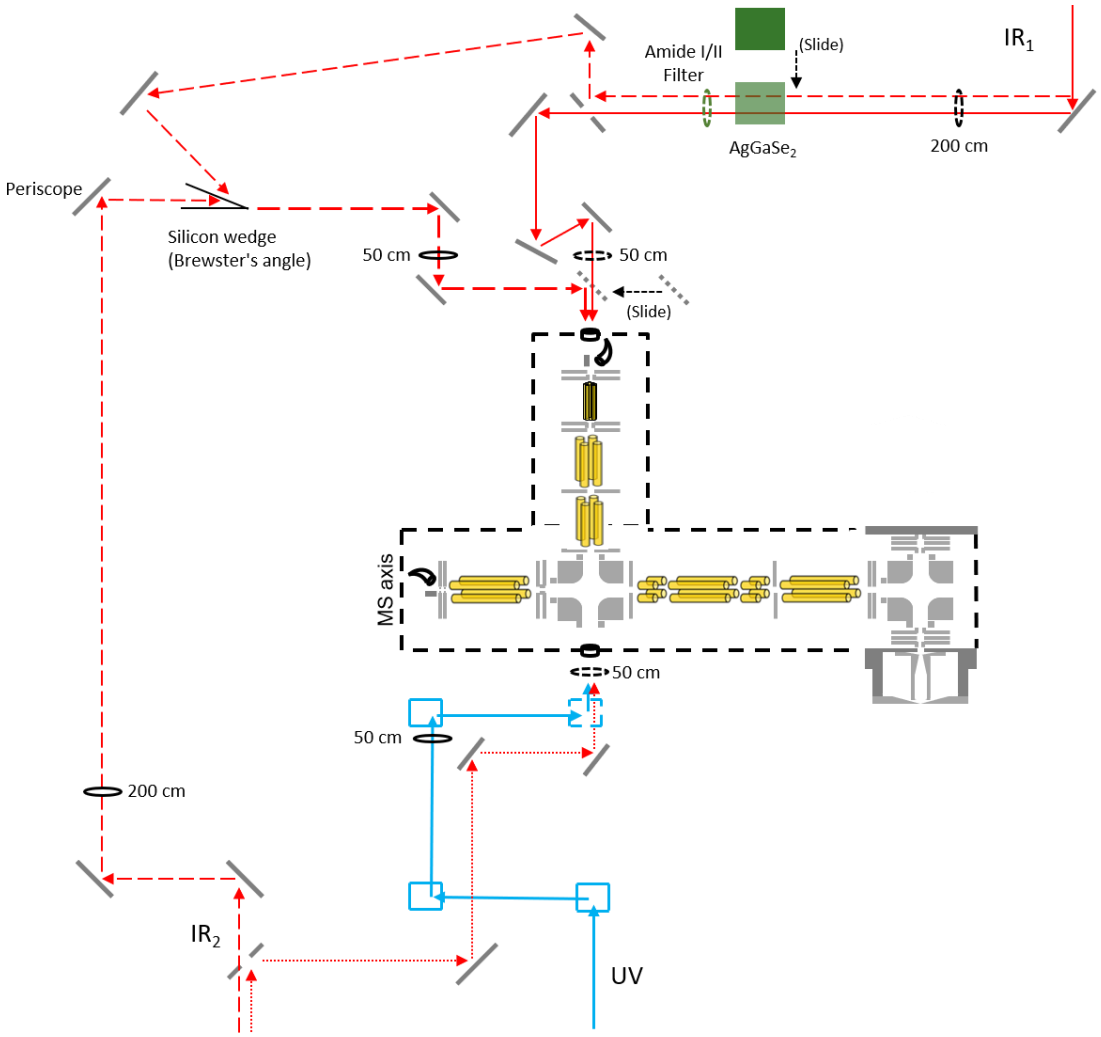


Figure 3.13: Laser paths and optics used for double and triple resonance experiments, see text for detail.

There are several beam paths available depending on the type of experiment being carried out. A schematic of the most commonly used beam paths is shown in Fig. 3.13. The location of each of three lasers relative to the instrument are designated based on their starting point in the figure. The blue arrow designates the UV light and blue boxes the UV prisms. The red arrows designate the IR light generated from either laser 1 ($IR_1$) or laser 2 ($IR_2$) and silver lines the silver mirrors. The long-dashed sliver and blue boxes represent mirrors or prisms that are on flipper mounts. The short-dashed sliver line represents a mirror that is on a sliding mount. The ovals represent focusing optics and are labelled with their focal length, the dashed ovals represent the

focusing optics that are on flipper mounts. The green oval represents an amide light filter. The IR light from either laser can be directed along multiple paths depending on the experiment being performed. The solid red arrow from $IR_1$ represents the beam path for double resonance experiments. The long-dashed line from both $IR_1$ and $IR_2$ represents the path for the triple resonance experiment, and the short-dashed line from $IR_2$ represents the beam path when performing the 2-Color IRMPD experiment.

### 3.8.1 UV Beam Path

The beam path for the UV laser in unchanged in all the experiments that require UV excitation. Once the UV light leaves the autotracker it is directed into the chamber using four prisms, the last of which is on a flipper mount. The flipper mount in this set up is such that the IR light from the second IR laser ($IR_2$) can be directed into the front of the chamber when performing the 2-Color IRMPD experiment. When directing the UV light into the chamber it is imperative to keep the fluence low enough such that multiple photon absorption does not occur. UV light from Rhodamine 6G is focused using a 50 cm $CaF_2$ focusing lens (solid oval) placed about 200 cm away from the center of the trap. The power of the UV light with this dye is generally very high and is best to have the light slightly divergent while going through the ion packet. Depending on the system being studied it is sometime necessary to use a 50 cm $CaF_2$ focusing lens (dashed oval) placed 50 cm away from the center of the trap when working with UV light produced from Coumarin 540 A.

### 3.8.2 $IR_1$ (Hydride Stretch and Amide I/II) Beam Path for IR-UV, UV-IR, and 2-Color IRMPD

In the IR-UV, UV-IR, and 2-Color IRMPD schemes (all double resonance schemes), the light from $IR_1$ is used and follows the path designated by the solid line. For the double resonance schemes, as the hydride stretch light leaves the $IR_1$ laser it is focused using a 200 cm $CaF_2$ focusing lens (dashed oval) placed about 250 cm away from the center of the trap. In order to direct this light into the trap $AgGaSe_2$ crystal is slide out, the dashed sliver mirrors are either flipped down or slide out, and the 50 cm focusing lens and the amide I/II filter is removed from the path. In all the double resonance experiments this laser is the scanning laser responsible for generating linear spectra. Under this focusing condition a linear regime is attained at moderate powers (10 mJ/pulse)

while at full power produces spectra that are saturated. The fluence however, is still low enough such that IRMPD cannot occur. This makes this beam path useful for the UV-IR experiments as well since a large signal enhancement is wanted without creating fragment signal from an IRMPD process. Furthermore, in the UV-IR experiment linear IR spectra can be recorded in the excited state as well, by adjusting the power to ~10 mJ/pulse. When directing amide light along this beam path for the double resonance experiments the 200 cm focus lens is removed, and the 50 cm focus lens is placed at the location indicated, which is about 50 cm away from the center of the trap. The $AgGaSe_2$ crystal and the amide I/II filter are both placed into the path. The dashed sliver mirrors are again left out of the beam path.

### 3.8.3 $IR_1$ (Amide I/II) Beam Path and Alignment Considerations for IR-UV-IR Triple Resonance

For the $IR_1$-UV-$IR_2$ triple resonance scheme the UV light follows the same beam path and is directing into the trap from the front of the chamber. The light from $IR_2$ is always in the hydride stretch region, while the light from $IR_1$ can either be in the amide I/II or the hydride stretch region. When $IR_1$ is set to use amide I/II light it follows the same path as is does for the double resonance scheme. The only difference in the amide I/II beam path is that the short dash sliver mirror is slid into place. When performing the triple resonance experiment in the amide region the sliver mirror is removed from the sliding mirror mount and is replaced with a $CaF_2$ window. The amide I/II filter is then removed from its original location, near the $AgGaSe_2$ crystal and placed in the same mount as the $CaF_2$ window on the side that is closest to the chamber. This filter is thin and will not fit into an adjustable mirror mount on it own, hence it is placed on top of a $CaF_2$ window allowing it to be mounted. The light from the second laser ($IR_2$) is directed along the path designated with the long-dashed arrow. The light is focused along the path using a 200 cm focusing lens to a periscope that first directs the light 90° up towards the lab ceiling then turns the light another 90° such that it is parallel to the lab floor. The light is then directed into a silicon wedge that is placed at Brewster's angle allowing it to pass through the wedge. The light is then directed towards the optic holder that was slid into place while being focused by a 50 cm lens. The amide I/II filter that was placed in the sliding mount on top of the $CaF_2$ window acts as a dichroic and reflects the hydride stretch light into the chamber. The amide I/II light that approaches the lens from the back passes through the filter allowing both colors of light to enter the chamber.

From a practical point of view, when aligning this set up it is easiest to first align the amide I/II light as if performing a double resonance experiment (amide filter back near the $AgGaSe_2$ crystal and the sliver mirror optic holder slide back out of the beam path). An important part of this alignment is to ensure that the UV light coming out of the back-chamber window is more or less hitting the center of the first solid silver mirror in its path. Once the UV and amide I/II light are aligned, ensured by seeing a depletion spectrum, the slidable silver mirror mount can be slid into place and loaded with the amide I/II filter and the $CaF_2$ window. It is important that the amide I/II filter is facing the chamber with the $CaF_2$ window behind. To ensure that the slidable mount is in line with the mirror behind it, slide it to a position such that the UV light is hitting the center of the amide filter. Ensure that the amide I/II light is still passing through the $CaF_2$ window and amide filter by tuning to a resonant transition and checking for depletion. Once this is true and the light is not hitting the mount, this path no longer has to be adjusted. To align the $IR_2$ light, use the UV light that is reflected off of the amide I/II filter to place the $IR_2$ light in the right position. By putting the Amide I/II filter in front of the $CaF_2$ window the UV light reflects directly from this surface. If the $CaF_2$ lens is place first with the Amide I/II filter behind it then there will be multiple reflections from each surface of the $CaF_2$ window and the Amide I/II filter. The adjustable mount that the amide I/II filter is sitting in and the silver mirror directly after the 50 cm focusing lens (on the $IR_2$) beam path can be used as near and far field adjustments for the $IR_2$ light. When making these adjustments the light should be blocked from entering the chamber to avoid hitting the trap. Once the beams are spatially overlapped the $IR_2$ beam can be let into the chamber while the $IR_1$ beam is blocked. Adjustments can be made to maximize the enhancement signal using the same near and far field adjustments. Once the $IR_1$ in unblocked there should be a dip in the enhanced signal. Ensure that the timings of all the laser are correct using a photodiode if there is no depletion.

### 3.8.4 $IR_1$ (Hydride Stretch) Beam Path and Alignment Considerations for IR-UV-IR Triple Resonance

When performing the triple resonance experiment in the hydride stretch region $IR_1$ is directed along the beam path indicated with a dashed arrow. In this path the 200 cm focusing lens, $AgGaSe_2$ crystal, and Amide I/II filter are removed from the beam path. The silver mirror represented with a long dash is flipped up and directs the beam towards the silicon wedge. In this set up both lasers are in the same wavelength regime, thus a dichroic will not work to place them on top of one

another, instead a silicon wedge placed at Brewster's angle is used. The wedge is placed such that it will reflect one polarization of light and pass the other. Since the two lasers are identical the light that comes out of them is of the same polarization. The periscope in the $IR_2$ beam path addresses this, as the polarization of the light is changed from s to p-polarized when it is directed up towards the ceiling. The two beams, now of different polarizations can be either passed or reflected off the silicon wedge and be placed on tope of one another. Once the beams are on top of each other they are both focused with a 50 cm focusing lens and directed to the slidable mirror mount holder. In this set up a silver mirror in placed in this mount and both beams are directed into the trap.

From a practical point of view, when aligning this set up it is best to first get the $IR_2$ beam aligned with the UV laser. To do this first ensure that when the silver mirror is slid into place the UV is hitting its center. The UV light reflected from this mirror can then be used to position the $IR_2$ beam. The slidable silver mirror and the mirror directly after the 50 cm focusing lens (on the $IR_2$ beam path) can be used as near and far field adjustments. As these adjustments are being made make sure that the $IR_2$ light does not enter the chamber to prevent hitting the trap. Once spatially overlapped the light can be let into the chamber but the power should be kept low, (25° on the OPA halfwave plate, OPO half waveplate set to max). Ensure the laser timings are correct and the $IR_2$ light is on a vibrational resonance in the excited state. If there is enhancement signal, make fine adjustments with the near and far field adjustments to maximize the signal. Once the signal is max at the lower power, the power can be turned up in order to see the full magnitude of the enhancement, should be at least a factor of 10x. If there is no enhancement signal at low power when the beam is initial let into the chamber, make fine adjustments with the near and far field adjustments. If there is still no enhancement signal block that laser and check the timings and spatial overlap once again. Avoid turning the laser power up if there is no initial enhancement as the laser will most likely hit the trap. Once the enhancement signal is achieved and maxed out block the $IR_2$ light at the chamber window.

To align $IR_1$ start by turning its power up to about half the max. Use the two mirrors prior to the silicon wedge to walk the beam onto the wedge. By the time the beam reaches this wedge it would have expanded quite a bit. Ensure that the bulk of the beam hits the wedge. Using a beam block, block the $IR_2$ from going through the wedge. Place a thermal paper about 10 cm in front of the wedge and ensure that the beam is being properly reflected off of the silicon wedge. The beam should appear as a square since it is slightly larger than the wedge. Make the necessary adjustments

to make the beam look full after reflecting off the wedge. Leaving the thermal paper in place unblock the $IR_2$ light and gauge how close the two beams are to one another. It will be helpful to reduce the power of $IR_2$ during these steps. Make the necessary adjustments using the furthest of the last two mirrors prior to the silicon wedge to walk the $IR_1$ light onto the $IR_2$ light. This is done iteratively by blocking and unblocking $IR_2$ and checking the two beams relative locations to one another after each adjustment. One the beams appear to be on top of one another 10 cm away from the wedge, place the thermal paper near the window of the chamber and check the overlap at this point. Adjust $IR_1$ such that it overlaps with $IR_2$ near the chamber window using the last mirror that independently adjust $IR_1$. Since both beams become focused at this point it is best to allow $IR_1$ to light the thermal paper up, and then block that beam. As the thermal spot begins to fade quickly unblock and block $IR_2$. In this manner it becomes easier to judge the overlap between the two lasers. Once the two beams are overlapped at the chamber window check the overlap near the wedge once more. Do this process iteratively until the beams are overlapped in both spots. Once spatially overlapped, tune $IR_1$ to a known transition. Unblock $IR_2$ and turn the power back up to see the full enhancement signal. Keep $IR_1$ at about half power and unblock this beam. If done properly there should be some depletion signal evident. Make fine adjustments with the last mirror that independently controls $IR_1$ to maximize the depletion signal. If there is no depletion signal when the $IR_1$ laser in unblocked, check the laser timings, and try scanning the wavelength of $IR_1$. If there is still no depletion, make fine adjustments using the last mirror that independently controls $IR_1$. If the signal begins to deplete as adjustments are being made to this mirror block the beam to ensure that the signal is restored. If it is, unblock the beam and keep adjusting. If it hasn't then the trap has been hit. Block the lasers and check the spatial overlap at the two point again. To avoid hitting the trap when blindly tuning the last $IR_1$ mirror it can be helpful to block $IR_2$ and adjust the timing of $IR_1$ such that it irradiates the ion pack after the UV laser. In this case the signal will be positive going when trying to get $IR_1$ in the correct position.

### 3.8.5 $IR_2$ (Hydride Stretch) Beam Path and Alignment Considerations for 2-Color IRMPD

The 2-Color IRMPD experiment is another double resonance experiment. In this case the $IR_1$ beam path is directed into the chamber through the back window as explained in section 3.8.2. The light from $IR_2$ is directed to the trap through the front window according to the short-dotted line. For this beam path a sliver mirror is flipped into the beam path (dashed sliver line) such that the

$IR_2$ light is directed towards the front of the chamber and the UV prism directly in from of the front window (dashed blue square) is flipped out. The $IR_2$ light is focused using a 50 cm focusing lens that is placed 50 cm away from the trap.

To align $IR_2$ in this beam path, it is best to used $IR_1$ or a laser diode that has been aligned through the center of the trap. The $IR_2$ light is rather large as it hits the last two mirrors prior to be turned into the chamber. When using $IR_1$ or a laser diode, it is possible to direct the $IR_2$ light on to the same beam path, however it may still be off due to the size of the $IR_2$ beam compared to the $IR_1$ or laser diode light after making its way through the chamber. The only way to assess if the $IR_2$ beam is aligned correctly is to turn the power all the way up such that fragmentation signal from IRMPD can be observed. If not, signal is observed when the power if turned up block the $IR_2$ light and check the precursor ion signal. If this signal has gone away, then the trap has been hit. If this signal still persists then fine adjustments can be made using the last mirror until fragmentation signal is observed. It may be best to perform this alignment during a heat cycle such that hitting the trap is a non-factor.

# CHAPTER 4. CONFORMATION-SPECIFIC INFRARED AND ULTRAVIOLET SPECTROSCOPY OF COLD $[YGPAA+H]^+$ AND $[YGPAA+H]^+$ IONS: A STEREOCHEMICAL "TWIST' ON THE B-HAIRPIN TURN

"Reprinted (adapted) from (DeBlase, A. F.; Harrilal, C. P.; Lawler J. T., Burke N. L.; McLuckey S. A., Zwier, T. S., Conformation-Specific Infrared and Ultraviolet Spectroscopy of Cold $[YAPAA+H]^+$ and $[YGPAA+H]^+$ Ions: A Stereochemical "Twist" on the β-Hairpin Turn. *J. Am. Chem. Soc.* ***2017****,* 139*, 5481-5493*). Copyright (2017) American Chemical Society."

## 4.1 Introduction

A powerful approach to directing the secondary structures of peptides involves insertion of one or more unnatural amino acids into the peptide sequence with well-defined structural consequences, thereby providing a basis for rational control of biological processes.[1] The current applications of peptides with designed $\beta$-hairpin folds include heme binding in micelles,[2] inhibition of $\beta$-amyloid aggregation,[3] drug delivery,[4] and model systems to study post-translational modifications.[5] A common approach to enforce a $\beta$-hairpin turn is the strategic insertion of the residues $^{D}$PX (where $^{D}$P is the unnatural stereoisomer of proline, and X = G or A) into a peptide sequence.[6] Other recent strategies include the incorporation of disulfide bonds[7] and the incorporation of charged side-chain groups.[8] However, the $^{D}$PX insertion method continues to be a robust approach to forming $\beta$-hairpins and the high folding propensity has been linked to the inherent conformational preferences of these two residues.

One important test for these structural preferences is to study model peptide sequences in the gas phase, where no intermolecular stabilization with solvent or other solutes is possible. To that end, in this work, we carry out conformation-specific IR and UV spectroscopy on two model protonated $^{D}$P and $^{L}$P diastereomers in the gas phase at 10 K, allowing us to probe the direct relationship between primary sequence and secondary structure. Recently, Rizzo and coworkers have employed such techniques to unravel the 3D structures of the $[GPGG+H]^+$ peptide[9] and the bradykinin 1-5 fragment ion,[10] which contains two proline residues. Our group has recently studied charge-stabilized $\beta$-hairpin formation in $[YGGFL+H]^+$ and $[YGGFL\text{-}OMe+H]^+$, providing

IR spectral signatures for its formation, and thereby serving as a useful foundation for the present studies.[11,12]

The $\beta$-hairpin is the tightest loop that tethers antiparallel $\beta$-sheets and is comprised of four residues (*i*, *i*+1, *i*+2, and *i*+3), where the *i*+1 and *i*+2 residues are involved in a Type I' or Type II' $\beta$-turn. The Ramachandran angles ($\phi$,*y*) for the (*i*+1) and (*i*+2) residues are either (+60°, +30°) and (90°, 0°), respectively (denoted Type I' $\beta$-turn), or (+60°,-120°) and (-80°,0°), respectively (denoted Type II' $\beta$-turn).[13,14] Three-dimensional modeling of these turns, informed by x-ray crystallography[13-16] and NMR measurements of the nuclear Overhauser effect (NOE),[6,17-19] reveals the presence of two critical hydrogen bonds that connect the opposite sides of a minimal four-residue $\beta$-hairpin.[20,21] One of these H-bonds forms an *i+3→ i+1* ten-membered ring when counting the N, H, O, and backbone atoms in the cycle (denoted a C10 H-bond), while the other forms a fourteen-membered *i→i+4* ring (C14 H-bond). The factors that contribute to the stability of the *b*-hairpin depend on a delicate balance of intramolecular and intermolecular interactions that vary with peptide sequence, solvent exposure, and protein environment.[22-24] It is well known that most short peptide sequences that fold into $\beta$-hairpins in native protein structures do not retain this structure as free peptides in solution,[25-28] which has motivated the use of $^{D}$PX (X = G or A) segments,[6,29-32] to ensure the presence and location of these turns. The $^{D}$P stereochemistry was required so that the correct "mirror-image" turn structure matched the handedness of the $\beta$-hairpin (*i.e.*, $^{D}$PX forms Types I' or II' turns while $^{L}$PX forms Types I or II turns).[6]

Diastereomer-specific effects involving proline are not confined to solution. In mass spectrometry studies in the gas phase, different conformational preferences for protonated $^{D}$P- *vs.* $^{L}$P-containing peptides have been postulated by Raulfs *et al.*,[33] who observed a modulation of the "proline effect" when studying the collision-induced dissociation (CID) of protonated AA$^{D}$PAA and AA$^{L}$PAA ions formed by electrospray ionization (ESI). The "proline effect" refers to the preference of Pro-containing peptides to fragment to $y_m^+$ over $b_n^+$ ions, which has been explained by the enhanced basicity of the nascent secondary amine intermediate in the $b_n^+$ ion,[34] which is neutralized by intramolecular proton transfer that results in formation of the $y_m^+$ ion, and by the ring strain[35] of the oxazolone intermediate/product in the $b_n^+$ species. Selective fragmentation such as the "proline effect" can often dominate the CID mass spectrum, subduing nonselective fragmentation channels that are useful for peptide sequencing.[33] Although the relative abundance of $b_n^+$ *vs.* $y_m^+$ fragments can be systematically altered by changing the substitution of the amine

(*i.e.*, changing the ring size from five-membered in proline to six-membered in pipecolic acid),[33,35] the domination of $b_n^+$ fragments in the case of AA$^{D}$PAA and $y_m^+$ fragments in the case of AA$^{L}$PAA is less intuitive because both $^{D}$P and $^{L}$P amino acids should have the same proton affinity and generate oxazolone rings with similar strain. As such, Raulfs *et al.*[33] proposed that the diastereomer-specific fragmentation implies that [AA$^{D}$PAA+H]$^+$ and [AA$^{L}$PAA+H]$^+$ peptides have different conformational preferences. A recent computational study by Abutokaikah *et al.*[36] using density functional theory (M06-2X/6-31+G**) showed a $b_4^+$-$y_1^+$ transition state [$\Delta G(298K) = 101.4$ kJ·mol$^{-1}$], which favored formation of the $b_4^+$ product for [AA$^{D}$PAA+H]$^+$ in agreement with the experiment. It is important to note that the relevant stereo-specific structural differences required to explain the CID behavior occur in the transition state and product ion structures. However, the hypothesis that the inherent conformational preferences of the reactant AA$^{D}$PAA *vs.* AA$^{L}$PAA also differ is intriguing in light of the aforementioned proclivity of $^{D}$PA to induce Type II' $\beta$-turns, which are not generally observed in the $^{L}$PA diastereomers.[6]

In this study, we explore the role of stereochemistry in determining the 3D structures of protonated gas-phase ions of the YA$^{D}$PAA and YA$^{L}$PAA diastereomers, close analogs of AAPAA that incorporate tyrosine at the N-terminus to provide a UV chromophore that enables UV photofragment spectroscopy. If the suggestion by Raulfs *et al.*[33] is correct, then the $^{D}$P- and $^{L}$P-containing peptides should have different conformational preferences that can be deciphered by conformation-specific infrared spectroscopy. As such, the objective of this study is to spectroscopically interrogate [YA$^{D}$PAA+H]$^+$ and [YA$^{L}$PAA+H]$^+$ in a cold gas-phase environment to better understand the inherent conformational preferences arising from the presence of $^{D}$P and $^{L}$P substitution in the absence of solute-solvent or other intermolecular interactions.[25-28]

Recently, photofragmentation spectroscopy of ESI-generated ions that are collisionally cooled to 10 K in an ion trap has emerged as a powerful approach to elucidate the 3D structures of large, gas-phase ions, including peptides,[12,37-42] clusters,[43-49] and catalytic complexes.[50-57] Furthermore, in the presence of a UV chromophore, IR-UV double resonance can be used to record single-conformation IR spectra. In cold [YGGFL+H]$^+$ and [YGGFL-OMe+H]$^+$ ions,[11] we observed $\beta$-turn formation with $G_3$ at the $i$+1 position and $F_4$ at the $i$+2 position, which is reinforced by additional ionic H-bonds from the -$NH_3^+$ group to the acid/ester carbonyl (C17 H-bond) and to the amide carbonyl at the $i$+2 position of the turn (C11 H-bond accepted by $G_3$). Assuming that $^{D}$PA encourages $\beta$-hairpin formation with $^{D}$P at the $i$+1 position and A at the $i$+2 position,[6] one might

expect a 3D structure similar to $[YGGFL+H]^+$ and $[YGGFL\text{-}OMe+H]^+$ because $G_3$ and $F_4$ are coincidentally replaced by $^{D}P_3$ and $A_4$ in $[YA^{D}PAA+H]^+$, placing the anticipated turn at the same location in both of these pentapeptides.

In the present study, we observe two predominant conformer families for $[YA^{D}PAA+H]^+$, which both differ from the single major conformer observed in $[YA^{L}PAA+H]^+$, confirming that the stereochemistry of the proline residues indeed shapes the potential energy landscape of these pentapeptides. We demonstrate that one of the $[YA^{D}PAA+H]^+$ conformers exists as a Type II' $\beta$-turn structure and that the "mirror image" Type II structure expected when $^{L}P$ is substituted for $^{D}P$ is destabilized because of an unfavorable steric interaction between $A_2$ and $P_3$, which we have confirmed by substituting $A_2$ with $G_2$ (*i.e.*, $[YGPAA+H]^+$). Our results test the purported conformational preferences of the $^{D}PA$ and $^{L}PA$ stereoisomers to deepen our understanding of the delicate rules that guide protein folding.

## 4.2 Methods

### 4.2.1 Experimental

All spectroscopic data were obtained using a custom instrument for cold ion spectroscopy, previously described.[12,58] The basic design is comprised of a triple quadrupole mass spectrometer with an orthogonally mounted "spectroscopy axis" between the second quadrupole (q2) and third quadrupole (q3). Ions are generated by nano-ESI (nESI) and isolated in q2 via an auxiliary broadband radio frequency (RF) notched chirp pulse excitation. A turning quad between q2 and q3 is used to steer the isolated precursor ions down the "spectroscopy axis" of the instrument, through which the ions are guided into an octupole ion trap[59] that is cooled to 5K via a closed-cycle helium cryostat (Sumitomo Heavy Industries, Tokyo, Japan). The ions are cooled to about 10 K[59] via collisions with the He buffer gas before spectroscopic interrogation by IR (Laservision OPO/OPA) and UV (ScanMatePro Lambda Physik frequency doubled by an Inrad Autotracker III) lasers. The resulting photofragments are extracted back down the "spectroscopy axis" and turned into a linear ion trap (LIT) for mass analysis (q3). This analysis quadrupole can be run in either of two modes. In "spectroscopy mode", the remaining precursor ions from the LIT are resonantly ejected using an auxiliary waveform calculated by the SX wave software,[60] with all remaining photofragments subsequently "dumped" onto a channeltron detector (4773G, Photonis USA). In

"mass spectrometry mode", the mass spectra of the photofragments are obtained by mass-selective axial ejection (MSAE).[61]

UV action spectra were obtained by recording the total photofragment signal as a function of the laser wavenumber. Having identified UV transitions that may be due to different conformational isomers, conformation-specific IR spectra were obtained by IR-UV double resonance. Following the technique employed by Rizzo and coworkers,[38] the UV laser is fixed on a transition, creating a steady-state photofragment ion signal. The tunable IR laser is timed to precede the UV laser by 200 ns. As the IR laser is scanned, IR absorption by the conformer of interest removes a fraction of its ground state population, resulting in a dip in the photofragment signal when the IR and UV lasers are both on resonance with the same conformer. In the case of conformer B of the $YA^{D}PAA$ species (defined in Section 3.1), we observed the depletions due to conformer B superimposed on gain signals from conformer A. These gain signals were generated by IR absorption of conformer A, which produces a broad UV absorption that is present at the UV wavelength used to monitor conformer B. Therefore, we obtained a conformation-specific IR spectrum of conformer B by subtracting a background gain spectrum with the UV laser tuned a few wavenumbers off resonance (lower wavenumber) from the IR-UV double resonance spectrum taken with the UV laser on resonance. This subtraction is shown in the Supporting Information (Section I).

To prepare $[YA^{D}PAA+H]^{+}$, $[YA^{L}PAA+H]^{+}$, and $[YGPAA+H]^{+}$ gaseous ions, we performed nESI of 200 $\mu$M solutions of $YA^{D}PAA$, $YA^{L}PAA$ (CPC Scientific), or YGPAA (GenScript) in 50:50 MeOH:$H_2O$. Mass spectra are provided in the Supporting Information (Section II), which includes CID and resonant UV photodissociation (UVPD) mass spectra that demonstrate a diastereomer-specific proline effect in YAPAA that is the same as the one observed in AAPAA by Poutsma and coworkers.[33] In fact, the UVPD mass spectra are not only diastereomer-specific but also conformer-specific with respect to the extent of -Tyr (-107) sidechain loss, a well-known loss channel from the electronic excited state of Tyr.[62-68] As shown in the Supporting Information (Section II), conformer A of $YA^{D}PAA$ exhibits photofragment signals associated with both -107 loss and the production of the $b_4^{+}$ fragment, while conformer B of $YA^{D}PAA$ fragments almost exclusively by -107 loss. We took advantage of this fact when recording the conformation-specific IR spectrum of conformer A of $YA^{D}PAA$ by recording the photofragment depletion exclusively from the $b_4^{+}$ channel, which was isolated by apex isolation. This obviated the need to subtract

spectra obtained with the UV laser on- and off- resonance (necessary in the case of conformer B) because the background was eliminated by discriminating against the loss channel associated with the other conformer.

### 4.2.2 Computational

Conformational searches were performed on gas-phase $[YA^{D}PAA+H]^{+}$ and $[YA^{L}PAA+H]^{+}$ structures via the Monte Carlo Multiple Minimum (MCMM) method implemented with the Amber* force field using the MACROMODEL software package.[69] The 100 lowest-energy structures from each search were further optimized at the M05-2X/6-31G* level of theory[70] using GAUSSIAN09 (Version E.01)[71] to refine the energetic ordering of these structures while accounting for dispersive interactions. Structures within 20 kJ/mol of the global minimum calculated at the M05-2X/6-31G(d) level of theory were optimized again using both the M05-2X functional and the B3LYP functional[72-74] with the Grimme's empirical dispersion correction (GD3BJ)[75] and a larger basis set [6-31+G(d)]. For this second tier of density functional theory (DFT) optimizations, tight convergence criteria and an ultrafine grid (int=grid=ultrafine) were used and harmonic level vibrational frequency calculations were also performed. The vibrational spectra reported herein are at the B3LYP-GD3BJ/6-31+G* level of theory with NH, CH, and H-bonded OH stretch frequencies were all scaled by 0.958 to account for anharmonicity. This scaling factor was chosen so that the calculated, scaled harmonic frequency of the free NH stretch fundamental in GPGG (3456 $cm^{-1}$) agrees with that obtained experimentally.[9] The free OH stretch transitions were scaled by 0.973 and the amide I (C=O) transitions were scaled by 0.981 for analogous agreement between the calculated and experimentally measured free acid OH (3567 $cm^{-1}$) and carbonyl (1786 $cm^{-1}$) transitions in $[GG+H]^{+}$, respectively.[76] Amide II transitions were scaled by 0.970 to achieve best agreement between the calculated and experimentally measured bands in the $[YGGFL+H]^{+}$ ion.[58] In this study, we use the B3LYP-GD3BJ/6-31+G(d) level of theory for vibrational frequency calculations because the M05-2X functional with a comparable basis set predicts an incorrect ordering of the free $-NH_3^{+}$ and C7 H-bonded NH stretches in protonated GPGG, as shown in the Supporting Information (Section III, Figure S.5).

The nomenclature used to label the structural families indicates the type of C*n* H-bond formed by the XH groups of each residue, separated by slashes (/) in order from N- to C-terminus. Multiple C*n* H-bonds formed by the same residue (e.g., $NH_3^{+}$) are distinguished by commas. The *cis vs.*

*trans* isomerization of the proline amide bond is also designated between slashes. All C*n* H-bonds involve NH groups unless explicitly indicated by an OH subscript (C$n_{OH}$). In the case of $[YGPAA+H]^+$ we denote the *cis vs. trans* orientation of the acid OH group, which formed two distinguishable families: C$n_{cOH}$ and C$n_{tOH}$. Likewise, $F_{NH}$ and $F_{OH}$ denote a free NH or OH, respectively. The $\pi$ symbol signifies the presence of an H-bond between an NH group and the $\pi$ cloud of the Tyr ring. For instance, the nomenclature for one of the assigned structures of $[YGPAA+H]^+$ is C14,C17,$\pi_{NH}$/$F_{NH}$/*trans*/C10/C7/C13$_{tOH}$, implying that the protonated $Y_1$ -$NH_3^+$ group forms three H-bonds (C14, C17, and to the π-cloud), the $G_2$ NH is free, the $G_2$-$P_3$ amide bond is *trans*, the $A_4$ NH forms a C10 H-bond, the $A_5$ NH forms a C7 H-bond, and the carboxylic acid OH forms a C13 H-bond that is *trans* with respect to the acid carbonyl.

## 4.3 Results and Analysis

### 4.3.1 UV Action Spectra

As a first step, cold UV action spectra of the two diastereomeric peptides, $[YA^DPAA+H]+$ and $[YA^LPAA+H]+$, were obtained by recording the total photofragment signal in all channels (b4+, y3+, and Tyr loss) as a function of the laser wavelength. Indeed, the UV spectra of the two diastereomers (Figure 1) are readily distinguishable from one another, confirming the ability of cryocooled ion spectroscopy to distinguish between them, and suggesting that different conformational preferences of the two diastereomers exist, as alluded to by Raulfs et al.[33] Although neither spectrum has a single simple Franck Condon progression, it is immediately apparent that the $[YA^DPAA+H]+$ spectrum is more complex than that of $[YA^LPAA+H]+$ because it contains irregularly spaced transitions with complicated intensity patterns.

To dissect this complicated spectrum, we turned to conformer-specific IR-UV double resonance spectroscopy to evaluate the contribution of different conformers to the spectra in Figure 1. Infrared spectra that were not conformation-specific were also obtained [Supporting Information Section IV, Figures S.6(a) and S.7(a)] by measuring the gain in resonant photofragment signal when the UV laser was not resonant with any particular conformer, thereby detecting only IR-excited ground state molecules, a method used previously to obtain spectra of the $[YGGFL+H]^+$ ion.[58] In the case of the $^D$P diastereomer, a comparison of the IR photofragment gain spectrum to the conformation-specific IR spectra demonstrated that, with the exception of the band labeled D

at 3435 $cm^{-1}$ [Fig. S.6(a)], all transitions in the non-specific infrared spectrum could be accounted for by combining the conformation-specific spectra obtained with the UV laser fixed on the A and B transitions, and to a much lesser extent, on the C transition.

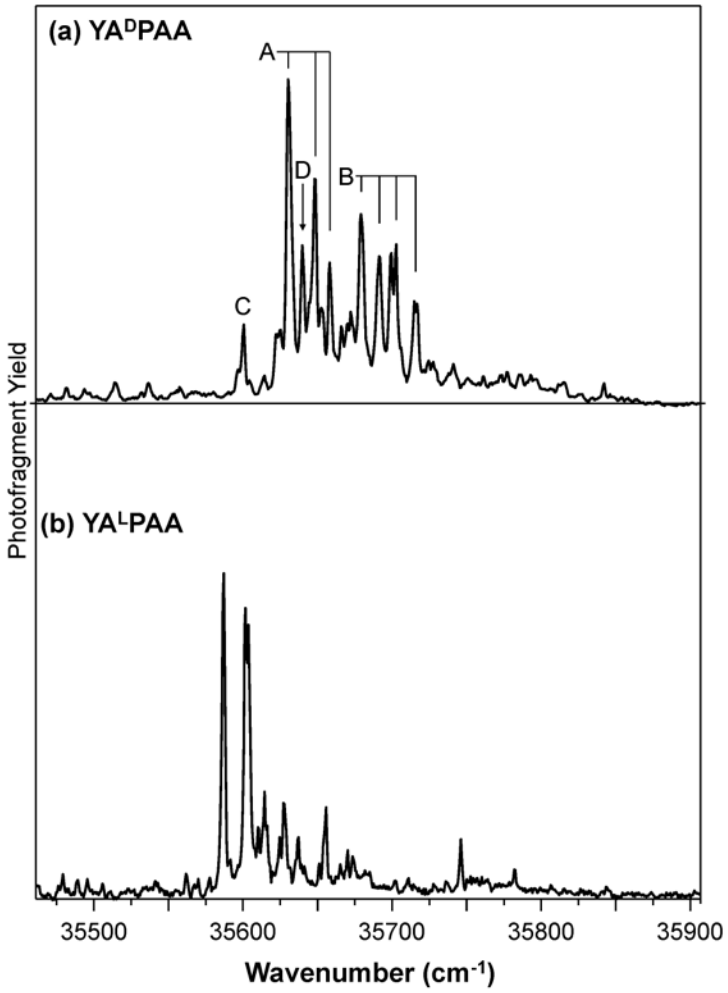


Figure 4.1: UV action spectra of $[YA^{D}PAA+H]^{+}$ (a) and $[YA^{L}PAA+H]^{+}$ (b). The origin bands of different conformers of the $[YA^{D}PAA+H]^{+}$ species are labeled A, B, C , and D in (a).

To assign the transitions in Fig. 4.1 to the different conformers (A-D), we obtained conformation-specific UV spectra by fixing the IR laser on the distinct vibrational transitions associated with each conformer, scanning the UV laser, and subtracting the UV spectra taken with the IR laser off from the spectra with the IR laser on resonance. These results are shown in the Supporting Information (Section IV, Fig. S.8).

In the $^{L}$P diastereomer (Figure S.7), both the non-specific and conformation-specific IR spectra were nearly identical, consistent with the presence of only one dominant conformer in $[YA^{L}PAA+H]^{+}$. Thus, we proceed in our analysis by considering the two major conformers in $[YA^{D}PAA+H]^{+}$ (A and B) and the single predominant conformer in $[YA^{L}PAA+H]^{+}$. In the Supporting Information (Section V), we discuss the vibrational spectra of the two minor conformers of $^{D}$P, which contain H-bonding networks that are similar to one of the major conformers.

### 4.3.2 YA$^{D}$PAA Conformation-Specific Infrared Spectra

The conformation-specific infrared spectra of conformers A and B of $[YA^{D}PAA+H]^{+}$ are given in Figure 4.2. The chemical structure of the sequence is included in Figure 3(a), which highlights the intramolecular H-bonds that are anticipated in the $\beta$-hairpin turn structure. The secondary structure formed by these H-bonds is shown in Figure 4.3(b), which is compared to the common biochemical illustration for this motif. Several of the sharp transitions in the spectra (Figure 4.2) are assignable based on previously published spectra of similar peptides. These include the Tyr free OH stretch transitions at 3648 cm$^{-1}$ in $[YGGFL+H]^{+}$ (3643 cm$^{-1}$),[12] and the free acid OH stretch at 3571 cm$^{-1}$ in conformer A, close to its value in $[GG+H]^{+}$ (3567 cm$^{-1}$).[76] Furthermore, both conformers have free NH stretch fundamentals at 3450 cm$^{-1}$, proximal to that found in $[GPGG+H]^{+}$ (3456 cm$^{-1}$).[9] A table of assignments for the infrared transitions in $[YA^{D}PAA+H]^{+}$, $[YA^{L}PAA+H]^{+}$, and $[YGPAA+H]^{+}$ are given in the Supporting Information (Section VI, Table S.1).

### 4.3.3 Conformer B

Upon comparing the spectra of conformers A and B to each other, we note the absence of the free acid OH stretch in conformer B and the emergence of a broad doubly peaked absorption near 3,000 cm$^{-1}$. Johnson and coworkers[76,77] have observed similar bands in this region of the spectrum, which were attributed to strong cyclic ionic H-bonds in which the X-H fundamental is strongly coupled to a dense manifold of background states.[78,79] In the case of YA$^{D}$PAA, the situation is further complicated by the fact that multiple -$NH_3^{+}$ NH stretches can be engaged in such interactions. Furthermore, the disappearance of the free acid OH stretch in the spectrum of

conformer B suggests that it, too, is engaged in a similarly strong H-bond that contributes to the absorption in this region.

In fact, we have observed strikingly similar broad, doubly peaked bands in the spectrum of $[YGGFL+H]^+$. Figure 2 compares these experimental spectra [traces (b) and (c)]. In the case of $[YGGFL+H]^+$, two $-NH_3^+$ stretching fundamentals as well as the acid OH stretch were calculated to account for these broad features in the experimental spectrum (red bands). Thus, we invoke a similar explanation in the case of conformer B of $[YA^DPAA+H]^+$. Upon further inspection of the experimental spectra, we note that the C10 and C14 H-bonded NH stretching fundamentals of $[YGGFL+H]^+$ are close in wavenumber to distinct transitions found in the $[YA^DPAA+H]^+$ conformer B spectrum.[12]

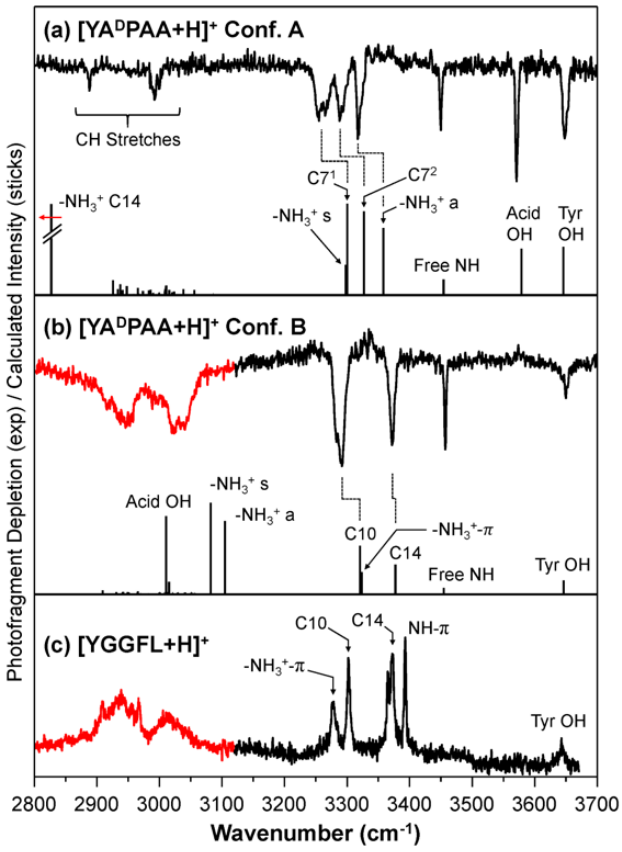


Figure 4.2: Conformation-specific IR spectra obtained using IR-UV double resonance spectroscopy for $YA^DPAA$ conformers A (a) and B (b). The corresponding calculated harmonic level stick spectra (see Section 2.2) are shown below the experimental spectra. The previously obtained[12] spectrum of $[YGGFL+H]^+$ is given in (c). The broad bands colored in red in (b) and (c) contain contributions from the acid OH and $-NH_3^+$ symmetric (s) and asymmetric (a) stretches, which are all engaged in strong H-bonding. The assigned structure for conformer A

contains a stronger -$NH_3^+$ ionic H-bond, which is expected to be anharmonically shifted below the scanned range as denoted by the red arrow. This transition is calculated to be 5.4 times more intense than the $C7^1$ transition and is not shown to scale (double hash through stick). All other labels correspond to those shown in the calculated structures in Figure 4.

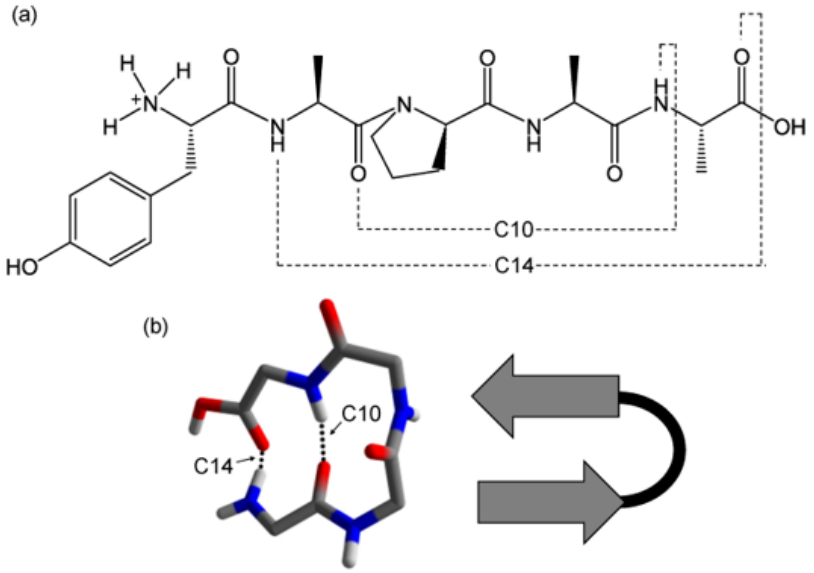


Figure 4.3: Chemical structure of $[YA^DPAA+H]^+$ (a) highlighting the expected C10 and C14 intramolecular H-bonds. A generic peptide backbone containing C10 and C14 H-bonds in a $\beta$-hairpin turn is shown in (b) beside of a cartoon representation of this secondary structure.

As anticipated, the lowest energy conformer with a $\beta$-hairpin motif similar to that of $[YGGFL+H]^+$, uniquely reproduces the experimental splittings and intensity patterns, leading to its assignment as conformer B. Figure 4.4 depicts all assigned structures discussed below and includes a comparison of the assigned structures for conformer B [Figure 4.4(b)] and $[YGGFL+H]^+$ [Figure 4.4(d)], which both comprise a charge-stabilized $\beta$-hairpin structure.

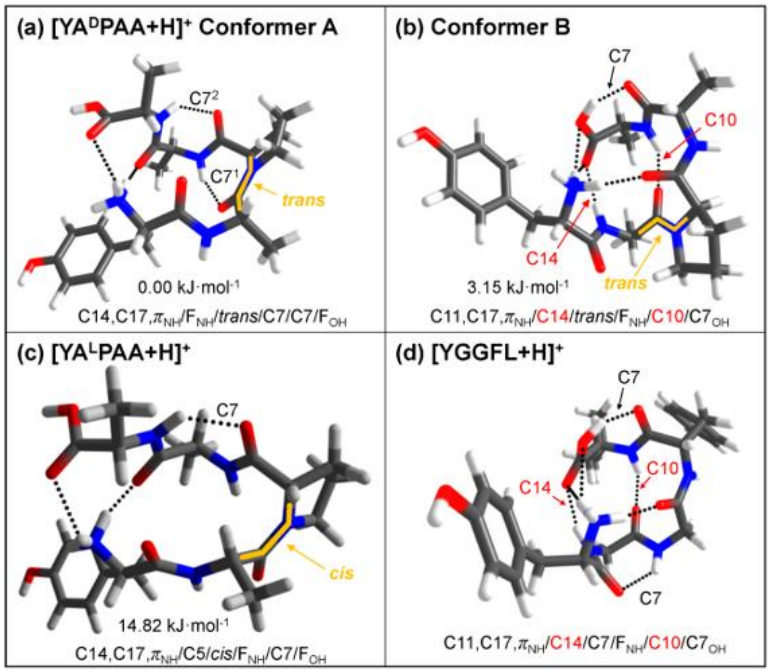


Figure 4.4: Illustrations of assigned 3D structures of $[YA^{D}PAA+H]^{+}$ conformer A (a), $[YA^{D}PAA+H]^{+}$ conformer B (b), $[YA^{L}PAA+H]^{+}$ (c), and $[YGGFL+H]^{+}$ (d). Critical C*n* bonds are labeled on the structures and an assignment of the conformational family is given beneath each structure along with the relative energy ($\Delta G_{rel}$ at 298 K) of the conformer. The *cis vs. trans* proline amide bond is also highlighted in orange.

### 4.3.4 Conformer A

With the assignment of the charge-stabilized $\beta$-hairpin in hand, we turn to conformer A, noting that it lacks the characteristic C14 H-bonded NH stretch at 3372 $cm^{-1}$. The global minimum of the conformational search, based on its relative Gibb's free energy at 298 K ($\Delta G_{rel} = 0$), contains a free acid OH stretch and two H-bonded backbone NH stretches (C14,C17,$\pi$/C14/*trans*/$F_{NH}$/C7/C7/$F_{OH}$). The appearance of the free -COOH transition at 3571 $cm^{-1}$ [Figure 4.2(a)] is accompanied by the disappearance of the broad transitions near 3000 $cm^{-1}$ in conformer B [Figure 4.2(c)], providing further support that the acid OH stretch contributes to the broad feature in conformer B. The weak, sharp transitions below 3000 $cm^{-1}$ are due to the alkyl CH stretch transitions, principally from the methyl groups of the Ala side chains.

In addition to capturing the free NH, acid OH, and Tyr OH stretch transitions, the calculated harmonic level spectrum shown in Figure 4.2(a) also closely reproduces the pattern of the three H-bonded NH stretch transitions. Furthermore, the breadth of the band near 3255 $cm^{-1}$ in the experimental spectrum is potentially explained by a convolution of the $-NH_3^+$ s and $C7^1$ NH stretch

transitions. In contrast to conformer B, conformer A appears to have only one dramatically shifted -$NH_3^+$ vibration, which is calculated to occur at 2827 $cm^{-1}$, 184 $cm^{-1}$ lower in wavenumber than the most strongly shifted -$NH_3^+$ stretching transition in conformer B. In addition, two -$NH_3^+$ stretches (-$NH_3^+$ s and -$NH_3^+$ a) are calculated to occur close in wavenumber to the neutral H-bonded NH stretches ($C_7^1$ and $C_7^2$). These -$NH_3^+$ stretch transitions are comprised of symmetric (s) and anti-symmetric (a) combinations of the NH oscillators involved in the C17 and NH-$\pi$ interactions, which are much less direct and, therefore, weaker H-bonds. As these two transitions approach the limit of a free -$NH_3^+$ group (3370 $cm^{-1}$ in $[GG+H]^+$)[76], the remaining NH oscillator is able to engage in a stronger H-bond, leading to a larger shift to lower wavenumbers. The same effect has been observed when sequentially binding $H_2$, $N_2$, and $CO_2$ tags to the nominally free -$NH_3^+$ groups in $[GG+H]^+$ (*i.e.*, weakening of the singly ionic H-bond as the free NH groups form progressively stronger interactions)[76] and has also been observed for the sequential solvation of protonated water clusters.[80] If the strong ionic H-bond in conformer A approaches the case of a singly ionic H-bond, it may appear shifted to as low as 2500 $cm^{-1}$, below the range scanned, and would become highly broadened, similar to $[GG+H]^+$.[76] As the present data was acquired above 2700 $cm^{-1}$, such a shifted and diffuse feature was not observed.

A comparison of calculated conformer energies, both zero-point corrected and corrected for Gibb's free energy at 298 K, are shown in the Supporting Information (Section VII, Table S.2). A common structural motif identified by the conformational search replaces the C10 and C14 H-bonds of the $\beta$-turn with two sequential C7 H-bonds (C7/C7). Although the calculated global minimum using the zero-point corrected energies at 0 K (C14,C17,$\pi$/C5/*trans*/C7/C7/$C16_{OH}$) also contained a C7/C7 motif, this structure can be ruled out as the assigned structure for conformer A based on the presence of a $C16_{OH}$ interaction rather than a free acid OH group. This structure is illustrated in Figure S.11 of the Supporting Information, which is 4.38 $kJ \cdot mol^{-1}$ lower in energy than conformer A at 0 K but 0.26 $kJ \cdot mol^{-1}$ higher in free energy than conformer A at 298 K. In the Supporting Information, we also provide further evidence for the structure of conformer A based on its vibrational spectrum in the electronic excited state (Section VIII, Figure S.12).

### 4.3.5 $[YA^{L}PAA+H]^{+}$ Conformation – Specific Infrared Spectrum

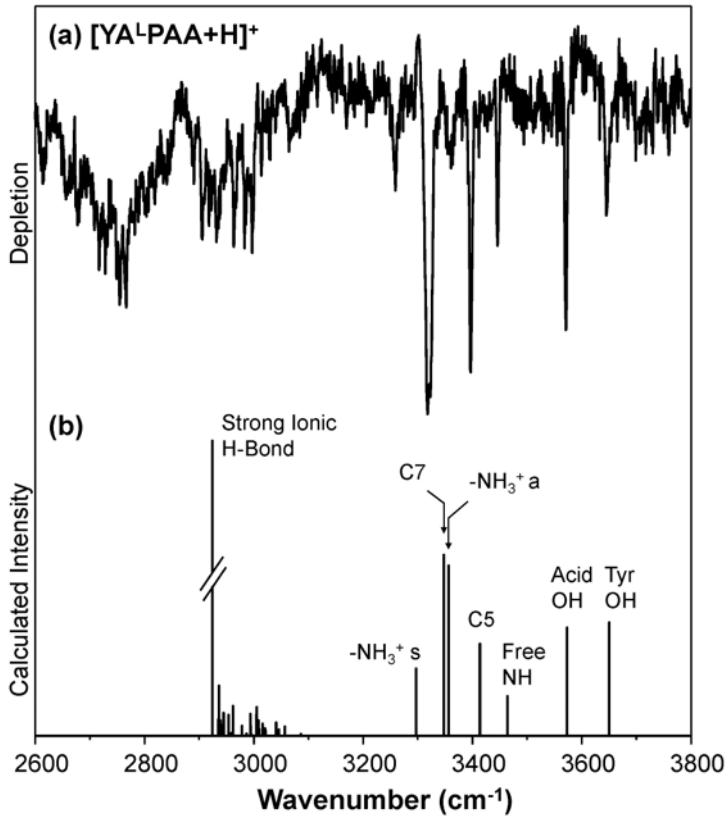


Figure 4.5: Comparison between the conformation-specific IR spectrum of $[YA^{L}PAA+H]^{+}$ taken by IR-UV double resonance (a) and the calculated harmonic level spectrum (see Section 2.2) of the lowest energy C14,C17,$\pi$ /C5/*cis*/$F_{NH}$/C7/$F_{OH}$ structure.

Having characterized the conformational preferences of the $[YA^{D}PAA+H]^{+}$ diastereomer, analogous spectra for the $[YA^{L}PAA+H]^{+}$ stereoisomer were obtained for comparison. Figure 5(a) presents the infrared photofragment depletion spectrum of the single observed conformer of the $^{L}$P diastereomer. As was the case for conformer A of $[YA^{D}PAA+H]^{+}$, the experimental spectrum in the NH and OH stretching region shows excellent agreement with the global minimum from the $[YA^{L}PAA+H]^{+}$ conformational search [Figure 4.4(c)] after including the Gibb's free energy correction at 298 K. The reasons for including the Gibb's free energy correction are considered further in the discussion section, where we present energy level diagrams from the conformational search. The assigned structure is labeled as C14,C17,$\pi$ /C5/*cis*/$F_{NH}$/C7/$F_{OH}$, and is held together largely by the $NH_3^+$ hydrogen bonds, with the turn built around a *cis*-proline conformation. The structure of the global minimum at 0 K (C11,C14,$F_{NH}$/C11/*cis*/$C5_N$/$F_{NH}$/$C13_{OH}$) is shown in the

Supporting Information (Section IX, Figure S.13), which can be ruled out based on the absence of a free Tyr OH stretch. Further evidence for the assigned structures of both the $^{D}$P and $^{L}$P diastereomers is gleaned from the carbonyl stretch region of the spectrum, which is also presented in the Supporting Information (Section X , Fig. S.14).

### 4.3.6 [YG$^{L}$PAA+H]$^{+}$ Conformation – Specific Infrared Spectra

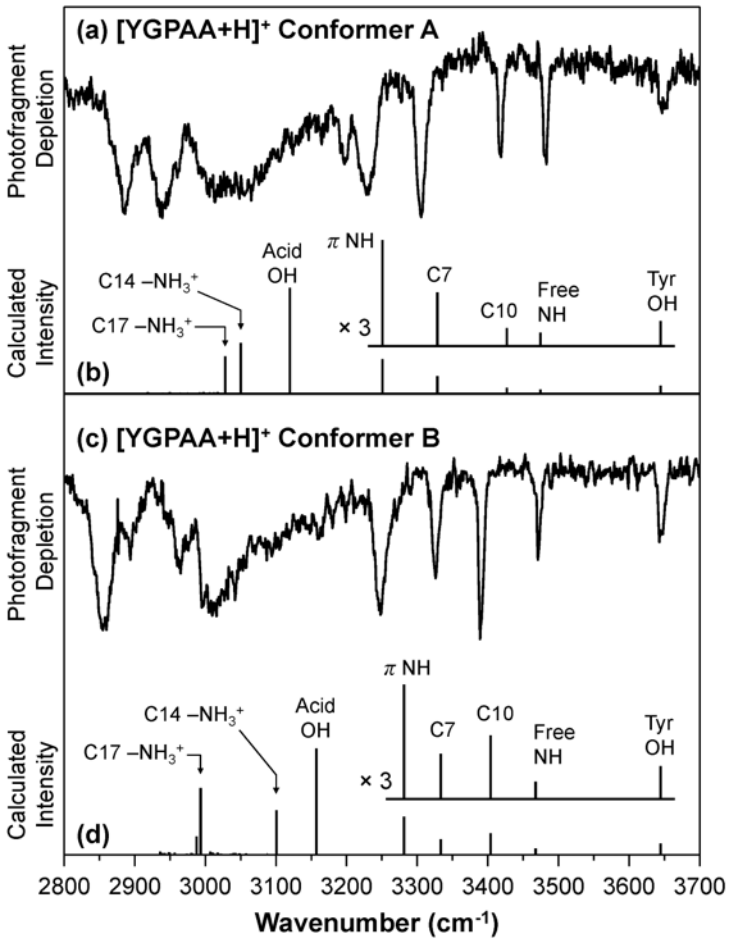


Figure 4.6: Conformation-specific IR spectra of the two major conformers of [YG$^{L}$PAA+H]$^{+}$ taken by IR-UV double resonance spectroscopy: A (a) and B (c). The calculated harmonic level spectra (see Section 2.2) for the lowest energy structures with *trans* and *cis* acid OH groups are given in (b) and (d), respectively

One of the striking aspects of the comparison between the observed conformers of [YA$^{D}$PAA+H]$^{+}$ and [YA$^{L}$PAA+H]$^{+}$ is that the $^{D}$P diastereomer has a *trans*($^{D}$P) configuration, while the $^{L}$P diastereomer is *cis*($^{L}$P). As further described in the Discussion Section, we postulated

that these differences could be attributed to steric hindrance between the $A_2$ and $^{L}P_3$ residues in the *trans*($^{L}$P) configuration. Thus, we hypothesized that removal of the methyl group of $A_2$ might alleviate such strain so that $^{L}$P *trans* structures become favored in $[YG^{L}PAA+H]^+$.

The UV photofragment spectrum of $[YG^{L}PAA+H]^+$ (Section XI, Figure S.15) shows well-resolved vibronic progressions due to three conformers, two of which (A and A') have identical IR spectra and thus most likely belong to two conformers differing only in the orientation of the Tyr OH. Single-conformer infrared spectra of the two distinct conformers of $[YG^{L}PAA+H]^+$ are shown in Figure 4.6. Both spectra show excellent agreement for the sharp transitions above 3200 $cm^{-1}$ with the calculated harmonic level vibrational frequencies and IR intensities of the two lowest energy conformers. The corresponding IR spectra in the Amide I and II regions are given in the Supporting Information (Section XI, Fig. S.16). Based on the hydride stretch region shown in Figure 4.6, the absence of an acid free OH stretch implies that that this oscillator is engaged in a hydrogen bond, unlike conformer A of $[YA^{D}PAA+H]^+$ and $[YA^{L}PAA+H]^+$. Furthermore, the agreement between the calculated and experimentally measured C10 NH stretch provides evidence for the presence of a $\beta$-turn.

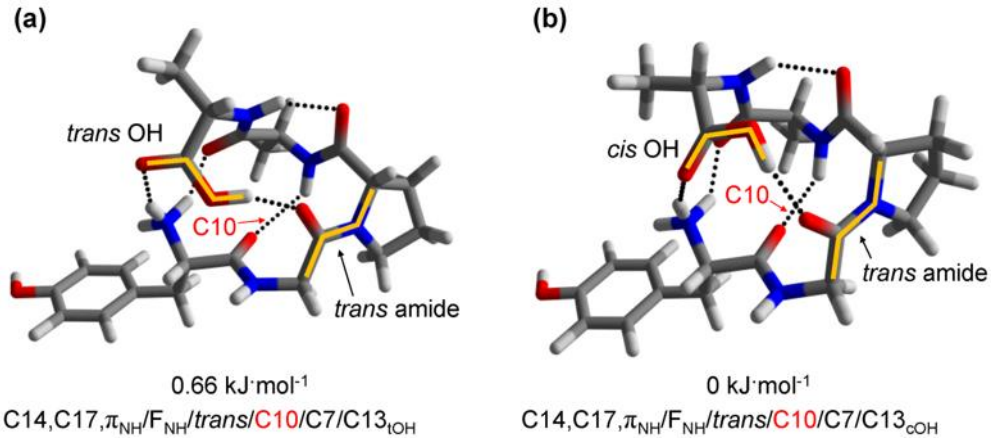


Figure 4.7: Assigned structures of conformers A (a) and B (b) of $[YGPAA+H]^+$. Critical C*n* bonds are labeled on the structures and an assignment of the conformational family is given beneath each structure along with the relative energy ($\Delta G_{rel}$ at 298 K) of the conformer. The *trans* proline amide bonds and *cis vs. trans* acid OH groups are also highlighted in orange.

The assigned structures for conformers A and B of $[YG^{L}PAA+H]^+$ are presented in Figure 4.7. In both of these structures, the same atoms are connected by H-bonds and the main difference is the orientation of the acid OH group, which is *trans* in conformer A but *cis* in conformer B relative

to the C=O group. As anticipated, the amide bond of the $^{L}$P residue occupies a *trans* configuration in both conformers, similar to $[YA^{D}PAA+H]^{+}$. Importantly, although both conformers contain a C10 H-bond, the location of the *β*-turn is shifted by one residue, with proline occupying the *i*+2 position of the turn in $[YG^{L}PAA+H]^{+}$ as opposed to the *i*+1 position in $[YA^{D}PAA+H]^{+}$ conformer B.

## 4.4 Discussion

Using IR-UV double resonance methods to record conformation-specific infrared spectra, we have characterized the two predominant conformers of $[YA^{D}PAA+H]^{+}$ and the single major conformer of $[YA^{L}PAA+H]^{+}$, and shown unequivocally that they are structurally unique from one another. Therefore, we confirm the suggestion of Raulf *et al.*[33] that protonated peptide sequences that are $^{D}$P and $^{L}$P diastereomers of one another have different conformational preferences in the gas phase, albeit in a close analogue containing a Tyr residue (YAPAA instead of AAPAA). Based on the close match between experiment and theory, it is clear that a common secondary structure, the Type II' *β*-turn, is present as one of two competing structural types in $[YA^{D}PAA+H]^{+}$ and that this motif is stereochemically switched off in $[YA^{L}PAA+H]^{+}$. This is in keeping with the use of $^{D}$PX to initiate *β*-turn formation in peptide sequences in solution.[6] Furthermore, the structure of the single observed conformer of $[YA^{L}PAA+H]^{+}$ contains a *cis* proline configuration rather than *trans*, providing deeper insight into the structural consequences of substitution of $^{D}$P for $^{L}$P in peptide sequences. This change in preference, *trans* ($^{D}$P) → *cis* ($^{L}$P), is reversed when Gly is substituted for Ala at position 2, as demonstrated from single-conformation studies of $[YG^{L}PAA+H]^{+}$. In the sections that follow, we probe more deeply these structural preferences and their implications.

### 4.4.1 Gas Phase β-Hairpin Formation: $[YA^{D}PAA+H]^{+}$ vs. $[YGGFL+H]^{+}$

As anticipated, one of the $[YA^{D}PAA+H]^{+}$ conformers (B) is folded into a *β*-hairpin structure comprising a type II' *β*-turn and the first 'rung' of an anti-parallel *β*-sheet (C14 H-bond), consistent with the solution phase studies that use $^{D}$PX to initiate anti-parallel *β*-sheet formation in solution.[6,21] The Ramachandran angles of the ideal Type II' *β*-turn are compared with those of $[YA^{D}PAA+H]^{+}$, $[YG^{L}PAA+H]^{+}$, and $[YGGFL+H]^{+}$ in Table 1. While the structural similarities

to the ideal turn are evident, both protonated ions show similar deviations that are driven by accommodation of the $-NH_3^+$ charge center in this protonated peptide. In fact, one of the important deductions of previous studies of $[YGGFL+H]^+$ and its methoxy-substituted analog, $[YGGFL\text{-}OMe+H]^+$,[11] was that the $-NH_3^+$ charge site stabilizes the $\beta$ -hairpin by forming H-bonds to C=O groups in the turn and at the C-terminus, which otherwise would not be stabilized by amide-amide H-bonds. This charge stabilization of the type II' $\beta$ -turn appears also to play a similar role in conformer B of $[YA^DPAA+H]^+$. Binding the $-NH_3^+$ to the C=O group of the C-terminal COOH group (via a C17 H-bond) also induces the OH group to reconfigure into a *trans* geometry in which it can form a strong, short C7 H-bond to the $A_4$ C=O group.

Table 4.1: Ramachandran Angles (°) of $[YADPAA+H]^+$ Conformer B and $[YGGFL+H]^+$ calculated at the B3LYP-GD3BJ/6-31+G* level of theory[a]

| | $\phi_{i+1}$ | $\psi_{i+1}$ | $\phi_{i+2}$ | $\psi_{i+2}$ |
|---|---|---|---|---|
| Ideal | 60 | -120 | -80 | 0 |
| $[YA^{D}\underline{P}AA+H]^+$ Conformer B | 47 | -136 | -68 | -18 |
| $[YG^{L}\underline{P}AA+H]^+$ | 59 | -135 | -86 | -11 |
| $[YG\underline{GF}L+H]^+$ | 54 | -131 | -64 | -19 |

[a]The underlined residues denote the *i*+1 and *i*+2 positions of the $\beta$-turn.

As $^DPX$ (X= G or A) segments are commonly used to induce the Type II' $\beta$-turn in solution, it is surprising that conformer A (C14,C17,$\pi_{NH}$/$F_{NH}$/*trans*/C7/C7/$F_{OH}$), which is not a $\beta$-hairpin loop, is also observed. This result is especially noteworthy because our previous studies showed that $[YGGFL+H]^+$ and its methyl esterified derivative[11,12] exclusively formed the Type II' $\beta$-turn. Because the turn occurs between $G_3$ and $F_4$, one might expect substitution of $^DP_3$ and $A_4$ for these residues in $[YA^DPAA+H]^+$ should only further enforce the same $\beta$-turn based on the aforementioned guidelines[6] from the solution-phase studies. In the case of $[YGGFL+H]^+$, the small achiral glycine residues minimize any sidechain-sidechain interactions; therefore, the folding of this sequence should be strongly influenced by the intramolecular solvation of the $-NH_3^+$ group. Furthermore, given the enhanced strength of these ionic H-bonds, one could argue that the structure of any protonated pentapeptide may coincidentally exhibit the Type II' $\beta$-turn. Such a

notion can be clearly ruled out in light of the unique structure for conformer A. Based upon the structural differences observed between $[YA^{D/L}PAA+H]^+$ and $[YG^{L}PAA+H]^+$, it is clear that both the stereochemistry of the proline group and the methyl substitution of the alanine groups influence the inherent conformational preferences of these pentapeptides.

The structure assigned to conformer A of $[YA^{D}PAA+H]^+$ differs from the $\beta$ -hairpin structure of conformer B first in having a different mode of self-solvating the $-NH_3^+$ group. The C11 $-NH_3^+$ H-bond to $G_3$ in conformer B is broken in $[YA^{D}PAA+H]^+$ conformer A. Instead, a strong C14 and weak C17 H-bond are formed by the $-NH_3^+$ group in conformer A with the carbonyl groups of $A_4$ and the C-terminus -COOH group ($A_5$), respectively [Figure 4(a)]. In so doing, the carboxylic acid OH no longer forms a H-bond with the $A_4$ C=O, and reverts to its otherwise lower-energy *cis* configuration. The C10 and C14 H-bonds that characterize the $\beta$ -hairpin in conformer B are replaced with a sequential pair of C7 H-bonds in conformer A, one of which reaches across the $^{D}P$ residue, creating a sharp $\beta$ -turn that enables the strong binding of the $-NH_3^+$ group with the $A_4$ C=O.

The close structural similarity between conformer B of $[YA^{D}PAA+H]^+$ and the sole observed conformer of $[YGGFL+H]^+$ exists despite the loss of a C7 H-bond ($G_3 \rightarrow Y$) when a proline residue is substituted at *i+3* position. It seems likely that the destabilization associated with loss of this H-bond may bring structures such as the one corresponding to conformer A into closer energetic proximity in $[YA^{D}PAA+H]^+$, where they can compete with $\beta$ -hairpin formation. Therefore, the present study provides evidence that the inherent conformational preferences of the individual amino acids and their differing abilities to form and accept neutral amide-amide H-bonds (*e.g.*, the nearest-neighbor C7 H-bond) modulate the relative stabilities of the various competing secondary structural motifs, despite the presence of the strong ionic H-bonds involving the $NH_3^+$ group.

### 4.4.2 The preference for cis-$^{L}$P in $[YA^{L}PAA+H]^+$

The single observed conformer of $[YA^{L}PAA+H]^+$ is that shown in Figure 4.4(c), a loosely-formed loop held together principally by the same $-NH_3^+ \leftrightarrow$C-terminal binding as in conformer A of $[YA^{D}PAA+H]^+$ (Figure 4.4a). In fact, the only amide-amide H-bonds in the structure are a C7 H-bond across $A_4$, which is also a structural element held in common with $[YA^{D}PAA+H]^+$ conformer A, and a weak C5 about $A_2$. The other C7 in $[YA^{D}PAA+H]^+$ (conformer A) is broken in $[YA^{L}PAA+H]^+$ in favor of forming a *cis*-Pro structure that sits at the apex of the loop.

The spectroscopic evidence for this *cis*-Pro structure is substantial. The distinct pattern of NH stretch fundamentals in the region above 3200 $cm^{-1}$ is matched remarkably well by the calculations (Figure 4.5), reflecting the open structure of the *cis*-Pro loop. Further support comes from the Amide I and II regions (Supporting Information Section X, Figure S.14) that show a higher wavenumber carbonyl transition (1769 $cm^{-1}$) in $[YA^{L}PAA+H]^{+}$, which is characteristic of the *cis*-Pro amide bond.

In typical peptide bonds, *trans*-amides are more stable than *cis*-amides by about 15 kJ/mol.[81] In Pro, this energy difference is reduced, rendering Pro as the only amino acid with feasible *cis-trans* isomerization about the amide bond. From a survey of the protein databank, 5% of Xaa-Pro bonds are *cis*, compared to only 0.03% of Xaa-nonPro bonds. The abundance of *cis* Xaa-Pro increases to about 30% for short peptides in solution at room temperature.[82]

Precedence for this isomerization also exists in the gas phase. For example, von Helden and coworkers[83] have provided evidence for the presence of *cis*-amide isomers in multiply charged ubiquitin ions based on ion mobility (larger collision cross-section for *trans*- than for *cis*-isomers). Isomerization about the Pro amide bond was also observed recently by Rizzo and coworkers, in their studies of the $[GPGG+H]^{+}$ ion[9] and the bradykinin 1-5 fragment.[10] In these studies, the *cis* and *trans* isomers were separated based on their ion mobilities and were distinguished using cryogenic ion vibrational predissociation spectroscopy. In contrast to $[GPGG+H]^{+}$, which contains both *cis* and *trans* isomers,[9] we observe only the *cis* amide bond in $[YA^{L}PAA+H]^{+}$ and only *trans* amide bonds in $[YA^{D}PAA+H]^{+}$. Note that the $^{L}P$ and $^{D}P$ stereoisomers of $[GPGG+H]^{+}$ are spectroscopically indistinguishable because the molecule only contains one chiral center.

### 4.4.3 Gas Phase Energy Level Diagrams for $[YA^{D}PAA+H]^{+}$ and $[YA^{L}PAA+H]^{+}$

To provide a more global framework in which to assess the origin of the stereochemical preference for the *cis* Pro amide bond in $[YA^{L}PAA+H]^{+}$ and the *trans* configuration in $[YA^{D}PAA+H]^{+}$, Figure 8 presents the relative free energies (298 K) of structures from the Monte Carlo conformational search at the DFT B3LYP-GD3BJ/6-31+G* level of theory. Results from the DFT M05-2X/6-31+G* level of theory are shown in the Supporting Information (Section XII, Figure S.17).

Based on the results in Figure 8, several critical deductions can be made. First, *trans* amide bonds are greatly preferred in the $^{D}P$ diastereomer, while structures containing *cis* amide bonds are

somewhat more favorable in the $^{L}$P diastereomer. Second, many of the $^{D}$P structures are lower in energy (as much as 15 kJ·mol$^{-1}$) than $^{L}$P global minimum, pointing toward an inherent stabilization of the [YA$^{D}$PAA+H]$^{+}$ diastereomer over [YA$^{L}$PAA+H]$^{+}$. Third, although conformer A of [YA$^{D}$PAA+H]$^{+}$ and the assigned structure of [YA$^{L}$PAA+H]$^{+}$ are global minima, there are ten structures that are lower in energy than the assigned structure of [YA$^{D}$PAA+H]$^{+}$ conformer B.

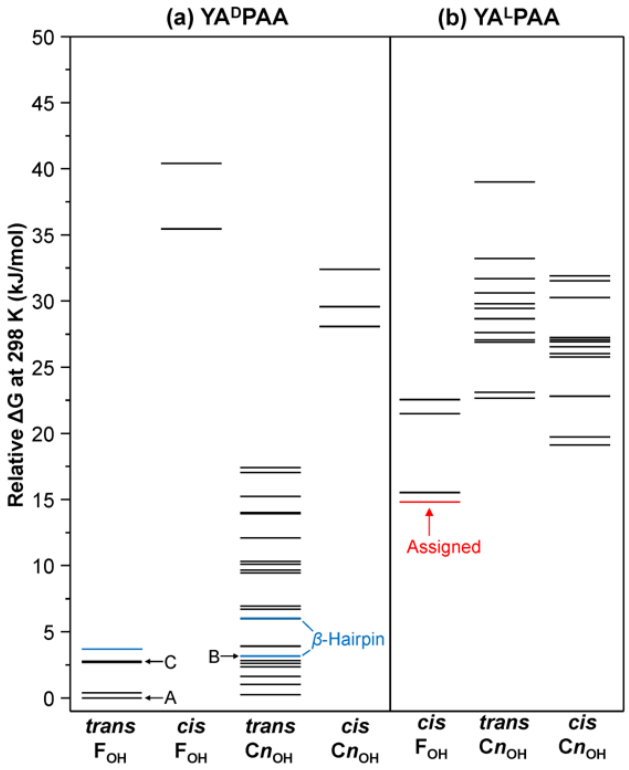


Figure 4.8: Potential energy diagrams for [YA$^{D}$PAA+H]$^{+}$ (a) and [YA$^{L}$PAA+H]$^{+}$ (b), which plot the calculated relative ΔG at 298 K for the various families of conformers at the B3LYP-GD3BJ/6-31+G* level of theory. These include *trans* and *cis* configurations of the AP amide bond as well as free (F$_{OH}$) *vs.* cyclic H-bonded (C*n*$_{OH}$) acid OH groups. The blue lines indicate structures in the *β*-hairpin turn family, similar to the one assigned in [YGGFL+H]$^{+}$. The red line in [YA$^{L}$PAA+H]$^{+}$ indicates the global minimum, which is also the assigned structure

Energy level diagrams at 0 K that were constructed using zero-point corrected energies are provided in the Supporting Information (Section XII, Figure S.18), which also tabulates the energies and H-bonding designations of the conformers (Section VII, Table S.2). It is important to note that all of the global minima structures based on thermal free energies at 298 K showed reasonable agreement between their calculated vibrational spectra and experimental spectra, while

this was not the case for the 0 K global minima obtained using zero-point corrected energies. Such a result is corroborated by the recent study of the bradykinin 1-5 fragment ion by Rizzo and coworkers,[10] who also showed better agreement between experiment and theory when considering the ΔG(298 K) free energy surface. However, based on the calculated energies in Figure 8, the observation of conformer B of $[YA^{D}PAA+H]^{+}$ seems to suggest that ΔG(298 K) is not the sole criteria in determining which conformers are observed in the experiment because it has been uniquely identified, ruling out the lower energy possibilities for this additional conformer. We do not believe that the observed structures were kinetically trapped from solution as has been observed in ion mobility studies[84] because the UV spectra taken under harsh ion funnel conditions and after the application of collisional activation in the linear ion trap prior to the cryo-cooled octupole trap were indistinguishable from the UV spectra of ions generated by softer conditions. The question of how the observed conformers at 10 K are selected during the collisional cooling process warrants further study. Therefore, we are currently calculating barrier heights to conformer interconversion and performing Rice–Ramsperger–Kassel–Marcus (RRKM) theory calculations to estimate the relative populations of low-energy conformers after the collisional cooling process and are also measuring these relative populations by IR-UV population transfer hole-filling spectroscopy.

In the Supporting Information (Section XIII, Figure S.19) we show a comparison between the calculated harmonic level spectrum of the lowest energy *trans* C$n_{OH}$ conformer (the global minimum at 0 K) and the experimental spectrum of conformer B. In addition to the poor agreement in the NH stretching region, the acid OH stretch is shifted to higher frequency than its location in the assigned structure (3316 vs. 3011 $cm^{-1}$). Such a shift is also observed in the other lower-energy structures and can be attributed to weakening of the C$n_{OH}$ H-bond, which is the strongest H-bond in the case of the assigned structure ($C7_{OH}$). As the broadened feature is quite similar between the experimental spectra of $[YGGFL+H]^{+}$ and $[YA^{D}PAA+H]^{+}$ conformer B [red bands in Figures 2(b) a 2(c)] and such a large calculated shift to higher frequency for $[YA^{D}PAA+H]^{+}$ conformer B would require emergence of a new band near the sharp NH stretches, it is highly unlikely that any of the remaining conformers are feasible. In both the previous assignment for $[YGGFL+H]^{+}$ and $[YA^{D}PAA+H]^{+}$ conformer B, the $C7_{OH}$ stretch is calculated to occur beneath the broadened bands [red in Figures 2(b) and 2(c)] in the experimental spectrum (2978 $cm^{-1}$ in $[YGGFL+H]^{+}$ and 3011 $cm^{-1}$ in $[YA^{D}PAA+H]^{+}$). The absence of the structures that are lower in energy than

conformer B suggests that there is some degree of funneling between these structures and conformer A as the ions are cooled. To better understand the origin of the observed structures by assessing the dynamics of the cooling process, future studies are needed quantify the relative populations of the conformers and estimate the barrier heights to isomerization using IR-induced population transfer schemes.[85,86]

### 4.4.4 Influence of Steric Hindrance on β-Hairpin Formation

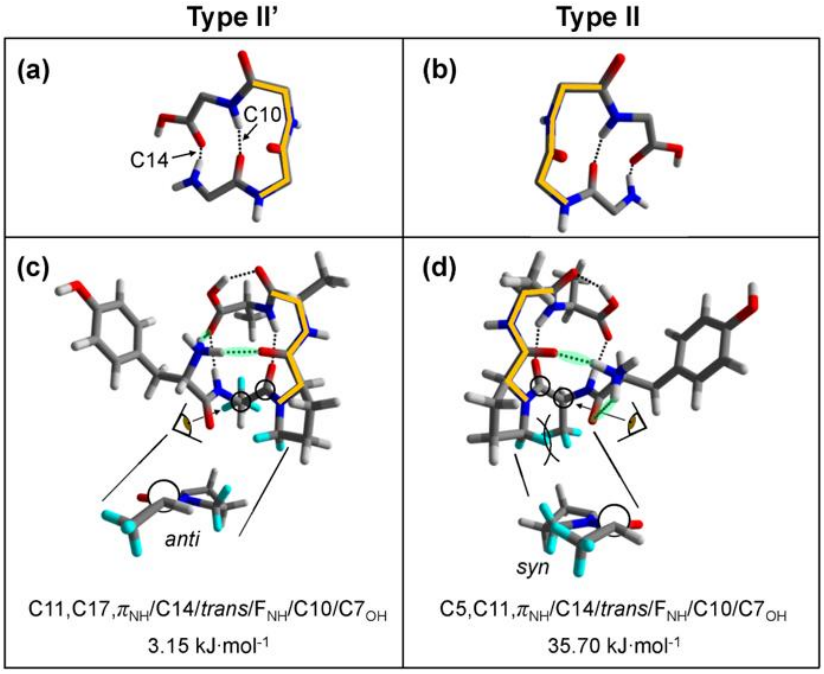


Figure 4.9: Comparison of the Type II' (left column) and Type II (right column) *β*-turn structures. The mirror image backbones are highlighted in orange. Ideal *β*-turns in generic pentapeptides with omitted sidechains are compared in (a) and (b), while the corresponding mirror-image backbone structures of $[YA^{D}PAA+H]^{+}$ and $[YA^{L}PAA+H]^{+}$ are compared in (c) and (d). Energies (ΔG at 298 K relative to the global minimum of each sequence: $[YAPAA+H]^{+}$ or $[YGPAA+H]^{+}$) are shown below each structure as well as conformational family assignments. Hydrogen atoms colored in turquoise denote the $A_2$ methyl and $P_3$ methylene groups, which are sterically encumbered in the *syn* (d) as opposed to the *anti* orientation (c). These interactions are highlighted by Newman projections down the $A_2$ $C_\alpha$-$C_{CO}$ bond (circled atoms). The perspectives of the Newman projections are illustrated by the arrows, which extend from the eyes in (c) and (d). H-bonds formed by the $-NH_3^{+}$ group are highlighted in green.

Having identified low-energy structures describing the different conformational preferences, we return to the original hypothesis of Gellman[6] that peptides containing $^{D}$PA favor Types I' or II'

$\beta$-turns that position adjoining amino acids on either side of the β-turn forming $\beta$-hairpins, while peptides containing $^{L}$PA favor Types I and II $\beta$-turns, that discourage $\beta$-hairpin formation. To further assess this possibility, we constructed a structure for $[YA^{L}PAA+H]^{+}$, which was similar to $[YA^{D}PAA+H]^{+}$ conformer B (Type II' $\beta$-turn), but reversed the sign of each of the Ramachandran angles ($\phi$ and $\psi$) to form a Type II $\beta$-turn. Figure 4.9 illustrates the Type II' (left column) and Type II (right column) classes of $\beta$-turns, comparing each of the calculated structures to the ideal backbone geometries given in Figures 4.9(a) or 4.9(b), where the turn is highlighted in orange. Figure 4.9(c) [same structure in Figure 4.4(b)] reproduces the structure of $[YA^{D}PAA+H]^{+}$ conformer B for comparison with the proposed Type II $\beta$-turn structure for $[YA^{L}PAA+H]^{+}$. Although changing the sign of the $\phi$ and $\psi$ angles forms structures that are pseudo "mirror images" with respect to the peptide backbones, these structures are not true mirror images because the true mirror images would change the chirality of all the non-Pro amino acid chiral centers from L- to D, while the structure in Figure 4.9(d) switches only the $^{D}$P to $^{L}$P but keeps all the others unchanged.

The structural difference between the Types II and II' structures of $[YA^{L}PAA+H]^{+}$ and $[YA^{D}PAA+H]^{+}$, respectively, arises from the changed orientation of the $A_2$ methyl and $P_3$ methylene groups, which are colored in turquoise in Figure 4.9. As the Newman projections illustrate, in the $[YA^{D}PAA+H]^{+}$ type II' $\beta$-turn, the $A_2$ methyl group is *anti* to the $P_3$ methylene with little steric interaction [Figure 4.9(c)], while in the $[YA^{L}PAA+H]^{+}$ counterpart, the type II $\beta$-turn has these groups *syn* with respect to the $A_2$ $C_\alpha$-$C_{CO}$ bond, where steric hindrance with between $A_2$ methyl and $P_3$ methylene is large. Indeed, while the neutral H-bonding patterns are nominally the same for these two diastereomers, $[YA^{D}PAA+H]^{+}$ conformer B [Figure 4.9(c)] places the $-NH_3^{+}$ group in C11 and C17 H-bonds, the optimized structure for the $[YA^{L}PAA+H]^{+}$ Type II $\beta$-turn analogue [Figure 4.9(d)] has the $NH_3^{+}$ group re-orienting to form C5 and C11 H-bonds. Because the H-bonding environment of the $-NH_3^{+}$ was equivalent in both starting geometries of the two diastereomers (C11 and C17), we surmise that addition of the $A_2$ methyl group alters the H-bonding environment of the $-NH_3^{+}$ after optimization. The combined result of this weaker ionic H-bonding interaction and/or increased steric hindrance between the $A_2$ methyl and $P_3$ methylene H-atoms raises the energy of the Type II $\beta$-turn structure for $[YA^{L}PAA+H]^{+}$ to 32.5 kJ·mol$^{-1}$ above that of its diastereomer counterpart. Note that the minimum H-H distance between the turquoise-colored H-atoms is 2.33 Å in Figure 4.9(d) [2.28 Å before optimization], which is less than twice

the H atom van der Waals radius (2.4 Å).[87] Our study of $[YG^{L}PAA+H]^{+}$ was motivated by the desire to test the hypothesis that steric hindrance destabilizes the Type II $\beta$-turn in $[YA^{L}PAA+H]^{+}$, by removing the methyl group responsible for such steric effects. However, as we have seen, the observed structures of $[YG^{L}PAA+H]^{+}$ show unexpected differences from $[YA^{L}PAA+H]^{+}$, which are interesting in their own right.

### 4.4.5 The preference of trans-$^{L}$P in $[YG^{L}PAA+H]^{+}$

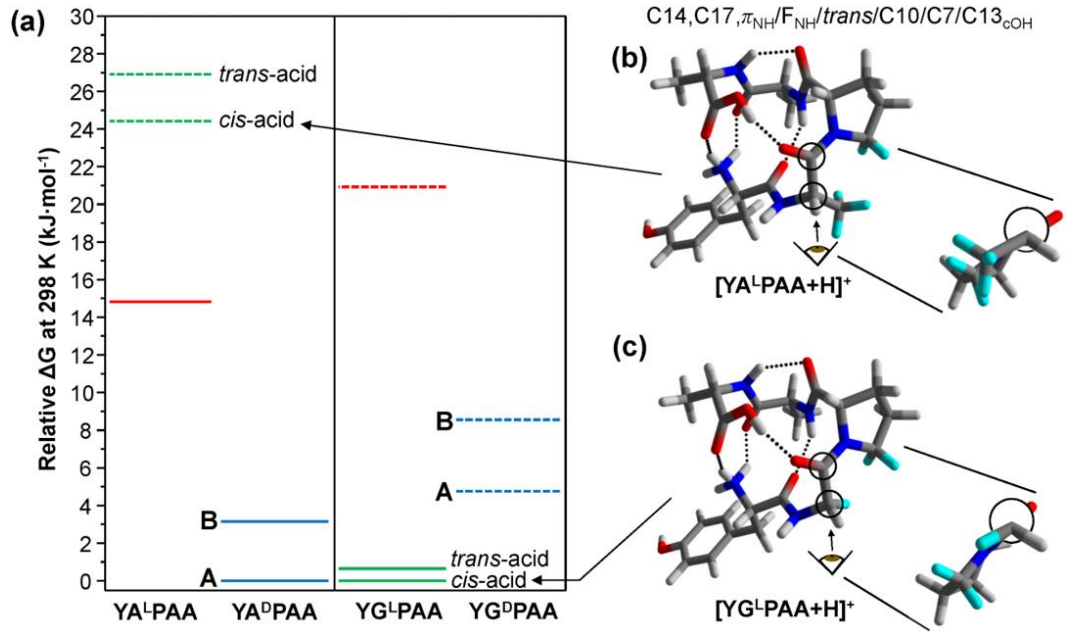


Figure 4.10: Energy level diagrams (a) showing the calculated relative free energies, ΔG(298K), of assigned $[YA^{D/L}PAA+H]^{+}$ and $[YG^{L}PAA+H]^{+}$ structures (solid lines) and relative energies of unassigned structures (dashed lines). The unassigned structures are similar to the assigned structures of $[YAPAA+H]^{+}$ and $[YGPAA+H]^{+}$ but remove and add the $A_2$ methyl group, respectively, to form analogous $[YGPAA+H]^{+}$ and $[YAPAA+H]^{+}$ structures. Corresponding methylated and demethylated structures are indicated by the groups of solid and dashed lines with the same color. Example structures are given in (b) and (c) that are based on the *cis*-acid global minimum conformer of $[YG^{L}PAA+H]^{+}$ (C14,C17,$\pi_{NH}$/$F_{NH}$/*trans*/C10/C7/$C13_{OH}$). Newman projections show steric interactions in the $[YA^{L}PAA+H]^{+}$ analogue that are absent in $[YG^{L}PAA+H]^{+}$.

The assigned structures for $[YG^{L}PAA+H]^{+}$ shown in Figure 4.7 are surprising in several ways. First, the conformer contains a *trans-Pro* ($G_2$-$P_3$) amide bond in contrast to the *cis*-Pro found in the single observed conformer of $[YA^{L}PAA+H]^{+}$. Second, the $\beta$-turn formed in either of the

assigned structures of $[YG^{L}PAA+H]^{+}$ occurs at the *i*+2 rather than the *i*+1 position. Finally, the $^{L}P$ residue is engaged in a Type II' rather than Type II $\beta$-turn. To understand why the anticipated Type II $\beta$-turn is not formed, we optimized each of the assigned $[YAPAA+H]^{+}$ structures after removal of the $A_2$ methyl group to form analogous $[YGPAA+H]^{+}$ structures, and the assigned $[YGPAA+H]^{+}$ structures after inclusion of this methyl group to form $[YAPAA+H]^{+}$ structures. The resulting energy diagram is given in Figure 4.10(a), where the solid lines denote the relative free energies of the assigned structures and the dashed lines indicate the free energies of structures that were modified to either include or exclude the methyl group. More complete energy level diagrams of $[YG^{L}PAA+H]^{+}$ are provided with further discussion in the Supporting Information (Section XIV, Figures S.20-S.21). Interestingly, the relative free energies of conformers A and B of $[YA^{D}PAA+H]^{+}$ (solid blue) and the assigned structure of $[YA^{L}PAA+H]^{+}$ (solid red) are largely retained in the analogous unobserved $[YG^{D}PAA+H]^{+}$ (dashed blue) and $[YG^{L}PAA+H]^{+}$ (dashed red) structures. What does change in dramatic fashion is the relative free energies of the two assigned $[YG^{L}PAA+H]^{+}$ structures (solid green), which are the two lowest in energy in $[YG^{L}PAA+H]^{+}$ but after addition of the methyl group their relative energies rise sharply to more than 10 kJ/mol above the global minimum of $[YA^{L}PAA+H]^{+}$ (dashed green).

Based on the calculated geometries, it appears that steric hindrance of the methyl group is responsible for this rise in energy of the $[YA^{L}PAA+H]^{+}$ structures relative to $[YG^{L}PAA+H]^{+}$. The steric interactions are highlighted in Figures 4.10(b) and 4.10(c) by the aqua colored H-atoms, which are shown in a Newman projection for both $[YA^{L}PAA+H]^{+}$ and $[YG^{L}PAA+H]^{+}$. Indeed, the $A_2$ methyl group, which is bulkier than the $G_2$ H-atom, leads to greater steric overlap despite the similar H-bonding environments in both diastereomers ($C14,C17,\pi_{NH}/F_{NH}/$*trans*$/C10/C7/C13_{OH}$). Furthermore, inspection of a typical Ramachandran plot[88] reveals that the $G_2$ residues of conformers A and B of $[YG^{L}PAA+H]^{+}$ have ($\phi,\psi$) angles [Supporting Information Section XIV, Table S.4: (59.1°,-134.7°) for A and (62.2°,-146°) for B] in what is essentially a forbidden region for pre-proline residues. Such an uncommon geometry is facilitated by the smaller and more flexible glycine residue, which reduces unfavorable steric interactions. Therefore, the shifted location of the $\beta$-turn in $[YG^{L}PAA+H]^{+}$, where $^{L}P$ is at the *i*+2 position, verifies the importance of the residues surrounding the $^{D/L}PA$ in shaping the secondary structure.

Going forward, it will be important to carry out similar studies on longer peptide sequences designed to favor $\beta$-hairpin formation, where additional H-bonds in the intra-strand anti-parallel $\beta$-sheet might stabilize the hairpin relative to competing possibilities. The present study has shown that, in protonated pentapeptides, formation of this minimalist $\beta$-hairpin is sensitively dependent on the stereochemistry (D versus L) and amide structure (*cis* vs. *trans*) of the Pro residue as well as the identity of the residues immediately adjacent to the turn (A versus G). In extending the $\beta$ -hairpin, the site of protonation would also be moved further from the $\beta$ -turn, changing the nature of its influence on the structural preferences. Comparison of such findings with solution-phase experiments can then be used to assess the intrinsic relationship between primary and secondary structures both in the absence and presence of solvent.

## 4.5 Conclusions

Using cold ion UV photofragment spectroscopy and conformation-specific IR spectroscopy carried out by IR-UV double resonance, we have probed the unique conformational preferences for the $^{L}$P and $^{D}$P diastereomers of $[YAPAA+H]^+$, a peptide just sufficient in length to form $\beta$-turns that take advantage of the $^{D}$PA proclivity to do so in solution.[6] In $[YA^{D}PAA+H]^+$, two dominant conformer families were observed. One of these was identified as the expected charge-stabilized Type II' $\beta$-turn similar to the one observed in $[YGGFL+H]^+$ and its methyl ester,[11] while the other contained two unanticipated sharp $\gamma$-turns involving sequential C7 nearest-neighbor H-bonds. Furthermore, we have shown that the $A_2$ methyl group plays a key role in the change in preference from *cis*- ($^{L}$P) to *trans*-amide ($^{D}$P) about the $A_2$-$P_3$ amide bond because removal of the methyl group in the case of $[YG^{L}PAA+H]^+$ results in reformation of a *trans*-amide bond. Thus, we have confirmed the hypothesis of Raulfs *et al.*[33] that the $^{D}$P and $^{L}$P–containing diastereomers of $[AAPAA+H]^+$ have different conformational preferences, which they suggest based on their differing fragmentation behavior in CID. This stereochemically-induced *cis*-*trans* isomerization was reversed by removing the methyl group from $A_2$, resulting in a *trans* $G_2$-$^{L}P_3$ amide bond in $[YG^{L}PAA+H]^+$. The present work that determines the inherent conformational preferences for these $^{D}$P- and $^{L}$P-containing peptides serves as an important foundation for subsequent studies of the dynamics of peptide folding on the gas-phase potential energy surface using IR photo-induced population transfer schemes.[85,86,89]

## 4.6 References


(1) Gellman, S. H. *Accounts Chem. Res.* **1998**, *31*, 173.

(2) Mahajan, M.; Bhattacharjya, S. *Angew. Chem. Int. Edit.* **2013**, *52*, 6430.

(3) Cho, P. Y.; Joshi, G.; Boersrna, M. D.; Johnson, J. A.; Murphy, R. M. *ACS Chem. Neurosci* **2015**, *6*, 778.

(4) Lee, J.; Kim, H.; Han, S.; Hong, E.; Lee, K. H.; Kim, C. *J. Am. Chem. Soc.* **2014**, *136*, 12880.

(5) Riemen, A. J.; Waters, M. L. *J. Am. Chem. Soc.* **2010**, *132*, 9007.

(6) Haque, T. S.; Gellman, S. H. *J. Am. Chem. Soc.* **1997**, *119*, 2303.

(7) Kier, B. L.; Anderson, J. M.; Andersen, N. H. *J. Am. Chem. Soc.* **2015**, *137*, 5363.

(8) Maynard, S. J.; Almeida, A. M.; Yoshimi, Y.; Gellman, S. H. *J. Am. Chem. Soc.* **2014**, *136*, 16683.

(9) Masson, A.; Kamrath, M. Z.; Perez, M. A. S.; Glover, M. S.; Rothlisberger, U.; Clemmer, D. E.; Rizzo, T. R. *J. Am. Soc. Mass Spectrom.* **2015**, *26*, 1444.

(10) Voronina, L.; Masson, A.; Kamrath, M.; Schubert, F.; Clemmer, D.; Baldauf, C.; Rizzo, T. R. *J. Am. Chem. Soc.* **2016**, *138*, 9224.

(11) Burke, N. L.; DeBlase, A. F.; Redwine, J. G.; Hopkins, J. R.; McLuckey, S. A.; Zwier, T. S. *J. Am. Chem. Soc.* **2016**, *138*, 2849.

(12) Burke, N. L.; Redwine, J. G.; Dean, J. C.; McLuckey, S. A.; Zwier, T. S. *Int. J. Mass Spectrom.* **2015**, *378*, 196.

(13) Lewis, P. N.; Momany, F. A.; Scheraga, H. A. *Biochim. Biophys. Acta* **1973**, *303*, 211.

(14) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170.

(15) Ueki, T.; Bando, S.; Ashida, T.; Kakudo, M. *Acta Crystallogr., Sect. B: Struct. Sci.* **1971**, *B 27*, 2219.

(16) Salveson, P. J.; Spencer, R. K.; Nowick, J. S. *J. Am. Chem. Soc.* **2016**, *138*, 4458.

(17) Blanco, F. J.; Jimenez, M. A.; Herranz, J.; Rico, M.; Santoro, J.; Nieto, J. L. *J. Am. Chem. Soc.* **1993**, *115*, 5887.

(18) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236.

(19) Espinosa, J. F.; Gellman, S. H. *Angew. Chem.-Int. Edit.* **2000**, *39*, 2330.

(20) Kopple, K. D.; Go, A.; Pilipauskas, D. R. *J. Am. Chem. Soc.* **1975**, *97*, 6830.

(21) Haque, T. S.; Little, J. C.; Gellman, S. H. *J. Am. Chem. Soc.* **1994**, *116*, 4105.

(22) Craven, T. W.; Cho, M. K.; Traaseth, N. J.; Bonneau, R.; Kirshenbaum, K. *J. Am. Chem. Soc.* **2016**, *138*, 1543.

(23) Kortemme, T.; Ramirez-Alvarado, M.; Serrano, L. *Science* **1998**, *281*, 253.

(24) Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342.

(25) Blanco, F.; Ramirez-Alvarado, M.; Serrano, L. *Curr. Opin. Struct. Biol.* **1998**, *8*, 107.

(26) Gellman, S. H. *Curr. Opin. Chem. Biol.* **1998**, *2*, 717.

(27) Lacroix, E.; Kortemme, T.; de la Paz, M. L.; Serrano, L. *Curr. Opin. Struct. Biol.* **1999**, *9*, 487.

(28) Griffiths-Jones, S. R.; Maynard, A. J.; Searle, M. S. *J. Mol. Biol.* **1999**, *292*, 1051.

(29) Kar, K.; Hoop, C. L.; Drombosky, K. W.; Baker, M. A.; Kodali, R.; Arduini, I.; van der Wel, P. C. A.; Horne, W. S.; Wetzelt, R. *J. Mol. Biol.* **2013**, *425*, 1183.

(30) Walters, R. H.; Murphy, R. M. *J. Mol. Biol.* **2011**, *412*, 505.

(31) Karle, I. L.; Awasthi, S. K.; Balaram, P. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 8189.

(32) Bandyopadhyay, A.; Misra, R.; Gopi, H. N. *Chem. Commun.* **2016**, *52*, 4938.

(33) Raulfs, M. D. M.; Breci, L.; Bernier, M.; Hamdy, O. M.; Janiga, A.; Wysocki, V.; Poutsma, J. C. *J. Am. Soc. Mass Spectrom.* **2014**, *25*, 1705.

(34) Schwartz, B. L.; Bursey, M. M. *Biol. Mass Spectrom.* **1992**, *21*, 92.

(35) Vaisar, T.; Urban, J. *J. Mass Spectrom.* **1996**, *31*, 1185.

(36) Abutokaikah, M. T.; Guan, S. S.; Bythell, B. J. *Journal of the American Society for Mass Spectrometry* **2017**, *28*, 182.

(37) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R. *Phys. Chem. Chem. Phys.* **2009**, *11*, 125.

(38) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *J. Am. Chem. Soc.* **2007**, *129*, 13820.

(39) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *Chimia* **2008**, *62*, 240.

(40) Stearns, J. A.; Guidi, M.; Boyarkin, O. V.; Rizzo, T. R. *J. Chem. Phys.* **2007**, *127*.

(41) Nagornova, N. S.; Guglielmi, M.; Doemer, M.; Tavernelli, I.; Rothlisberger, U.; Rizzo, T. R.; Boyarkin, O. V. *Angew. Chem.-Int. Edit.* **2011**, *50*, 5383.

(42) Wassermann, T. N.; Boyarkin, O. V.; Paizs, B.; Rizzo, T. R. *J. Am. Soc. Mass Spectrom.* **2012**, *23*, 1029.

(43) Wolke, C. T.; Fournier, J. A.; Miliordos, E.; Kathmann, S. M.; Xantheas, S. S.; Johnson, M. A. *J. Chem. Phys.* **2016**, *144*.

(44) DePalma, J. W.; Kelleher, P. J.; Johnson, C. J.; Fournier, J. A.; Johnson, M. A. *J. Phys. Chem. A* **2015**, *119*, 8294.

(45) Fournier, J. A.; Wolke, C. T.; Johnson, M. A.; Odbadrakh, T. T.; Jordan, K. D.; Kathmann, S. M.; Xantheas, S. S. *J. Phys. Chem. A* **2015**, *119*, 9425.

(46) Wolke, C. T.; Menges, F. S.; Totsch, N.; Gorlova, O.; Fournier, J. A.; Weddle, G. H.; Johnson, M. A.; Heine, N.; Esser, T. K.; Knorke, H.; Asmis, K. R.; McCoy, A. B.; Arismendi-Arrieta, D. J.; Prosmiti, R.; Paesani, F. *J. Phys. Chem. A* **2015**, *119*, 1859.

(47) Fournier, J. A.; Wolke, C. T.; Johnson, C. J.; Johnson, M. A.; Heine, N.; Gewinner, S.; Schollkopf, W.; Esser, T. K.; Fagiani, M. R.; Knorke, H.; Asmis, K. R. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 18132.

(48) Fournier, J. A.; Johnson, C. J.; Wolke, C. T.; Weddle, G. H.; Wolk, A. B.; Johnson, M. A. *Science* **2014**, *344*, 1009.

(49) Heine, N.; Fagiani, M. R.; Rossi, M.; Wende, T.; Berden, G.; Blum, V.; Asmis, K. R. *J. Am. Chem. Soc.* **2013**, *135*, 8266.

(50) Menges, F. S.; Craig, S. M.; Totsch, N.; Bloomfield, A.; Ghosh, S.; Kruger, H. J.; Johnson, M. A. *Angew. Chem.-Int. Edit.* **2016**, *55*, 1282.

(51) Ingram, A. J.; Wolk, A. B.; Flender, C.; Zhang, J. L.; Johnson, C. J.; Hintermair, U.; Crabtree, R. H.; Johnson, M. A.; Zare, R. N. *Inorg. Chem.* **2014**, *53*, 423.

(52) Wolk, A. B.; Leavitt, C. M.; Fournier, J. A.; Kamrath, M. Z.; Wijeratne, G. B.; Jackson, T. A.; Johnson, M. A. *Int. J. Mass Spectrom.* **2013**, *354*, 33.

(53) Duffy, E. M.; Marsh, B. M.; Voss, J. M.; Garand, E. *Angew. Chem.-Int. Edit.* **2016**, *55*, 4079.

(54) Marsh, B. M.; Zhou, J.; Garand, E. *Phys. Chem. Chem. Phys.* **2015**, *17*, 25786.

(55) Garand, E. *Abstr. Pap. Am. Chem. Soc.* **2014**, *248*.

(56) Marsh, B. M.; Duffy, E. M.; Soukup, M. T.; Zhou, J.; Garand, E. *J. Phys. Chem. A* **2014**, *118*, 3906.

(57) Marsh, B. M.; Zhou, J.; Garand, E. *J. Phys. Chem. A* **2014**, *118*, 2063.

(58) Redwine, J. G.; Davis, Z. A.; Burke, N. L.; Oglesbee, R. A.; McLuckey, S. A.; Zwier, T. S. *Int. J. Mass Spectrom.* **2013**, *348*, 9.

(59) Boyarkin, O. V.; Kopysov, V. *Rev. Sci. Instrum.* **2014**, *85*.

(60) Han, H. L.; Londry, F. A.; Erickson, D. E.; McLuckey, S. A. *Analyst* **2009**, *134*, 681.

(61) Londry, F. A.; Hager, J. W. *J. Am. Soc. Mass Spectrom.* **2003**, *14*, 1130.

(62) Kang, H.; Jouvet, C.; Dedonder-Lardeux, C.; Martrenchard, S.; Gregoire, G.; Desfrancois, C.; Schermann, J. P.; Barat, M.; Fayeton, J. A. *Phys. Chem. Chem. Phys.* **2005**, *7*, 394.

(63) Gregoire, G.; Jouvet, C.; Dedonder, C.; Sobolewski, A. L. *J. Am. Chem. Soc.* **2007**, *129*, 6223.

(64) Lucas, B.; Barat, M.; Fayeton, J. A.; Perot, M.; Jouvet, C.; Gregoire, G.; Nielsen, S. B. *J. Chem. Phys.* **2008**, *128*.

(65) Perot, M.; Lucas, B.; Barat, M.; Fayeton, J. A.; Jouvet, C. *J. Phys. Chem. A* **2010**, *114*, 3147.

(66) Dehon, C.; Soorkia, S.; Pedrazzani, M.; Jouvet, C.; Barat, M.; Fayeton, J. A.; Lucas, B. *Phys. Chem. Chem. Phys.* **2013**, *15*, 8779.

(67) Kumar, S. S.; Lucas, B.; Soorkia, S.; Barat, M.; Fayeton, J. A. *Phys. Chem. Chem. Phys.* **2012**, *14*, 10225.

(68) Zabuga, A. V.; Kamrath, M. Z.; Boyarkin, O. V.; Rizzo, T. R. *J. Chem. Phys.* **2014**, *141*.

(69) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440.

(70) Zhao, Y.; Schultz, N. E.; Truhlar, D. G. *J. Chem. Theory Comput.* **2006**, *2*, 364.

(71) 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.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J.; Gaussian, Inc.: Wallingford, CT, USA, 2009.

(72) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.

(73) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(74) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200.

(75) Grimme, S.; Ehrlich, S.; Goerigk, L. *J. Comput. Chem.* **2011**, *32*, 1456.

(76) Leavitt, C. M.; DeBlase, A. F.; Johnson, C. J.; van Stipdonk, M.; McCoy, A. B.; Johnson, M. A. *J. Phys. Chem. Lett.* **2013**, *4*, 3450.

(77) DeBlase, A. F.; Bloom, S.; Lectka, T.; Jordan, K. D.; McCoy, A. B.; Johnson, M. A. *J. Chem. Phys.* **2013**, *139*.

(78) Wolke, C. T.; DeBlase, A. F.; Leavitt, C. M.; McCoy, A. B.; Johnson, M. A. *J. Phys. Chem. A* **2015**, *119*, 13018.

(79) Johnson, C. J.; Dzugan, L. C.; Wolk, A. B.; Leavitt, C. M.; Fournier, J. A.; McCoy, A. B.; Johnson, M. A. *J. Phys. Chem. A* **2014**, *118*, 7590.

(80) Olesen, S. G.; Guasco, T. L.; Roscioli, J. R.; Johnson, M. A. *Chem. Phys. Lett.* **2011**, *509*, 89.

(81) Baldoni, H. A.; Zamarbide, G. N.; Enriz, R. D.; Jauregui, E. A.; Farkas, O.; Perczel, A.; Salpietro, S. J.; Csizmadia, I. G. *Theochem-J. Mol. Struct.* **2000**, *500*, 97.

(82) Neale, C.; Pomes, R.; Garcia, A. E. *J. Chem. Theory Comput.* **2016**, *12*, 1989.

(83) Warnke, S.; Baldauf, C.; Bowers, M. T.; Pagel, K.; von Helden, G. *J. Am. Chem. Soc.* **2014**, *136*, 10308.

(84) Pierson, N. A.; Chen, L. X.; Valentine, S. J.; Russell, D. H.; Clemmer, D. E. *J. Am. Chem. Soc.* **2011**, *133*, 13810.

(85) Dian, B. C.; Clarkson, J. R.; Zwier, T. S. *Science* **2004**, *303*, 1169.

(86) Seaiby, C.; Zabuga, A. V.; Svendsen, A.; Rizzo, T. R. *J. Chem. Phys.* **2016**, *144*.

(87) Bondi, A. *J. Phys. Chem.* **1964**, *68*, 441.

(88) Read, R. J.; Adams, P. D.; Arendall, W. B.; Brunger, A. T.; Emsley, P.; Joosten, R. P.; Kleywegt, G. J.; Krissinel, E. B.; Lutteke, T.; Otwinowski, Z.; Perrakis, A.; Richardson, J. S.; Sheffler, W. H.; Smith, J. L.; Tickle, I. J.; Vriend, G.; Zwart, P. H. *Structure* **2011**, *19*, 1395.

(89) Clarkson, J. R.; Baquero, E.; Shubert, V. A.; Myshakin, E. M.; Jordan, K. D.; Zwier, T. S. *Science* **2005**, *307*, 1443.

# CHAPTER 5. SALT BRIDGE NETWORKS MEDIATED BY TYROSINE: INFRARED AND UV SPECTRAL SIGNITURES FROM CRYO-COOLED PEPTIDE IONS

## 5.1 Introduction

Intramolecular electrostatic interactions such as hydrogen bonding and salt bridge interactions are important structural elements for peptides and proteins.[1] Salt-bridge interactions, which are much stronger than hydrogen bonds, are formed between two oppositely charged groups in which at least one pair of heavy atoms is within hydrogen bonding distance.[2] Interactions involving three or more oppositely charged groups are referred to as salt bridge networks.[2,3] Salt bridge interactions can influence the secondary and tertiary structures and are relevant in protein-protein and protein-ligand interfaces.[2,4] Furthermore, salt bridges can also play a role in both stabilizing and destabilizing a given conformation depending on environmental factors and position within the primary sequence.[5–7] Although significant progress has been made, modeling and predicting salt-bridges remains challenging. This is partly caused by difficult in estimating the energy cost of dehydrating a basic residue, the energy gained by forming a strong ion paring, and the stringent geometrical criteria imposed by the hydrogen bonding between the residues.[2,8] Nevertheless, characterizing salt bridge interactions persist as an import task as they can be exploited from a design perspective to encourage the formation of specific secondary structure or to position residues in particular geometrical locations.[9]

Several factors influence whether or not a salt bridge can be formed as well geometry of the resulting interaction. The main amino acids involved in salt bridge formation are Arg, Lys, His, Asp and Glu. Depending on which positively charged amino acid participates in salt bridge formation the preferred geometry and type of bonding will be affected. Salt bridges involving Arg are by far the most common.[2] As both Asp and Glu carry their negative charge on the carboxylate moiety there is little variability in the type of salt bridge formed when either are incorporated. While a deprotonated C-terminus presents itself as an alternative hydrogen bonding partner for the negatively charged moiety there is very little analysis of crystal structures that take this bonding type into consideration. This may be in part that most studies interrogate the role of salt bridge formation towards stabilizing secondary structure where the interactions take place towards the center of the sequence. Distances between oppositely charged residues also appears as a critical

factor in determining the formation of salt bridges.[2] Most often binding partners appear close in sequence, within a 5-residue spacing. These types of interactions are referred to as local salt bridges. While not as common salt-bridges can occur between amino acids that are separated by much greater distances. Interestingly, the hydrogen bonding pattern of salt bridges formed under these conditions have different preferences for binding and are more similar to the binding patterns of intermolecular salt bridges. Salt-bridge networks are also prevalent but are overall less observed.

A majority of crystal structures show that most salt-bridges exist in buried hydrophobic region of proteins with limited surface solvent accessibility.[2] The stability that these interactions provide is highly debated as there is a large desolvation energy associated with burying a charged residue in the interior of the protein. On the other hand, it has been suggested that the strong electrostatic interactions formed by the oppositely charged interacting residues aid in offsetting the large dehydration penalty. There are no experimental methods available to directly measure the free energy gain by forming salt bridge interaction however, there are mutation and unfolding approaches as a function of $pk_a$ to make reasonable estimates.[10] Many studies still shown conflicting evidence, however. [2,5,7,11]

As stated previously, most salt bridges are analyzed based on their incorporation in well-defined secondary structures. These results show for example, that in alpha helices there is a naturally reoccurring spacing of oppositely charge residues 3 or 4 units apart (one turn of the helix). This spacing is considered optimal for salt bridge formation.[2,6] The higher frequency of this spacing in naturally forming sequences is thought to occur because of the structural reinforcing features that they have on the alpha helix motif. The same type of pattern can be found in beta sheet motifs as well. Under these circumstances it is logically that these interactions occur with such frequency because help preserve the secondary structure.[11] Experimental studies found that placement of oppositely charges amino acids at spacings other than the optimal position in alpha helices were found to aid in the destabilization of the structure and even sped up unfolding.[6] These results suggest that the role of salt bridge formation is highly variable for structural stability and truly depends on the primary sequence.

The role of salt bridge formation in non-polar environments is also of interest to researchers that commonly use mass spectrometry to study the conformational shapes and unfolding patterns of gas phase ions. Similar to the above arguments the roles of salt bridges in the gas phase have been highly debated based on the energy differences between the gas phase zwitterionic and

protonated versions of amino acids.[12] For each amino acid the zwitterionic form is largely higher in energy compared to the protonated version. Recently, however, there has been growing experimental evidence that salt bridge interactions are highly prevalent in gas phase ion structure. Initial experiments performed using infrared multiple photon dissociation[13] (IRMPD) on mass selected gas phase ions provided the first direct evidence that amino acids could be found in their zwitterionic state when complexed to metal cations[14], an electron[15], or water molecules[16]. Studies have gone as far as interrogating dipeptides and how the arrangement of charges[17] and relative basicity[18] of competing protonation sites effect salt-bridge formation, which were the first studies to demonstrate that protonated peptides could exist as salt-bridges in the gas phase.[17,18] Other methods capable of directly probing for the presence of carboxylate moieties in overall positively charged peptide and protein ions were performed by Zubarev[19] and Julian[20] using 157 and 266 nm laser light to induce the formation of photofragments unique to the presence of a carboxylate. Recently, we have developed an ion/ion reaction[21] that directly probes for the presence of carboxylate moieties in positively charged ions, the results of which agree with the spectroscopic probes mentioned above.

These studies provide direct evidence for the presence of salt bridges in the gas phase which is difficult due to the fact that there is no mass-to-charge distinction between a zwitterionic peptide (e.g., with two protonated sites and one deprotonated anion) and a protonated peptide of the same over-all charge state (e.g., with a single protonated site). Various activation techniques such as tandem mass spectrometry, IMS[22], and HDX[23] provide indirect evidence of their presence. While it has become increasingly clear that salt bridges are quite ubiquitous in the gas phase there is still a lack of information about the effect that the primary sequence has on the formation of these interactions.[24–28] This largely parallels the information known about the salt-bridges in crystal structures of proteins.

In this study we take a more candid approach in analyzing salt bridge formation. As stated before the analysis of salt bridge formation is often looked at in context of their incorporation into well-defined secondary structure of proteins that can be crystalized.[2] Here we aim to study how the surrounding primary sequence influences the formation of salt bridges in absence of a well-defined secondary structure that may advantageously place them in optimal or suboptimal locations for binding. In this approach folding is driven by salt bridge formation or lack thereof which may effectively decouple the free energy gain by forming a well-defined structure from

forming a salt bridge. Furthermore, if folding is driven entirely based on salt bridge formation then the solvation effects of the surrounding residues in stabilizing this motif can be understood.

For these experiments we take advantage of high-resolution infrared spectroscopy performed on cryogenically cooled peptide ions.[29,30] In this approach we study the gas phase ions of several short peptide sequences that incorporate advantageous functional groups that would encourage salt bridge formation. The peptides YGRAR, FGRAR, and YGRDR in various charge states are all interrogated with this method. The advantage of performing structural studies on cryocooled ions is several-fold. In this method ions are brought into the gas phase via nano-electrospray ionization (nESI), desolvated, and guided into a mass spectrometer. The mass spectrometer provides a non-polar environment, similar to that found in a hydrophobic pocket of a protein.[31] The lack of solvent further re-creates the natural environment of buried salt bridges and allows for solvent induced broadening of the IR spectra to be avoided.[32] Cryo-cooling essential freezes the ions in the conformations populated at room temperature[33] and allows for highly resolved vibrational spectra to be recorded that appear sensitive to the ions three-dimensional shape.[30,32] Through the use of quantum mechanical calculations structures can be matched to the resulting experimental spectrum.[34] The combination of cryogenic ion spectroscopy and mass spectroscopy as described here has long been demonstrated as proven technique to provide unambiguous structural assignments.

Results show that the plus one charge state of YGRAR readily forms a salt bridge between the central protonated Arg residue ($ArgH^+(3)$) and the deprotonated C-terminus. The binding motif is classified as a bidentate side-on interaction that contains a similar hydrogen bonding geometry to $Arg_iAsp_{i+2}$ commonly observed at the N-cap of helices in crystal structures. However, because the salt bridge incorporates the deprotonated C-terminus the salt bridge is classified as an $Arg_iArg_{i+2}$ or more specifically $Arg_iC\text{-terminus}_{i+2}$. Due to the short sequence a well-defined secondary strucutre in not formed and it is apparent that the rest of the sequence folds such that the salt bridge interaction can be stabilized. In this case the role of the primary sequence in solvating the salt bridge interaction can be observed. The structure reveals that an ion-dipole interaction is formed between the tyrosine OH and carboxylate and positioned such it bridges the C-terminal protonated guanidinium group of $ArgH^+(5)$ to the carboxylate. This binding motif with the C-terminus has not been observed in crystal structures although it is highly similar to that formed with the Asp/Glu side chains. The interaction between the permanent dipole of the Tyr OH however, is quite unique.

The role of ion-dipole interactions towards salt bridge stabilization has not be thoroughly investigated. A similar type of ion-dipole interaction has been found to have stabilizing effects on protonated lysine side chains in various protein structures.[35] In this context we show the substitution of Tyr for a Phe, effectively removing the permanent dipole, destabilizes the salt bridge all together. This suggests that the role of nearby dipoles may aid in minimizing the energy required for salt bridge formation in highly non-polar environments. Additional studies performed on the plus two charge state of YGRAR indicate the absence of salt bridges formation.

## 5.2 Methods

### 5.2.1 Experimental

All spectroscopic data was recorded using a custom instrument for cold ion spectroscopy as previously described.[36] The instrument is comprised of a triple quadrupole mass spectrometer and a spectroscopy axis mounted between the second and third quadrupoles. Briefly, ions are generated through nESI and guided and stored in the second quadrupole. The mass of interest is then isolated and steered down the spectroscopy axis by a turning quadrupole mounted between the second and third quadrupole ion traps. Ions are then guided into an octopole ion trap held at 5 K via a closed cycle helium cryostat (Sumitomo Heavy Industries, Tokyo Japan). As ions enter the trap they collide with cold helium buffer gas pulsed into the cryo-trap prior to their arrival, reaching vibrational temperatures of ~10 K. Spectroscopic interrogation is carried out using IR (Laservision OPO/OPA) and UV lasers (ScanMatePro dye laser frequency doubled by an Inrad Autotracker III). Resulting photofragments are then extracted back down the spectroscopy axis and turned into a linear ion trap (Q3) for mass analysis. This quadrupole can be operated in two modes. In spectroscopy mode, residual precursor ions are ejected from the trap using a supplemental AC waveform calculated by the SX wave software[37] and the remaining photofragments are extracted on to the detector. In mass spectrometry mode, mass spectra can be recorded using mass selective axial ejection (MSAE).[38]

UV action spectra are recorded by monitoring the total fragmentation signal as a function of UV wavelength. Electronic transitions belonging to different conformations are then probed using IR-UV double resonance. In this scheme, the wavelength of the UV laser is fixed on the electronic transition of interest while the IR laser is timed to irradiate the ion packet 200 ns prior to the UV

laser. When the IR photon is resonant with a vibrational transition due to the conformation being monitored by the UV laser, a fraction of the ground state population is removed resulting in a depletion of the photofragmentation signal from the UV laser. Monitoring the depletion signal as a function of the IR wavelength results in a conformer specific IR spectrum.

In a similar manner it is possible to record the IR spectrum after UV excitation. In this scheme the IR laser is timed to irradiate the ion packet after (+5 ns) UV irradiation. As the IR laser comes into resonance with vibrational transitions present in the excited state an increase in the photofragmentation signal will be observed. As the UV laser only excites a single conformation the excited state IR spectrum remains conformer specific.

$IR_1$-UV-$IR_2$ triple resonance schemes were also employed when the fragmentation signals were low. The second IR laser is timed to irradiate the ion packet ~10 ns after absorption of a UV photon.[39] Subsequent absorption of multiple IR photons after a UV absorption event results in a large increase in fragmentation signal (often by a factor of 25 or more). Precaution was taken to ensure that no fragmentation was observed when $IR_2$ irradiated the ion packet in absence of a UV photon. $IR_2$ spectra of the UV excited conformations were recorded in order to determine the appropriate wavelength for maximum fragmentation with $IR_2$ in the triple resonance scheme. The excited state spectra for most conformations show a broad absorption near 3300 $cm^{-1}$, providing a non-specific wavelength to increase the fragmentation yields across all conformations. By fixing the wavelength of $IR_2$ and scanning the UV it is possible to record an IR enhanced UV spectrum. The weak UV intensity of $[FGRAR+H]^+$ benefited from use of the IR-enhancement when recording the UV photofragment spectrum. Finally, by keeping the relative timing between $IR_1$ and the UV laser the same as in the double resonance scheme, it is possible to record conformer specific spectra by depleting the $IR_2$ enhanced UV photofragment signal. All IR spectra in the amide I/II region were recorded in this manner as was the hydride stretch spectrum of $[FGRAR+H]^+$.

### 5.2.2 Computational

Conformational searches for all peptides systems were performed as gas phase ions structures via the Monte Carlo multiple minimum method implemented with the OPL3 force field using the MACROMODEL software package.[39] Iterative searches were performed for the $[YGRAR+H]^+$ salt bridge configuration with protonation at both Arg residues and a deprotonated C-terminus,

singly protonated $[YGRAR+H]^+$ (CS) with the protonation site at either the central or C-terminal arginine, $[FGRAR+H]^+$ with protonation sites at either Arg residue. Searches were considered converged when there were no new structures found or there were no new structures found within 10 kJ/mol of the force field global minimum structure.

All structures that were further optimized and vibrational frequencies calculated for were done so using DFT with the B3LYP[40] functional with the Grimme's dispersion correction (GD3BJ)[41] and [6-31+G(d)][42] basis set. Scaling factors for the OH stretches were 0.973, amide NH stretches 0.958, guanidinium NH's 0.96, and 0.981 for the amide I/II regions. These scale factors were chosen based on commonly used literature values. The value for the guanidinium NH's was derived from scaling the calculated free $NH_2$ asymmetric stretch to its experimental free value. The resulting salt bridge structures of YGRAR were filtered by only considering geometries that contained a hydrogen bound tyrosine hydroxyl group based on the experimental IR spectrum. The 139 filtered structures, along with 25 of the lowest force field energy structures that contained an unbound (free) tyrosine OH were further optimized using DFT. The geometries with the unbound tyrosine OH were calculated to get a reasonable estimate on difference in $\Delta G_{298\,K}$ between the two distinct geometries. The energies of the singly protonated CS version of $[YGRAR+H]^+$ was also calculated. 50 of the lowest energy structures in terms of force field energy for each protomer were further optimized to compare the difference in $\Delta G_{298\,K}$ between the CS and salt bridge structures. Only the CS versions of FGRAR were considered based on the experimental IR spectra. Initially, 50 of the lowest energy structures of each protomer was calculated. It was found that the C-terminal protonation site was much more favored, the lowest energy $ArgH^+(3)$ protonated FGRAR structures was 22 kJ/mol higher in terms of $\Delta G_{298\,K}$. Thus, 50 additional $ArgH^+(5)$ protomers were ran. A search for a suitable assignment is still on going as many of the optimized structures fail to accurately re-create the experimental spectrum. The next steps involve filtering the force field searches based on the experimental constraints imposed by the IR spectrum, the main one being a hydrogen bound acid OH.

Time dependent density functional theory (TDDFT) was used to calculate the vertical excitation energies of the assigned $[YGRAR+H]^+$ structure and other model compounds. This was performed using M052X/6-31+G* level of theory. The results were scaled by 0.837 which placed the vertical excitation energy of bare para cresol in good agreement with it experimental observed $S_1 \leftarrow S_0$ transition.

## 5.3 Differentiation between Zwitterionic and Non-Zwitterionic Structures

YGRAR was initially investigated in the +1 and +2 charge states, as $[M+H]^{+1}$ and $[M+2H]^{+2}$. Figure 1 shows the formal charge arrangements for the +1 (left column) and +2 (right column) charge state. Since there are three basic sites (two Arg and N-terminal amine), we show just one of the possible charge configurations. The top row shows a possible charge distribution for the CS versions of each charge state while the bottom row indicates the zwitterionic forms. Since CS and zwitterionic forms of the same charge state appear at the same m/z value, distinguishing between the two forms is not possible based simply on mass-spectral data. In particular, for the +1 state, the CS form will contain a single protonation site at either arginine residue, combined with a neutral C-terminus, while the zwitterionic form will incorporate two protonated basic sites and a deprotonated C-terminus. Based on the relative gas phase basicity between arginine and the N-terminal amine (35 kcal/mol)[43], it is likely that both arginine residues would be protonated prior to the N-terminus. In the plus two charge state, the CS form again will have both arginine residues protonated and a neutral C-terminus, while the zwitterionic form will have all three basic sites protonated (including the N-terminal amine), with a deprotonated C-terminus.

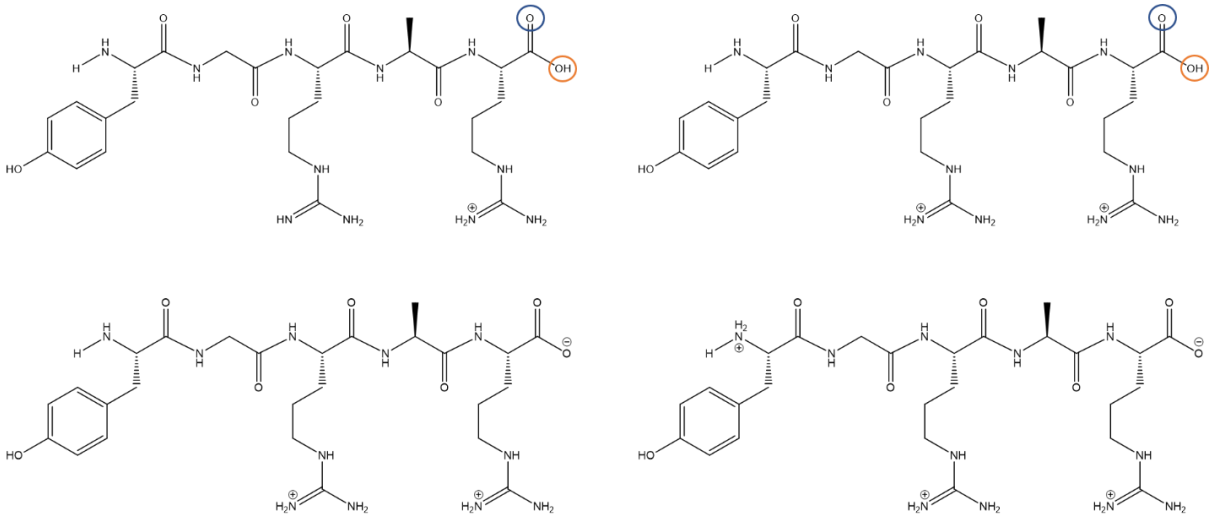


Figure 5.1: Differences in the protonation sites of the charge solvated (CS) (Top) and zwitterionic forms (bottom) of YGRAR in the plus one (left column) and plus two (right column) charge state. The functional groups which provide distinct IR spectral signatures for the (CS) forms are circle in orange and blue

With the use of IR spectroscopy, it is possible to distinguish between the CS and zwitterionic forms quite readily. Each CS form has a neutral C-terminal carboxylic acid group, which gives rise to unique IR absorptions in the hydride stretch and C=O stretch regions.[33,36,44] When the carboxylic acid OH group (circled in orange in Scheme 1) is not involved in a hydrogen bond, its OH stretch fundamental consistently appears around 3572 $cm^{-1}$. However, when this OH group is in a strong H-bond, as it often is based on previous work, its OH stretch fundamental appears as a rather broad absorption near 3000 $cm^{-1}$.[36,44] In the amide I/II region, the carboxylic acid carbonyl stretch (circled in blue) appears as the highest frequency stretch. Its exact frequency depends on its local H-bonding environment, ranging between 1725 and 1800 $cm^{-1}$.[17,36,44] The absence of clear OH stretch and C=O stretch fundamentals in these characteristic ranges' points clearly to the zwitterionic form being present.

The vibrational absorptions unique to the carboxylate anion ($COO^-$) can also be identified. When the carboxylate anion is in a symmetric environment, its two stretching modes are delocalized, producing antisymmetric and symmetric stretch fundamentals, which appear near 1600 $cm^{-1}$ and 1400 $cm^{-1}$, respectively.[15,17,18,45] Fig. 5.2 shows the ranges of expected frequencies for $COO^-$ and COOH throughout the amide I/II and hydride stretch regions. By identifying these transitions, it is possible to make unambiguous assignments that differentiate between the CS and zwitterionic forms.

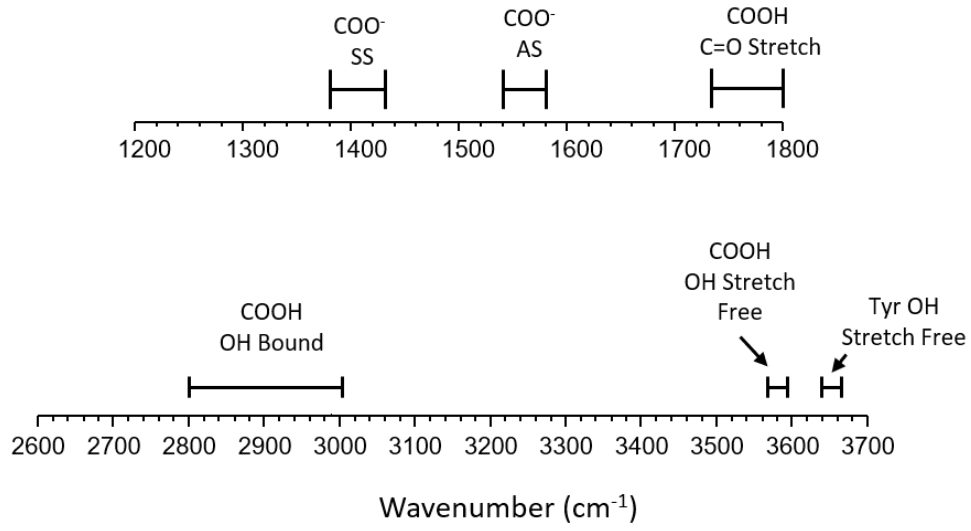


Figure 5.2: Frequency ranges of the neutral COOH and anionic $COO^-$ functionalities in the amide I/II (top) and hydride stretch (bottom) regions.

## 5.4 Results – $[YGRAR+H]^+$

### 5.4.1 Support for Zwitterionic Structure in $[YGRAR+H]^+$

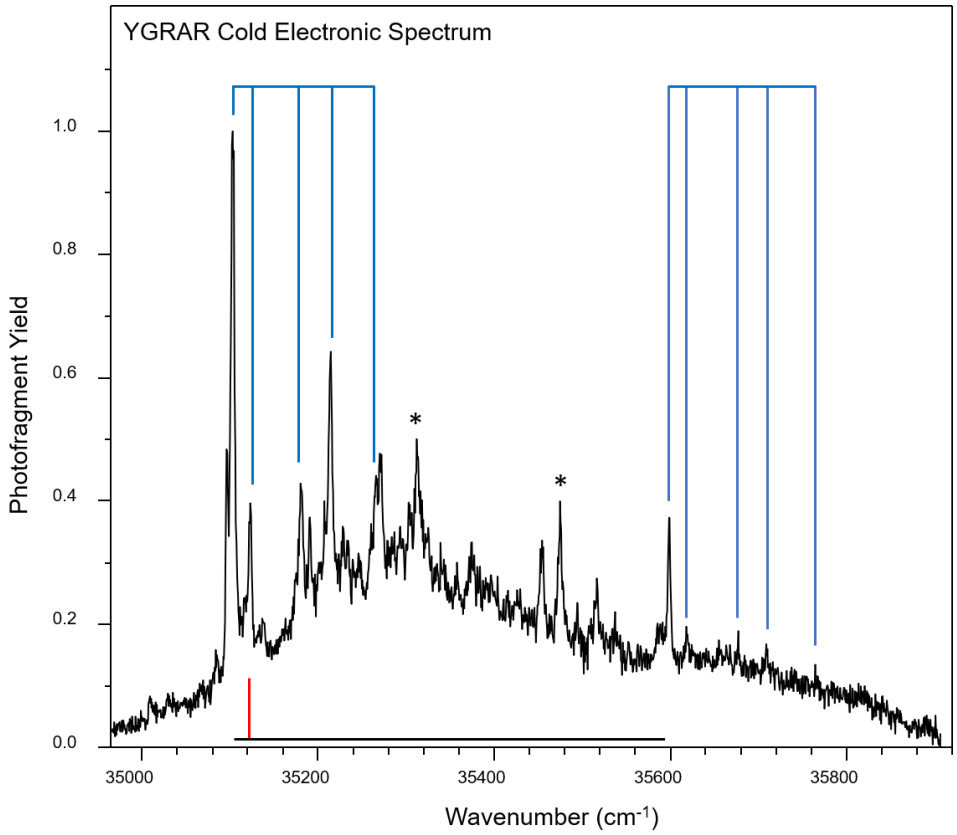


Figure 5.3: 10 K electronic spectrum of $[YGRAR+H]^+$ and TDDFT calculated vertical excitation energy of the assigned strucutre (red line).

Figure 5.3 presents the UV photofragment spectrum of $[YGRAR+H]^+$ over the $S_0$-$S_1$ origin region of the Tyr chromophore. The origin transition is at 35,103 $cm^{-1}$, with short Franck-Condon progressions in several low-frequency modes built off it. Note that this origin is 300-400 $cm^{-1}$ red-shifted from most Tyr-containing peptides, whether neutral or protonated.[36,46] The sharp transitions rise out of a broad background that is also quite unusual compared to previous Tyr-containing peptides, especially under cryo-cooled conditions. We will return to these aspects of the spectroscopy after considering the infrared spectroscopy.

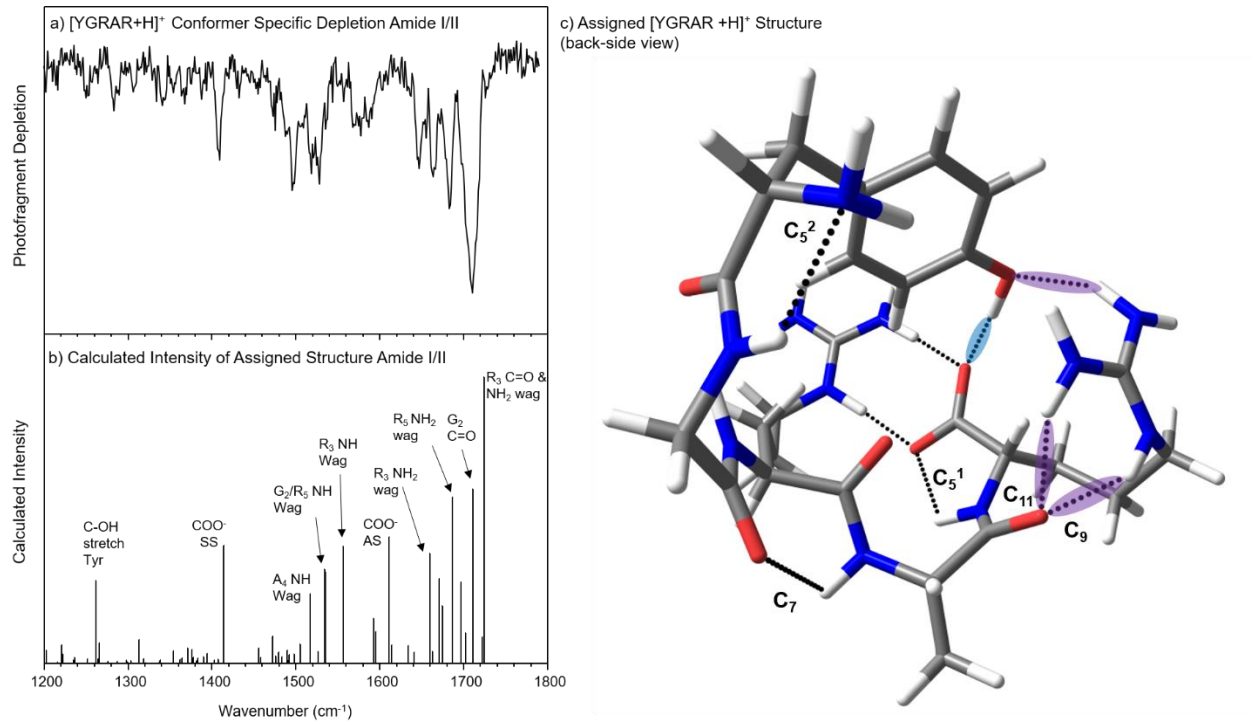


Figure 5.4: Conformer specific depletion spectra of $[YGRAR+H]^+$ (a) and the calculated intensity pattern (b), of the assigned strucutre in the amide I/II region. A back-side view of the assigned $[YGRAR+H]^+$ structure, highlighting the solvation of the C-terminal arginine residue.

The $S_0$-$S_1$ origin transition served as monitor transition for IR-UV double resonance scans. These IR depletion spectra in the mid-IR (1200-1800 $cm^{-1}$) and hydride stretch regions (2600-3700 $cm^{-1}$) are displayed in Figures 5.4a and Figure 5.5 b, respectively. Qualitatively, these spectra provide a strong initial indication that $[YGRAR+H]^+$ has a salt bridge present. First, in Fig. 5.4a, a clear spectral signature for the carboxylate anion is observed at 1410 $cm^{-1}$, due to the symmetric stretch fundamental. Several previous studies focusing on peptide[17,47] and model carboxylate zwitterions[45] have located and assigned this transition, as it appears in a comparatively uncluttered region of spectrum well outside the amide I/II vibrations. Missing transitions characteristic of a neutral COOH group provide further confirmation of salt bridge formation. In the spectrum in Figure 5.4a, the highest frequency transition appears at 1710 $cm^{-1}$, followed by a large number of transitions at 1700 $cm^{-1}$ and below. As stated above in experimental spectra on a range of protonated peptides, the acid carbonyl stretch fundamental appears near 1800 $cm^{-1}$ when free, and shifts down to as low as 1725 $cm^{-1}$ with increasing strength of the H-bond(s) to it.[44] Indeed, even under the unconstrained situation present in a guanidinium ion-$CH_3COOH$ complex, calculations

predict a lower bound of 1725 $cm^{-1}$ for the C=O stretch fundamental (see SI Fig. 1 for more details). We note that the acid C=O stretch is calculated to appear near 1700 $cm^{-1}$ in the unconstrained $CH_3COOH$-$NH_3CH_3$ ion complex, Fig. S2, however this has not been experimentally observed for the range of peptides studies and evidence from the hydride stretch region supports that the N-terminus is not protonated. Instead, transitions that appear near 1700 $cm^{-1}$ and below are ascribed to backbone carbonyl stretches and the $NH_2$ bending modes of the guanidinium group.[17]

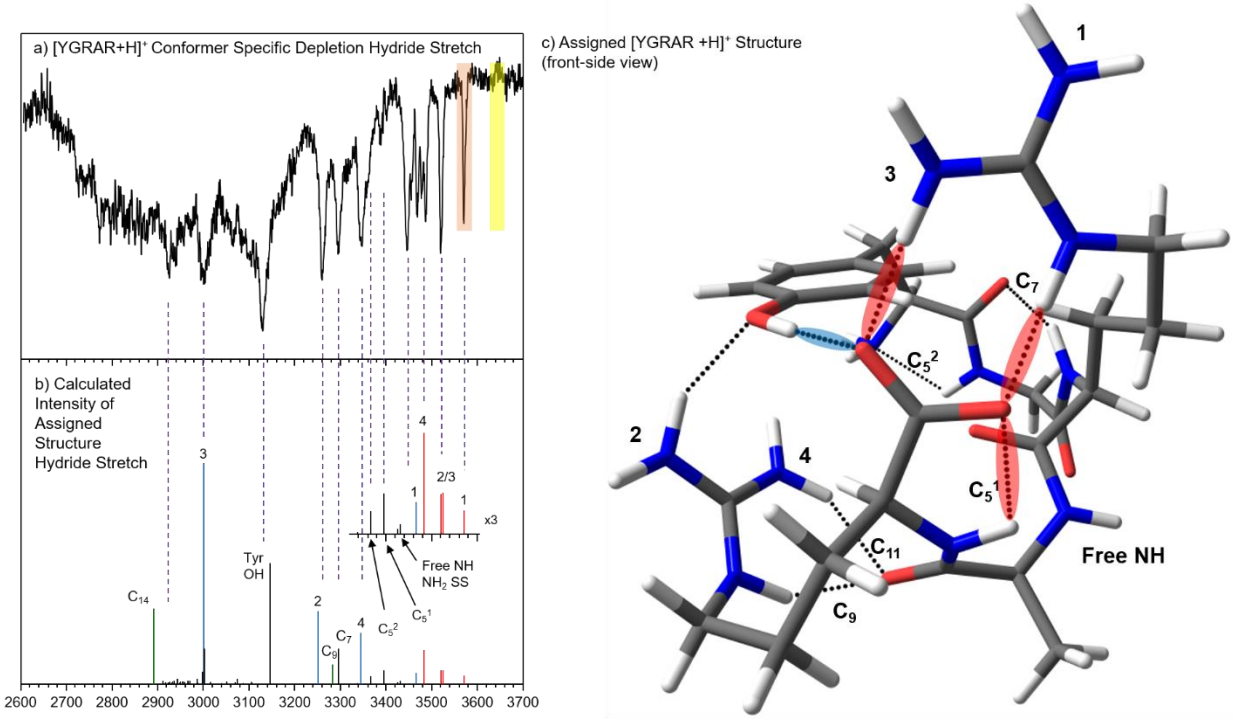


Figure 5.5: Conformer specific depletion spectra of $[YGRAR+H]^+$ (a) and the calculated intensity pattern of the assigned strucutre (b), in the hydride stretch spectral region. A front-side view of the assigned $[YGRAR+H]^+$ structure. (c) See text for details regarding color codes and numbering.

A consistency check that $[YGRAR+H]^+$ exists in zwitterionic form requires that no transition can be assigned to the carboxylic acid OH stretch. However, in Fig 5.5a, the peak shaded in orange at 3572 $cm^{-1}$ is close to that observed previously for a free carboxylic acid OH stretch.[79] On the other hand, the asymmetric stretches of the guanidinium $NH_2$ groups also will appear in a region between 3500 and 3600 $cm^{-1}$, making the assignment of this transition ambiguous. To distinguish between these possibilities, the IR spectrum of $^{18}O$ labeled $[YGRAR+H]^+$ was recorded. A 10 $cm^{-1}$ frequency shift is expected for the OH functional group upon $^{18}O$ labeling. As both oxygens on

the C-terminus readily exchange, the spectrum was recorded for the peptide mass corresponding to two $^{18}O$ exchanges, supplementary Fig. S3. The heavy labeled spectrum is identical to the $^{16}O$ peptide spectrum (Fig. S4), indicating that none of the vibrational bands are associated with a carboxylic acid OH stretch, further verifying the presence of a zwitterionic structure.

Due to the sheer number of transitions in the hydride stretch region, it is difficult to make predictions about the 3-D structure without the aid of an extensive computational conformational search. One key transition that helps in narrowing the conformational space however, is the tyrosine OH stretch. When this OH is free, as it typically is[29,33,36,44], its OH stretch fundamental appears consistently at 3650 $cm^{-1}$, shaded in yellow in Figure 3a. However, it is missing in the spectrum. The absence of this stretch allowed us to filter the initial conformational search to only output candidate structures in which the Tyr OH is bound. Furthermore, as the tyrosine is the N-terminal amino acid it is the third potential protonation site. The fact that the side chain however, can fold in towards the peptide backbone to engage in a hydrogen bond reaffirms that the N-terminus is neutral. Otherwise, the sidechain would be forced away from the backbone in effort for the charged N-terminus to become solvated by the backbone carbonyl groups and side-chains.

Another distinguishing feature in the hydride stretch region that supports the presence of a zwitterionic structure is the broad absorption between 2650 and 3200 $cm^{-1}$. In other protonated peptides studied to date,[36,44] a broad absorption down in this region indicates the presence of one or more strong H-bonds due either to the $NH_3^+\cdots O{=}C$ or $OH\cdots O{=}C$ due to the neutral carboxylic acid group. However, the intensity and over-all breadth of the absorption is significantly greater in $[YGRAR+H]^+$ than in those cases (Fig. 3b), extending all the way down to 2600 $cm^{-1}$. We postulate that NH stretch transitions due to protonated Arg forming zwitterionic bonds to the carboxylate anion could be responsible for this broad absorption as that would give rise to even stronger ionic hydrogen bonds.[48]

Taken as a whole, the IR spectral signatures present in $[YGRAR+H]^+$ provide evidence that the major conformation adopted contains a salt bridge. These findings are reinforced by the presence of the spectral signature at 1410 $cm^{-1}$ that appears where the $COO^-$ symmetric stretch is known to appear along with the absence of any transition that can be associated with a COOH neutral group in both regions of the spectrum.

### 5.4.2 Conformational Search and Assigned $[YGRAR+H]^+$ Structure

Analysis of the IR spectra for $[YGRAR+H]^+$ supports a salt bridge structure that contains a bound tyrosine and a neutral N-terminal group. Using these constraints, an initial conformational search was carried out using standard force fields, in which both arginine residues were protonated, and the C-terminus deprotonated. Once a fully converged search was obtained, a subset of structures that contained a bound tyrosine OH group were submitted to higher-level DFT calculations. For completeness sake, similar force field and DFT calculations of the 50 lowest energy structures of each possible CS form were carried out along with 25 low energy salt bridge structures that contain a free tyrosine. Energy level diagrams for structures of each type that are within 10 kJ/mol of the minimum for their charge distribution are shown in the supplementary material (Figure S5). As anticipated, all the lowest energy DFT optimized structures are zwitterionic with the two Arg groups protonated and the C-terminus deprotonated. The lowest energy charge-solvated structure has the proton on the C-terminal Arg, and is 61 kJ/mol higher in energy than the global minimum, while the singly protonated central Arg, $ArgH^+(3)$, is at +67 kJ/mol. The proton affinity of the N-terminal amine is known to be more than 145 kJ/mol smaller than arginine's basic site, so no calculations were carried out on structures with the N-terminus protonated. The lowest energy salt bridge structure with a free Tyr OH appears 24 kJ/mol higher in free energy that the assigned global minimum.

The structure of the global minimum salt bridge conformation is shown with a front-side view Fig. 5.5a. Its calculated vibrational frequencies and IR intensities in the amide I/II and hydride stretch regions are compared with experiment in Fig. 5.4b and Fig. 5.5b, respectively. The labels on the calculated IR transitions correspond to the groups acting as the major carriers of the IR transitions in Fig. 5.5. In both regions, the calculated spectrum is in close agreement with experiment, leading us to assign this structure as the carrier of the experimental spectrum.

Despite the simplification that arises from recording the spectrum of a single conformation of an isolated, cryo-cooled ions, the spectra are still quite congested due to the sheer number of IR active vibrations involved. As anticipated, the calculations predict a strong transition at 1414 $cm^{-1}$, Fig. 5.4b, due to the carboxylate symmetric stretch fundamental. The experimental spectrum, Fig. 5.4a, has a transition at 1410 $cm^{-1}$ that is free from interference from other transitions, and is thus clearly assigned to the carboxylate anion, confirming and strengthening the deduction that $[YGRAR+H]^+$ contains a salt bridge. The corresponding transition due to the carboxylate

antisymmetric stretch, Fig. 5.4b, which is calculated at 1610 $cm^{-1}$, is less easily assigned, but is a likely contributor to the broadened band just below 1600 $cm^{-1}$, Fig. 5.4a.

In the amide I region, Fig. 5.4a, most of the transitions between 1625 and 1710 $cm^{-1}$ can be assigned to the set of backbone carbonyl stretch fundamentals or $NH_2$ wagging motions of the guanidinium groups. There are ten or more transitions in this region, which are only partially resolved, but the experimental pattern is captured by the calculation. In the amide II region, the double peak between 1450 and 1550 $cm^{-1}$ can be accounted for by the backbone amide NH groups and the $ArgH^+$(3) NH bend that is directly bound to the carboxylate oxygen.

The hydride stretch region, Fig. 5.5b, has contributions from NH stretch transitions due to the two charged guanidinium groups, which contribute ten NH stretch fundamentals to the spectrum. These are in addition to the four amide NH stretch fundamentals of the peptide backbone, the two weak NH stretch transitions of the N-terminal $NH_2$ group, and the Tyr OH stretch. Each protonated guanidinium ion possesses two $NH_2$ groups, which are labeled 1-4 on the global minimum structure displayed in Fig. 5.5. Each $NH_2$ group gives rise to two transitions: an asymmetric (AS) and a symmetric stretch (SS), which are labeled as red and blue sticks, respectively, in Fig. 5.5b. The four highest frequency stretches are all AS $NH_2$ stretches. The highest frequency of them (3571 $cm^{-1}$) is due to $NH_2$(1), which does not participate in any hydrogen bonds (Figure 5.5). As such, the SS of this $NH_2$ group (blue line labelled 1 in Fig. 5.4b) is predicted to be 110 $cm^{-1}$ lower in frequency than its counterpart AS on the same $NH_2$ group. $NH_2$(2) and $NH_2$(3) both participate in a single H-bond. If this H-bond is strong, it will largely decouple the SS and AS modes, making a H-bonded NH stretch fundamental and a free NH stretch, the latter of which would appear approximately half way between the free $NH_2$ SS and AS transitions; that is, ~ 50 $cm^{-1}$ lower in frequency than the AS of the $NH_2$(1) group.[35] This is precisely what occurs, with the free NH stretch transitions NH(2) and NH(3) nearly degenerate at ~3520 $cm^{-1}$. $NH_2$(4) forms two hydrogen bonds, one a weak $\pi$ H-bond to the aromatic chromophore and the other closing an 11-membered H-bonded ring ($C_{11}$) to the alanine carbonyl group. As this group engages in two hydrogen bonds it appears as the lowest of the four AS's. In the experimental spectrum, this transition is close to the $NH_2$(1) SS, making a total of four transitions appearing between 3460-3525 $cm^{-1}$, as observed experimentally. That the experimental transitions are shifted slightly lower in frequency may be an effect of the presence of the carboxylate anion not fully captured by the calculations.

The triad of transitions between 3250-3350 $cm^{-1}$ in the experimental spectrum are matched well by the NH···OH and NH···O=C H-bonds of $NH_2$(2) and $NH_2$(4), respectively, and the C7/C9 H-bonded NH stretch fundamentals, which are unresolved experimentally. Finally, the broad absorption stretching from 2650-3200 $cm^{-1}$ has three broadened but resolved peaks at 2925, 3000, and 3130 $cm^{-1}$ that we assign to the Tyr OH, and the two NH groups ($NH_2$(3) and C14) that form the strong bidentate H-bonds to the carboxylate oxygens. Note that the Tyr OH group also H-bonds to the back side of the carboxylate anion, serving as a bridge between the other guanidinium ion (via $NH_2$(2)) and the carboxylate. The Tyr OH is thus shifted almost 440 $cm^{-1}$ from its free value (3572 $cm^{-1}$). Thus, the three furthest shifted hydride stretch fundamentals are due to the three groups that bind directly to the carboxylate anion. Of these, the C14 NH stretch is furthest shifted, as it is the sole means of binding to one of the oxygens, while the other oxygen has its electron density shared by the guanidinium $NH_2$(3) and the Tyr OH, leading to a weakening of both, probably through an anti-cooperative effect.

Further experimental verification that the assigned structure is truly the correct geometry can be attained by recording the IR spectrum of the excited state, see methods for details. In the excited NH oscillators near the aromatic chromophore will experience a red-shit (shift lower wavenumber) due to the localize excitation on the chromophore.[39,44] The excited state spectra for YGRAR are shown in Fig S6. As expected, the frequencies of the AS/SS of the $NH_2$ groups labelled 2 and 4 all shift (see supporting information for more detail), providing further experimental evidence that the structure assigned is indeed the structure observed.

In summary, the conformation shown in Figure 5.5c is the global free energy minimum, meets the experimental structural criteria, and provides a detailed fit of the experimental spectra far superior to all competing alternatives. In terms of zero-point energy, the assigned strucutre appears only 0.5 kJ/mol higher than the global minimum structure. The zero-point energy global minimum structure differs primarily in that the aromatic ring is closer to the central guanidium group allowing weak $NH_2$-π hydrogen bonds to be formed. This structure, while the global minimum in terms of zero-point energy appears 5 kJ/mol above the assigned structure at 298 K suggesting that it will not be populated at room temperature.[32,33,50]

The electronic and infrared spectra of the +2 charge state of YGRAR is shown in Fig. S7 and S8. The IR spectra on a majority of UV transitions suggest that this charge state exist in a charge solvated configuration.

### 5.4.3 Effects of $[YGRAR+H]^+$ Structure on the Electronic Spectra

Initially, it was hypothesized that $[YGRAR+H]^+$ would form a salt-bridge network in which the carboxylate would interact directly with both charged arginine residues, much as in the proposed $[bradykinin+H]^+$ and assigned $[ArgArg+H]^+$ structures.[17,51] However, the assigned structure (Fig. 5) has the carboxylate anion involved in bidentate binding to the two side-on NH groups of the central protonated arginine. The protonated C-terminal arginine does not bind directly to the carboxylate, but instead forms hydrogen bonds to the tyrosine oxygen, the aromatic ring, and the carbonyl group of the Ala(4) residue. Since only two oppositely charged groups engage directly with one another, this structure is classified as a salt-bridge rather than a salt-bridge network. The Tyr OH group serves as a neutral bridge between the C-terminal $ArgH^+$ and the 'back' of the carboxylate anion. In this way, the Tyr OH can further stabilize the carboxylate anion, and provide an indirect binding mechanism for the C-terminal $ArgH^+$ to the strong and stable bidentate ion pair. We will return to discuss this binding mode in the Discussion section.

Besides adding structural stability, the bridging Tyr OH group places the UV chromophore in an unusual environment that also impacts the electronic spectrum. Indeed, of the several previous studies of Tyr-containing protonated peptides[29,33,36,44], this is the first example in which the Tyr OH was not free. Generally, when phenol, the Tyr chromophore, engages in a hydrogen bond as donor, its electronic origin transition shifts to lower frequency. For instance, the $S_0$-$S_1$ origins of phenol-$H_2O$[52] and phenol-$NH_3$[53] are shifted 350$cm^{-1}$ and 650 $cm^{-1}$ from the phenol monomer origin.[53] Consistent with this structure, the $S_0$-$S_1$ origin transition appears at 35,103 $cm^{-1}$, shifted down in frequency by ~400 $cm^{-1}$ from where peptide ions of this size with an unbound Tyr OH group generally absorb. Since the tyrosine OH is also acting as an acceptor in a H-bond with $ArgH^+$(5), the magnitude of the red shift is less than if it were solely a donor. Again, this is in keeping with what is known for PhOH-$(H_2O)_n$ clusters[52], in which PhOH first acts as a H-bond donor to $H_2O$ when n is small but becomes an acceptor-donor in larger clusters (starting with n=3), shifting the cluster transition back toward the monomer origin.

The experimentally observed origin of isolated para cresol in the gas phase is 35,343 $cm^{-1}$.[46] The calculated vertical excitation energy for isolated para cresol, and para cresol bound to a carbonyl and carboxylate functional group was calculated at the TDDFT M052x/6-31+G*. A scale factor of 0.837 was applied to all the calculated vertical energies which places the calculated vertical energy for isolated para cresol in good agreement with the experimental value.[81] The

calculated scaled values are provided in Fig S9. The red line in Figure 5.2 denotes the scaled vertical excitation energy of the assigned structure in Fig. 5.2, calculated at the same level of theory. This value is in excellent agreement with the observed excitation energy, further confirming the environment of the tyrosine chromophore.

The blue tie-line attached to the origin transition identifies the low-frequency fundamentals with Franck Condon activity built off the origin. A similar pattern of transitions is built off another fundamental which is +495 $cm^{-1}$ from the origin transition. The transitions marked with asterisks were all used as monitor transitions for IR-UV double resonance scans, and determined to have the same IR spectrum, and thus were determined to be the same conformation as that shown in Fig. 5.5. Since most of the transitions in the electronic spectrum can be accounted for as due to a single conformation, it is apparent that $[YGRAR+H]^+$ is found mainly in the assigned salt bridge structure.

A second obvious consequence of the unique electronic environment of the bridging Tyr is the broad absorption that spans the entire spectral region scanned. In order to test whether this broad absorption was due to saturation effects, we carried out UV photofragment scans at lower UV powers, attenuated by a factor of 2, with no change in the relative contribution from this background. We also tested for complete cooling of the ions in the cryo-trap by varying the wait time between ions entering the trap and UV excitation, with no change observed at times much longer than the known cooling time of 1 ms for ions of this size.[33]

In what follows, we test what aspects of the YGRAR sequence are responsible for salt bridge formation and for the unusual broadening in the UV spectrum by carrying out cold-ion spectroscopy on a series of close analogs that modified the peptide sequence in well-defined ways. In $[FGRAR+H]^+$, we remove the OH group from Tyr by replacing it with Phe, a seemingly innocuous change that leads to loss of zwitterion formation. In $[YGRAR\text{-}OMe+H]^+$, we remove the possibility of carboxylate formation, forcing a CS structure. Finally, in $[Ac\text{-}YGRAR+H]^+$, the N-terminal amine is acylated, removing any chance of protonation there.

## 5.5 Results – $[FGRAR+H]^+$

### 5.5.1 Electronic Spectrum

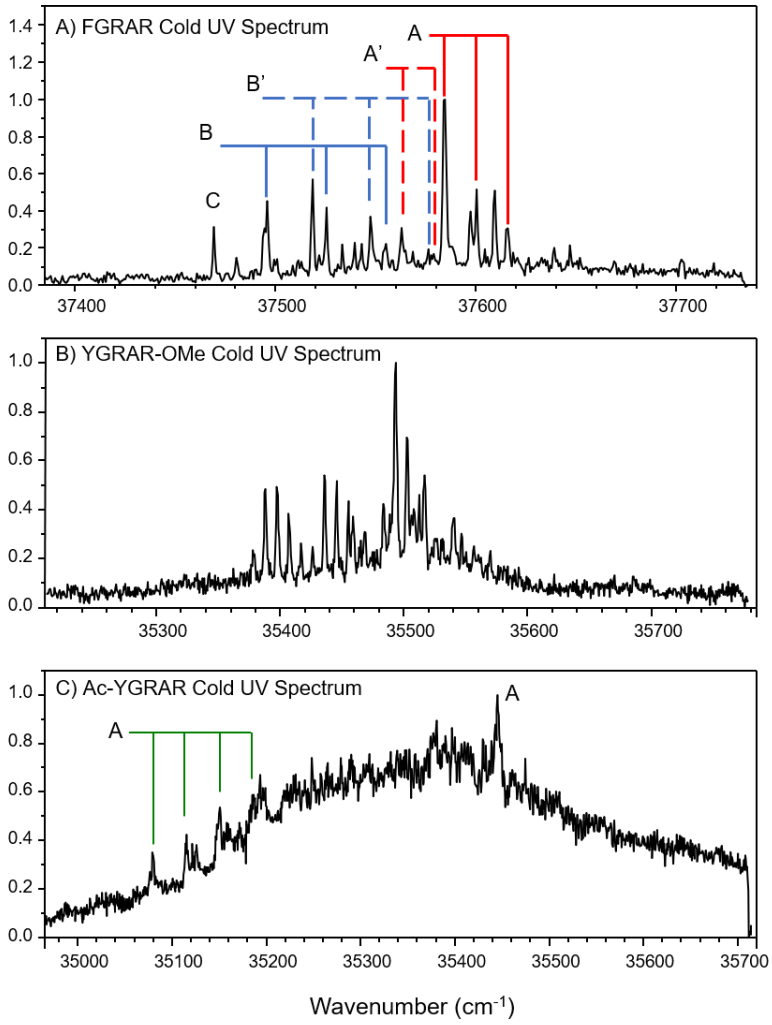


Figure 5.6: Normalized electronic spectra of cryo-cooled $[FGRAR+H]^+$ (a), $[YGRAR\text{-}Ome]^+$ (b), and $[Ac\text{-}YGRAR+H]^+$ (c).

The protonated peptide $[FGRAR+H]^+$ is a close analog to $[YGRAR+H]^+$ differing only in the removal of the OH group from the aromatic chromophore. The UV photofragment spectrum of $[FGRAR+H]^+$ in the $S_0$-$S_1$ origin region of the Phe chromophore is shown in Figure 5.6a. This peptide was found to adopt three major (A, B, B') and two minor (A', C) conformations, already a substantial change from the single conformation observed in $[YGRAR+H]^+$. Transition C is

without an obvious partner transition. Due to their weaker intensity, transitions C and A' were not characterized further.

Transitions A and A' share similar Franck-Condon activity in a 16 $cm^{-1}$ vibration, while B and B' both possess a progression in a 29$cm^{-1}$ mode. Furthermore, the splitting between the electronic origins of the A/A' or B/B' pair is quite similar (22 $cm^{-1}$). We postulate that the two members of each pair share the same backbone conformation but differ in the phenyl ring orientation, much as occurred in previous studies of Tyr-containing peptides.[33] It is important to note that the UV photofragment spectrum of $[FGRAR+H]^+$ is completely free of the background absorption present in $[YGRAR+H]^+$.

### 5.5.2 Infrared Spectra

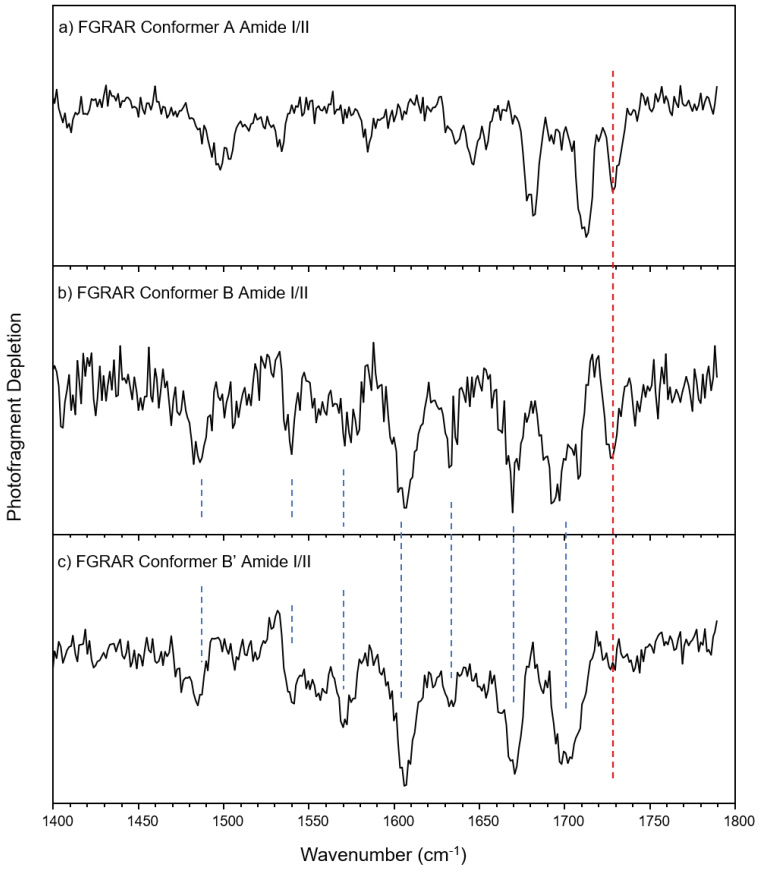


Figure 5.7: Conformer specific spectra of $[FGRAR+H]^+$ conformers A (a), B (b), and B' (c) in the amide I/II spectral region.

In order to test our hypothesis regarding the presence of conformer pairs, and to probe whether $[FGRAR+H]^+$ is zwitterionic or CS in character, we recorded IR-UV DR spectra (Figure 7a-c) while monitoring the electronic origin transitions of each of the three major conformers, A, B, and B', respectively**.** Since our primary interest here is in zwitterion structures, we focused our study initially on the 1400-1800 $cm^{-1}$ region that contain transitions diagnostic of whether zwitterionic or charge-solvated structures are present. A red dashed line going through all the spectra at 1730 $cm^{-1}$ represents the position of the hydrogen bound acid carbonyl stretch in the assigned protonated YGGFL structure.[29,36] In the spectra of $[FGRAR+H]^+$ it is clear that both conformers A and B share this transition providing a strong evidence that neither are salt bridge structures. In fact, a transition at 1730 $cm^{-1}$ agrees well with the predicted frequency of the acid carbonyl stretch when receiving a hydrogen bond from the charged arginine, as shown in Figure S1**.** This transition is much weaker for the spectrum of $[FGRAR+H]^+$ B', however, as previously state it is expected that B' has a very similar geometry as B and only differs in the rotation of the phenyl group. The blue drop lines between the spectra of B and B' show that they share all the same transitions supporting this claim. The weaker absorption at 1730 $cm^{-1}$ for B' is most likely due to the fact that there is some gain superimposed on the depletion signal at the UV wavelength used to record the spectrum of B'.[44] Secondly, we note that there is no evidence for a symmetric $COO^-$ stretch near 1400 $cm^{-1}$ as in $[YGRAR+H]^+$.

Based on the combined data from the UV and IR-UV DR spectra, we conclude that $[FGRAR+H]^+$ has two pairs of conformers that make up the vast majority of the population, none of which are zwitterionic in character. Thus, removal of the OH group from Tyr has tipped the balance from zwitterionic to charge-solvated.

### 5.6 Results – Capping $[YGRAR+H]^+$

Figures 5.6b, c present UV photofragment spectra of $[YGRAR\text{-}OMe+H]^+$ and $[Ac\text{-}YGRAR+H]^+$, respectively. As already mentioned, methylation of the C-terminus removes the possibility of forming a zwitterion structure, while acylation of the N-terminus prevents protonation of the N-terminal amine. As expected, the UV spectrum of the former (Fig. 5.6b) displays electronic origins in the 35,400-35,500 $cm^{-1}$ region where conformers with a free Tyr OH normally absorb. There also is at most a small broad absorption in these necessarily CS structures.

By contrast, the UV photofragment spectrum of $[Ac\text{-}YGRAR+H]^+$ has an intense, broad absorption that stretches across hundreds of wavenumbers, with sharp transitions superimposed on top of it. The set of four transitions marked by a tie line is a Franck-Condon progression with a spacing of 36 $cm^{-1}$ built off the electronic origin at 35,079 $cm^{-1}$. This red-shift of the origin relative to CS structures is strongly reminiscent of the shift present in $[YGRAR+H]^+$ (whose origin is at 35108 $cm^{-1}$) and suggests that the Tyr OH is acting in a donor-acceptor bridge, much as it is in $[YGRAR+H]^+$.

Consistent with this, the IR-UV DR spectrum of $[Ac\text{-}YGRAR+H]^+$ in both the hydride stretch, shown in Figure S10, and amide I/II region, Fig. 8, are nearly identical to that of $[YGRAR+H]^+$, consistent with formation of the same Tyr-mediated salt bridge network. The first transition in Fig. 5.8 appears at 1710 $cm^{-1}$, again outside the range of where the hydrogen bound carboxylic acid carbonyl appears. Indeed, the transition near 1400 $cm^{-1}$ is diagnostic of the $COO^-$ group. In the NH stretch region there is a similar broad absorption to that in $[YGRAR+H]^+$, spanning from 2600-3400 $cm^{-1}$. Evidence for a free acid or tyrosine were also missing in the hydride stretch region, confirming that this sequence is also zwitterionic. Thus, acylation of the N-terminus has at most a minor perturbation on the zwitterionic structure shown in Figure 5.5.

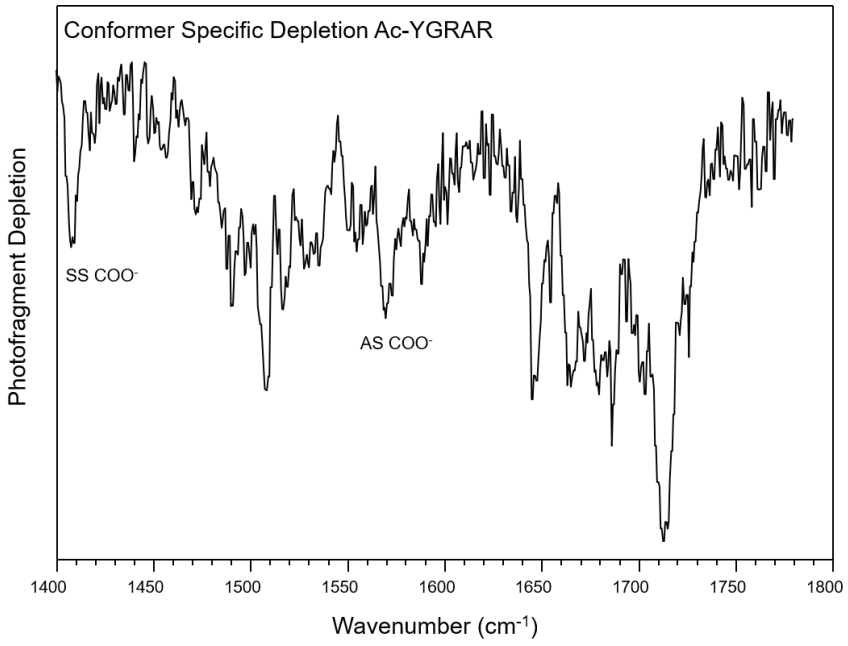


Figure 5.8: Conformer specific spectra of $[Ac\text{-}YGRAR+H]^+$ in the amide I/II spectral region.

## 5.7 Discussion

### 5.7.1 Hydrogen Bonding Patterns in $ArgH^{+}\cdots COO^{-}$ Salt Bridges

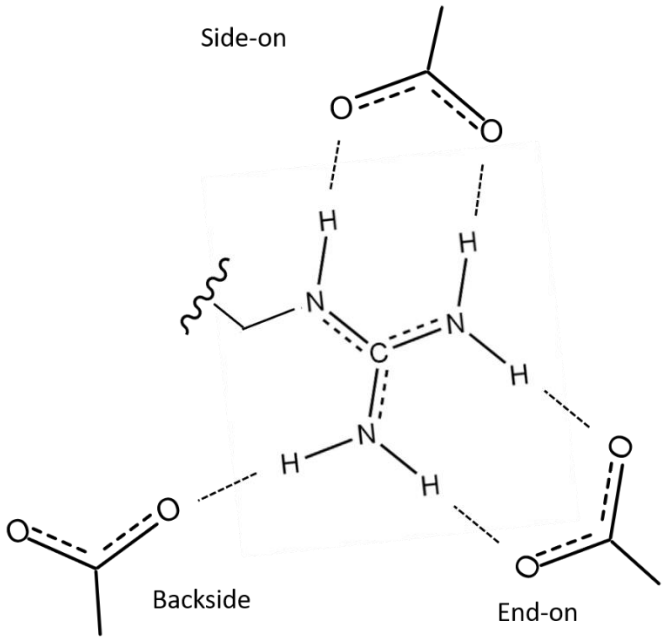


Figure 5.9: Different hydrogen bonding patterns for salt bridges involving guanidinium and carboxylate functionalities

The charge distribution around a protonated guanidinium group makes it such that there are three potential binding sites for a carboxylate moiety. The geometries are referred to as backside, end-on, and side-on and are illustrated in Fig. 9. The backside geometry only allows for a single hydrogen bond to be made to one of the hydrogens on the $NH_2$ group. This form of binding, as expected, is the least favored. The end- and side-on interaction both allow for bidentate hydrogen bonds to be formed and are more commonly observed in crystal structures. At short distances, less than 10 angstroms between the α-carbons of the oppositely charged binding partners, the side-on interaction appears highly favored over the end-on geometry, while at longer distances the bias is shifted to the latter geometry. The differences in binding geometry as a function of distance can be rationalized based on the angle of approach between closely spaced residues. Because of the ability to form bidentate interactions most salt bridges involve Arg over Lys and His. Lys and His both can only form a single hydrogen bond to the carboxylate moiety at a time.

### 5.7.2 Bidentate Salt Bridge Binding $[YGRAR+H]^+$

The observed structure for $[YGRAR+H]^+$ shown in Figure 5a has as its dominant structural feature a salt bridge between $ArgH^+(3)$ and the C-terminal $CO_2^-$ anion. Strong, in-plane, side-on H-bonds are formed with the two carboxylate oxygens, with N-H···O distances of 1.68 and 1.75 Å and N···O distances of 2.72/2.78 Å.[54] A similar bidentate binding is found in the crystal structure of arginine acetate complex, but in that case, binding occurs as the end-on, involving the two $NH_2$ groups of guanidyl ($R_{N..O}$ = 2.830/2.824 Å), compared to $NH_2$/NH side-on in $[YGRAR+H]^+$. The best-fit structure shows a good match between the spatial separation of the two NH bonds and the pair of carboxylate oxygens.[55] The near-symmetric binding to the two carboxylate oxygens keeps the stretching modes as symmetric (SS) and antisymmetric (AS) stretches, with the former providing the marker transition of carboxylate's presence at 1410 $cm^{-1}$. The AS calculated at 1610 $cm^{-1}$ has the potential to be heavily mixed with wagging modes of the NH and $NH_2$ groups of the $ArgH^+$ and amide groups. This is the likely reason that the experimental spectrum in this region is not well resolved experimentally.

While the bidentate binding is nearly symmetric, some asymmetry is induced by the side-on binding of the Tyr OH to the carboxylate O to which the $NH_2$ group of $ArgH^+(3)$ is already bound (Fig. 5b shaded in blue). This interaction is significant, shifting the Tyr OH stretch down to 3140 $cm^{-1}$, almost 500 $cm^{-1}$ below the free OH fundamental of Tyr (3640 $cm^{-1}$). The OH···O H-bond distance is calculated to be 1.68 Å and the heavy-atom separation $R_{O...O}$ = 2.67 Å. One would anticipate an anti-cooperative effect between the two H-bonds to the same carboxylate oxygen, and indeed, this is the likely reason for the longer distance of this NH···O H-bond compared to the other leg of the bidentate bonding.

In regard to intramolecular salt bridges involved in proteins this side-on form of binding agrees with the hydrogen bonding geometry expected for binding partners that are spatially close to one another. The distance between the α-carbons of both arginine residues are ~ 5 Å. In many cases the effect of the secondary structure on the salt bridge geometry is studied. These studies reveal that in helical structures the separation between binding partners is ideally 3 or 4 residues, while at the N-caps of helices or between secondary elements a spacing of 2 residues is ideal. In the case of the latter, it is common to observed both $Asp/Glu_{i-2}Arg_i$ and $Arg_iAsp/Glu_{i+2}$, where the negatively charged partner can be located either before or after the arginine residue in the primary sequence. The latter of these two sequences is most like what is present in the assigned

$[YGRAR+H]^+$ structure, $Arg_iArg_{i+2}$, or more accurately put $Arg_iC\text{-}term_{i+2}$. While markedly different from what is commonly observed, as this geometry involves the C-terminus rather that a carboxylate side chain, the binding interactions are highly similar to that of $Arg_iAsp/Glu_{i+2}$. The side-on bidentate hydrogen bond and even the hydrogen bond formed between the carboxylate oxygen and the amide nitrogen of the C-terminal amino acid are all highly similar. These hydrogen bonds are highlighted in red on the assigned structure in Fig. 5.5c. In some ways it may be expected that a salt bridge involving the C-terminus be of similar geometry to that formed when the side chains of Asp/Glu are incorporated, as the negative functional groups are all carboxylates. It would be of interest however, to mine the protein data bank in order to see how often the C-terminus is involved in salt bridge interactions as this is not commonly considered. Furthermore, from a design perspective it becomes clear that the C-termini of peptides can be capped with a salt bridge motif as the basicity of the C-terminus is not dependent on the amino acid side chain. That is to say a sequence YGRAA may be just as likely to form this same salt bridge. Initial studies have been performed on this sequence however, it is clear that that need for a positive residual charge (N-terminal protonation) to make the ion detectable greatly influences the geometry of the sequence as both charge site need to be sufficiently stabilized.

### 5.7.3 Tyrosine Mediated Salt Bridge Networks

An interesting feature in the assigned $[YGRAR+H]^+$ structure is the role of the tyrosine hydroxyl group as both a hydrogen bond donor and acceptor. The structure shown in Figure 5.5c shows a 'front-side' view of the two protonated arginine residues, the carboxylate group, and the Tyr OH. Figure 5.4c shows a 'back-side' view that highlights $ArgH^+(5)$ and its binding to the Tyr oxygen and the Ala(4) C=O group, all highlighted in purple. These cut-away views show clearly the role the Tyr OH plays as a donor-acceptor bridge between $ArgH^+(5)$ and the carboxylate anion, forming what might be termed a Tyr-mediated salt bridge network.

Donald et al. have recently carried out a detailed analysis of salt bridges and salt-bridge networks found in the protein data bank.[2] All the more common networked salt bridges incorporate two carboxylates and one $ArgH^+$ in a ( - + -) structure, that takes advantage of the two separate binding pockets afforded by $ArgH^+$. Based on this, it would appear that two $ArgH^+$ groups can't bind effectively to a single carboxylate anion, in contrast of the assigned structure of the dipeptide $[ArgArg+H]^+$. We see here that the Tyr OH can serve as an effective neutral bridge that can bind

surreptitiously to the back end of the carboxylate without perturbing the bidentate binding. The OH also serves as another acceptor group to which $ArgH^+(5)$ can bind, but at a further distance and opposite side from $ArgH^+(3)$.

The importance of this bridge is highlighted by the precipitous change between zwitterionic and CS forms that accompanies the removal of the OH group on Tyr in changing from $[YGRAR+H]^+$ to $[FGRAR+H]^+$. Not only is the bridged network no longer possible in $[FGRAR+H]^+$, but the $ArgH^+(3)\cdots CO_2^-$ ion pair also is not retained. It would seem, then, that the Tyr bridge plays a crucial role in stabilizing the zwitterion structure relative to CS alternatives. Formation of the ion pair in the absence of the Tyr OH bridge is still anticipated to be quite favorable; however, this can only occur in $[FGRAR+H^+]$ if $ArgH^+(5)$ is sufficiently stabilized. While it is undoubtedly true that the tyrosine OH group stabilizes both the carboxylate and $ArgH^+(5)$ in forming a bridge between them, the removal of this group must also set off a cascade of effects that destabilize the protonated C-terminal Arg. By removing a potential hydrogen bonding site for the charged C-terminal arginine, the only other stabilization for $ArgH^+(5)$ is the interaction with the Ala C=O group, which must not be able to sufficiently stabilize the excess charge. To better understand how the potential energy surface changes by substituting Tyr for Phe a conformational search was performed on a salt bridge configuration of FGRAR. In total 304 structures were found after the search converged. 4 representative structures we taken for to make comparisons to the assigned YGRAR structure (see supporting information Fig. S11 and 12 for more details)

In order to form a salt-bridge network composed of three charges, not only must the three charges be arranged in close proximity within a peptide sequence or spatially within the tertiary structure, but the asymmetry in number of charge sites puts particularly stringent criteria on how the three charges arrange in space so as to lead to net stabilization. The present study on $[YGRAR+H]^+$ shows an alternative strategy that may be relevant in proteins, but not as easily recognized to date. By positioning a polar side chain between the 'extra' charge and the main zwitterion pair, the polar group can orient its dipole to serve as a bridge between the opposite charges. In the present case, the bridge involves the Tyr OH, but one could imagine Ser, Gln, or Asn playing a similar role. We are currently performing an analysis of the protein data bank to determine the extent to which such bridged networks are formed.

Furthermore, the ion dipole interaction between the tyrosine hydroxyl group and the carboxylate moiety may aid in lowering the dehydration penalty when in hydrophobic regions/non-polar environments. This can aid in making the formation of traditional salt bridges more favorable. Such stabilizing ion dipole interactions have been found to occur between protonated lysine side chain $NH_3^+$ and backbone carbonyls. In this arrangement carbonyl oxygens occur in 3 fold symmetry around the lysine $NH_3^+$ group, while a fourth carbonyl group approaches the lysine Nζ from below. The resulting hydrogen bond from the 4$^{th}$ carbonyl appears linear with respect to the lysine $C^{\varepsilon}$-$N^{\zeta}$···O angle. Calculations show that combination of these interactions can contribute up to 2 kcal/mol of stabilization energy to a protein. Synonymous, the interaction between the tyrosine OH and $COO^-$ is likely to play a similar stabilizing role and can be expected to play a larger role in offsetting the dehydration penalty.[5,56] This role of this interaction is quite evident by in the lack of observed structures of FGRAR in which the C-terminal Arg seems sufficiently stabilized and the interaction between $ArgH^+$(3) remains similar to that in YGRAR. We are currently screening the protein data bank to for the frequency of polar side chains near carboxylate moieties.

### 5.7.4 Effects of Salt Bridge Network on Electronic Spectra

One of the striking correlations present in our data is that those peptides that form Tyr-mediated salt bridge networks have a broad and rather intense background absorption in the UV that starts near the $S_0$-$S_1$ origin of the Tyr chromophore (Figures 5.3, 5.6b). This broad background seems to derive its oscillator strength from the $\pi\pi^*$ state (in that it begins near the $S_0$-$S_1$ origin), but have as its source another state, since the $S_0$-$S_1$ transitions themselves are still present and reasonably sharp.

As we have already stated, one of the striking aspects of the tyrosine's local environment is the very large electric field it imposes on the aromatic ring and OH group by virtue of its serving as a donor-acceptor bridge between $ArgH^+$(5) cation and the $CO_2^-$ anion. It would be of interest to understand the difference of photostability between salt bridges that involve tyrosine to those that simple have a tyrosine nearby in sequence.

## 5.8 Conclusion

Through the use of IR-UV and IR-UV-IR double and triple resonance schemes it is possible to structurally resolve the major conformations of YGRAR and FGRAR in the plus one charge states. By making use of a short peptide sequence it was possible to encourage the folding of the peptide backbone to be driven by the formation or lack of formation of a salt bridge. In this approach it has been possible to understand the necessary solvation effects provided by the backbone in order to stabilize salt bridge binding. This is in contrast to studying the efficiency of salt bridge formation between residues within well-defined secondary structures. With this approach we see that the plus one charge state of YGRAR forms a dominate zwitterionic conformation, the largest spectroscopically resolved to date. The assigned structure shows a salt bridge binding geometry that is similar to the binding geometry of $Arg_iAsp/Glu_{i+2}$ however, is different in the sense that the carboxylate moiety is provided by a deprotonated C-terminus rather than a deprotonated acidic side chain. It is clear that there is an essential hydrogen bond formed between the tyrosine hydroxyl group and carboxylate. The engagement of the tyrosine in this manner gives rise to unique absorption properties. It is clear that the large absorption is worsened when the N-terminus is capped with a acetyl group, however based on the spectra, the Ac-YGRAR is expected to have a very similar geometry to YGRAR. The broad absorption as but vanishes when the C-terminus is capped, preventing the formation of a salt bridge. The photo stability of tyrosine mediated salt bridges seems as a potential interest from a photo physical standpoint. In quite a unique manner the tyrosine OH serves to bridge the (+ - +) charges through a neutral bridge, a motif that may not be well described in the existing literature. Furthermore, while the tyrosine hydroxyl oxygen seems to play a large role in solvating the ‘extra’ charge it is quite possible that in absence of the C-terminal arginine the presence of the neutral dipole can serve to stabilize salt bridging interaction in highly hydrophobic protein regions. This is supported by the loss of salt bridge formation in FGRAR, after removing a single polar group on the peptide chain. The role of ion-dipole interactions may be ubiquitous in stabilizing buried salt-bridges.

## 5.9 Supporting Information

### 5.9.1 Calculated Frequencies of Acetic Acid --- Guanidinium/Ammonium Complexes

The vibrational frequencies of unrestrained ion-molecule complexes between acetic acid and guanidinium/ammonium are shown in Fig. S1 and S2, respectively. In both cases the *cis* and *trans* geometry for the COOH group was taken into consideration. All calculations were performed at the same level of theory and scaled according to the computational methods section.

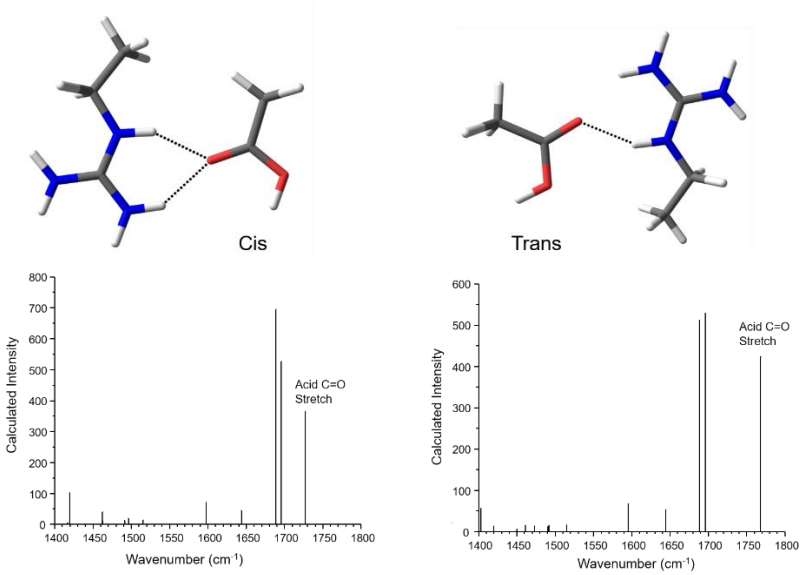


Figure 5.10: Scaled calculated vibrational frequencies for the guanidinium and acetic acid complex in the *cis* and *trans* COOH geometries.

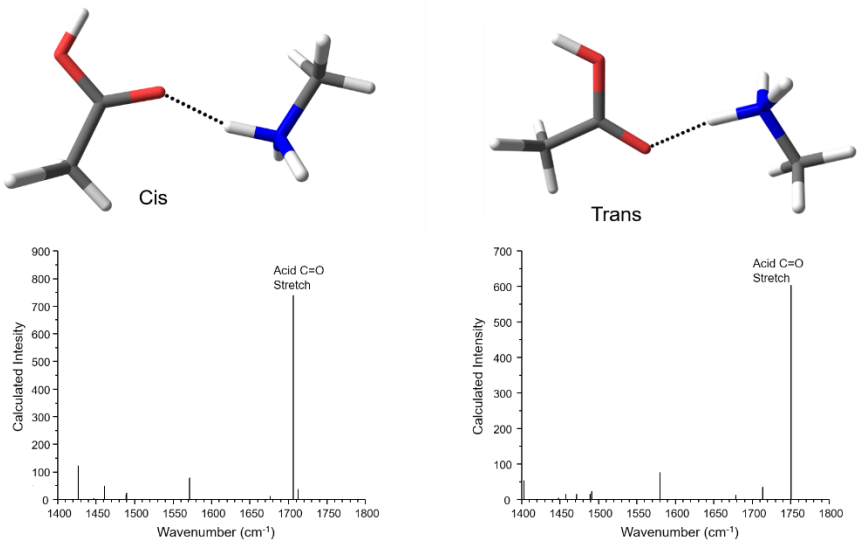


Figure 5.11: Scaled calculated vibrational frequencies for the ammonium and acetic acid complex in the *cis* and *trans* COOH geometries

### 5.9.2 [YGRAR+H]$^+$ O$^{18}$ Labelled Experiments

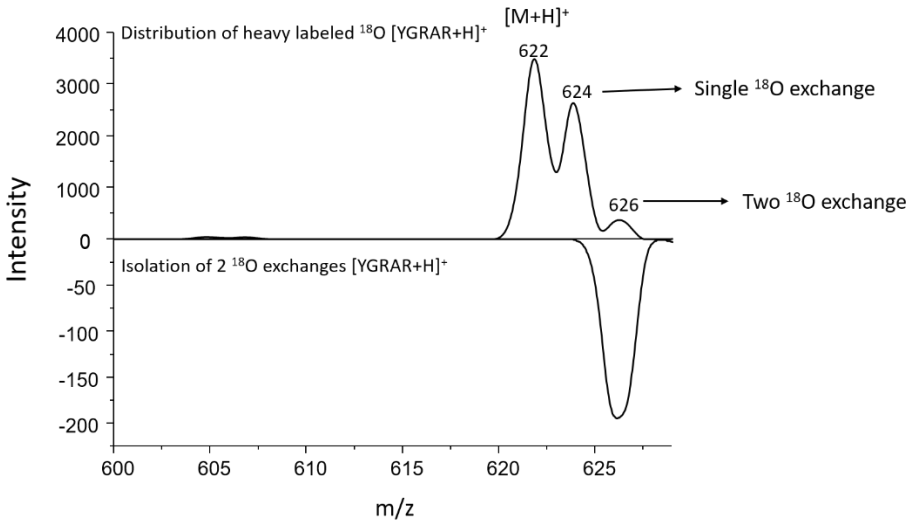


Figure 5.12: Mass spectrum of non-isolated heavy labelled [YGRAR+H]$^+$ (top), mass isolation of peptide corresponding to two heavy oxygen exchanges (bottom)

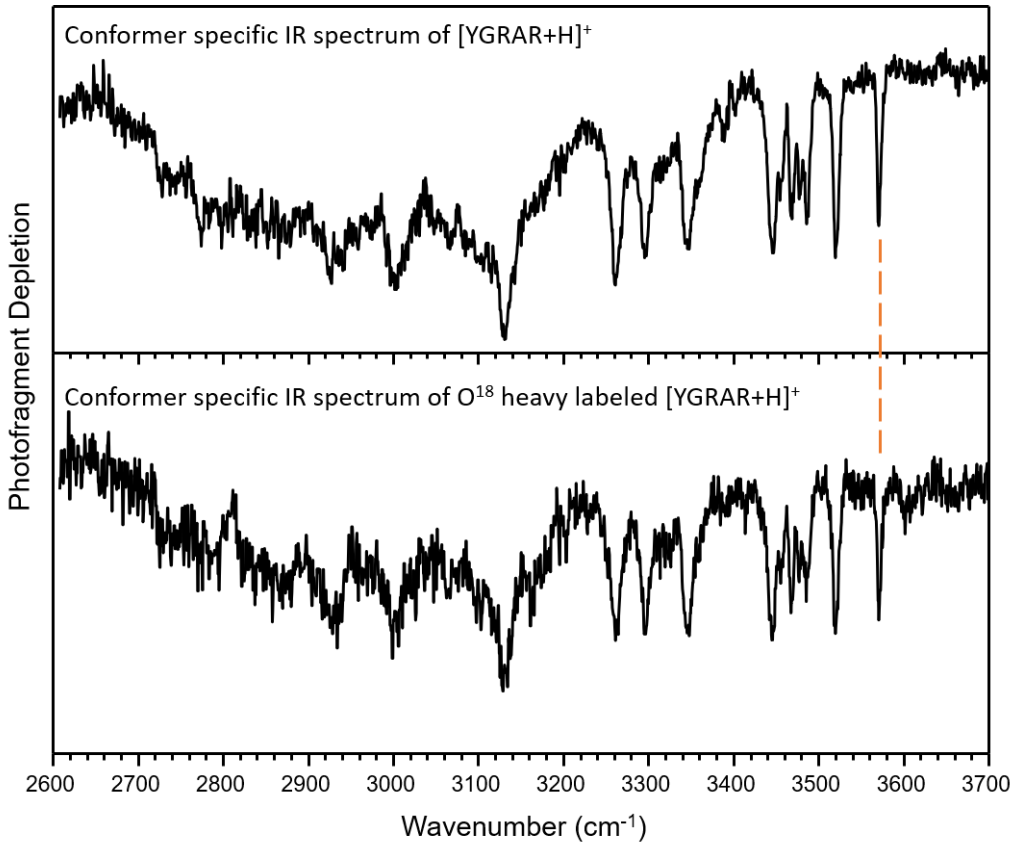


Figure 5.13: Conformer specific spectra recorded on the $S_1 \leftarrow S_0$ origin of non-heavy (top) and heavy (bottom) labelled $[YGRAR+H]^+$

### 5.9.3 Energy Level Diagrams of $[YGRAR+H]^+$ Protomers

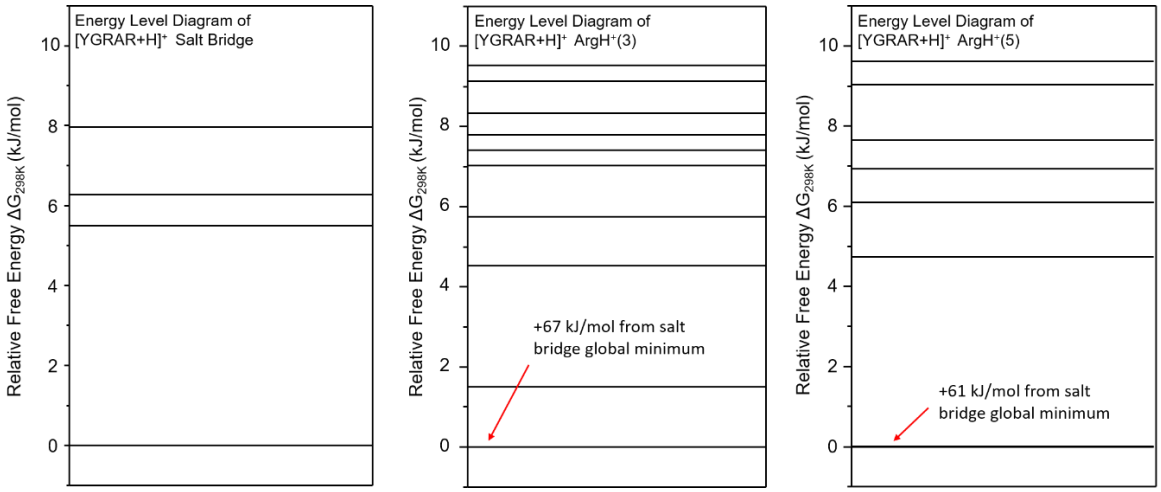


Figure 5.14: Energy level diagrams in terms of $\Delta G_{298K}$ for salt bridge configuration (left) and single protonated charge solvated versions with either the central arg protonated (middle) or the C-terminal arginine protonated (right). All energy diagrams are in relative energy to the global minimum of their respective protomer family. The salt bridge global minimum is the most stable across all protomer families.

### 5.9.4 Comparison of $S_0$ and $S_1$ IR spectra for $[YGRAR+H]^+$

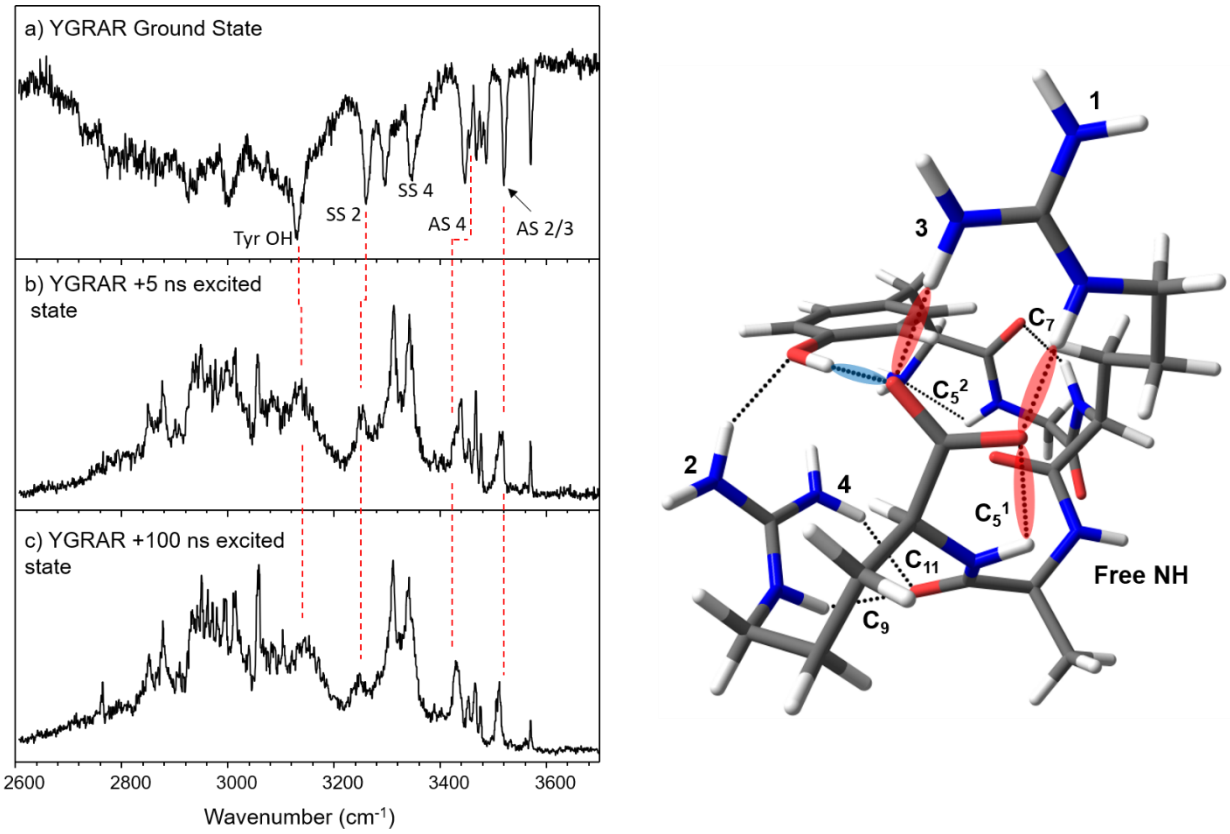


Figure 5.15: IR spectrum of $[YGRAR+H]^+$ recorded at the origin transition in the ground state (a), 5 ns after UV excitation (b) and 100 ns after UV excitation (c). Assigned $[YGRAR+H]^+$ structure (left)

An experimental verification that can be taken to ensure that the assigned structure is correct is to record the IR spectrum in the excited state. NH oscillators near the excited UV chromophore will respond to the local excitation on the chromophore and appear at a lower frequency in the excited state. From the assigned structure the transitions associated with the NH2 group labelled 2 and 4 would be expected to be the most responsive as they are nearest the tyrosine ring. The AS stretch of NH2(2) group was calculated to be degenerate nearly degenerate with the AS stretch of NH2(3). As such in the experimental spectrum, the single transition at 3522 cm-1 was labelled to be responsible to both the NH2 (2) and (3)'s AS fundamental. In the excited state this transition is seen to broaden to a much greater degree compared to the surrounding transitions. This is interpreted as the NH2(2) AS being shifted down in frequency which would cause a this transition

to appear broadened. In a similar but more dramatic manner the NH2(2) SS near 3261 cm-1 appears shifted to a lower frequency by ~10 cm-1. The intensity of this transition in the excited state is also seen to be dramatically reduced. The AS stretch of the NH2(4) appears near the center of the cluster of transitions between 3430 and 3500 cm-1. In the excited state the lowest frequency stretch in this cluster is seen to broaden. This is most likely due to the AS NH2(4) transition red-shifted although it is less clear due to the congestion. The SS stretch of the NH2(4) appears unperturbed in the excited state. The transition labelled as the tyrosine OH also appears with less intensity in the excited state further confirming the assignment of this band. The trends from the excited state at +5 ns appear to hold event at +100 ns.

### 5.9.5 UV and IR spectra of the $[YGRAR+2H]^{2+}$

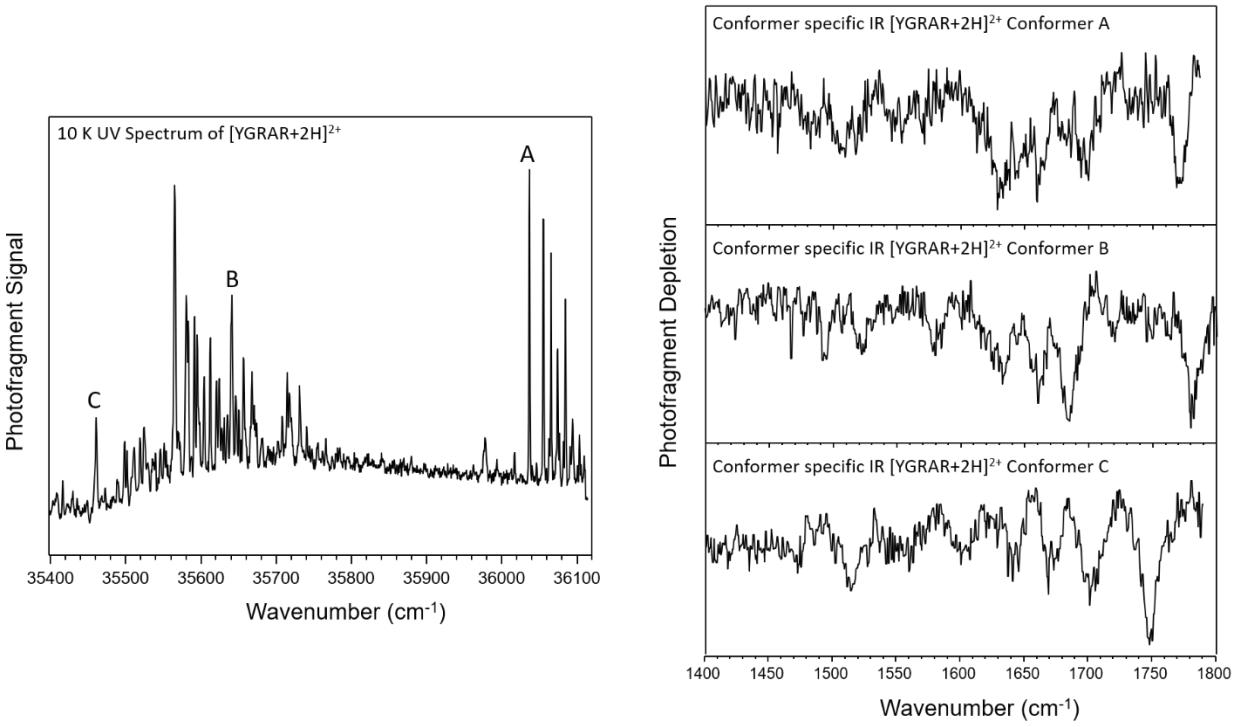


Figure 5.16: Electronic spectrum of $[YGRAR+2H]^{2+}$ with the origin transitions the conformations (A-C) further probed with IR-UV double resonance (left). Conformer specific spectra of conformer A-C (right).

The IR spectra of the three conformations labelled A-C all show a transition at 1750 $cm^{-1}$ or higher, indicative of a carboxylic acid carbonyl stretch. In order to probe the dense set of UV

transitions between 35400 and 35750 $cm^{-1}$ IR-UV hole burning was employed. In this scheme the IR laser is fixed on a vibrational transition of interest, in this case the COOH carbonyl stretch at 1780 $cm^{-1}$, and the UV laser is scanned. The time delay is between the two laser is 200 ns, with the IR laser firing first. As the UV is scanned, any conformation that has a vibrational transition at 1780 $cm^{-1}$ will absorb an IR photon. This causes the UV transitions associated with those conformers to become broadened and appear less intense. The difference between the UV spectra with and without IR irradiation produces a hold burn scan and identifies which UV transitions are associated with the conformations that contain the particular vibrational transition chosen. By fixing the laser on 1780 $cm^{-1}$ it is possible to scan over a region and identify all the UV transitions associated with the charge solvated structures.

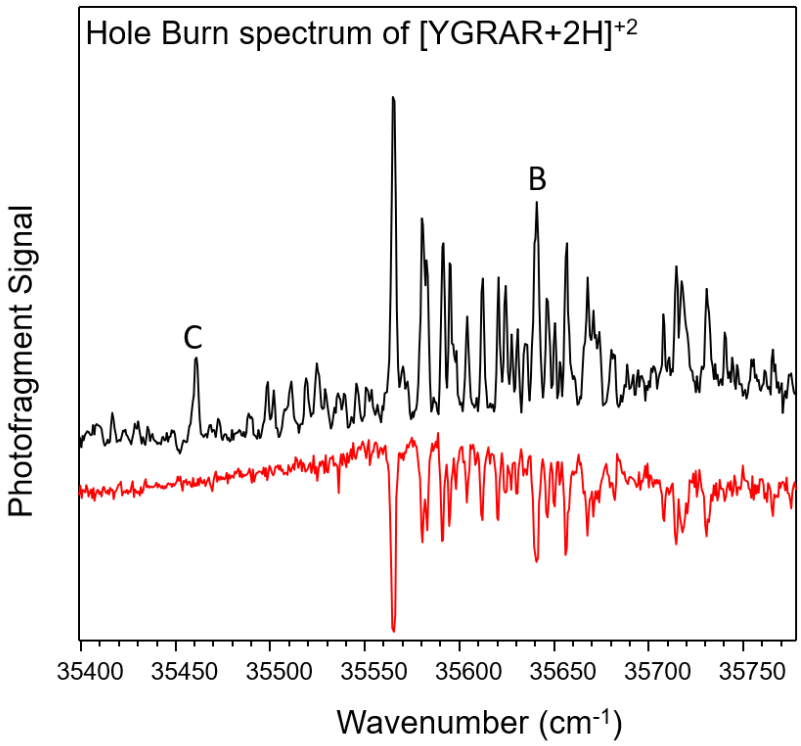


Figure 5.17: Original UV spectrum of $[YGRAR+2H]^+$ (black trace), hole burn scan with IR laser fixed at 1780 $cm^{-1}$ (red)

The transition between 35,400 and 35,750 $cm^{-1}$ were all burned out when the IR laser was fixed on the vibrational transition associated with the carboxylic acid carbonyl stretch. The only transition that was not burned out was conformer C, however, the conformer specific spectrum

shown in Fig. 6.17 reveals that this structure has the carboxylic acid carbonyl at a slightly lower frequency. The transitions UV transitions at higher frequency (36,000 – 36,100 $cm^{-1}$) were not further probed as they are most likely vibrionic bands built off the transitions near 35,500 – 35,700 $cm^{-1}$.

### 5.9.6 TD-DFT calculations

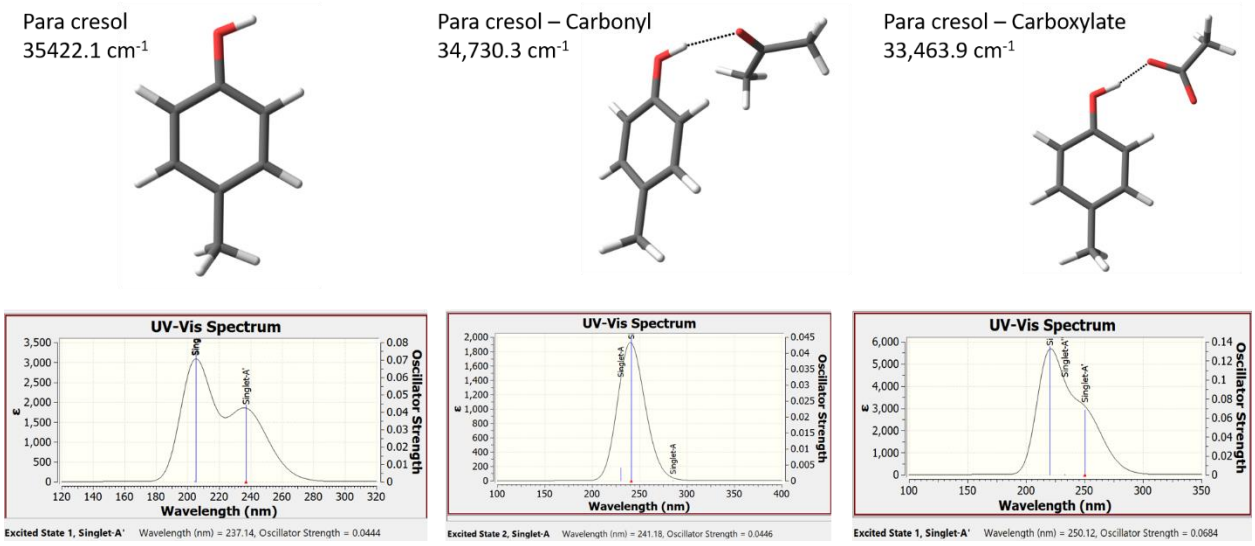


Figure 5.18: TDDFT calculates on bare para cresol (left), bound to a carbonyl functionality (middle), and to a carboxylate functionality (right). The scaled vertical excitation energies for each is given under the label.

### 5.9.7 IR spectrum of [Ac-YGRAR+H]$^+$ in hydride stretch region

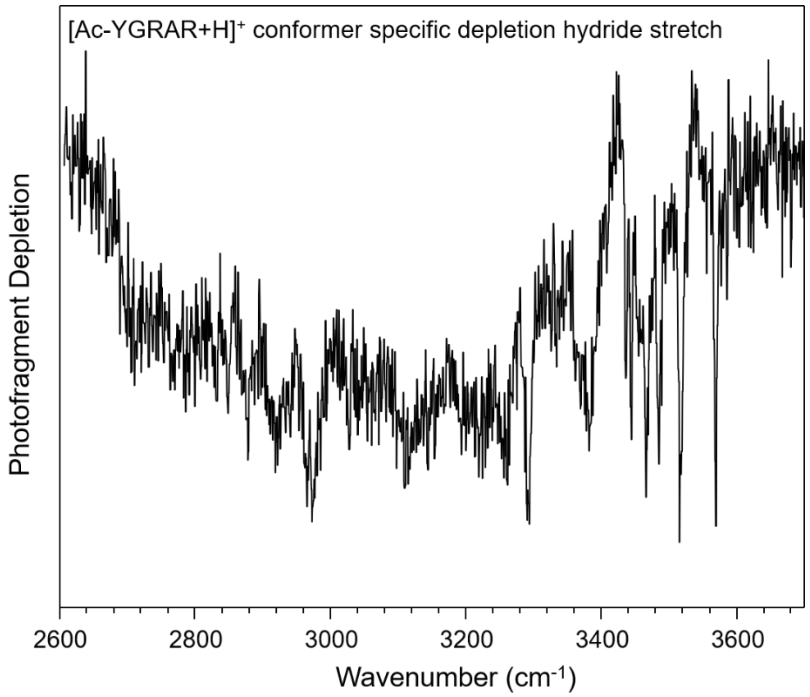


Figure 5.19: Conformer specific IR spectrum of [Ac-YGRAR+H]$^+$ in the hydride stretch region. This spectrum is very similar to that of [YGRAR+H]$^+$ in that there is a broad absorbance between 2600 to 3400 cm$^{-1}$. Many of the sharp vibrational transition also line up well with those of [YGRAR+H]$^+$.

### 5.9.8 Representative Structures of salt bridge versions of [FGRAR+H]$^+$

To better understand the structural changes induced by substituting Tyr for Phe, a conformational search was performed on the salt bridged version of [FGRAR+H]$^+$. While it is experimental established that this sequence does not adopt a salt bridge configuration it is of interest to see how the resulting structures fail to solvate the zwitterionic interaction. To this end a conformational search was performed and resulting in 310 structures after converging. To simplify analysis the structures were grouped into 5 clusters based on their root mean square deviation (RMSD) along the backbone atoms. 4 of the 5 clusters resulting in highly linear backbone geometries in which the C-terminal arginine only made a single hydrogen bond. Due to the limited solvation of the C-terminal arginine were not considered any further. The remaining cluster contained 210 structures and was further clustered into 5 groups, this time based on RMSD of all heavy atoms. A representative structure from 4 of the 5 structures are shown in Fig. 6.20 The 5$^{th}$

cluster resulted in structures in which the C-terminal arginine only formed a single bond and were thus not included in the following discussion.

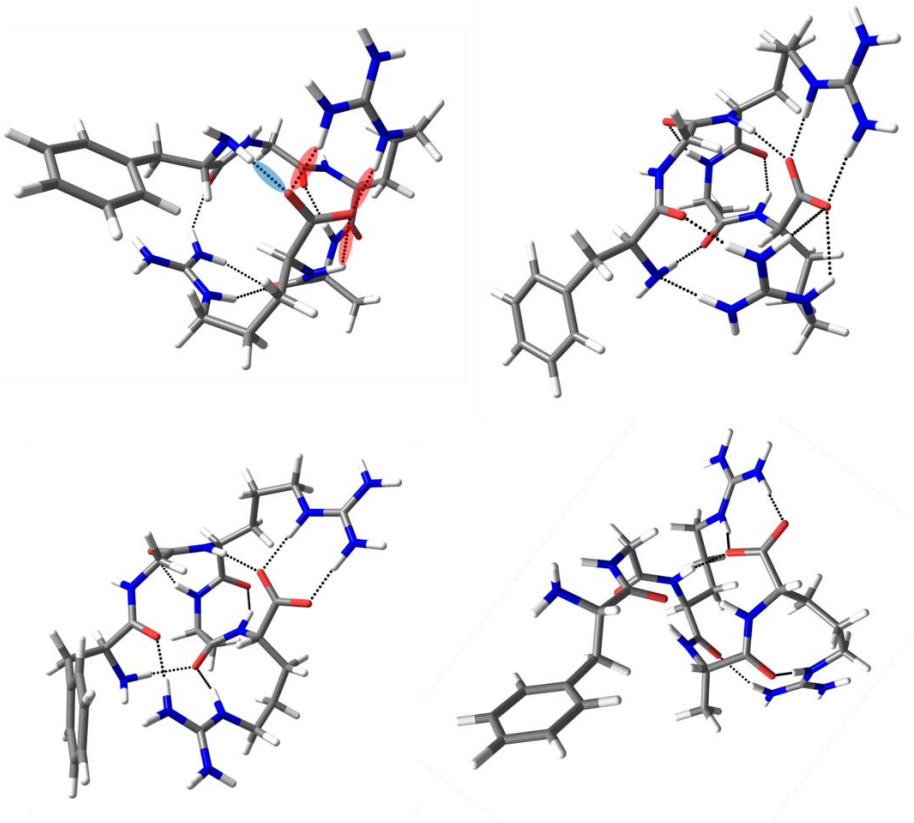

Figure 5.20: Four representative structures of the salt bridged version of $[FGRAR+H]^+$

The structure on the top left shows the same hydrogen bonding pattern between $ArgH^+(3)$ and the $COO^-$. The similar bidentate bonds between the $COO^-$ and $ArgH^+(3)$ and the C5 between the $COO^-$ and $ArgH^+(5)$ amide nitrogen are highlighted in red. The hydrogen bond in $[YGRAR+H]^+$ that was formed between the tyrosine OH and $COO^-$ is now formed using the N-terminal $NH_2$, H-bond shaded in blue. Even $ArgH^+$ makes the same number of hydrogen bonds as in YGRAR, albeit it is between the carbonyl groups rather than to a tyrosine hydroxyl group. The $ArgH^+(5)$ in this case is also missing the interaction with the π cloud, which one would not expect to add much solvation energy towards solvating the guanidinium group. The fact that this structure is not populated is quite surprising and suggests either the N-terminal $NH_2 \cdots COO^-$ hydrogen bond is not

as strong as that formed by the tyrosine hydroxyl group, or that two hydrogen bonds formed by a protonated arginine is simply not enough to stabilize the positive charge.

The structure on the top right shows a different binding geometry between the arginine residues and carboxylate. In this strucutre both arginine residues interaction directly with the $COO^-$, $ArgH^+(3)$ still makes a bidentate hydrogen bond, while $ArgH^+(5)$ forms a bifurcated hydrogen bond to a $COO^-$ oxygen. In this case both arginine residues seem well solvated, however, the proximity of the two guanidinium groups may severely destabilize this structure.

The bottom two structure both make similar hydrogen bonding interactions to the $COO^-$, a bidentate hydrogen bond to $ArgH^+(3)$ along with a hydrogen bond formed between a backbone amide NH and the $COO^-$ much like the hydrogen bond between the tyrosine OH and $COO^-$ in the assigned $[YGRAR+H]^+$ structure. This hydrogen bonding patterns appears quite often which may indicate its importance in the YGRAR structure. In each other the bottom structures the $ArgH^+(5)$ make only 2 hydrogen bonds which seems quite insufficient based on how the $ArgH^+(5)$ is solvated in the YGRAR strucutre. The two strucutre on top even show that there is a larger capacity to solvate this arginine further.

Based on the analysis of the structures shown here in combination with the assigned $[YGRAR+H]^+$ structure, it seems as if the peptide backbone is too short to fully solvate the two protonated arginine residues. The $ArgH^+(5)$ seems to run out of hydrogen bonding partners and the ridged side on bidentate hydrogen bond formed between the $COO^-$ and $ArgH^+(3)$ may prevent the rest of the backbone carbonyls from reaching around to solvate the positive charge. As a consequence, the it is more energetically favorable for each arginine to share the ionizing proton as is most likely the case in the actual $[FGRAR+H]^+$ structures. It would be quite interesting to see how additional polar side chains in close proximity to $ArgH^+(5)$ would affect the salt bridge formed in YGRAR.

## 5.10 References


(1) Kumar, S.; Nussinov, R. Close-Range Electrostatic Interactions in Proteins. *ChemBioChem* **2002**, *3* (7), 604–617. https://doi.org/10.1002/1439-7633(20020703)

(2) Donald, J. E.; Kulp, D. W.; DeGrado, W. F. Salt Bridges: Geometrically Specific, Designable Interactions. *Proteins* **2011**, *79* (3), 898–915. https://doi.org/10.1002/prot.22927.

(3) Basu, S.; Mukharjee, D. Salt-Bridge Networks within Globular and Disordered Proteins: Characterizing Trends for Designable Interactions. *J Mol Model* **2017**, *23* (7), 206. https://doi.org/10.1007/s00894-017-3376-y.

(4) Elston, T.; Wang, H.; Oster, G. Energy Transduction in ATP Synthase. *Nature* **1998**, *391* (6666), 510. https://doi.org/10.1038/35185.

(5) Waldburger, C. D.; Schildbach, J. F.; Sauer, R. T. Are Buried Salt Bridges Important for Protein Stability and Conformational Specificity? *Nat Struct Mol Biol* **1995**, *2* (2), 122–128. https://doi.org/10.1038/nsb0295-122.

(6) Meuzelaar, H.; Tros, M.; Huerta-Viga, A.; van Dijk, C. N.; Vreede, J.; Woutersen, S. Solvent-Exposed Salt Bridges Influence the Kinetics of α-Helix Folding and Unfolding. *J Phys Chem Lett* **2014**, *5* (5), 900–904. https://doi.org/10.1021/jz500029a.

(7) Pylaeva, S.; Brehm, M.; Sebastiani, D. Salt Bridge in Aqueous Solution: Strong Structural Motifs but Weak Enthalpic Effect. *Sci Rep* **2018**, *8* (1), 1–7. https://doi.org/10.1038/s41598-018-31935-z.

(8) Department of Biotechnology, University of Burdwan, Burdwan, West Bengal, India; Bandyopadhyay, A. K.; Ul Islam, R. N.; Mitra, D.; Banerjee, S.; Goswami, A. Stability of Buried and Networked Salt-Bridges (BNSB) in Thermophilic Proteins. *Bioinformation* **2019**, *15* (1), 61–67. https://doi.org/10.6026/97320630015061.

(9) Call, M. E.; Pyrdol, J.; Wiedmann, M.; Wucherpfennig, K. W. The Organizing Principle in the Formation of the T Cell Receptor-CD3 Complex. *Cell* **2002**, *111* (7), 967–979. https://doi.org/10.1016/S0092-8674(02)01194-7.

(10) Bosshard, H. R.; Marti, D. N.; Jelesarov, I. Protein Stabilization by Salt Bridges: Concepts, Experimental Approaches and Clarification of Some Misunderstandings. *Journal of Molecular Recognition* **2004**, *17* (1), 1–16. https://doi.org/10.1002/jmr.657.

(11) Ban, X.; Lahiri, P.; Dhoble, A. S.; Li, D.; Gu, Z.; Li, C.; Cheng, L.; Hong, Y.; Li, Z.; Kaustubh, B. Evolutionary Stability of Salt Bridges Hints Its Contribution to Stability of Proteins. *Computational and Structural Biotechnology Journal* **2019**, *17*, 895–903. https://doi.org/10.1016/j.csbj.2019.06.022.

(12) Ding, Y.; Krogh-Jespersen, K. The Glycine Zwitterion Does Not Exist in the Gas Phase: Results from a Detailed Ab Initio Electronic Structure Study. *Chemical Physics Letters* **1992**, *199* (3), 261–266. https://doi.org/10.1016/0009-2614(92)80116-S.

(13) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. Gas-Phase Infrared Multiple Photon Dissociation Spectroscopy of Mass-Selected Molecular Ions. *International Journal of Mass Spectrometry* **2006**, *254* (1), 1–19. https://doi.org/10.1016/j.ijms.2006.05.009.

(14) Dunbar, R. C.; Polfer, N. C.; Oomens, J. Gas-Phase Zwitterion Stabilization by a Metal Dication. *J. Am. Chem. Soc.* **2007**, *129* (47), 14562–14563. https://doi.org/10.1021/ja076131i.

(15) Skurski, P.; Rak, J.; Simons, J.; Gutowski, M. Quasidegeneracy of Zwitterionic and Canonical Tautomers of Arginine Solvated by an Excess Electron. *J. Am. Chem. Soc.* **2001**, *123* (44), 11073–11074. https://doi.org/10.1021/ja011358d.

(16) Blom, M. N.; Compagnon, I.; Polfer, N. C.; von Helden, G.; Meijer, G.; Suhai, S.; Paizs, B.; Oomens, J. Stepwise Solvation of an Amino Acid: The Appearance of Zwitterionic Structures. *J. Phys. Chem. A* **2007**, *111* (31), 7309–7316. https://doi.org/10.1021/jp070211r.

(17) Prell, J. S.; O'Brien, J. T.; Steill, J. D.; Oomens, J.; Williams, E. R. Structures of Protonated Dipeptides: The Role of Arginine in Stabilizing Salt Bridges. *J. Am. Chem. Soc.* **2009**, *131* (32), 11442–11449. https://doi.org/10.1021/ja901870d.

(18) Prell, J. S.; Demireva, M.; Oomens, J.; Williams, E. R. Role of Sequence in Salt-Bridge Formation for Alkali Metal Cationized GlyArg and ArgGly Investigated with IRMPD Spectroscopy and Theory. *J. Am. Chem. Soc.* **2009**, *131* (3), 1232–1242. https://doi.org/10.1021/ja808177z.

(19) Kjeldsen, F.; Silivra, O. A.; Zubarev, R. A. Zwitterionic States in Gas-Phase Polypeptide Ions Revealed by 157-Nm Ultra-Violet Photodissociation. *Chemistry – A European Journal* **2006**, *12* (30), 7920–7928. https://doi.org/10.1002/chem.200600248.

(20) Bonner, J.; Lyon, Y. A.; Nellessen, C.; Julian, R. R. Photoelectron Transfer Dissociation Reveals Surprising Favorability of Zwitterionic States in Large Gaseous Peptides and Proteins. *J. Am. Chem. Soc.* **2017**, *139* (30), 10286–10293. https://doi.org/10.1021/jacs.7b02428.

(21) Pitts-McCoy, A. M.; Harrilal, C. P.; McLuckey, S. A. Gas-Phase Ion/Ion Chemistry as a Probe for the Presence of Carboxylate Groups in Polypeptide Cations. *J. Am. Soc. Mass Spectrom.* **2019**, *30* (2), 329–338. https://doi.org/10.1007/s13361-018-2079-7.

(22) Eiceman, G. A.; Karpas, Z.; Jr, H. H. H.; Karpas, Z.; Jr, H. H. H. *Ion Mobility Spectrometry*; CRC Press, 2013. https://doi.org/10.1201/b16109.

(23) Konermann, L.; Pan, J.; Liu, Y.-H. Hydrogen Exchange Mass Spectrometry for Studying Protein Structure and Dynamics. *Chem Soc Rev* **2011**, *40* (3), 1224–1234. https://doi.org/10.1039/c0cs00113a.

(24) Zhang, Z.; Vachet, R. W. Gas-Phase Protein Salt Bridge Stabilities from Collisional Activation and Electron Transfer Dissociation. *International Journal of Mass Spectrometry* **2017**, *420*, 51–56. https://doi.org/10.1016/j.ijms.2016.09.010.

(25) Zhang, Z.; Browne, S. J.; Vachet, R. W. Exploring Salt Bridge Structures of Gas-Phase Protein Ions Using Multiple Stages of Electron Transfer and Collision Induced Dissociation. *J. Am. Soc. Mass Spectrom.* **2014**, *25* (4), 604–613. https://doi.org/10.1007/s13361-013-0821-8.

(26) Skinner, O. S.; McLafferty, F. W.; Breuker, K. How Ubiquitin Unfolds after Transfer into the Gas Phase. *J. Am. Soc. Mass Spectrom.* **2012**, *23* (6), 1011–1014. https://doi.org/10.1007/s13361-012-0370-6.

(27) Covey, T.; Douglas, D. J. Collision Cross Sections for Protein Ions. *J Am Soc Mass Spectrom* **1993**, *4* (8), 616–623. https://doi.org/10.1016/1044-0305(93)85025-S.

(28) Jurchen, J. C.; Cooper, R. E.; Williams, E. R. The Role of Acidic Residues and of Sodium Ion Adduction on the Gas-Phase H/D Exchange of Peptides and Peptide Dimers. *J Am Soc Mass Spectrom* **2003**, *14* (12), 1477–1487. https://doi.org/10.1016/j.jasms.2003.08.005.

(29) Burke, N. L.; Redwine, J. G.; Dean, J. C.; McLuckey, S. A.; Zwier, T. S. UV and IR Spectroscopy of Cold Protonated Leucine Enkephalin. *International Journal of Mass Spectrometry* **2015**, *378*, 196–205. https://doi.org/10.1016/j.ijms.2014.08.012.

(30) Rizzo, T. R.; Boyarkin, O. V. Cryogenic Methods for the Spectroscopy of Large, Biomolecular Ions. In *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*; Rijs, A. M., Oomens, J., Eds.; Topics in Current Chemistry; Springer International Publishing: Cham, 2015; pp 43–97. https://doi.org/10.1007/128_2014_579.

(31) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. Spectroscopic Studies of Cold, Gas-Phase Biomolecular Ions. *International Reviews in Physical Chemistry* **2009**, *28* (3), 481–515. https://doi.org/10.1080/01442350903069931.

(32) Boyarkin, O. V. Cold Ion Spectroscopy for Structural Identifications of Biomolecules. *International Reviews in Physical Chemistry* **2018**, *37* (3–4), 559–606. https://doi.org/10.1080/0144235X.2018.1547453.

(33) 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* (8), 2096–2107. https://doi.org/10.1021/acs.jpca.7b12751.

(34) DeBlase, A. F.; Dziekonski, E. T.; Hopkins, J. R.; Burke, N. L.; Sheng, H.; Kenttämaa, H. I.; McLuckey, S. A.; Zwier, T. S. Alkali Cation Chelation in Cold β-O-4 Tetralignol Complexes. *J. Phys. Chem. A* **2016**, *120* (36), 7152–7166. https://doi.org/10.1021/acs.jpca.6b06942.

(35) Rogacheva, O. N.; Izmailov, S. A.; Slipchenko, L. V.; Skrynnikov, N. R. A New Structural Arrangement in Proteins Involving Lysine NH 3 + Group and Carbonyl. *Sci Rep* **2017**, *7* (1), 1–8. https://doi.org/10.1038/s41598-017-16584-y.

(36) Burke, N. L.; DeBlase, A. F.; Redwine, J. G.; Hopkins, J. R.; McLuckey, S. A.; Zwier, T. S. Gas-Phase Folding of a Prototypical Protonated Pentapeptide: Spectroscopic Evidence for Formation of a Charge-Stabilized β-Hairpin. *J. Am. Chem. Soc.* **2016**, *138* (8), 2849–2857. https://doi.org/10.1021/jacs.6b00093.

(37) Han, H.; Londry, F. A.; Erickson, D. E.; McLuckey, S. A. Tailored-Waveform Collisional Activation of Peptide Ion Electron Transfer Survivor Ions in Cation Transmission Mode Ion/Ion Reaction Experiments. *Analyst* **2009**, *134* (4), 681–689. https://doi.org/10.1039/b821348h.

(38) Londry, F. A.; Hager, J. W. Mass Selective Axial Ion Ejection from a Linear Quadrupole Ion Trap. *J Am Soc Mass Spectrom* **2003**, *14* (10), 1130–1147. https://doi.org/10.1016/S1044-0305(03)00446-X.

(39) Zabuga, A. V.; Kamrath, M. Z.; Boyarkin, O. V.; Rizzo, T. R. Fragmentation Mechanism of UV-Excited Peptides in the Gas Phase. *J. Chem. Phys.* **2014**, *141* (15), 154309. https://doi.org/10.1063/1.4897158.

(40) Lee, C.; Yang, W.; Parr, R. G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37* (2), 785–789. https://doi.org/10.1103/PhysRevB.37.785.

(41) Grimme, S.; Ehrlich, S.; Goerigk, L. Effect of the Damping Function in Dispersion Corrected Density Functional Theory. *J Comput Chem* **2011**, *32* (7), 1456–1465. https://doi.org/10.1002/jcc.21759.

(42) Becke, A. D. Density-functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98* (7), 5648–5652. https://doi.org/10.1063/1.464913.

(43) Bleiholder, C.; Suhai, S.; Paizs, B. Revising the Proton Affinity Scale of the Naturally Occurring α-Amino Acids. *Journal of the American Society for Mass Spectrometry* **2006**, *17* (9), 1275–1281. https://doi.org/10.1016/j.jasms.2006.05.010.

(44) DeBlase, A. F.; Harrilal, C. P.; Lawler, J. T.; Burke, N. L.; McLuckey, S. A.; Zwier, T. S. Conformation-Specific Infrared and Ultraviolet Spectroscopy of Cold [YAPAA+H]+ and [YGPAA+H]+ Ions: A Stereochemical "Twist" on the β-Hairpin Turn. *J. Am. Chem. Soc.* **2017**, *139* (15), 5481–5493. https://doi.org/10.1021/jacs.7b01315.

(45) Habka, S.; Very, T.; Donon, J.; Vaquero-Vara, V.; Tardivel, B.; Charnay-Pouget, F.; Mons, M.; Aitken, D. J.; Brenner, V.; Gloaguen, E. Identification of Ion Pairs in Solution by IR Spectroscopy: Crucial Contributions of Gas Phase Data and Simulations. *Phys. Chem. Chem. Phys.* **2019**, *21* (24), 12798–12805. https://doi.org/10.1039/C9CP00700H.

(46) King, G. A.; Devine, A. L.; Nix, M. G. D.; Kelly, D. E.; Ashfold, M. N. R. Near-UV Photolysis of Substituted Phenols. *Phys. Chem. Chem. Phys.* **2008**, *10* (42), 6417–6429. https://doi.org/10.1039/B809250H.

(47) Smith, Z. M.; Steinmetz, V.; Martens, J.; Oomens, J.; Poutsma, J. C. Infrared Multiple Photon Dissociation Spectroscopy of Cationized Canavanine: Side-Chain Substitution Influences Gas-Phase Zwitterion Formation. *International Journal of Mass Spectrometry* **2018**, *429*, 158–173. https://doi.org/10.1016/j.ijms.2017.08.009.

(48) Leavitt, C. M.; DeBlase, A. F.; Johnson, C. J.; van Stipdonk, M.; McCoy, A. B.; Johnson, M. A. Hiding in Plain Sight: Unmasking the Diffuse Spectral Signatures of the Protonated N-Terminus in Isolated Dipeptides Cooled in a Cryogenic Ion Trap. *J. Phys. Chem. Lett.* **2013**, *4* (20), 3450–3457. https://doi.org/10.1021/jz401681y.

(49) Douberly, G. E.; Walters, R. S.; Cui, J.; Jordan, K. D.; Duncan, M. A. Infrared Spectroscopy of Small Protonated Water Clusters, H+(H2O)n (n = 2−5): Isomers, Argon Tagging, and Deuteration. *J. Phys. Chem. A* **2010**, *114* (13), 4570–4579. https://doi.org/10.1021/jp100778s.

(50) High Susceptibility of Histidine to Charge Solvation Revealed by Cold Ion Spectroscopy - Pereverzev - 2017 - Angewandte Chemie International Edition - Wiley Online Library https://onlinelibrary.wiley.com/doi/full/10.1002/anie.201709437 (accessed Nov 3, 2019).

(51) Strittmatter, E. F.; Williams, E. R. Structures of Protonated Arginine Dimer and Bradykinin Investigated by Density Functional Theory: Further Support for Stable Gas-Phase Salt Bridges. *J. Phys. Chem. A* **2000**, *104* (25), 6069–6076. https://doi.org/10.1021/jp000038y.

(52) Watanabe, T.; Ebata, T.; Tanabe, S.; Mikami, N. Size-selected Vibrational Spectra of Phenol-$(H_2O)_n$ ($n$=1–4) Clusters Observed by IR–UV Double Resonance and Stimulated Raman-UV Double Resonance Spectroscopies. *The Journal of Chemical Physics* **1996**, *105* (2), 408–419. https://doi.org/10.1063/1.471917.

(53) Solgadi, D.; Jouvet, C.; Tramer, A. Resonance-Enhanced Multiphoton Ionization Spectra and Ionization Thresholds of Phenol-(Ammonia)n Clusters. *The Journal of Physical Chemistry* **1988**, *92* (12), 3313–3315. https://doi.org/10.1021/j100323a001.

(54) Suresh, C. G.; Vijayan, M. X-Ray Studies on Crystalline Complexes Involving Amino Acids and Peptides. *International Journal of Peptide and Protein Research* **1983**, *21* (3), 223–226. https://doi.org/10.1111/j.1399-3011.1983.tb03098.x.

(55) Salunke, D. M.; Vijayan, M. X-Ray Studies on Crystalline Complexes Involving Amino Acids and Peptides. IX. Crystal Structure of l-Ornithine l-Aspartate Hemihydrate. *International Journal of Peptide and Protein Research* **1983**, *22* (2), 154–160. https://doi.org/10.1111/j.1399-3011.1983.tb02080.x.

(56) Kumar, S.; Nussinov, R. Salt Bridge Stability in Monomeric Proteins11Edited by J. M. Thornton. *Journal of Molecular Biology* **1999**, *293* (5), 1241–1255. https://doi.org/10.1006/jmbi.1999.3218.

# CHAPTER 6. INFRARED POPULATION TRANSFER SPECTRSCOPY OF CRYO-COOLED IONS: QUANTITATIVE TEST OF THE EFFECTS OF COLLISIONAL COOLING ON THE ROOM TEMPERATURE CONFORMER POPULATIONS



## 6.1 Introduction

The field of chemical dynamics builds on a foundation of experimental studies involving molecules small enough that significant control can be exercised over the translational, rotational, and vibrational states of the reactants and for which such state-specific characterization of products is possible.[1-7] Among unimolecular reactions, conformational isomerization is of paramount importance, involving no breakage of chemical bonds, but a rearrangement of torsional degrees of freedom over modest-sized barriers.[8-11] In peptides and peptidomimetics, the large-molecule limit constitutes protein folding, which continues to challenge experimental characterization and computational prediction.[12, 13]

Between these two extremes of small and large molecule lies an interesting size regime worthy of examination. These molecules are large enough that they isomerize on potential energy surfaces of a complexity such that chemical intuition alone is of little value to predict the conformational preferences and isomerization pathways involved.[14] Yet, they are small enough that the powerful experimental tools and computational methods of gas phase physical chemistry can still be brought to bear to gain significant insight to the conformational preferences, competing H-bonded networks, mechanisms, and timescales for isomerization between the low-lying conformational minima.[15]

By bringing these molecules into the gas phase and collisionally cooling them to within a few degrees of absolute zero, the molecules are brought to the vibrational zero-point levels of the low-lying conformational minima on the potential energy surface, where conformational isomerization no longer occurs. Using double-resonance laser spectroscopy, single-conformation

spectra in the ultraviolet and infrared can be carried out.[16] The single-conformation infrared spectra are extraordinarily useful, since the hydride stretch, amide I and amide II regions are sensitive to the number, strength, and type of hydrogen bonds involved. By comparing the experimental spectra to the predictions of theory, it is often possible to assign the observed conformations to specific hydrogen-bonding architectures, providing benchmark spectra for testing force field and *ab initio* methods.[17]

This foundation of single-conformation spectra opens up new experimental approaches to study conformational isomerization. When the molecules of interest are neutral molecules containing an aromatic chromophore, a variation of IR-UV double resonance spectroscopy, called Infrared-induced Population Transfer (IR-PT) spectroscopy, was developed.[14, 18] When carried out in a supersonic expansion, the infrared and ultraviolet laser beams are spatially separated, and a temporal delay inserted between them. IR excitation of a single conformation early in the expansion initiates isomerization, followed by collisional re-cooling to the zero-point levels of 'reactant' and 'products' before UV interrogation of the fractional population changes so induced. The focus of many of the studies using IR-PT to date are capped amino acids and their derivatives, with molecular size of about 20 heavy atoms, where isomerization and re-cooling occur on the microsecond timescale.

Led by the pioneering work of Rizzo and co-workers, IR-UV double resonance methods have been extended more recently to cryo-cooled, gas-phase ions.[19] As with their neutral counterparts, the single-conformation IR and UV spectra of the cold ions provides a level of detail to structural characterization equivalent to that available from the expansion-cooled neutrals. One significant benefit of this extension to ions is that electrospray ionization can be used to bring molecules of much larger size into the gas phase, extending our reach into the 'complexity gap' that separates small from large. Thus, studies of the conformational preferences of larger peptides and their synthetic foldamer counterparts are possible, as protonated or metallated ions, enabling characterization of prototypical secondary structures and probing the influence of charge on their formation.[20-22] This has included studies of protonated peptides 10-15 amino acids long, necessitating the use of isotopic substitution and challenging even the best of computational methods for characterizing the observed conformations.[23, 24] In these studies, the ions are cooled to vibrational temperatures of about 10 K by collisional cooling with helium introduced into the

cryo-cooled ion trap where they are pre-cooled by collisions with the walls of the trap, held at 5 K. [19]

Just as in neutrals, this foundation of single-conformation spectroscopy opens up conformation-specific studies of conformational isomerization dynamics. A recent study by Seaiby *et al*.[25] introduced the ion-specific analog of IR-PT spectroscopy using the cryo-cooled ion trap, showing that IR-excited ions could be re-cooled, leading to measurable population changes following conformation-specific excitation.

The present manuscript describes a combined experimental and computational characterization of the conformational preferences and IR-induced conformational isomerization of the cryo-cooled, gas-phase protonated pentapeptide Tyr-Gly-Pro-Ala-Ala, $[YGPAA+H]^+$. The single-conformation spectroscopy of this ion was reported in preliminary fashion in a recent study, where it was used as a comparison with $[YAPAA+H]^+$.[26] Here we provide a more complete description of the single-conformation spectroscopy and extend the work to include IR-PT studies of its conformational isomerization.

As with most studies of cryo-cooled ions, interrogation of the gas-phase ions occurs after cooling to 10 K by collisions with cold helium introduced into the trap prior to the arrival of room temperature ions. In all these studies, an important and largely unanswered question is the relationship between the final observed conformational populations and those present at room temperature in the gas phase before the cooling process. At the one extreme of fast cooling compared to isomerization, the final populations are frozen in at room temperature values. At the other extreme in which isomerization is fast compared to cooling, significant redistribution will occur towards those structures favored at low temperature. In order to answer this global question, there are other questions that undergird them. First, what are structures of the observed conformers? Second, given these structures, can we identify isomerization pathways between them? Third, what are the measured fractional abundances of the cryo-cooled ion conformers? Fourth, how do these measured abundances compare with the calculated room temperature distributions? Finally, what are the timescales for isomerization and collisional cooling as a function of internal energy, and how do they compare with one another?

Protonated YGPAA is a model system in which just two principal conformational families are present, with these conformations differing in structure along well-defined isomerization pathways that are computationally accessible. We develop a model for the isomerization and

collisional cooling constrained by the experimental results, and apply this model to predict the initial room temperature populations in order to see how they compare with the measured distributions in the cryo-cooled ion trap. We shall see that the conformational populations at room temperature are modified only slightly by the collisional cooling, with near-quantitative agreement between experiment and theory. Thus, significant kinetic trapping can occur in ions in this size regime. The good correspondence between experiment and theory strengthens confidence in the model, enabling extrapolation of the present results to a broader range of circumstances where isomerization is a multi-step process, and to its extension to even larger ions.

## 6.2 Methods

### 6.2.1 Experimental

All spectroscopic data were obtained using a custom instrument for cold ion spectroscopy, previously described.[27] The design comprises a triple quadrupole mass spectrometer in which a spectroscopy axis is mounted between the second ($Q_2$) and third ($Q_3$) quadrupoles. Briefly, ions generated by nano-ESI are trapped and isolated in $Q_2$ via RF DC isolation. A turning quadrupole between $Q_2$ and $Q_3$ is used to steer the mass selected ions down the spectroscopy axis. Ions are then guided into an octupole ion trap which is cooled to 5 K via a closed cycle helium cryostat (Sumitomo Heavy Industries, Tokyo, Japan). Ions are cooled to ~10 K via collisions with the He buffer gas before spectroscopic interrogation by IR (Laservision OPO/OPA) and UV lasers (ScanMatePro Lambda Physik frequency doubled by an Inrad Autotracker III). The resulting photofragments are then extracted back down the spectroscopy axis and turned into a linear ion trap (LIT) for mass analysis ($Q_3$). This quadrupole can be operated in two modes. In spectroscopy mode, remaining precursor ions are resonantly ejected using a supplemental auxiliary waveform calculated by the SX wave software[28] and the remaining ions dumped onto a channeltron detector. In mass spectrometry mode, mass spectra can be recorded using mass-selective axial ejection (MSAE).[29]

UV action spectra are obtained by collecting total photo-fragmentation signal as a function of UV wavelength. Transitions belonging to different conformers are identified and further probed using IR-UV double resonance. In this scheme, the UV laser is fixed on a transition of interest. A tunable IR laser is timed to precede the UV laser by 200 ns. As the IR laser is scanned, IR

absorption by the conformer of interest removes a fraction of its ground state population resulting in a depletion in the total photofragment signal. Monitoring this depletion as a function of the IR wavelength results in a conformer specific IR spectrum. In the case of conformer B, we see depletions in photofragment signal when the UV and IR lasers are both on resonance with the transitions associated with B, superimposed on gains signals from conformer A. These gain signals are generated by IR absorption of conformer A, which produces a broad UV absorption that is present at the UV wavelength used to monitor conformer B. To obtain a conformer specific IR spectrum of conformer B we subtract a background gain spectrum with the UV laser tuned off resonance of all conformers as shown in Figure S1.

In this work, IR-PT spectroscopy is used to study the process of conformation-specific conformational isomerization and to measure the fractional populations in the cryocooled ion trap. This technique, first employed in studies of neutrals, has been extended to ions by Rizzo and co-workers, and is described by those authors.[25] Since a clear understanding of the method is essential to this work, a brief description follows.

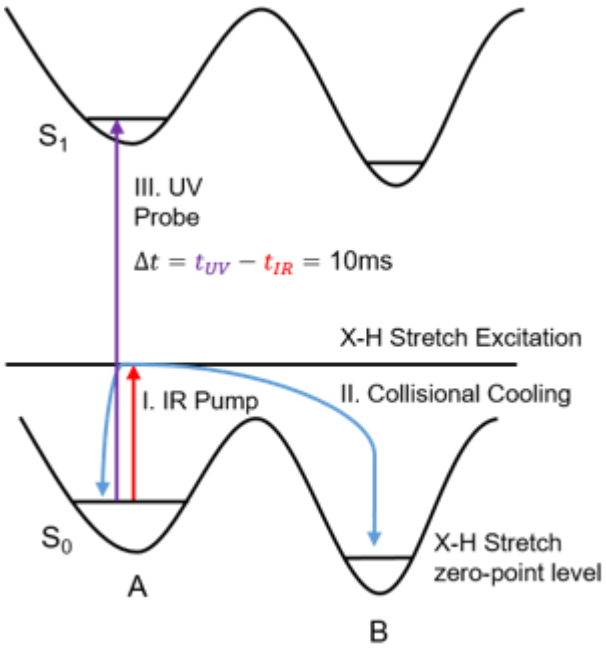


Figure 6.1: Population transfer scheme. Selective IR excitation (I) imparts energy to a single conformer or family of conformers. If the imparted energy exceeds the barriers to isomerization, then as the excited population is collisionally re-cooled (II) it can either isomerize into another available minimum or relax into its initial well. The amount of population lost or gained by the

isomerization process can be monitored by conformer selective UV excitation (III) after all the IR excited population is re-cooled.

In IR-PT spectroscopy, shown schematically in Figure 6.1, the wavelength of the UV laser is fixed on a transition belonging to a particular conformer (e.g., conformer A)**.** IR excitation is timed to precede the UV laser by 10 ms (see Sec. 3.3), thus giving sufficient time for the IR-excited ions to isomerize and collisionally cool prior to conformation-specific interrogation via its UV photofragment ion signal. Isomerization may occur if the energy associated with vibrational excitation exceeds the barrier to isomerization. The increased time between IR and UV laser pulses provides sufficient time for ions excited by the IR photon(s) to cool back down to the zero-point level. In this scheme, depletion in photofragment signal is a result of population from the UV probed conformer isomerizing into other accessible conformations. Similarly, gains in signal are a result of other conformations isomerizing into the conformer interrogated by the UV laser. Here we operate the IR laser at 5 Hz and the UV at 10 Hz. In order to treat the population, transfer signal, $I_{PT}(\tilde{\nu})$, as a fractional population change, the difference between successive laser pulses is plotted as a ratio to the photofragment signal in the absence of IR excitation,

$$I_{PT}(\tilde{\nu}) = \frac{I_{UV}(\tilde{\nu}) - I_{UV+IR}(\tilde{\nu})}{I_{UV}(\tilde{\nu})} \qquad (1)$$

where $I_{UV}$ is the total photofragment signal, and $I_{UV+IR}$ is the resulting photofragment signal upon infrared absorption 10 ms prior to UV irradiation. Further details are given later in the text.

### 6.2.2 Computational

Conformational searches were performed via the Monte Carlo multiple minimum method implemented with the Amber* force field using the MACROMODEL software package.[30] The 199 resulting structures were further screened by optimization at the DFT M05-2X/6-31G* level of theory[31] using GAUSSIAN09 (Version E.01)[32]. Finally, those structures within 20 kJ/mol of the global minimum at the M05-2X level were re-optimized using DFT B3LYP/6-31+G* level of theory[33, 34] with Grimme's empirical dispersion correction (GD3BJ).[35] A tight convergence criteria and an ultrafine grid (int=grid=ultrafine) were used for the optimizations performed at this level

of theory. All harmonic-level vibrational frequency calculations were also calculated at the B3LYP/6-31+G* GD3BJ level in which all hydride stretch frequencies are scaled by 0.958 to account for anharmonicity. Free OH stretches are scaled by 0.973 and the amide I (C=O) transitions were scaled by 0.981. Further details of choice of scale factors and level of theory can be found elsewhere.[26] Transition states for isomerization pathways were identified using QST3 with Gaussian09. Vertical electronic transition energies were calculated using time dependent density function theory (TDDFT) at the M052X/6-31+G* level of theory and scaled by 0.84 to bring the calculated $S_0$-$S_1$ energy differences into range of the experimentally observed origin bands.

## 6.3 Results and Analysis

### 6.3.1 Cold UV Action Spectrum

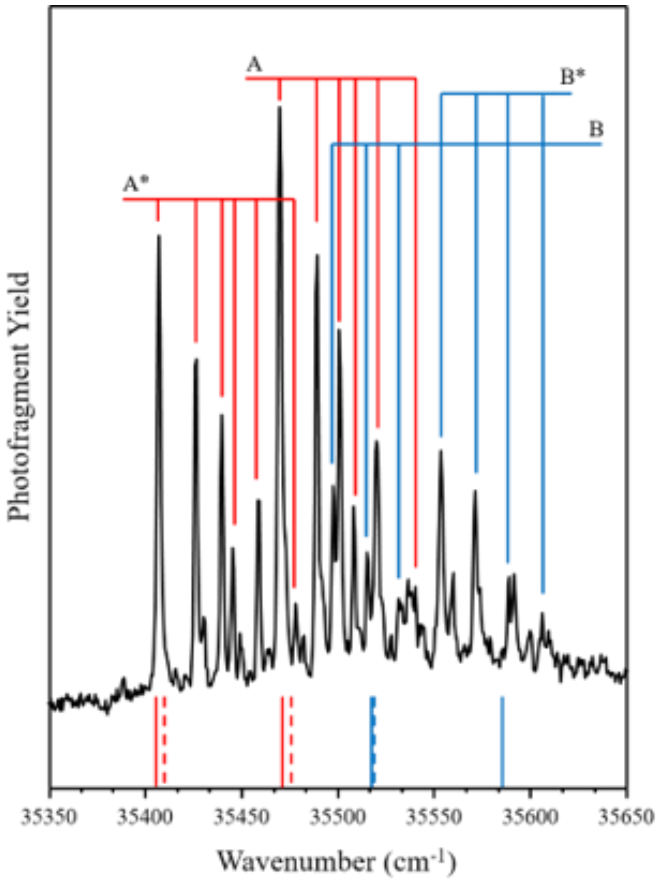


Figure 6.2: Cold UV action spectrum of [YGPAA+H]+. Transitions due to the two conformational families A and B are labeled with red and blue lines, respectively. Transitions labeled as A* and B* arise from conformations differing only in the orientation of the tyrosine hydroxyl group of the assigned structures. Calculated, scaled vertical excitation energies of the assigned structures are shown below the UV. Solid red and blue lines are the vertical excitation energies of the respective tyrosine hydroxyl rotamers, while dashed lines denote a second structural change involving proline ring puckering.

The cold UV action spectrum was taken by monitoring the combined photofragment ion intensity from all available loss channels (-tyr, $b_4^+$, and $y_3^+$) of $[YGPAA+H]^+$, and is shown in Figure 6.2. UV transitions corresponding to the two conformational families observed, are indicated with red (A) and blue (B) lines, respectively. Representative structures for each family are shown in Figure 6.3 (a) and (b). IR-UV double resonance on the transitions labeled A and A*

resulted in identical infrared spectra (Figure S2). This indicates that the structures giving rise to these different Franck-Condon progressions have the same or very similar structures. The frequency shift of the A and A* electronic origins is attributed to the effect of the tyrosine hydroxyl group orientation on the electronic transition. Figure 6.3 (c) and (d) show members of family A which incorporate this group (circled in blue) in different orientations. In addition, the proline ring can also pucker either up or down, as shown with a red circle and arrow in Figure 3 (c) and (d). Time Dependent DFT (TDDFT) calculations at the M052X/6-31+G* level of theory was used to predict the vertical excitation energies of the assigned structures and their respective tyrosine/proline pucker isomers. The vertical excitation energies of the tyrosine OH rotamers of family A and B are shown with red and blue solid line below the UV spectrum in Figure 6.2. The relative splitting between the rotamers agrees with the experimental splitting of the two origin bands for both families, allowing a tentative assignment of the tyrosine rotamer responsible for each origin. The proline pucker has very little effect on the vertical excitation energies as shown with dashed red and blue lines and are likely unresolvable in the experimental spectra.

### 6.3.2 Assigned Structures and Energy Level Diagrams

The infrared spectra in the hydride stretch (2800-3700 $cm^{-1}$) and amide I & II (1400-1800 $cm^{-1}$) regions for families A and B are shown in Figure S3 (a) and (b), respectively, with a representative calculated harmonic-level stick spectrum beneath each experimental trace. Structural assignments were made previously[26] and will not be discussed in detail here. Figure 6.3 (a) and (b) show the lowest energy representative structures for families A and B, respectively. Note that the major difference between the two families is the *cis* vs. *trans* configuration of the carboxylic acid group. Other members of family A which are shown in Figure 6.3 (c) and (d) exemplify the structural permutations which exist in each family. Since the tyrosine hydroxyl group is a hindered rotor, and the different orientations of the proline puckers are not significant enough to affect the hydrogen bonding network, only structures which differ by a *trans* vs. *cis* carboxylic acid are distinguishable in the infrared. Thus, the lowest energy structures from each family are representative of the spectral differences between the observed families.

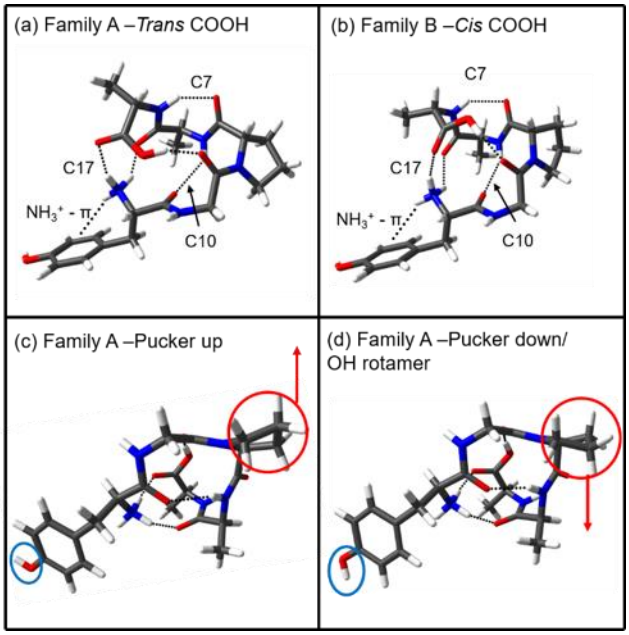


Figure 6.3: Lowest energy representative structures of family A and B, (a) and (b), respectively. Other members which belong to family A, yet do not produce distinct vibrational spectra (c) and (d). These structures differ either by the proline pucker orientation, circled in red, or the tyrosine hydroxyl group, circled in blue. These structural permutations are also observed in family B.

The spectroscopic evidence presented in Figure 6.2 for the presence of both families is already an indication that the cooling process is not truly adiabatic, which is not surprising. Previous studies of cryo-cooled ions have reported better agreement when the experimental spectra are assigned to structures that are lowest in free energy at 298 K rather than $\Delta E_0$.[36] This implies that structures favored at room temperature retain significant population during the cooling process, which would be expected if the cooling rate was faster than the rate of isomerization to the global minimum.

In $[YGPAA+H]^+$, we are offered an opportunity to explore the physical parameters that govern kinetic trapping in more quantitative detail. An energy level diagram depicting the members within each family showing their zero-point corrected relative energies (Figure 6.4 (a)) and the Gibbs free energy at 298 K (Figure 6.4 (b)). Red lines depict the structures with a *trans* carboxylic acid (family A) while the blue lines are associated with structures having a *cis* carboxylic acid (family B). Tyrosine rotamers are designated by columns labeled as Tyr (1)/ Tyr (2). A further structural permutation found in family A is the rotation of the tyrosine group (Figure S4), such that there is no $NH_3^+$-π hydrogen bond as depicted in Figure 6.3 (a). The energy levels

drawn in darker and lighter shades label structures that have the proline pucker in the up or down position, as shown pictorially in Figures 6.3 (c) and (d). We note that the lowest-energy structures for each family differ not only by the carboxylic acid configuration but also by the proline pucker orientation. On either energy scale, the lowest energy conformer from family A or B is calculated to be the global minimum from the initial set of DFT optimized structures.

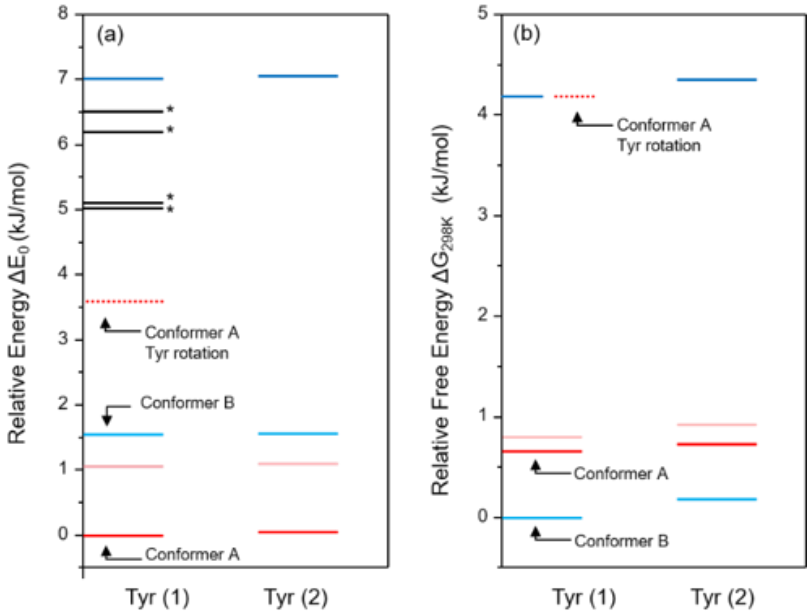


Figure 6.4: . Energy level diagrams of the zero-point corrected relative energies (a) and Gibb's free energies at 298 K (b), of members of families A and B of $[YGPAA+H]^+$. Red and blue lines represent *trans* and *cis* structures, respectively. Darker and lighter shades designate the proline pucker orientation as up or down. Columns labeled as Tyr (1)/(2) differentiate between tyrosine OH rotamers while dashed lines represent different positions of the tyrosine ring. Energy levels marked in (a) with an asterisk have $\Delta G_{298} > 5$ kJ/mol and therefore have insignificant starting populations in the simulations.

Interestingly, as Figure 6.4 shows, the relative ordering of the lowest-energy structures of conformer families A and B changes from *trans* to *cis* when entropic effects are taken into account in going from $\Delta E_0$ (Figure 6.4 (a)) to $\Delta G_{298}$ (Figure 6.4 (b)). If only these two lowest-energy structures are considered, such a re-ordering would lead to a cross-over in which conformer is most populated as temperature increases. However, in the present case, the pair of *cis*-COOH proline pucker up conformers is high enough in energy that they don't capture a significant fraction of the population, even at 298 K (Figure 6.3 (b)). As a result, since the four *trans* COOH conformers are

lower in zero-point corrected energy, *trans*-COOH (A) conformers are favored over *cis*-COOH (B), with fractional populations that approach 100% A as T approaches 0 K. By contrast, at 298 K, the four *trans* and two *cis* conformers will all be populated significantly at thermal equilibrium. Thus, even though the two *cis* conformers are lowest in free energy, the larger number of low free-energy *trans*-conformers leads to a predicted 298 K distribution (58% A, 42 % B) that still favors *trans*-COOH (A) over *cis* (B).

By using the peptide ion $[YGPAA+H]^+$ as a model system, with its two conformers with firm conformational assignments, it is possible to investigate the parameters that govern the populations that are ultimately observed after the cooling process. To do so requires measuring the population distribution after the ions are cooled to 10 K so it can be compared to the predicted equilibrium distribution at room temperature. This measurement will provide insight to how far the equilibrium distribution shifts, if at all, before becoming kinetically trapped.

### 6.3.3 Determining the Experimental Population Distribution

Figure 6.5(a) and (c) show the population transfer spectra for families A and B, respectively, of $[YGPAA+H]^+$ along with their respective IR-UV depletion spectra shown in (b) and (d). The horizontal dashed lines on the population transfer spectra identify zero change in signal intensity. In each spectrum, the negative going portions of the trace are indicative of isomerization *out of* the family being probed and correlate to the vibrational transitions unique to each family (A or B), as is obvious by comparison with the IR-UV spectra in Figures 6.5(b), (d). On the other hand, positive going signals represent a gain in intensity due to isomerization *into* the conformer being probed. For example, the positive signal in the population transfer spectrum of family A, Figure 6.5 (a), correlates to the vibrational transitions of family B, Figure 6.5(d). The population transfer spectra show how the excited population redistributes itself for each IR transition. If no new minima are populated and population is not lost in the excitation process, then the weighted sum of the population transfer spectra for the existing conformations should be zero at every IR wavenumber:

$$\Delta N_{tot}(\tilde{\nu}) = F_A I_A^{PT} + F_B I_B^{PT} = 0 \quad (2)$$

where $\Delta N_{tot}$ is the change in total population, $F_X$ is the initial fractional abundance of conformer X, and $I_X^{PT}$ is fractional population change for conformer X at IR wavenumber $\tilde{\nu}$. The criterion in eqn. (2) is met when the population transfer spectra of family A and B are weighted by $F_A$=

0.65±.04 and $F_B$= 0.35±.04, respectively, where the error bars represent two standard deviations on the mean. The sum of the weighted spectra is shown in Figure 6.5 (e), with red dashed lines locating two standard deviations away from zero. This measurement effectively quantifies the distribution of the population that exists in the trap after the cooling process. Since all members in each family (A or B) have the same infrared spectrum, the population transfer measurement effectively probes the population contribution from the whole family. Thus, 65% of the population is found with the carboxylic acid in the *trans* configuration (family A), while the remaining 35% populates family B and has the carboxylic acid in the *cis* configuration. Population transfer spectra were only recorded for the UV transitions associated with A and B* in Figure 6.2.

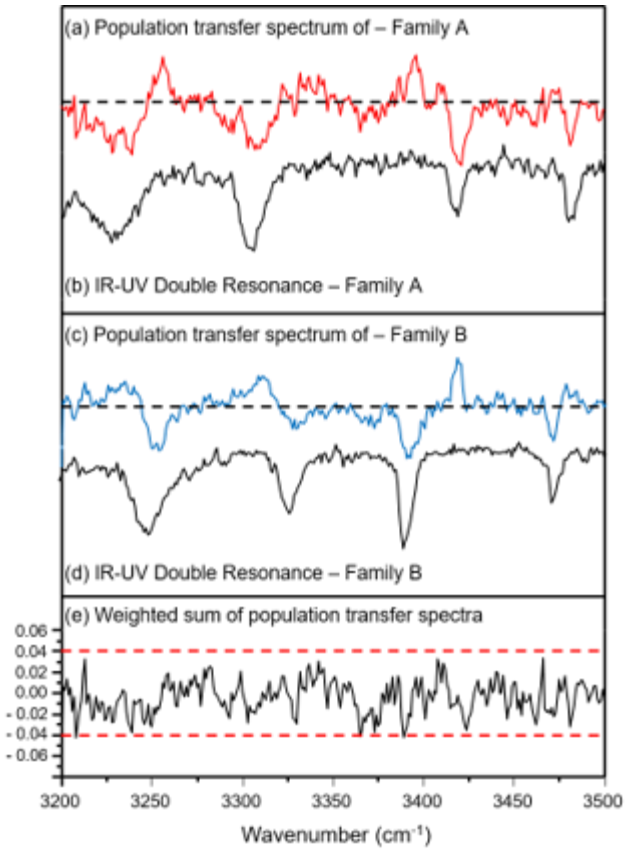


Figure 6.5: (a and c): Population transfer spectra of Family A (red) and B (blue). Black dashed lines mark zero change in signal. (b and d): IR-UV double resonance of Family A and B, respectively. (e): Weighted sum of the two population transfer spectra. Red dashed lines represent two standard deviations away from zero.

The fractional abundance of a particular conformer *i*, $F_i$, at equilibrium as a function of temperature is given by:

$$F_i = \frac{\exp\left(\frac{-\Delta G_i}{k_b T}\right)}{\sum_{i=0}^{j} \exp\left(\frac{-\Delta G_i}{k_b T}\right)} \qquad (3)$$

where $\Delta G_i = k_b T ln Q_v(i)$, is the relative free energy and $Q_v(i)$ is the vibrational partition function of conformer *i*. The unscaled harmonic level frequencies for each conformer are used to calculate ΔG as a function of temperature.

At 300 K the predicted equilibrium distribution is ~58:42 *trans*:*cis* and ~100% *trans* at 10 K. The calculated distribution accounts for the contributions of all members in each family. As

expected, the distribution shifts in favor of the lowest energy conformer as internal energy is removed from the system. Comparatively, the measured distribution 65:35 *trans*:*cis* lies in between the two extremes and indicates that the distribution does begin to shift to the *trans* family. This analysis implies that isomerization to the global minimum is dependent on the collisional cooling rate. To better understand this relationship, the cooling rate must be established and compared to the rates of isomerization.

### 6.3.4 Establishing Cooling Rates

In order, to establish the timescale for collisional cooling of the ions in the cryocooled ion trap, the UV spectrum was taken as a function of ion arrival time into the cryogenic ion trap. Rather than completely emptying the second quadrupole of trapped ions, as is done under normal operating conditions, we extracted a small packet (300 μs in duration), dense enough to produce stable photofragment signal. This allowed the flow of ions into the trap to be stopped while the measurement was being performed. In order to establish a stable reference time, the UV laser wavelength was fixed off-resonance (35377 $cm^{-1}$), where the photofragment signal resulting from the presence of warm ions could be monitored as a function of time. The maximum in this warm-ion signal was used as reference time $t_0$, against which the delays were measured. As Figure 6 (a-e) shows, within 1.0 millisecond of this reference time, the ions are completely vibrationally cooled to their conformational zero-point levels. Figure 6.6(f) shows the UV spectrum after 10 ms of cooling, confirming that no further cooling occurs after the initial 1 ms period. Interestingly, after 300 μs of cooling a portion of the population appears to be fully cold. The presence of these cold ions is attributed to the ion packet arriving at the cold trap distributed in space/time. Ions which enter the trap first will begin to cool while warm ions towards the back of the packet continue to enter. We discriminate against the ions that arrive early by starting the measurement, $t_0$, when there is a maximum in the off resonance warm signal. Consequently, we measure the cooling time of the ions that arrive in the trap later in time such that few warm ions continue to enter the trap. Thus, the cold ions observed at 300 μs could result from a portion of the ion packet which begins the cooling process prior to $t_0$. Nevertheless, this measurement explicitly shows that the ions entering the cold trap are fully cooled after 1 ms under our experimental conditions, a timescale somewhat faster than that obtained in a similar measurement of cooling in an octupole trap by the Rizzo group.[25, 37]

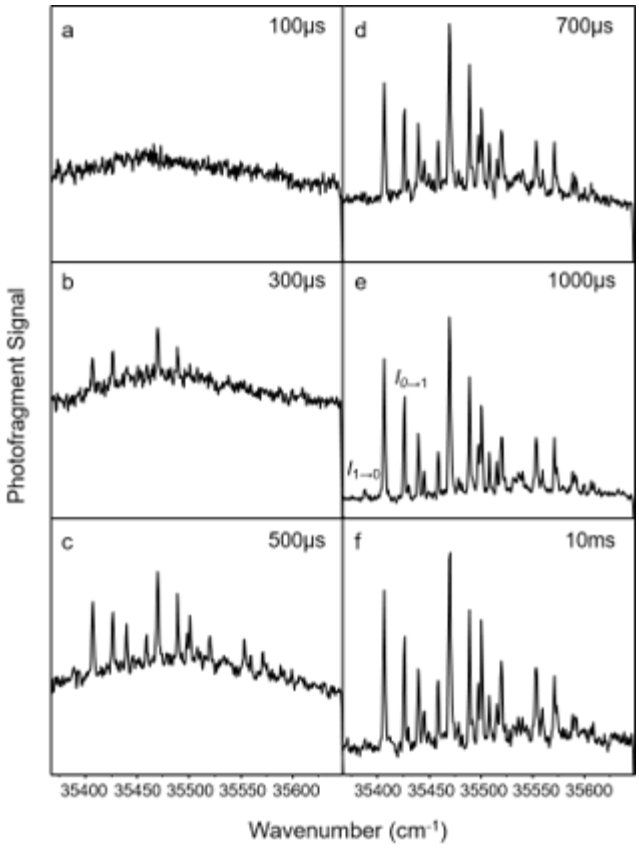


Figure 6.6: UV Spectrum of $[YGPAA+H]^+$ as a function of arrival time in the cryogenic ion trap.

The final vibrational temperature of the ions can be calculated by comparing the intensities of the v"=0→v'=1 cold band with that of the v"=1→v'=0 hot band, labeled in Figure 6 (e). The Franck-Condon factors for these transitions are identical in a harmonic analysis, thus the ratio of their intensities can be related to a vibrational temperature through:

$$\frac{I_{1\to0}}{I_{0\to1}} = exp\left(\frac{-\Delta E}{k_b T}\right) \quad (4).$$

A vibrational temperature of ~10 K is derived from this analysis. Assuming that the ions enter at room temperature (~300 K), leads to an average cooling rate of ~0.29 $K\cdot\mu s^{-1}$, obtained by taking the temperature difference between the initial and final temperature and dividing by the total cooling time.

The average internal energy of the $[YGPAA+H]^+$ ions as a function of temperature T is given by:

$$< E_v > = \sum_{n=1}^{3n-6} \frac{hc\tilde{\nu}_n}{e^{\frac{hc\tilde{\nu}_n}{k_b T}} - 1} \qquad (5).$$

Using the harmonic frequencies for each of the observed conformers, eqn. (5) predicts an average vibrational energy at 300 K of ~6700 $cm^{-1}$, which is collisionally removed in 1.0 ms, leading to an average cooling rate of dEv/dT = 6.7 $cm^{-1} \cdot \mu s^{-1}$.

### 6.3.5 Isomerization Pathways and Reactions Rates

As a protonated pentapeptide with 66 total atoms, the potential energy surface on which isomerization occurs in $[YGPAA+H]^+$ is highly multi-dimensional. However, the change in geometry between the two observed families is relatively simple. Figure 6.7 presents a two-dimensional surface composed of reaction coordinates associate with the carboxylic acid *cis-trans* isomerization (vertical axis) and proline ring pucker (horizontal). The four relevant conformational minima are shown schematically at the four corners, with the carboxylic acid group undergoing the *cis-trans* isomerization shaded in blue. Each isomerization step is accompanied by a relatively simple change in geometry for which a single transition state was identified computationally. We note that this is a simplified potential surface connecting the observed conformers because the tyrosine OH rotamer structures are omitted. However, doing so does not compromise the integrity of the model because the rotamer structures are nearly isoenergetic. Therefore, during the cooling process the population distribution between rotamers remains ~50:50 (Figure S5). Note that the hindered rotation of the entire tyrosine ring (short dashed line in Figure 6.4) is not included explicitly in the 2D surface. However, this structure is included in the kinetic modeling described below.

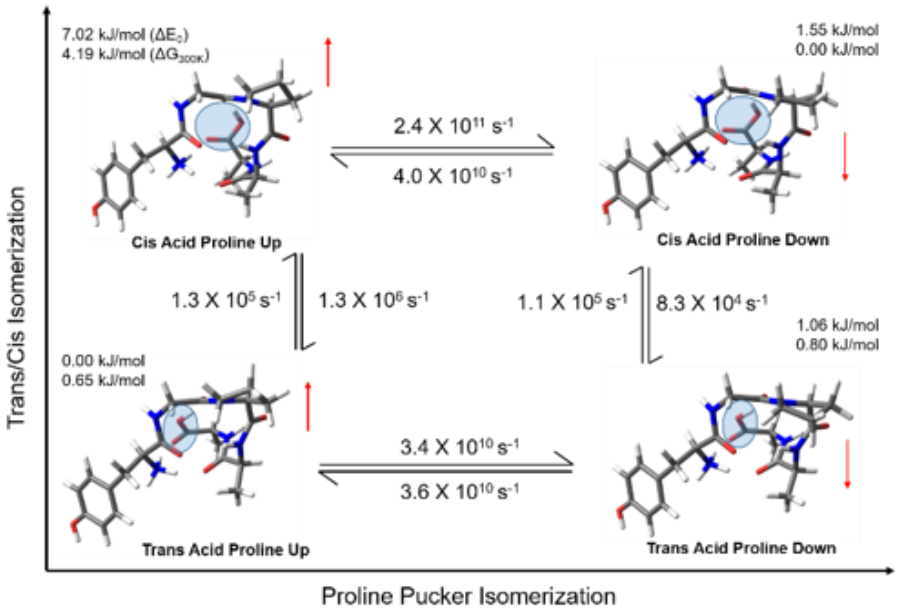


Figure 6.7: . Schematic 2D potential energy surface for the isomerization between members of families A (top) and B (bottom). The orientation of the proline pucker in each structure is indicated with a red arrow, while the carboxylic acid geometry is shaded in blue. Relative rates associated with each isomerization step at 6700 $cm^{-1}$ internal energy are indicated next to the arrows. The relative energy in terms of $\Delta E_0$ and the $\Delta G_{300}$ are also provided above each structure.

The structures, energies, and vibrational frequencies of the first-order transition states on the surface were used to calculate the rate constants for isomerization as a function of internal energy via Rice-Ramsperger-Kassel-Marcus (RRKM) theory[38] using equation (6),

$$k(\varepsilon) = \frac{\sigma W^{\neq}(\varepsilon - \varepsilon_o)}{h\rho(\varepsilon)} \qquad (6)$$

where σ is the reaction degeneracy, $W^{\neq}$ is the number of vibrational states available to the transition state with barrier height $\varepsilon_o$, $h$ is Planck's constant, and $\rho$ is the density of states at energy $\varepsilon$. The density of states and number of states were approximated using the Beyer-Swinehart algorithm[39] using 1 $cm^{-1}$ energy bins. If the cooling maintains a Boltzmann distribution of internal energies with well-defined $T_{vib}$, then isomerization rate constants can also be computed using Activated Complex Theory[40] (ACT):

$$k(T_v) = \frac{k_b T_v}{h}\frac{q^{\neq}}{q} exp\left(\frac{-\varepsilon_0}{k_b T_v}\right) \qquad (7)$$

where $k_b$ is the Boltzmann constant, $T_v$ is the vibrational temperature, $h$ is Planck's constant, $q^{\neq}$ and $q$ are the vibrational partition functions of the transition state and reactant, respectively, and $\varepsilon_0$ is the energy barrier to isomerization.

The zero-point corrected barrier heights out of the global minimum well were found at the B3LYP/6-31+G(d) level of theory to be 1095 $cm^{-1}$ for proline pucker, 3327 $cm^{-1}$ for the carboxylic acid isomerization, and 1106 $cm^{-1}$ for the tyrosine ring rotation. The rate constants at 298 K for the forward and backward reactions are indicated along each isomerization pathway in Figure 6.7. As the peptide ion is cooled, the isomerization rates decrease quickly. Once the rate-limiting COOH isomerization rate constant becomes comparable to the effective rate constant for cooling ($10^3$ $s^{-1}$), we can expect interconversion between the families to slow and kinetic trapping to occur.

### 6.3.6 Collisional Cooling Simulation

Simulations were performed both in terms of internal energy (beginning with 6700 $cm^{-1}$ internal energy) and temperature (beginning at 300 K). Details of the procedure are given in the supplementary material. The energy simulation stopped when the highest *cis-trans* barrier height is reached, while a final temperature of 20 K serves as endpoint for the temperature simulation. The hard sphere collision model was used to estimate the number of collisions the ions experience during the cooling process. The hard sphere collisional cross sections of the assigned ion structures were computed using the trajectory method in MOBCAL[41, 42], and a pressure estimate of 1 x $10^{-2}$ mbar was used to calculate the number density of the 4 K helium buffer gas while the ions are in the trap.[37] We note that the helium is pulsed into the cold trap prior to ions arriving and is continuously pumped out as the ions cool. To check whether the number density of helium does not change significantly during the 1 ms cooling process, the delay time between the helium pulse and the arrival time of ions into the cold trap was increased to 73 ms. With the increased delay a sufficient amount of helium was still present that ions could still be trapped and collisionally cooled to their vibrational zero-point level. This indicates that the pump out time is significantly longer than the cooling time. Figure S6 compares the UV photofragment spectra of $[YGGFL+H]^+$ taken under the increased delay time to that under normal conditions. Thus, assuming the pressure change in 1 ms is negligible, an ion experiences about 500 collisions during the cooling process, which, when divided by the starting internal energy of the ion at room temperature results in a ~13 $cm^{-1}$ energy loss on average per collision. The average time between collisions is estimated to be

1.9 μs. Since the total time required to cool the ion packet was measured experimentally, the approximations used in the hard sphere model do not affect the average cooling rate but instead influence the calculated number of collisions within the defined cooling period, and therefore, the amount of internal energy lost per collision. These collision-specific quantities are not required to simulate the population change as a function of internal energy/temperature.

Using the full reaction network shown in Figure S7, the time-dependent populations were calculated as a function of internal energy or temperature during the cooling process, comparing them with the results assuming full thermal equilibration of the conformer populations at each step. As a starting point, a simple model was implemented in which the temperature was changed by $\Delta T = -1$ K at a rate of -0.29 K/μs. Alternatively, in terms of internal energy, a constant amount of internal energy (13 $cm^{-1}$) was removed each collision (1.9 μs), independent of internal energy. The results for the temperature simulation are shown in Figure 6.8 (a), while those in terms of internal energy are shown in Figure 6.8 (b). The red and blue traces in Figure 6.8 depict the total *trans* (Family A) and *cis* populations (Family B), respectively. The solid trace follows the fully equilibrated distribution as a function of internal energy or temperature, while the dashed lines represent the cooling simulation performed using the average cooling rate. The total cooling time is also plotted in each graph.

Since previous studies of vibration-to-translation energy transfer in collisions of molecules with atoms have shown that the amount of energy transferred per collision scales with the internal energy of the vibrationally excited partner (i.e., $\Delta E/E$ = constant), we also carried out a simulation assuming this cooling behavior (short dashed line in Figure 6.8b).[43] This yielded very similar results to those from the constant-loss simulation (Table S1). Figure S8 compares the time-dependence of the internal energy in the two cooling models. A more detailed discussion of the kinetic scheme is also included in the supplemental material.

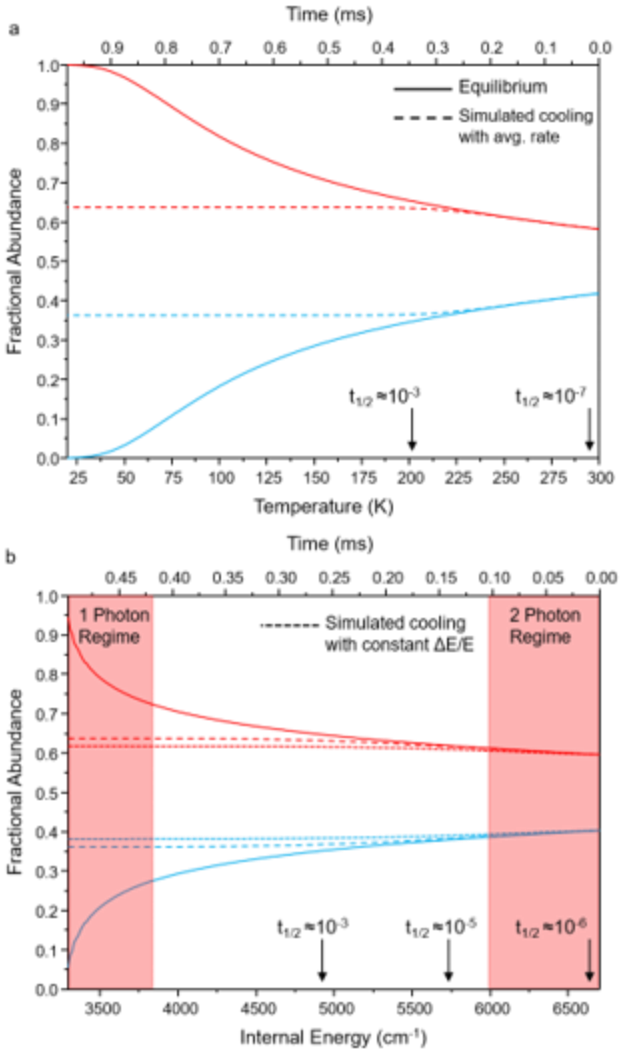


Figure 6.8: Simulations of the fractional abundances ( ) *trans* and ( ) *cis* COOH populations for simulated cooling of the initially 300 K Boltzmann distribution in terms of temperature and average internal energy, removing 1 K at the average cooling rate of 0.29 K/μs or 13 $cm^{-1}$ per collision (every 1.9 μs) (longer dashed lines). Simulated results corresponding to a constant fractional loss of ΔE/E (short dashed lines) are also shown in (b). Solid lines depict the population distribution assuming full thermal equilibration of the population at every step in the cooling process. The half-life of the rate-limiting step of *cis-trans* isomerization is shown at various cooling times.

Modeling the population evolution based on energy or temperature yielded very similar results, each with a final population percentage of ~ 64:36 *trans:cis*, which agrees remarkably well with the experimentally measured distribution of 65:35 *trans*:*cis*.

The half-life of the rate-limiting *cis-trans* isomerization is given at various internal energies and temperatures on the plots in Figure 6.8. Based on the average cooling time it is apparent that

isomerization becomes much slower than the cooling rate (~$10^3$ $s^{-1}$) after ~ 300 μs of cooling or conversely at 4700 $cm^{-1}$ or ~200 K. After this point in the cooling process the populations don't change significantly (<1%), leaving the distribution kinetically trapped. Note that the internal energy at which isomerization is quenched is about 1400 $cm^{-1}$ above the highest *cis-trans* isomerization barrier.

While not the primary focus of the simulation, the calculations also shed light on the IR-induced isomerization process itself. As the red shaded regions in Figure 8 indicate, the energy of one IR photon in the amide NH stretch region puts energy into a given conformer just above the calculated rate-limiting barrier to *cis-trans* COOH isomerization, with a calculated isomerization rate constant far too slow to account for the observed isomerization in the IR-PT experiment. One can surmise on this basis that the molecules undergoing isomerization have absorbed two IR photons, where $t_{1/2}$ ~ 1 μs, and is therefore fast compared to the cooling time. Note that this internal energy is close to the average vibrational energy at 300 K (6700 $cm^{-1}$). Since our only assumption when determining the fractional abundances is that the total change in population at every IR wavelength is zero, any non-linear effects from multi-photon absorption do not affect the weighted sum of the population transfer spectra, and thus not the fractional abundances extracted from them.

### 6.4 Discussion

IR-UV double resonance spectroscopy has been used to identify two distinct conformational families in the protonated peptide ion $[YGPAA+H]^+$. The distinctive difference between families is the *cis* vs. *trans* configuration of the carboxylic acid group on the C-terminus, changing the way the COOH group interacts with the $NH_3^+$ group and peptide backbone. Each family contains several members, which are structurally distinct, yet still produce the same infrared spectrum, since they don't materially change the hydrogen-bonded network. Specifically, family members differ in either the orientation of the OH group on the tyrosine ring, or in the pucker of the proline ring. Using the relative energies of the observed conformers at 10 K, the equilibrium distribution at 10 K is predicted to be essentially 100 % *trans.* The observation of multiple conformations is an indication that structures of importance at room temperature do not completely isomerize to the single global minimum structure that is preferred at 10 K. This is not surprising, since many conformation-specific studies of both neutrals and ions find multiple conformers, only one of

which is the global minimum, indicating that some level of kinetic trapping is common in supersonic expansions or cryocooled ion traps.[18, 36]

Using population transfer spectroscopy the population distribution after the cooling process has been directly measured and compared to the distribution which is predicted at 300 K. Experimental results show that the distribution is 65:35 trans:cis while the distribution predicted by B3LYP/6-31+G(d) GDB3J calculations at room temperature is 58:42 trans:cis. The larger trans conformer population measured experimentally is consistent with some transfer of population from cis to trans during the cooling process; however, as previously mentioned by Voronina et al.[36], this process occurs in competition with the collisional cooling in the trap. Simulations based on the experimentally established cooling rate and calculated rates of isomerization were used to model the evolution of the population distribution as a function of internal energy and temperature. The theoretical simulations correlate well both in terms of internal energy and temperature with the measured distribution, providing evidence that significant kinetic trapping of the conformer populations does occur during the cooling process in the cryocooled ion trap as the ions, initially at 300 K, cool to a final vibrational temperature near 10 K.

This kinetic trapping occurs when the rate of isomerization slows to the point that the cooling rate successfully competes with isomerization to quench further population changes. We have experimentally determined that cooling is complete in $[YGPAA+H]^+$ within 1 ms, so that isomerization competes with cooling at internal energies where the isomerization rate is $>10^3$ $s^{-1}$. The RRKM isomerization rate k(E) for a given single-step process involves a competition between the number of states available to the transition state for that process as a function of energy in excess of the barrier, and the density of states in the reactant well at the reactant's total energy (eqn. 6). Figure 6.9 (a) is a semi-log plot of the total number of states available to the transition state for $[YGPAA+H]^+$ as a function of total internal energy up to 6700 $cm^{-1}$, for a series of barrier heights ranging from 1000-4000 $cm^{-1}$ (12 – 36 kJ/mol). In making these plots, calculations were carried out in two ways. The dashed traces for barriers of 1095 and 3327 $cm^{-1}$ used the full set of calculated harmonic vibrational frequencies for the conformational minima and transition states for the proline pucker (TS=1095 $cm^{-1}$) and *cis-trans* COOH isomerization's (TS=3327 $cm^{-1}$), respectively, and are shown with dashed lines. The solid-line plots are for four representative barriers of 1000, 2000, 3000, and 4000 $cm^{-1}$ assuming that the 3N-7 vibrational frequencies of the transition states for model pucker/COOH isomerization reactions are identical to those of the

minima, with one mode of 200 $cm^{-1}$ removed. For the pucker and *cis-trans* isomerizations we have checked that this simple approximation gives rate constants less than a factor of 2 different than those that use the actual transition states and vibrational frequencies.

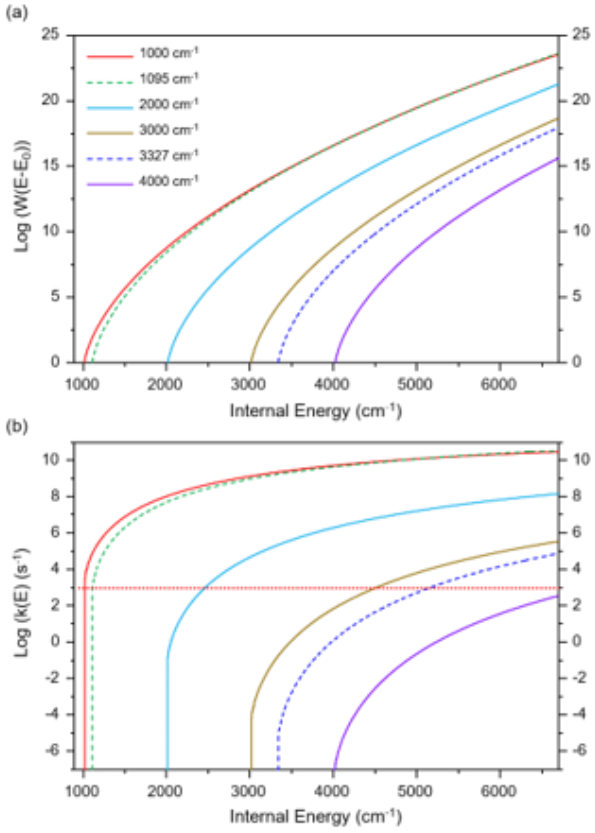


Figure 6.9: (a) Semi-log plot of the number of states available to the transition state at total energy E, $W^{\neq}(E-E_0)$, for barrier heights of 100, 1095 (pucker), 200, 3000, 3327 (*cis-trans* COOH), and 4000 $cm^{-1}$. (b) Energy-dependent RRKM rate constants for single-step isomerizations for the same set of barrier heights as a function of internal energy of the reactant. See text for detail.

The accuracy of this approximation for the proline pucker and cis-trans isomerizations indicates that, in molecules of this size, first-order transition states will often have structures and vibrational frequencies that are close to those of the minima they connect. In $[YGPAA+H]^+$, the five lowest energy minima and the transition states connecting them all have about 50 of the 192 vibrational modes below 500 $cm^{-1}$, with a very similar frequency distribution (see Tables S2 and S3). These low-frequency modes dominate the density of states and number of states in the k(E) calculation. Furthermore, at the barrier, the density of states of the protonated pentapeptide

reactant is already of gigantic proportions for a barrier of 1000 $cm^{-1}$ (□(E) = 2 $X10^7$ states/$cm^{-1}$), and increases by about 3 orders-of-magnitude for each 1000 $cm^{-1}$ increase in barrier height.

Figure 6.9 (b) shows the analogous plots of log k(E) for energies beginning at the barrier height up to an average internal energy of 6700 $cm^{-1}$ (<$E_v$(300 K)>), plotted as a function of total internal energy. In order to understand the effects of collisional cooling in the cryo-cooled ion trap on the initial Boltzmann distribution of ions at room temperature, a horizontal red dashed line is shown on the plot at a cooling rate of $1X10^3$ $s^{-1}$, based on the experimentally measured timescale for cooling, which is complete in the trap within 1 msec.

For isomerization steps occurring over relatively small barriers (~1000 $cm^{-1}$), Figure 6.9 (b) shows that the isomerization rate remains faster than the cooling rate until energies within a few tens of wavenumbers above the barrier. Thus, for low barrier isomerization processes, redistribution of population will continue down to temperatures with average internal energies near the barrier height. In $[YGPAA+H]^+$, the distribution among ring-puckering conformers (with barriers near 1000 $cm^{-1}$) would be characterized by temperatures of ~70 K, but this has little consequence for the population distribution between the two families we measure experimentally.

Conversely, for larger barriers, the density of states in the conformer well is so large that at threshold the isomerization rate is exceedingly slow (e.g., $10^{-4}$ $s^{-1}$ for $E_{barrier}$ = 3000 $cm^{-1}$). As a result, the isomerization rate for larger barrier processes slows precipitously as the internal energy of the molecule is removed, allowing the cooling rate to become dominant at energies well in excess of the barrier height. These conditions favor kinetic trapping of entropically favored structures.

The minimum barrier height that will cause kinetic trapping is critically dependent on the molecular size, which dictates the dimensionality of the surface, the nature of the isomerization barriers, the vibrational frequencies, the average internal energy of the ions at room temperature, and the typical number of steps linking 'reactant' to 'product'. The present study has clearly shown that, while the cooling rates are ~$10^3$ slower in the cryo-cooled ion trap (ms timescale) than in a supersonic expansion (μs timescale), cryo-cooled ions the size of the protonated peptide $[YGPAA+H]^+$ have typical isomerization rates slower by a similar factor, so that they are once again in competition with cooling at this slower rate. Starting from an average internal energy (6700 $cm^{-1}$) well above the rate-limiting barrier for *cis-trans* COOH isomerization (3327 $cm^{-1}$), the conformational populations are 'frozen in' nearly 1400 $cm^{-1}$ above this barrier, with an average

internal energy of ~4700 $cm^{-1}$ (56 kJ/mol) and an effective conformational temperature of T(conformers) ≅ 250 K.

As the molecular size increases still further, the number of low frequency modes will continue to grow quickly, leading to further steep increases in the density of states of the conformational minima as a function of energy. While the average internal energy also increases, the barrier heights for individual steps (e.g., involving breaking a H-bond) will remain similar in size, so that the timescales for individual isomerization steps will continue to slow, while the cooling rate in the ion trap should stay roughly constant. Thus, it is likely that the minimum barrier height for significant kinetic trapping in the ion trap will, if anything, slowly drop as the size of the ion increases. It is also important to note that processes with low barriers are associated with subtle geometry changes such as the hindered rotation of a methyl group or free OH group (e.g., the Tyr OH), with little consequence for understanding the potential energy landscape for peptide folding. For isomerization processes that result in spectroscopically distinct conformers, hydrogen bonds will typically need to be broken and re-formed, with barriers as large or larger than the *cis-trans* COOH isomerization that is rate-limiting in the present study. Furthermore, the number of steps will also increase, further slowing the net rate of inter-conformer conversion.

## 6.5 Conclusion

The major inference of the present work is that ions the size of pentapeptides and greater, the rate of isomerization between conformers with distinct H-bonding architectures will often be slow enough that significant kinetic trapping will occur during the cooling process in the cryocooled ion trap. The observed conformers will thus reflect a conformational population distribution close to the room temperature Boltzmann distribution present in the room temperature, collision-dominated ion sources and traps used for ion production and mass selection. This enables those using techniques of cryo-cooled ion spectroscopy to study structures that are important at room temperature, yet still benefit from the high resolution afforded by vibrational cooling down to 10 K.

One issue this study does not directly address is the degree to which the observed conformers reflect the conformational preferences of the ion in the solution from which electrospray ionization occurs. This topic is of significant current interest to those engaged in native mass spectrometry, where inferences regarding solution-phase structures are drawn from the gas phase mass

spectrometry studies.[44] Annealing techniques can be used to test whether the probed structures are the same ones that are prepared when excess energy is given to the ion before thermal equilibration. It will also be fascinating to search for and spectroscopically characterize both gas-phase and solution-phase structures of ions using single-conformation spectroscopy.

## 6.6 Supporting Information

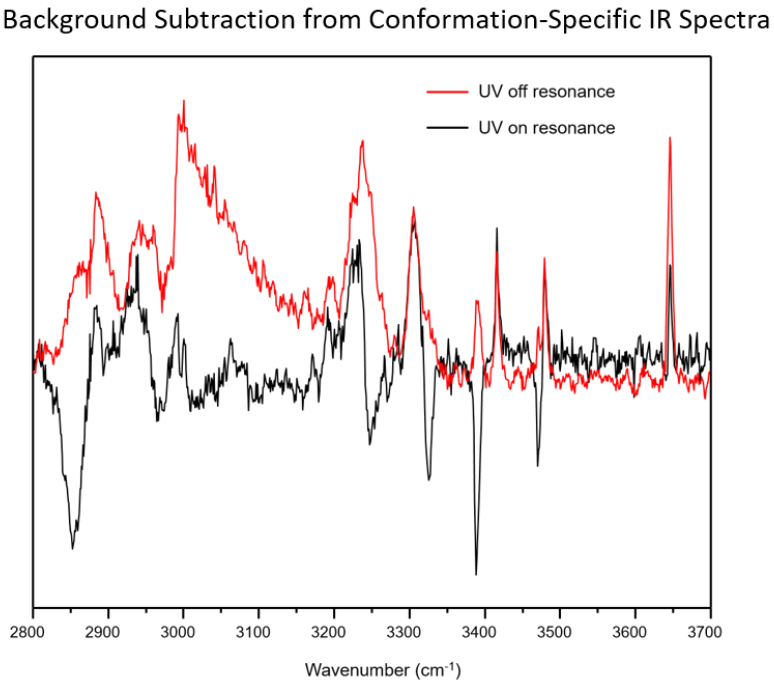


Figure 6.10: Non-confomer specific gain spectra (red) taken off resonance at 35548.6 $cm^{-1}$ overlaid on raw conformation-specific infrared spectra of conformer B at 35553.7 $cm^{-1}$. The gain spectra were normailized to the unique NH stretch of conformer A in the on-resonance signal at 3308 $cm^{-1}$. The resulting subtraction in given in Figure S.3 (b).

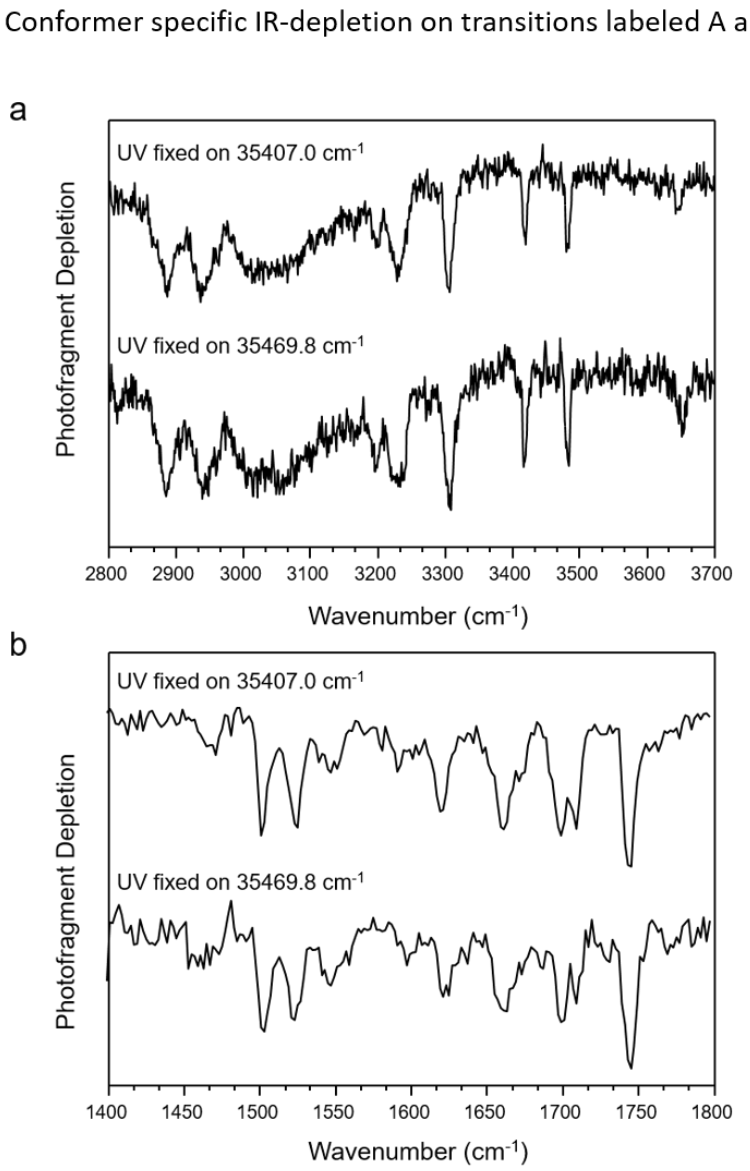


Figure 6.11: Conformer specific IR depletion spectra taken on distinct UV transitions 35407.0 $cm^{-1}$ and 35469.8 $cm^{-1}$ in the hydride stretch (a) and amide I/II (b) regions. The UV transitions reported here are tentatively assigned to tyrosine rotamers (A and A*). Resulting IR spectra are identical, indicating conformations giving rise to the distinct UV transitions possess similar 3D structure.

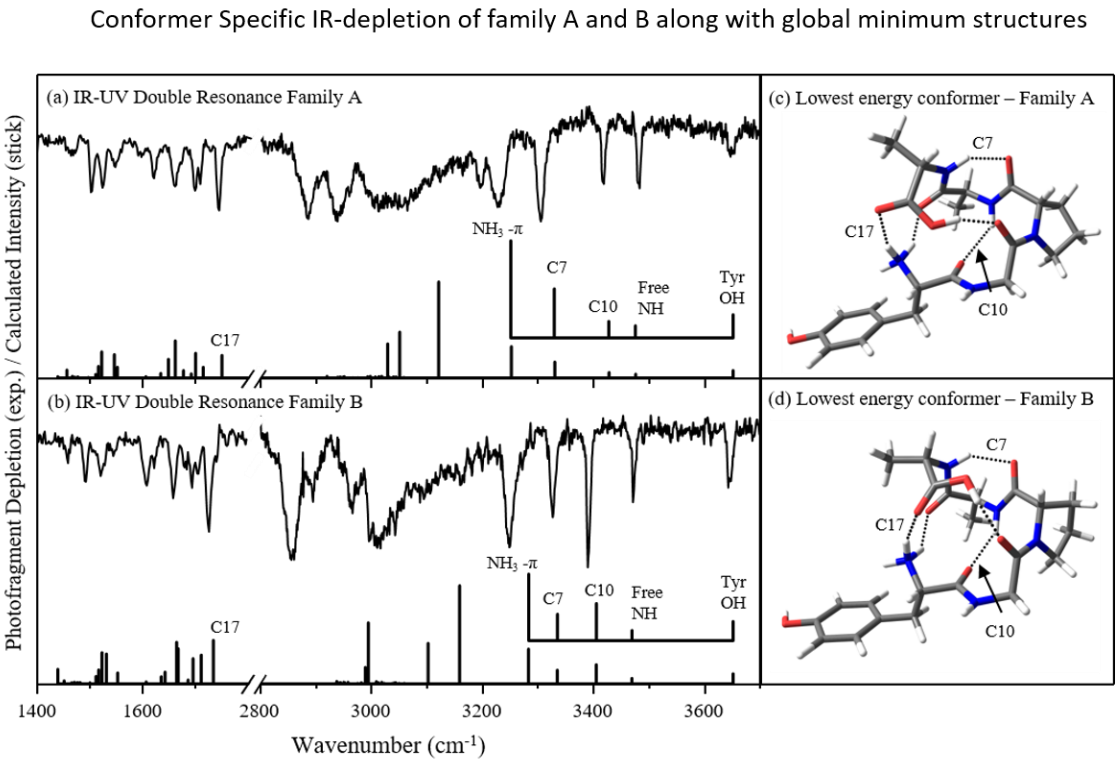


Figure 6.12: Conformer specific IR spectra in amide I/II and hydride stretch regions of family A (a) and B (b). Lowest energy representative structures for families A and B are shown in (c) and (d), respectively.

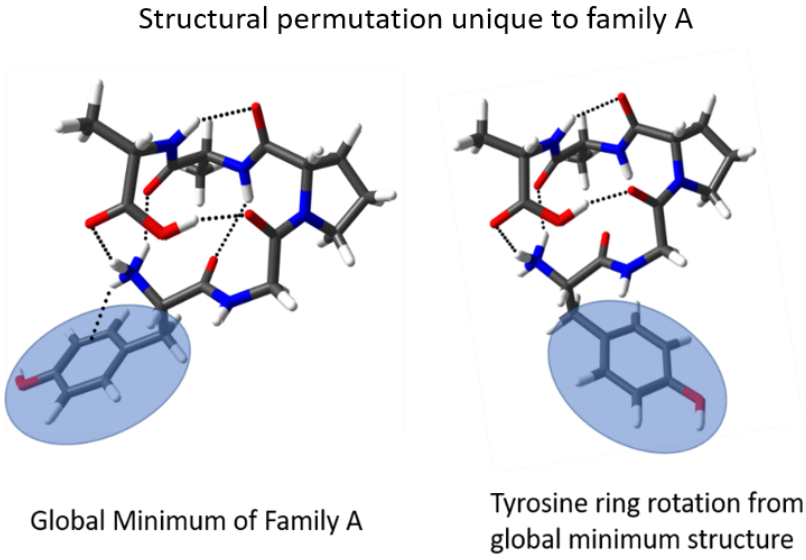


Figure 6.13: Comparison of the global minimum *trans* structure to the conformation differing only in the tyrosine hindered rotation, labelled with a short dashed red line in Figure 6.4. There are no further structural permutations to the structure pictured on the right.

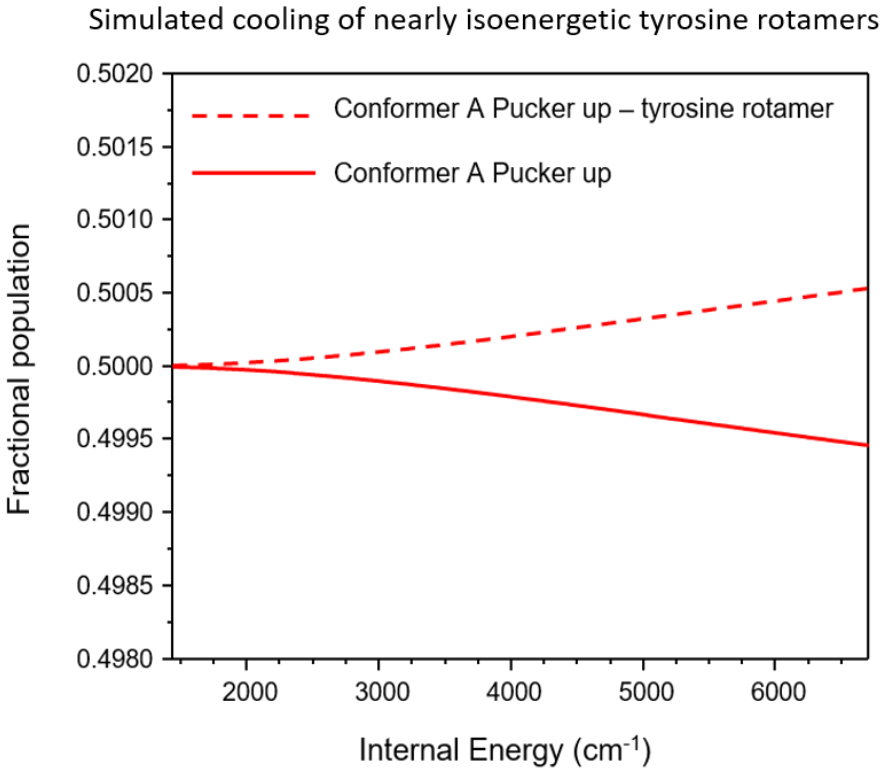


Figure 6.14: Equilibrium distribution as a function of internal energy between tyrosine OH rotamers of Conformer A with the proline puckered in the up position. Note the scale on the fractional populations, which are maintained near 0.500:0.500 throughout. Since the two conformations are nearly degenerate in free energy as a function of temperature there is no driving force to populate one over the other as internal energy is removed, thus it is possible to omit them from the kinetic cooling scheme.

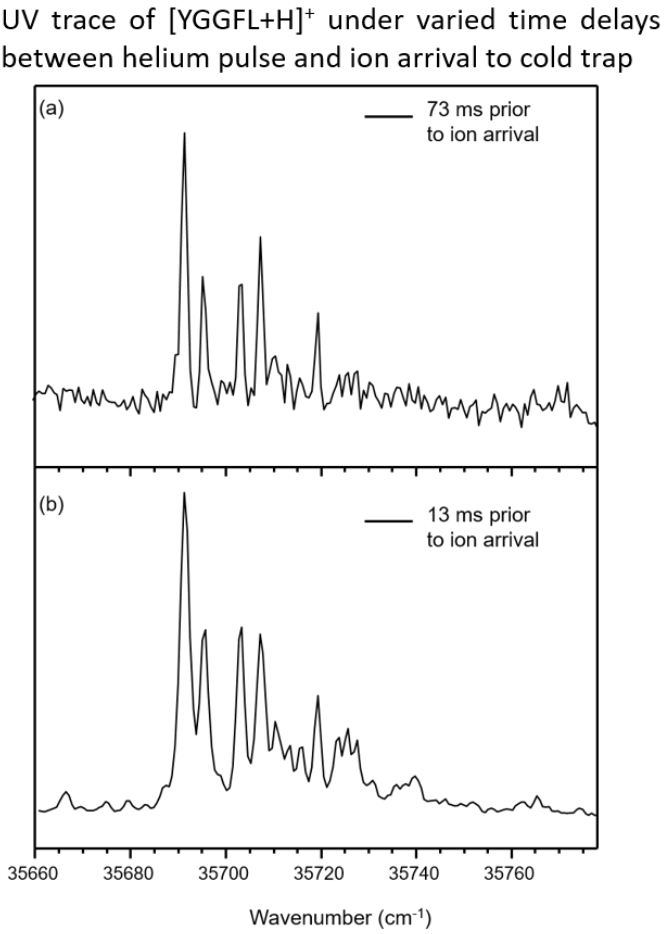


Figure 6.15: UV trace of $[YGGFL+H]^+$ under an increased delay between the helium pulse and ion arrival to the cold trap (a) compared to the delay under normal operations.

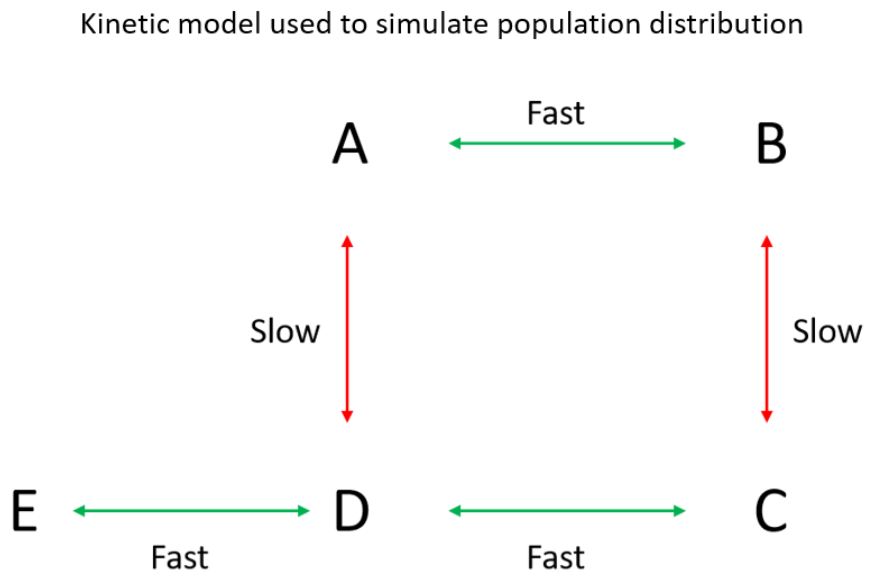


Figure 6.16: A simplified reaction scheme that shows the relative isomerization rates for all species involved in the kinetics scheme used. A and B represent *cis* species while D, C and E represent *trans* species. More details are in the following text.

### 6.6.1 Details of Cooling Simulation

Figure 6.16 shows the full reaction coordinate used to simulate the population distribution as a function of temperature and internal energy. For simplicity, letters are used to represent the various conformers which isomerize to one another. A = *Cis* acid proline pucker up, B = *Cis* acid proline pucker down, C = *Trans* acid proline pucker down, D = *Trans* acid proline pucker up, E = *Trans* acid proline pucker up/tyrosine ring rotation. Prior to all simulations the equilibrium distribution was calculated for the 5 species as a function of temperature/internal energy. Each isomerization reaction considered can be grouped as a slow/fast reaction. The pucker isomerization, as well as the tyrosine ring rotation are both considered fast reactions as their isomerization rates are orders of magnitude faster than the cooling rate ($10^{10}$ $s^{-1}$ *vs.* $10^{3}$ $s^{-1}$) and essentially remain so up until a few tens of wavenumbers above their respective barrier heights. On the other hand, the *cis-trans* isomerization reactions have initial isomerization rate constants at $\langle E_v \rangle$ = 6700 $cm^{-1}$ of k(E) ~$10^{6}$ $s^{-1}$, and slow significantly prior to reaching their respective barrier heights. Thus, the reaction coordinate was broken down into a series of fast and slow steps as. To begin the simulation a set amount of temperature/internal energy was removed and the two slow steps were adjusted first. When only considering the two slow steps, they are independent of one

another and are treated as two separate reactions of opposing equilibria (A↔D and B↔C). To simulate the population change within a given time period the following equation for opposing equilibria was used:

$$[A(t)] = A_e + x_e \exp[-(k_1 + k_{-1})t] \quad \text{(S1)}$$

$$\text{where } x_e = A_0 - A_e$$

$A_e$ is the equilibrium population for species A at the given temperature/internal energy, $x_e$ is the difference between the starting concentration of A, ($A_0$), and its final concentration $A_e$, $k_1$ and $k_{-1}$ are the forward and backward rates at the specified temperature/internal energy, and (t) is the time period between temperature/internal energy down steps. As temperature/internal energy is removed, the fully equilibrated population distribution among the 5 species changes. Using eqn. S1 it is possible to account for the shift in population between A and D as well as B and C, under kinetic constraints, representing the *cis-trans* isomerization. The time allotted for these reactions are determined by the average time for the temperature of the ions to drop by 1 K (1 μs) or to lose 13 $cm^{-1}$ (1.9 μs) of internal energy. After the distribution between A /D and B/C are adjusted the fast steps are considered. In considering only the fast steps the isomerization between A and B is independent of the isomerization between C, D, and E. Since the rates associated with the fast isomerization processes are much greater than the cooling rate, it is assumed that the species connected by these fast steps will maintain an equilibrated distribution throughout the cooling process. Thus, in this model, A/B maintain a near-equilibrium distribution throughout the cooling process while A/D and B/C do not. In order to make this adjustment the ratio of the new population between A and B (resulting populations after the slow step) is compared to the ratio of A to B from the fully equilibrated distribution. The population distribution between A and B is then adjusted such that new populations are in equilibrium. The same process is done for the *trans* species (C, D and E). Once this is done another down step in temperature/internal energy is taken and the process is repeated. In carrying out the simulation in this manner the change in the population distribution between the *cis* and *trans* species can be effectively model under kinetic constraints. A further constraint on the simulation is that the sum of the populations must equal one before the next down step is taken.

### 6.6.2 Details of Cooling Models

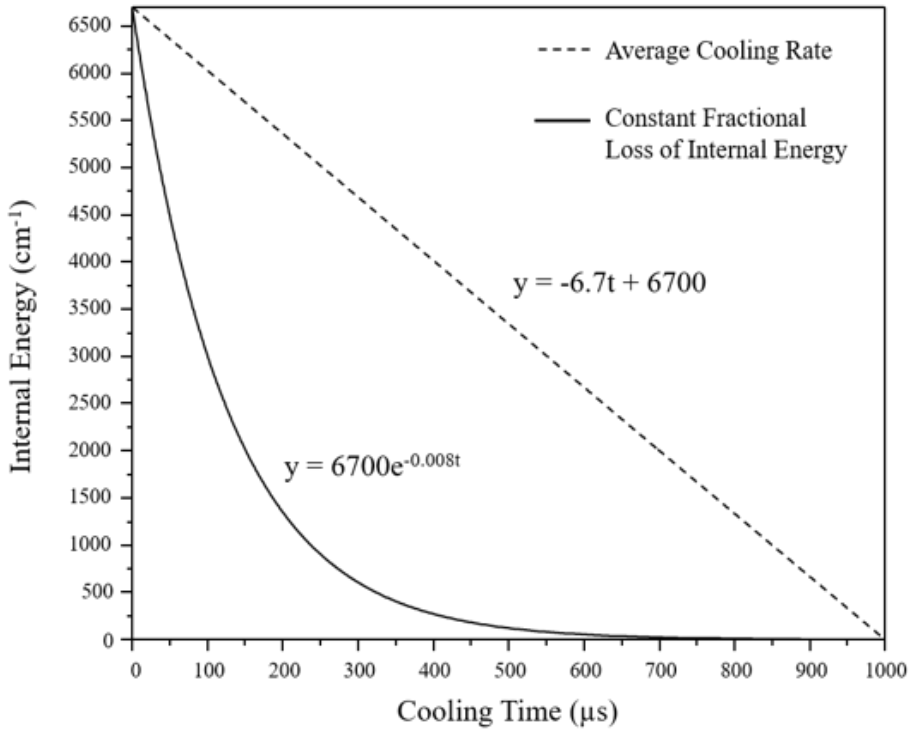


Figure 6.17: Comparison of different cooling models employed to simulate population distributions vs. internal energy

A comparison of the two different cooling models employed to carry out the cooling simulations. Using the average cooling rate, it is assumed that the ion loses internal energy at a constant rate ($\Delta E/\Delta t$ = 6.7 $cm^{-1}/\mu s$) as a function of total internal energy (dashed line). Alternatively, vibrational cooling is often modeled as a constant fraction of total internal energy ($\Delta E/E$ = constant). Figure 6.6 in the text demonstrates that ions entering the cold trap cool to a vibrational temperature of 10 K in 1 ms. Assuming that ions enter the trap at a vibrational temperature equivalent to room temperature, it is possible to fit a cooling rate of $\Delta E/E$ = -0.008, leading to an exponential decay in internal energy with time (solid line). Simulations run with either energy loss function resulted in ending population distributions nearly identical to one another (Table 6.1). This insensitivity to cooling model is due to a slow change with temperature of the family populations. The proline pucker down structures shown in Figure 6.7 contain a majority of the total population (~60%) at room temperature. These structures remain relatively close in terms of free energy as a function of internal energy, thus the total *cis* to *trans*

isomerization along this coordinate is rather small. While the proline pucker up structures have a large difference in energy between them, the higher member of this pair contains only 6% of the total population at room temperature. Thus, even though there is a larger driving force to isomerize from *cis* to *trans* along the pucker coordinate, there is very little population to be transferred. Consequently, the different rates at which energy is removed at early times has a small effect on the overall final distribution.

Table 6.1: Final simulated population distribution for each model

| | **Population Distribution** | |
|---|---|---|
| **Cooling Model** | *Cis* | *Trans* |
| Average Cooling Rate | 38% | 62% |
| Constant Fractional Loss | 36% | 64% |

Table 6.2: Distribution of low frequency modes in the 5 minima relevant for cooling simulations

| | Minima | | | | |
|---|---|---|---|---|---|
| Frequency Range ($cm^{-1}$) | Cis Pucker up | Trans Pucker up | Trans Pucker down | Cis Pucker up | Trans Tyr. Rotation |
| 0-100 | 11 | 10 | 10 | 11 | 10 |
| 100-200 | 10 | 11 | 11 | 10 | 11 |
| 200-300 | 9 | 10 | 10 | 9 | 9 |
| 300-400 | 10 | 8 | 8 | 10 | 10 |
| 400-500 | 7 | 8 | 8 | 7 | 8 |

Table 6.3: Distribution of low frequency modes in the 5 transition states that connect the minima used for cooling simulations

| | Transitions States | | | | |
|---|---|---|---|---|---|
| Frequency Range | Cis Pucker Isomerization | Trans Pucker Isomerization | Cis-Trans Isomerization Pucker up | Cis-Trans isomerization Pucker Down | Trans Tyr. Rotation |
| 0-100 | 10 | 10 | 11 | 11 | 9 |
| 100-200 | 11 | 11 | 9 | 9 | 11 |
| 200-300 | 8 | 9 | 11 | 11 | 10 |
| 300-400 | 10 | 8 | 9 | 9 | 9 |
| 400-500 | 7 | 8 | 7 | 7 | 8 |

## 6.7 References


(1) Chichinin, A. I.; Gericke, K.-H.; Kauczok, S.; Maul, C., Imaging chemical reactions - 3D velocity mapping. *Int. Rev. Phys. Chem.* **2009,** *28*, 607-680.

(2) Neumark, D. M., Slow electron velocity-map imaging of negative ions: applications to spectroscopy and dynamics. *J. Phys. Chem. A.* **2008,** *112*, 13287-13301.

(3) Neuhauser, D.; Judson, R. S.; Kouri, D. J.; Adelman, D. E.; Shafer, N. E.; Kliner, D.; Zare, R. N., State-to-state rates for the D+ H2 (v = 1, j = 1) -> HD(v'j') + H Reaction: Predictions and Measurements. *Science* **1992,** *257*, 519-522.

(4) Bethlem, H. L.; Berden, G.; Meijer, G., Decelerating Neutral Dipolar Molecules. *Phys. Rev. Lett.* **1999,** *83*, 1558-1561.

(5) Prozument, K.; Park, G. B.; Shaver, R. G.; Vasiliou, A. K.; Oldham, J. M.; David, D. E.; Muenter, J. S.; Stanton, J. F.; Suits, A. G.; *et. al.*, Chirped-pulse millimeter-wave spectroscopy for dynamics and kinetics of pyrolysis reactions. *Phys. Chem. Chem. Phys* **2014**, *16*, 15739-15751.

(6) Kaiser, R. I., Experimental investigation on the formation of carbon-bearing molecules in the interstellar medium via neutral-neutral reactions. *Chem. Rev.* **2002,** *102*, 1326-1330.

(7) Balucani, N.; Asvany, O.; Chang, A.; Lin, S.; Lee, Y.; Kaiser, R.; Bettinger, H.; Schleyer, P. v. R.; Schaefer III, H., Crossed beam reaction of cyano radicals with hydrocarbon molecules. I. Chemical dynamics of cyanobenzene (C 6 H 5 CN; X 1 A 1) and perdeutero cyanobenzene (C 6 D 5 CN; X 1 A 1) formation from reaction of CN (X 2 Σ+) with benzene C 6 H 6 (X 1 A 1g), and d 6-benzene C 6 D 6 (X 1 A 1g). *J. Chem. Phys.* **1999,** *111*, 7457-7471.

(8) Zheng, J.; Kwak, K.; Xie, J.; Fayer, M. D., Ultrafast carbon-carbon single-bond rotational isomerization in room-temperature solution. *Science* **2006,** *313*, 1951-1955.

(9) Dian, B. C.; Brown, G. G.; Douglass, K. O.; Rees, F. S.; Johns, J. E.; Nair, P.; Suenram, R. D.; Pate, B. H., Conformational isomerization kinetics of pent-1-en-4-yne with 3,330 cm− 1 of internal energy measured by dynamic rotational spectroscopy. *Proc. Natl. Acad. Sci.* **2008,** *105*, 12696-12700.

(10) Dian, B. C.; Brown, G. G.; Douglass, K. O.; Pate, B. H., Measuring picosecond isomerization kinetics via broadband microwave spectroscopy. *Science* **2008,** *320*, 924-928.

(11) Pierson, N. A.; Clemmer, D. E., An IMS–IMS threshold method for semi-quantitative determination of activation barriers: Interconversion of proline cis↔ trans forms in triply protonated bradykinin. *Int. J. Mass Spectom.* **2015,** *377*, 646-654.

(12) Marks, D. S.; Hopf, T. A.; Sander, C., Protein structure prediction from sequence variation. *Nat. Biotechnol.* **2012,** *30*, 1072-1080.

(13) Tian, P.; Best, R. B., How Many Protein Sequences Fold to a Given Structure? A Coevolutionary Analysis. *Biophys. J.* **2017,** *113*, 1719-1730.

(14) 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.

(15) Evans, D. A.; Wales, D. J.; Dian, B. C.; Zwier, T. S., The dynamics of conformational isomerization in flexible biomolecules. II. Simulating isomerizations in a supersonic free jet with master equation dynamics. *J. Chem. Phys.* **2004,** *120*, 148-157.

(16) Lubman, D. M.; Jordan, R. M., Design for improved resolution in a time-of-flight mass spectrometer using a supersonic beam and laser ionization source. *Rev. Sci. Instrum.* **1985,** *56*, 373-376.

(17) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V., Highly resolved spectra of gas-phase gramicidin S: a benchmark for peptide structure calculations. *J. Am. Chem. Soc.* **2010,** *132*, 4040-4041.

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

(19) Boyarkin, O. V.; Mercier, S. R.; Kamariotis, A.; Rizzo, T. R., Electronic spectroscopy of cold, protonated tryptophan and tyrosine. *J. Am. Chem. Soc.* **2006,** *128*, 2816-2817.

(20) DeBlase, A. F.; Dziekonski, E. T.; Hopkins, J. R.; Burke, N. L.; Sheng, H.; Kenttämaa, H. I.; McLuckey, S. A.; Zwier, T. S., Alkali Cation Chelation in Cold β-O-4 Tetralignol Complexes. *J. Phys. Chem. A.* **2016,** *120*, 7152-7166.

(21) Burke, N. L.; DeBlase, A. F.; Redwine, J. G.; Hopkins, J. R.; McLuckey, S. A.; Zwier, T. S., Gas-Phase Folding of a Prototypical Protonated Pentapeptide: Spectroscopic Evidence for Formation of a Charge-Stabilized β-Hairpin. *J. Am. Chem. Soc.* **2016,** *138*, 2849-2857.

(22) Inokuchi, Y.; Kaneko, M.; Honda, T.; Nakashima, S.; Ebata, T.; Rizzo, T. R., UV and IR Spectroscopy of Cryogenically Cooled, Lanthanide-Containing Ions in the Gas Phase. *Inorg. Chem.* **2016,** *56*, 277-281.

(23) Roy, T. K.; Kopysov, V.; Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V.; Gerber, R. B., Conformational Structures of a Decapeptide Validated by First Principles Calculations and Cold Ion Spectroscopy. *ChemPhysChem* **2015,** *16*, 1374-1378.

(24) Stearns, J. A.; Seaiby, C.; Boyarkin, O. V.; Rizzo, T. R., Spectroscopy and conformational preferences of gas-phase helices. *Phys. Chem. Chem. Phys.* **2009,** *11*, 125-132.

(25) Seaiby, C.; Zabuga, A. V.; Svendsen, A.; Rizzo, T. R., IR-induced conformational isomerization of a helical peptide in a cold ion trap. *J. Chem. Phys.* **2016,** *144*, 014304.

(26) DeBlase, A. F.; Harrilal, C. P.; Lawler, J. T.; Burke, N. L.; McLuckey, S. A.; Zwier, T. S., Conformation-Specific Infrared and Ultraviolet Spectroscopy of Cold [YAPAA+ H]+ and [YGPAA+ H]+ Ions: A Stereochemical "Twist" on the β-Hairpin Turn. *J. Am. Chem. Soc.* **2017,** *139*, 5481-5493.

(27) Redwine, J. G.; Davis, Z. A.; Burke, N. L.; Oglesbee, R. A.; McLuckey, S. A.; Zwier, T. S., A novel ion trap based tandem mass spectrometer for the spectroscopic study of cold gas phase polyatomic ions. *Int. J. Mass Spectom.* **2013,** *348*, 9-14.

(28) Han, H.; Londry, F. A.; Erickson, D. E.; McLuckey, S. A., Tailored-waveform collisional activation of peptide ion electron transfer survivor ions in cation transmission mode ion/ion reaction experiments. *Analyst* **2009,** *134*, 681-689.

(29) Londry, F.; Hager, J. W., Mass selective axial ion ejection from a linear quadrupole ion trap. *J. Am. Soc. Mass. Spectrom.* **2003,** *14*, 1130-1147.

(30) Mohamadi, F.; Richards, N. G.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C., MacroModel—an integrated software system for modeling organic and bioorganic molecules using molecular mechanics. *J. Comput. Chem.* **1990,** *11*, 440-467.

(31) Zhao, Y.; Schultz, N. E.; Truhlar, D. G., Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory Comput.* **2006,** *2*, 364-382.

(32) 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 09, Revision E.01. Gaussian, Inc., Wallingford CT, 2013.

(33) Lee, C.; Yang, W.; Parr, R. G., Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B.* **1988,** *37*, 785.

(34) BecNe, A., Densityüfunctional thermochemistry. III. The role of exact exchange. *J. Chem. Phys* **1993,** *98*, 5648-5652.

(35) Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *J. Comput. Chem.***2011,** *32*, 1456-1465.

(36) Voronina, L.; Rizzo, T. R., Spectroscopic studies of kinetically trapped conformations in the gas phase: the case of triply protonated bradykinin. *Phys. Chem. Chem. Phys.* **2015,** *17*, 25828-25836.

(37) Zabuga, A. V.; Kamrath, M. Z.; Boyarkin, O. V.; Rizzo, T. R., Fragmentation mechanism of UV-excited peptides in the gas phase. *J. Chem. Phys.* **2014,** *141*, 154309.

(38) Baer, T.; Hase, W. L., *Unimolecular reaction dynamics: theory and experiments*. Oxford University Press on Demand: 1996.

(39) Beyer, T.; Swinehart, D., Algorithm 448: number of multiply-restricted partitions. *Comm. ACM* **1973,** *16*, 379.

(40) Eyring, H., The activated complex in chemical reactions. *J. Chem. Phys.* **1935,** *3* (2), 107-115.

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

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

(43) Barker, J. R., Direct measurements of energy-transfer involving large molecules in the electronic ground state. *J. Chem. Phys.* **1984,** *88*, 11-18.

(44) Shi, L.; Holliday, A. E.; Bohrer, B. C.; Kim, D.; Servage, K. A.; Russell, D. H.; Clemmer, D. E., "Wet" Versus "Dry" Folding of Polyproline. *J. Am. Soc. Mass. Spectrom.* **2016,** *27*, 1037-1047.

# CHAPTER 7. 2-COLOR IRMPD ON CONFORMATIONALLY COMPLEX IONS: PROBING COLD ION STRUCTURE AND HOT ION UNFOLDING

## 7.1 Abstract

Two-Color infrared multiple photon dissociation (2C-IRMPD) spectroscopy is a technique that mitigates spectral distortions due to nonlinear absorption that is inherent to one-color IRMPD. We use a 2C-IRMPD scheme that incorporates two independently tunable IR sources, providing considerable control over the internal energy content and type of spectrum obtained by varying the trap temperature, laser wavelengths, which laser is scanned, their fluences and the pump-probe time delay. We demonstrate the application of this variant of 2C-IRMPD to conformationally-complex peptide ions. The 2C-IRMPD technique is used to record near linear action spectra of ions with temperatures ranging from 10-300 K, is applied to both cations and anions and single conformers in a conformational mixture. Furthermore, the technique has the capability to explore conformational unfolding by recording spectra while significantly varying the internal energy of the ion. We demonstrate these capabilities through studies of the protonated peptide ions YGGFL ($NH_3^+$-Tyr-Gly-Gly-Phe-Leu, Leu-enkephalin) and YGPAA ($NH_3^+$-Tyr-Gly-Pro-Ala-Ala).

## 7.2 Results and Discussion

Infrared Multiple Photon Dissociation (IRMPD),[1–5] messenger tagging,[6–10] and IR-UV double resonance (IR-UV DR)[11–14] are important action-based techniques used to record IR spectra of gas phase ions. The resulting vibrational spectra provide insights to the hydrogen-bonded networks that play key roles in determining the preferred conformations of biological ions. Each H-bond donor (e.g., N-H) or acceptor (e.g., C=O) group has the frequency of its vibrational fundamental shifted by the strength and orientation of the H-bond(s) in which it is involved, and the couplings between them. Thus, the resulting experimental spectrum reflects the three-dimensional structure of the ion. When coupled with theory, the experimental spectrum allows for explicit structural assignments to be made.[13] IRMPD is applicable to ions of all temperatures and places no restrictions on the types of ions that can be studied. However, the spectra that result lack conformer specificity and can have frequency shifts and intensity distortions associated with the multiple

photon absorption event.[2,3] On the other hand, messenger tagging and IR-UV double resonance spectroscopies both result in highly resolved, conformer specific, linear IR spectra with minimal thermal broadening.[13,15,16] Such high-quality spectra are obtainable because ions are cryo-cooled prior to being probed.[11]

2C-IRMPD, shown in Fig. 7.1 (b), is an alternative, but less widely employed technique introduced by Y.T. Lee and coworkers in the late 80's.[17] The basis of this technique is to divide the IRMPD process into two irradiation steps. The first laser, which is tunable, is set to a lower fluence and used to vibrationally excite ions while the second laser, set to high fluence conditions, induces dissociation. Early variations of the technique employed a continuous wave (cw), tunable laser for the initial excitation step, and a high-power, line tuned $CO_2$ laser to induce dissociation. Initial applications were to protonated water clusters[17] and other small molecular ions.[18,19] Variations of this technique are used currently where the light from a line tuned $CO_2$ laser is used to enhance the fragmentation signal of the ions excited by the tunable IR source. Studies that have employed this variant of the technique have been directed towards metal ion complexes[20,21] as well as the ions formed after peptide fragmentation.[22,23] A more recent 2C-IRMPD scheme, introduced by Niedner-Schatteburg and co-workers,[24] uses two OPO/OPA laser systems and provides additional versatility, since the second laser can also be scanned. To date, the dual OPO laser-based 2C-IRMPD technique has been applied to small non-covalent protonated water[25–27] and metalated ion complexes.[24,28]

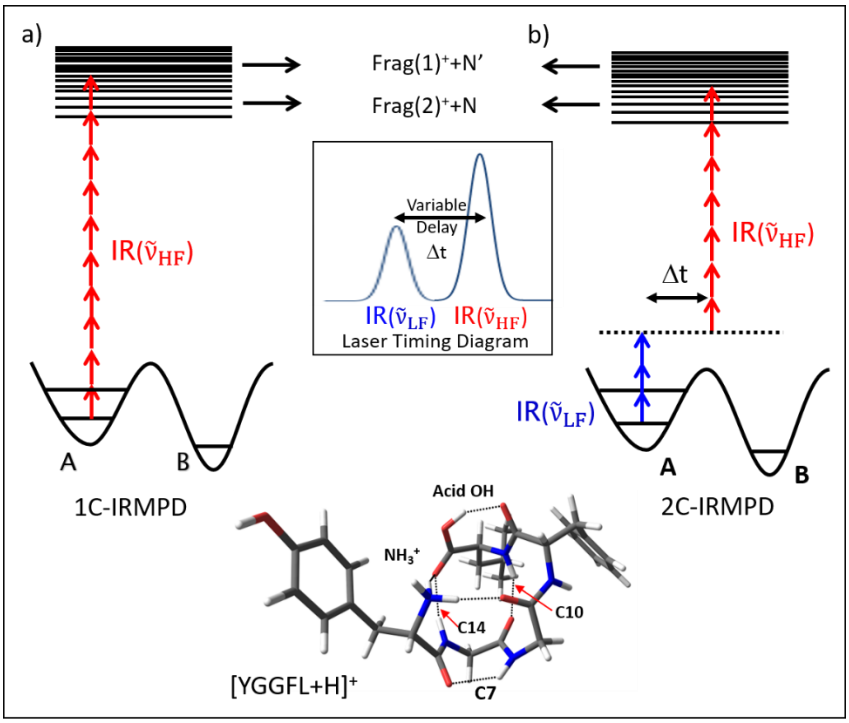


Figure 7.1: Schematic diagram of (a) one-color IRMPD, and (b) two-color IRMPD. The inset shows a timing diagram for 2C-IRMPD. Variation of which laser is scanned, the fluences of the two IR lasers, and the time delay between them can be used to obtain different kinds of spectra, as explained further in the text. The assigned structure of $[YGGFL+H]^+$ serves as an example of the conformationally complex ions to be studied.

In the present work, we build upon the dual OPO laser 2C-IRMPD technique and apply it to significantly larger, conformationally complex peptide ions. This study is performed on the custom-built apparatus for ion spectroscopy at Purdue[29], as described in more detail in the experimental section. We demonstrate several of the principal attributes of the method when applied to conformationally complex ions. Primary among them is the ability to record near linear spectra, which is crucial for structural characterizations[30] over a range of initial ion temperatures (10-300 K). The method is applied to ions of both polarities without the need for a UV chromophore. We evaluate conditions under which 2C-IRMPD can be used to record single-conformer IR spectra. Finally, and perhaps most interestingly, we apply 2C-IRMPD to study conformational 'unfolding' as the ion's internal energy is varied over wide ranges.

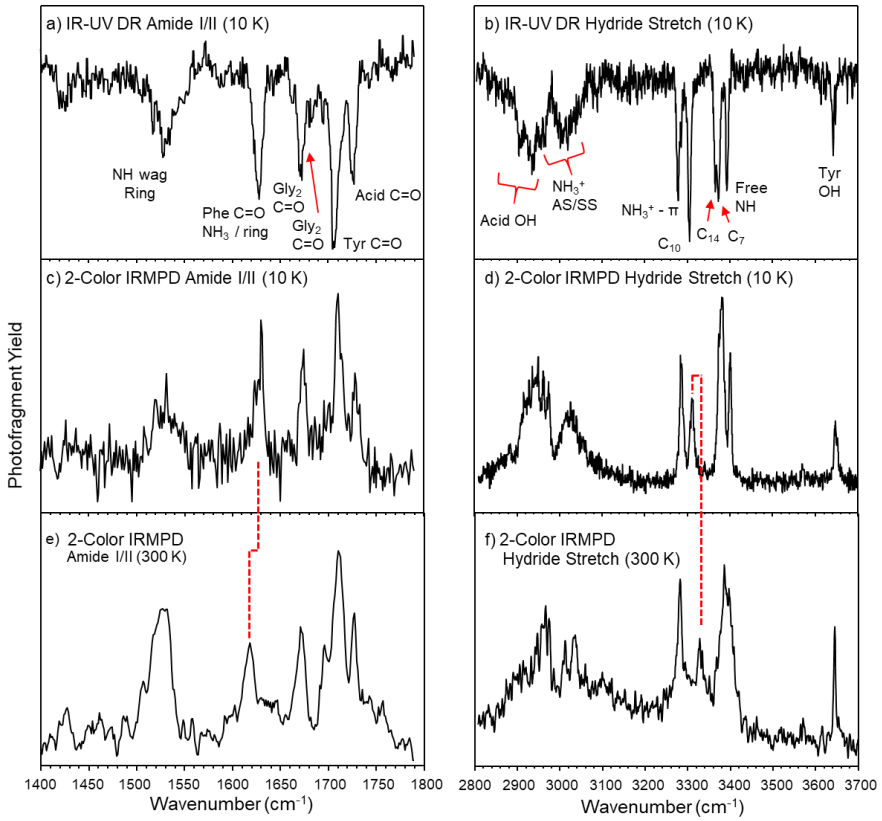


Figure 7.2: IR spectra of protonated YGGFL in the amide I/II and hydride stretch regions recorded with IR-UV DR (a) and (b), with 2-Color IRMPD at 10 K (c and d), and with 2-Color IRMPD at 300 K (e and f) with the laser scheme $[\tilde{\nu}_{LF}(\text{tuned})|\Delta t=30\text{ns}|\tilde{\nu}_{HF}(3420\text{cm}^{-1})]$

$[YGGFL+H]^+$ has previously been studied by our group using IR-UV DR. A single conformer was found to be present at 10 K and assigned by comparison with calculations to a charge stabilized β-turn, as shown in the inset to Fig. 7.1.[31,32] The spectra of the cryo-cooled ion in the amide I/II and hydride stretch regions taken with IR-UV DR are reproduced in Fig. 7.2 (a) and 2 (b), respectively. These spectra serve as references for comparison with the 2-color IRMPD technique as they are generated with single photon absorption events starting from 10 K ions.[33–35] The spectra in both regions show well-resolved IR transitions that can be assigned by comparison with theory to NH stretch and amide I/II fundamentals associated with particular NH/C=O groups. These modes and hydrogen bonds are labeled in the spectrum as 'Cn' where 'n' stands for the number of atoms in the H-bonded ring closed by the NH···O=C H-bond. The C14 and C7 NH stretch fundamentals are close in frequency and are barely resolvable in the hydride stretch region at 10 K. The relatively broad absorptions observed in the 2800-3100 $cm^{-1}$ region are associated with the

ammonium ($NH_3^+$) NH stretch and C-terminal carboxylic acid OH stretches. When these functional groups form strong hydrogen bonds, they couple to many lower frequency modes, broadening the transitions even at 10 K.[33,36]

Cryo-cooling ions to vibrational temperatures of 10 K aids in generating conformer specific linear action spectra when using IR-UV DR. Furthermore, the low temperatures provide well-resolved vibrational spectra. Figures 7.2 (c) and 2 (d) display 2C-IRMPD action spectra in the amide I/II and hydride stretch regions, respectively, with the ion trap held at 5 K ($T_{vib}$(ions) ~ 10 K).[11,37] In both regions the spectra closely resemble the linear action spectra displayed above them, demonstrating that the benefits of cryo-cooling are retained when using 2C-IRMPD, even for ions in this size regime. By comparison, these benefits could be lost if a traditional IRMPD approach is used, due to the nature of the multiple photon absorption process.[34] Furthermore, because the 2C-IRMPD technique is dependent only on IR absorption, it shares the ability to study anionic and non-aromatic systems with messenger tagging and IR+$CO_2$ 2-color schemes. The spectra of deprotonated YGGFL (recorded recently using messenger tagging[34]) and the hexapeptide GAIDDL, are shown in supplementary material (Fig. S1a,b).

Among the benefits of employing a second tunable IR laser is in making possible the acquisition of (i) near-linear IR spectra over a range of ion temperatures (10-300 K here), and (ii) single-conformer IR spectra in a conformational mixture. These are attractive characteristics for a broadly useful technique. [35,38] The ability to record near-linear spectra at 300 K is demonstrated by the spectra of $[YGGFL+H]^+$ displayed in Fig. 7.2 (e) and (f). In both the hydride stretch and amide I/II regions the spectra recorded at 300 K are quite similar to the corresponding spectrum recorded at 10 K, indicating that the conformational population continues to reside in the same single conformer of $[YGGFL+H]^+$ at 300 K. Not surprisingly, some of the transitions in the hydride stretch region broaden due to the presence of sequence bands out of vibrational levels that are now populated at 300 K. In particular, the set of transitions near 3400 $cm^{-1}$ (C14, C7, and free NH) are no longer resolved. The transitions associated with the H-bonded $NH_3^+$ NH stretch fundamentals near 3050 $cm^{-1}$ are also significantly broadened and support a high-frequency tail in the 3100-3250 $cm^{-1}$ region, demonstrating that the frequencies of these fundamentals are especially sensitive to internal energy content.

The spectra do show evidence for small structural changes within the conformational well that accompany warming to 300 K. In the hydride stretch region, the C10 NH stretch undergoes a 20

$cm^{-1}$ shift to higher frequency, signaling that the C10 NH···O=C H-bond that closes the β-turn becomes weaker at 300 K. Concomitantly, in the amide I/II region, the band at 1630 $cm^{-1}$, which is due to the Phe C=O, actually undergoes a 15 $cm^{-1}$ shift to lower frequency. Interestingly, this C=O group accepts a H-bond from the carboxylic acid OH, suggesting a modest strengthening of the OH···O=C H-bond as the β-turn weakens. Both these frequency shifts are indicated by red dashed lines in Fig. 7.2c↔e and Fig 7.2d↔f. Aside from these minor changes, it is clear that the same single conformation still dominates the ion population even at room temperature, consistent with recent studies[37,39,40] that show small changes in the conformational population during the collisional cooling process in the cold trap. As will be seen these minor frequency shifts are related to structural changes as a function of energy that can cause significant distortions in the IRMPD spectra and to some extent the 2C-IRMPD spectra. In summary, Fig. 7.2(e) and (f) show that near linear spectra can be recorded without the need for cryo-cooling, while at the same time avoiding significant non-linearity in the spectral intensities. The ability to record near linear spectra at 300 K is an attractive feature of 2C-IRMPD technique and combines some of the most attractive aspects of the cryogenic techniques and traditional IRMPD.

In order to record 2-color spectra like those in Fig. 7.2(e-f), it is first necessary to obtain a single laser IRMPD (1C-IRMPD) (Fig. 1(a)) spectrum, as detailed in the experimental section. By comparison to the 2C-IRMPD spectra in Fig. 7.2, the 1-color scan, shown in Fig. 3(a) and Fig S2 displays large distortions in transition intensities at both temperatures, with some transitions completely missing from the spectrum. As we will show, the distortions and missing bands in the 1C-IRMPD spectrum are due to isomerization/unfolding that takes place as the peptide ion approaches internal energies near the dissociation threshold.

Using the scheme, $[\tilde{\nu}_{\mathrm{LF}}(\mathrm{tuned})|\Delta t = +30\mathrm{ns}|\tilde{\nu}_{\mathrm{HF}}(3420\ \mathrm{cm}^{-1})]$ (see experimental section for details), shown schematically in Fig. 7.1(b), it is possible to record a 2C-IRMPD spectrum of $[YGGFL+H]^+$ free from distortions due to multiphoton absorption processes.[3] Due to the fast IVR rates in ions of this size, the resonances present in the vibrationally preheated ions are not expected to show mode specificity. In this manner, the high fluence laser becomes resonant after the ion internal energy is raised by absorption from the low fluence laser, enabling acquisition of an undistorted spectrum. This approach is similar to that taken by Johnson and co-workers in their studies of ion clusters.[25–27] The 2C-IRMPD spectra of $[YGGFL+H]^+$ shown in Figures 7.2 (c-f) were all recorded using this approach, with a 30 ns delay between $\tilde{\nu}_{\mathrm{LF}}$ and $\tilde{\nu}_{\mathrm{HF}}$.

The importance of recording linear spectra when making conformational assignments is highlighted by comparing the 1C-IRMPD spectra (Fig. 7.3 (a), Fig. S2 (a)) to their counterparts recorded using 2C-IRMPD (Fig. 7.2 c-f). As anticipated, large discrepancies in peak intensities and frequency shifts can affect 1C-IRMPD spectra when it is applied to structures containing H-bonded networks. Additionally, some transitions present in the 2C-IRMPD spectrum (Fig. 7.2 d) are completely absent in the 10 K 1C-spectrum (Fig. S2 a), most notably the $NH_3^+$-π and C10 NH stretch fundamentals (positions marked with black circles). At 300 K, Fig 7.3 (a), fewer IR photons are required to reach the energies at which fragmentation is facile, resulting in some of the bands appearing weakly. Even with the additional starting internal energy at 300 K the C10 NH stretch fundamental in still absent (Fig. 7.3 a). Recall that this fundamental was sensitive to internal energy, shifting ~20 $cm^{-1}$ between 10 and 300 K (Fig 7.2 d and f), and therefore is anticipated to broaden and shift away from its ground state resonance with further increase in ion internal energy. Similarly, the Phe C=O stretch that was shifted to a lower frequency at 300 K appears only weakly in a previously published 1C-IRMPD spectrum[41] indicating that structural rearrangements as a function of internal energy also govern the appearance of the IRMPD spectra. Such intensity distortions are anticipated to occur frequently in flexible ions that form many hydrogen bonds. By contrast, the spectra of smaller, rigid ions often agree well with those predicted by harmonic vibrational frequency calculations.[3,42–44] More generally, we postulate that, if the structure from which the ion eventually fragments is quite different than its starting, ground state structure, then bands that reflect those structural changes are likely to be weak or missing in the 1C-IRMPD spectrum. This is a similar scenario to that proposed by Yacovitch *et al.*[45]

One of the intriguing prospects made possible by independent control of the two IR sources is to obtain IR spectra of vibrationally hot ions well above the threshold energy for fragmentation, probing the spectroscopy of the unfolding ions. A straight-forward yet effective means for doing so is to simply switch the order of the high-fluence and low-fluence IR pulses, as shown in Figure 3(c) and described in the experimental section below. The scheme employed here is $[\tilde{\nu}_{HF}(3396\ cm^{-1})|\Delta t = 30ns|\tilde{\nu}_{LF}(tuned)]$.

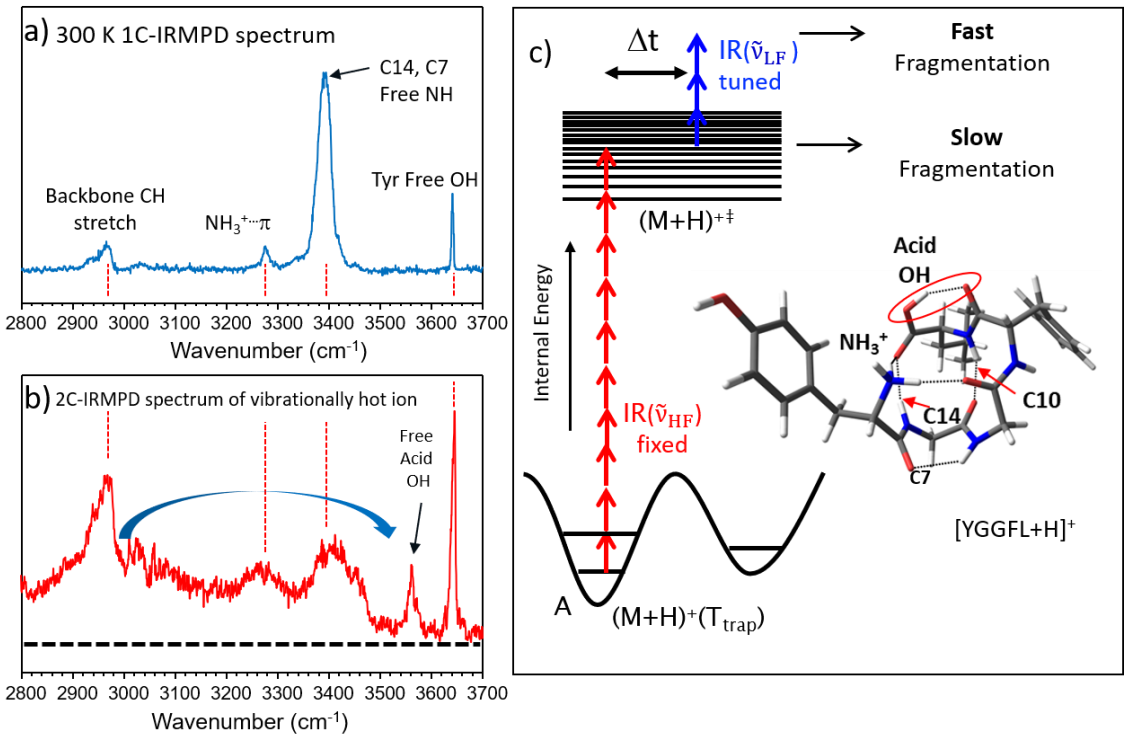


Figure 7.3: **(**a) 1C-IRMPD spectrum of $[YGGFL+H]^+$ at 300 K, for comparison with (b) 2C-IRMPD spectrum of the vibrationally hot ion formed by IR excitation at 3396 $cm^{-1}$ after a time delay of 30 ns. The dotted black line represents zero signal with respect to absorption from $\tilde{\nu}_{LF}$. The signal increase from $\tilde{\nu}_{LF}$ is about a factor of 2. (c) Schematic diagram of the experiment used to record the IR spectrum in 3(b).

The most striking feature in the 'hot ion' spectrum in Fig. 7.3 b is the new band that appears at 3550 $cm^{-1}$. Based on its frequency, this new band is assigned to the fundamental of a free carboxylic acid OH stretch.[46] The presence of this free OH stretch transition is especially interesting, as the starting structure shown in Fig. 7.3, is one in which the COOH group is *trans*, with the OH engaged in a strong C7 H-bond with the neighboring C=O group. The blue arrow in Fig. 7.3b denotes this shift in the frequency of the COOH stretch fundamental. The strong absorption peaked at 2950 $cm^{-1}$ that remains is due to the alkyl CH fundamentals that do not shift with internal energy.[23]

The calculated transition state for breaking this H-bond to form *cis* COOH is about 0.53 eV (4240 $cm^{-1}$) above the ground state (see Fig. S4 for details), while the lowest energy threshold for fragmentation is 1.14 eV.[47] Fragmentation becomes observable at ~5 eV internal energy, where its rate occurs on the millisecond time scale of our experiment. For reference, at energies where the

fragmentation rates are ~1 $s^{-1}$ (~3 eV), the OH isomerization rate is already on the order of ~4.0 x $10^8$ $s^{-1}$ (see Fig. S3 and associated discussion for more detail). Together, the lower energy threshold for isomerization and the experimental observation of the free acid OH stretch demonstrate that single-step isomerization rates are considerably faster than the ion fragmentation rate. Therefore, the hot ion spectrum in Fig. 7.3(b) is of a partially unfolded $[YGGFL+H]^+$ ion following IR excitation well above the threshold for breaking chemical bonds in the ion.

The spectrum that results is an interesting mixture of absorption types. Some of the transitions have a clear counterpart in the 300 K 1C-IRMPD experiment shown above it. The vertical red dashed lines between Fig. 7.3 (a and b) connect several transitions in the 1C-IRMPD scan ($NH_3^{+\cdots}\pi$, C14, C7, free NH, and Tyr OH fundamentals) with peaks that appear in broadened form in the vibrationally hot 2C-IRMPD spectrum. As we have argued, these absorptions are those that do not shift or broaden significantly with internal energy, and therefore more efficiently absorb IR photons in the first step. However, the intensities are far different in the hot ion spectrum than in the 300 K 1C spectrum, indicating that any enhancement due to a lack of frequency shift at these internal energies is quite small.

At the same time, the hot-ion spectrum has broadened absorptions stretching from 2800-3480 $cm^{-1}$. The high frequency edge at 3480 $cm^{-1}$ has a rather sharp cut-off that reflects the fact that free amide NH stretch transitions appear in the 3450-3480 $cm^{-1}$ region, with no means of shifting to higher wavenumber. The fact that the absorption intensity in this free amide NH region is small indicates that the vibrationally hot ion that we interrogate has not completely unfolded. Indeed, the absorptions that stretch down to 2800 $cm^{-1}$ indicate that some of the strongest H-bonds are still intact, at least to some degree. In the 10 K spectrum of Fig. 7.2(d), the low frequency bands below 3200 $cm^{-1}$ have contributions from the H-bonded COOH, the two $NH_3^{+\cdots}O{=}C$ H-bonds, in addition to alkyl CH stretches. It seems likely, then, that the broad absorption that fills in the 2800-3200 $cm^{-1}$ region is largely due to the $NH_3^+$ group, which sits atop the β-turn and binds to the C-terminal C=O group and the Gly(3) C=O at the center of the turn. The distribution of structures probed by the low-fluence IR laser have loosened, but not broken, these strong ion-carbonyl interactions.

One of the expectations for spectroscopic studies of large biomolecular ions in the gas phase is that many of these ions will exist in more than one conformation. In this sense $[YGGFL+H]^+$ is more the exception than the rule. Thus, conformer specificity is a highly attractive feature for any spectroscopic technique, including 2C-IRMPD. This is routine in IR-UV DR as the UV spectra of

the different conformations are generally resolved and thus have transitions that can serve as monitor transitions for specific conformations. 2C-IRMPD requires the presence of IR transitions unique to each conformation, much as in messenger tagging. Here we are able to achieve conformer specificity in a unique manner, which relies on the intermediate state associated with the IR-excited conformation being longer-lived than the time delay between the two lasers.

To demonstrate both the success and challenges of obtaining conformer specific spectra, we use the singly-protonated YGPAA pentapeptide, $[YGPAA+H]^+$ as an example. It is known to adopt two major conformations based on previous IR-UV DR studies by our group.[37,46] These two conformational families, with structures shown in Fig. S5, differ primarily in the *cis* vs. *trans* geometry of the carboxylic acid OH. These changes in structure also modulate the strengths of the hydrogen bonds involving this group, leading to conformation-specific shifts in the IR spectrum.[37,46] The conformer-specific IR-UV DR spectra for conformer B (*cis*) and A (*trans*) are reproduced in Fig. 7.4 (a and b), respectively.

The conformer specific spectrum of conformer A recorded using 2C-IRMPD is displayed in Fig. 7.4(c). To our knowledge, this is the first single-conformer spectrum obtained exclusively via an IRMPD approach. To record this spectrum, the high fluence laser ($\tilde{\nu}_{HF}$) was fixed ~5 $cm^{-1}$ above the C10 transition (3426 $cm^{-1}$) that is unique to conformer A and delayed by 5 ns from the tuned low-fluence laser ($\tilde{\nu}_{LF}$). The short time delay between the two laser pulses was chosen so that isomerization from B into A does not contribute significantly to the spectrum of the IR-excited ion. In Fig. S6(c) the time delay was increased to 30 ns, leading to a spectrum in which absorptions from B begin to appear, in addition to those from A. In Fig. 7.4(d), a 2C-IRMPD spectrum that is not conformer specific is displayed, obtained by fixing $\tilde{\nu}_{HF}$ just off resonance (3634 $cm^{-1}$) from an IR transition shared by the two conformers (3640 $cm^{-1}$).

Attempts to record a conformer specific spectrum of conformer B were unsuccessful largely due to the fact that the 1C-IRMPD spectrum of $[YAPAA+H]^+$ is missing any IR bands that are unique to conformer B (see supporting information for more detail). It is noteworthy that 1C-IRMPD would not even recognize the presence of two conformers in this sample. The intriguing differences between conformers A and B that enable conformer-specific infrared spectra of A but not B, will stimulate future studies aimed at understanding the conformer-specific dynamics that precedes fragmentation.

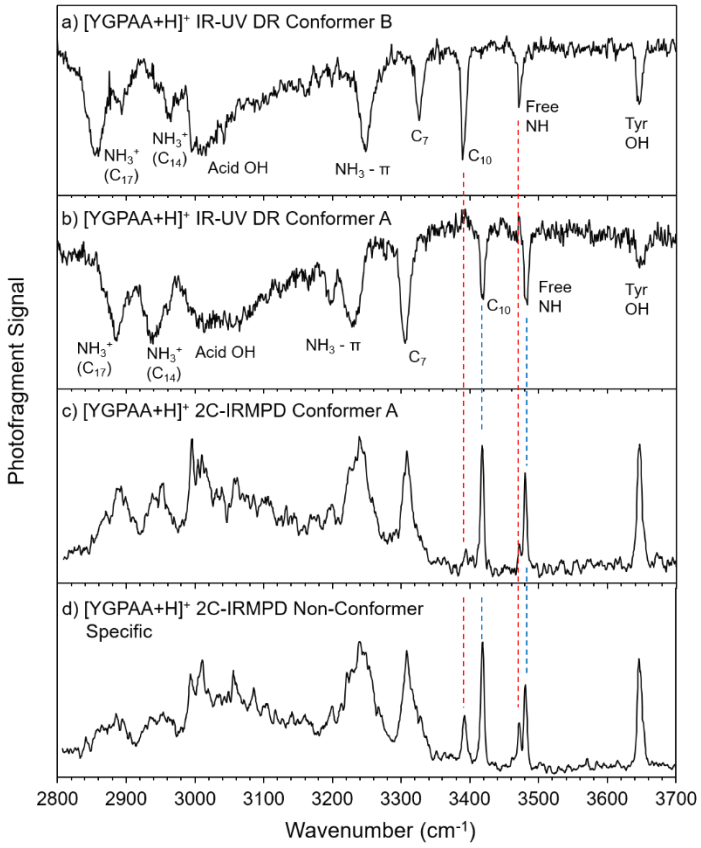


Figure 7.4: Conformer specific IR spectra recorded using IR-UV DR of conformer B and A of $[YGPAA+H]^+$ (a and b). Conformer specific spectrum of conformer A ($[\tilde{\nu}_{LF}(tuned)|\Delta t = +5ns|\tilde{\nu}_{HF}(3426\ cm^{-1})]$) and composite spectrum ($[\tilde{\nu}_{LF}(tuned)|\Delta t = +30ns|\tilde{\nu}_{HF}(3634\ cm^{-1})]$) recorded using 2C-IRMPD (c and d), respective laser conditions are given in parentheses.

The present work has established the versatile capabilities of the dual OPO variant of 2C-IRMPD as a technique for studying the infrared spectroscopy of large, conformationally-complex ions. We have demonstrated that linear IR spectra can be obtained at cryogenic temperatures, with all its attendant benefits of sharp spectra that can be used to make 3D structural assignments. Under favorable conditions, it is possible to record conformer specific spectra. Since the spectra can be generated entirely with IR light, ions of both polarities that do not possess a UV chromophore can be studied without the need for a tag. Furthermore, the ability to record undistorted IR spectra at variable trap temperatures enables structural changes as a function of temperature to be probed.

Finally, an intriguing aspect of the 2C-IRMPD method is its' potential for dynamical studies in which the fluence of the two IR sources and the pump-probe time delay can be varied. In the present work, we have shown the extreme example of a spectrum of $[YGGFL+H]^+$ following IR

excitation with the high-fluence laser, where the ion internal energy is well above the threshold for fragmentation. This spectrum shows evidence for substantial unfolding on the tens of nanoseconds timescale, pointing the way for future studies in which a range of peptide ion sizes, internal energies and time delays are explored to obtain snapshots of the ion as it unfolds in response to increases in internal energy. The tunability of both lasers also provides the potential to perform these experiments in a conformer selective manner.

### 7.3 Experimental Methods

All spectra were recorded on a custom-built apparatus that is comprised of a tandem triple quadrupole mass spectrometer on one axis with an orthogonal spectroscopy axis mounted between the second and third quadrupole. The methods for recording UV and IR-UV DR spectra are described elsewhere.[32] Briefly, ions are generated via nano ESI and guided into the second quadrupole (Q2). The ion of interest is mass isolated via RF DC isolation and then guided via a turning quadrupole down the spectroscopy axis where the ions are trapped in a cryogenically held octupole ion trap. The trap is held at 5 K via a close cycle helium cryostat. Ions are cooled to ~10 K via collisions with the He buffer gas before spectroscopic interrogation by the IR (Laservision OPO/OPA) lasers. The IR-induced photofragments are extracted back down the spectroscopy axis and turned into the third quadrupole (Q3). The residual precursor ions are ejected from the trap via a supplemental auxiliary waveform calculated with the SX wave software[48] and the remaining photofragments are extracted onto a channeltron detector.

The dual OPO 2C-IRMPD scheme allows for a variety of spectra to be recorded based on which laser is tuned and their relative timings. 1C-IRMPD spectra are obtained by recording the resulting photofragment signal as a function of the wavelength $\tilde{\nu}_{HF}$ recorded at 300 K (Fig. 7.3a) and at 10 K (Fig. S2). The high-fluence laser beam (~30 mJ/pulse) is focused to the center of the octupole trap using a 50 cm focal lens. The vibrational bands that appear in 1C-IRMPD guides the choice of fixed laser frequency in the following 2C-IRMPD schemes. In all the 2C-IRMPD schemes the low-fluence laser ($\tilde{\nu}_{LF}$) is counter propagated to $\tilde{\nu}_{HF}$ and re-sized to ~7 mm dia. as it passes through the ion packet ( ~15 mJ/pulse).

To generate the near linear 2-color spectra displayed in Fig. 7.2 c-f, $\tilde{\nu}_{HF}$ is fixed at 3420 $cm^{-1}$, just off resonance of the free NH fundamental and timed to irradiate the ion packet 30 ns after $\tilde{\nu}_{LF}$, which is scanned across the spectral region. This is given the short-hand notation

$[\tilde{\nu}_{LF}(tuned)|\Delta t = +30ns|\tilde{\nu}_{HF}(3420\ cm^{-1})]$, where Δt indicates the time delay of the second listed laser to that of the first. As $\tilde{\nu}_{LF}$ comes into resonance with a vibrational transition the vibrational bands present in the 1C-IRMPD spectra broaden and create a resonance for $\tilde{\nu}_{HF}$ which induces photofragmentation. The 2C-IRMPD spectra in Fig. 7.2 c-f kept all laser conditions unchanged with changing trap temperature to minimize non-linear effects. Corresponding conditions for $[YGPAA+H]^+$ are given in caption of Figure 7.4.

By switching which laser is tuned $[\tilde{\nu}_{LF}(3396\ cm^{-1})|\Delta t = 30ns|\tilde{\nu}_{HF}(tuned)]$, it is possible to demonstrate the degree of spectral broadening after the peptide ion has absorbed a few photons (< 3) from $\tilde{\nu}_{LF}$. These spectra are referred to as $\tilde{\nu}_{LF}$ pre-heated spectra and are displayed with red traces in Fig S2 a and b, 10 K and 300 K, respectively.

IR spectra of vibrationally hot ions, Fig. 7.3b, is obtained by switching the order of the two lasers. Here the $\tilde{\nu}_{HF}$ is fixed on resonance with a vibrational band present in the 1C-IRMPD spectrum and $\tilde{\nu}_{LF}$ is scanned at 30 ns delay, $[\tilde{\nu}_{HF}(3396m^{-1})|\Delta t = 30ns|\tilde{\nu}_{LF}(tuned)]$. By fixing $\tilde{\nu}_{HF}$ on resonance, 3396 $cm^{-1}$, a fraction of the ion population is promoted above the dissociation threshold, producing a steady-state photofragment signal (the black dashed line in Fig. 7.3 (b)). At wavenumber positions where the hot ion absorbs, the additional internal energy enhances the fragmentation rate, leading to fragmentation that successfully competes with collisional cooling in the trap. The result is an IR spectrum of the vibrationally hot ions produced by $\tilde{\nu}_{HF}$.

## 7.4 Supporting Information

### 7.4.1 2C-IRMPD Spectra of Anions

In order to demonstrate the versatility of the 2C-IRMPD approach to IR spectroscopy of conformationally complex ions, we have recorded infrared spectra of anions in addition to the protonated ions already discussed in the main text. Figure S1a,b presents the 10 K 2C-IRMPD spectra of deprotonated YGGFL and GAIDDL anions, respectively. The spectrum of the deprotonated YGGFL anion recently recorded by Schinle *et al.*[1] with messenger tagging is nearly identical to that in Figure S1a, demonstrating that 2C-IRMPD can be used to obtain linear spectra comparable to those from the tagging approach. The spectrum of the deprotonated hexapeptide, $[GAIDDL-H]^-$, provides a demonstration of a near linear IR spectrum of a peptide without an aromatic chromophore.

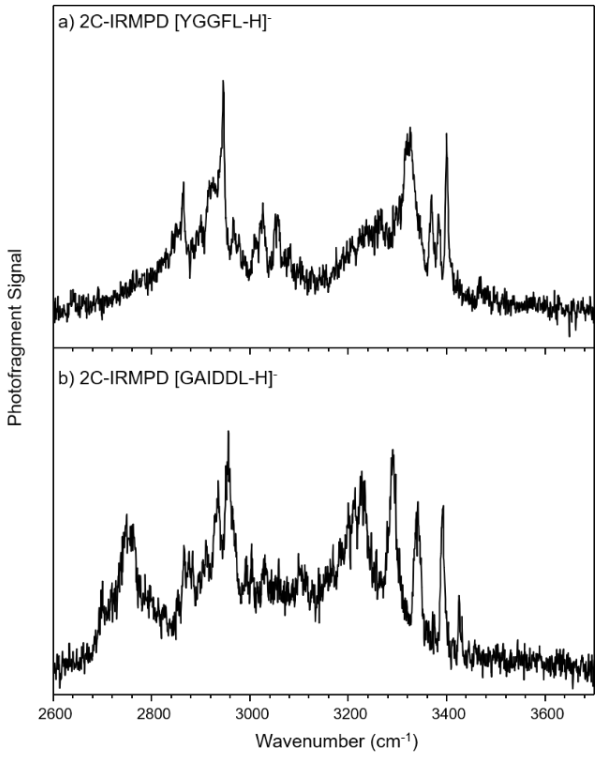


Figure 7.5: 2C-IRMPD spectra of 10 K deprotonated $[YGGFL-H]^-$ and $[GAIDDL-H]^-$ (a) and (b), respectively.

### 7.4.2 Comparison of 1C-IRMPD Spectra of [YGGFL+H]$^+$ with 2C-IRMPD, , $\tilde{\nu}_{LF}$ Pre-heated Spectra

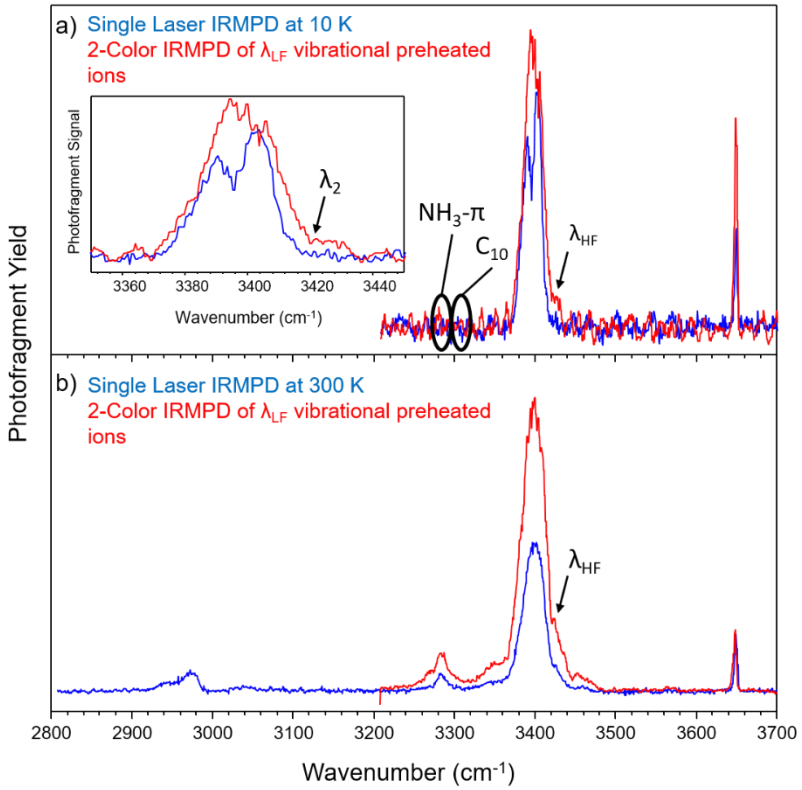


Figure 7.6: 1C-IRMPD scan recorded at 10 or 300 K (blue traces) (a) and (b), respectively. 2-Color spectrum ($[\tilde{\nu}_{LF}(3396\ cm^{-1})|\Delta t = 30ns|\tilde{\nu}_{HF}(tuned)]$) of ions that have been initially preheated with $\tilde{\nu}_{LF}$ (red trace) starting at 10 and 300 K (a) and (b). $\tilde{\nu}_{HF}$ indicates the position of the high fluence laser when recording a linear 2-Color spectrum.

### 7.4.3 RRKM Calculations of k(E) for the *trans-cis* COOH Isomerization in $[YGGFL+H]^+$

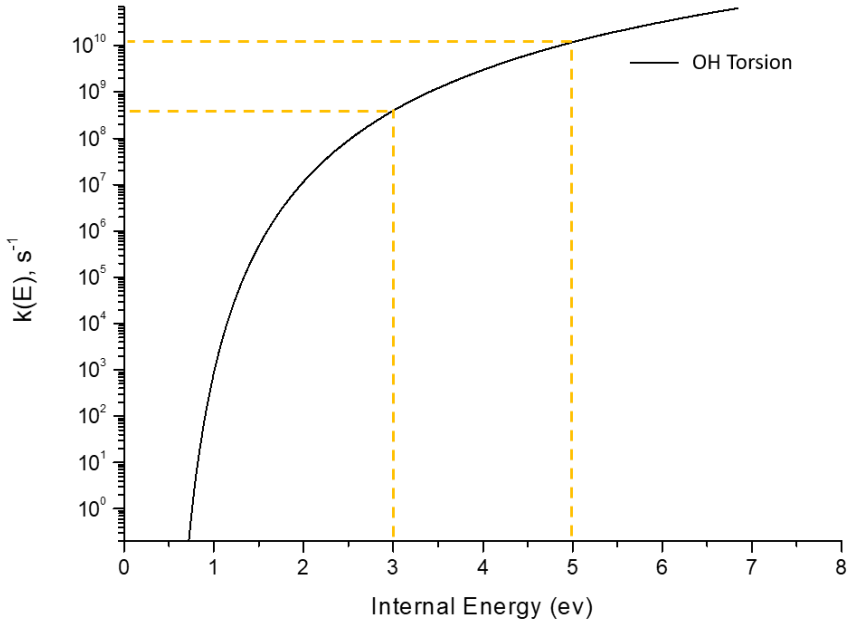


Figure 7.7: Rates of OH trans-cis isomerization (moving from bound to free position) as a function of internal energy.

At internal energies (3 eV) where the fragmentation rate is about 1 $s^{-1}$ the single step isomerization reaction is already 8 orders of magnitude faster. At internal energies (5 ev) where the fragmentation rate is on the millisecond time scale the single step isomerization rate is 1.0 x $10^{10}$ $s^{-1}$. In a single color IRMPD experiment a sufficient amount of energy must be deposited into the ion at a single IR wavelength to induce dissociation. Hydrogen bonded modes that are sensitive to the internal energy content will shift their frequencies and/or undergo isomerization prior to reaching the energies where fragmentation can occur. In the peptide ions studied here, this is significantly less than the amount of energy necessary to induce dissociation.

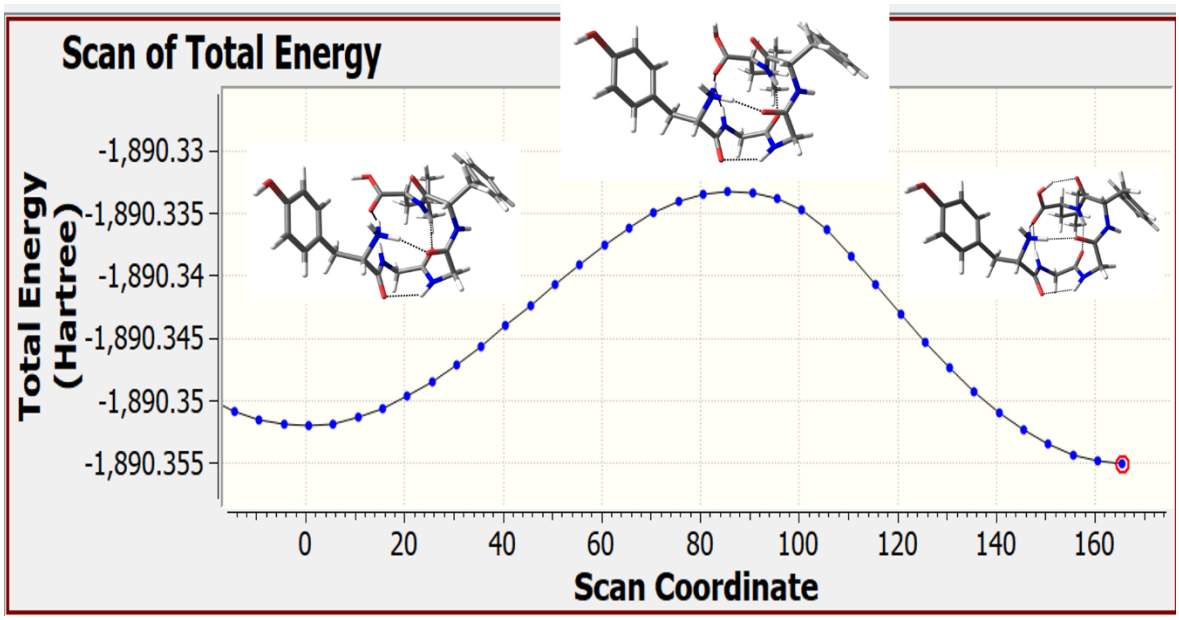


Figure 7.8: Relaxed potential energy scan of the carboxylic acid OH dihedral angle.

The maximum and minima structures were used as inputs for QST3 calculations at the B3LYP/6-31+G* level of theory with Grimme's empirical dispersion (GD3BJ). A tight convergence criterion and ultrafine grid were used for optimizations. The product structure (cis acid OH) was individually optimized at the same level of theory. The relative zero-point corrected energies between the minimum ground state and transition state structures were used to calculate the barrier for the trans to cis isomerization. The relative zero-point corrected energies for each structure are listed in Table S1.

Table 7.1: Relative Energies of Structures Involved in Carboxylic Acid Isomerization

| Structure | Relative ΔE (kJ/mol) | Relative ΔE ($cm^{-1}$) |
|---|---|---|
| $[YGGFL+H]^+$ (*trans* acid) | 0.00 | 0.00 |
| $[YGGFL+H]^+$ (transition state) | 50.73 | 4240 |
| $[YGGFL+H]^+$ (*cis* acid) | 7.29 | 610 |

### 7.4.4 Known Structures for Conformers A and B of $[YGPAA+H]^+$

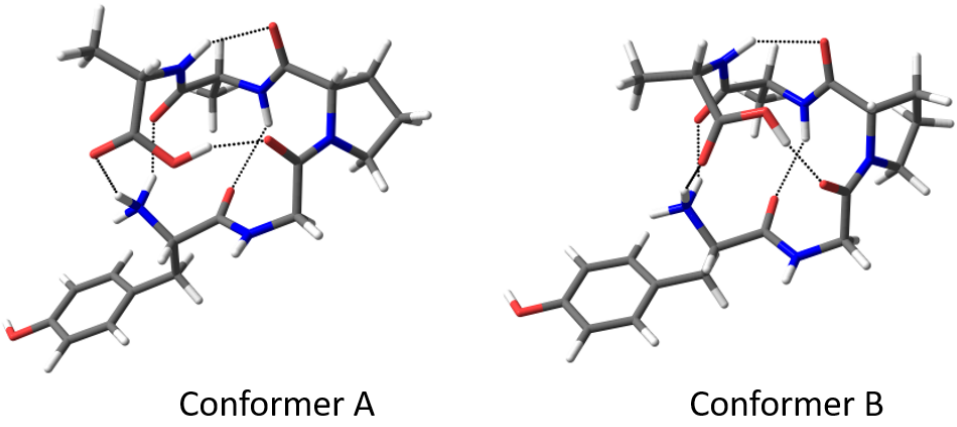


Figure 7.9: Conformers A and B of $[YGPAA+H]^+$ taken from Ref. 29 in main text. Both structures are beta turns mainly differing in the geometry of the COOH group.

### 7.4.5 Understanding the Inability to Record a Conformation-specific IR Spectrum of Conformer B of [YGPAA+H]$^+$

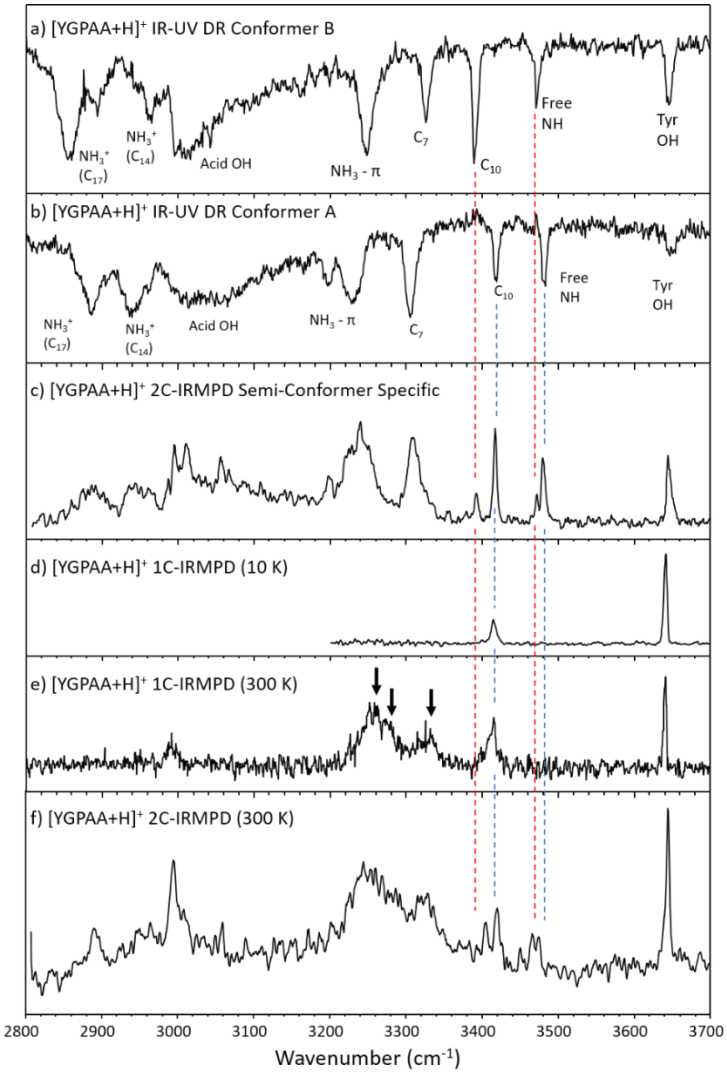


Figure 7.10: Conformer specific IR-UV DR spectra of [YGPAA+H]$^+$ conformers B and A, (a and b), respectively. 2C-IRMPD semi-conformer specific spectra of conformer A ($[\tilde{\nu}_{LF}(tuned)|\Delta t = +30ns|\tilde{\nu}_{HF}(3426\ cm^{-1})]$) (c), 1C-IRMPD spectrum taken at 10 K (d) and 300 K (e). 2-C IRMPD spectrum at 300 K ($[\tilde{\nu}_{LF}(tuned)|\Delta t = +30ns|\tilde{\nu}_{HF}(3634\ cm^{-1})]$) (f).

As discussed in the main text the 1C-IRMPD spectra of flexible peptide ions in this size regime are composed of transitions whose frequencies do not change significantly with ion internal energy, so that many photons can be absorbed at a single IR wavelength. For these bands, we have a well-defined prescription for fixing the wavelength of $\tilde{\nu}_{HF}$ such that only the conformer responsible for the absorption will absorb further photons and dissociate. At 10 K the 1C-IRMPD

spectrum, Figure S6 (d), contains only 2 bands, the C10 NH stretch that is unique to conformer A and the tyrosine OH band that is shared by both conformers. The one-color spectrum recorded at 300 K is shown in Fig. S6 (e). At 300 K fewer photons are required to observe fragmentation, making possible the observation in the spectrum of other bands between 3200-3380 $cm^{-1}$ and near 3000 $cm^{-1}$. Attempts to record conformer specific spectra were made by fixing $\lambda_{HF}$ at the frequencies indicated by the 3 arrows in Fig. S6 (e) while the ions were cooled to 10 K. The resulting spectra from each location were representative of conformer A rather than the expected conformer B spectrum.

The inability to record a conformer specific spectrum of B implies that the structure from which fragmentation occurs is significantly different from the conformer B ground state structure. Furthermore, structures populated in route to dissociation starting from the ground state of B have several transitions in common with conformer A. In this sense the ability to track the unfolding process between the ground and fragmenting structure will provide a basis from which the relationship between the two structures can be established. The 2-color spectrum recorded at 300 K, Fig. S6 (f), reassures that both conformations are present at room temperature. The C10 band belonging to conformer B is shifted 15 $cm^{-1}$ higher in frequency from its position at 10 K, which accounts for its absence in the 1 color spectra at both temperatures, Fig. S6 (d) and (e). Furthermore, the Free NH's for both conformations are at a very similar wavenumber position at 300 and 10 K. Their absence in the 1 color spectra is indicative that these bands shift as the internal energy is raised beyond its room temperature value.

## 7.5 References


(1) Woodin, R. L.; Bomse, D. S.; Beauchamp, J. L. Multiphoton Dissociation of Molecules with Low Power Continuous Wave Infrared Laser Radiation. *J. Am. Chem. Soc.* **1978**, *100* (10), 3248–3250. https://doi.org/10.1021/ja00478a065.

(2) Polfer, N. C. Infrared Multiple Photon Dissociation Spectroscopy of Trapped Ions. *Chem. Soc. Rev.* **2011**, *40* (5), 2211–2221. https://doi.org/10.1039/C0CS00171F.

(3) Oomens, J.; Sartakov, B. G.; Meijer, G.; von Helden, G. Gas-Phase Infrared Multiple Photon Dissociation Spectroscopy of Mass-Selected Molecular Ions. *International Journal of Mass Spectrometry* **2006**, *254* (1), 1–19. https://doi.org/10.1016/j.ijms.2006.05.009.

(4) *Gas-Phase IR Spectroscopy and Structure of Biological Molecules*; Rijs, A. M., Oomens, J., Eds.; Topics in Current Chemistry; Springer International Publishing: Cham, 2015; Vol. 364. https://doi.org/10.1007/978-3-319-19204-8.

(5) Jašíková, L.; Roithová, J. Infrared Multiphoton Dissociation Spectroscopy with Free-Electron Lasers: On the Road from Small Molecules to Biomolecules. *Chemistry – A European Journal* **2018**, *24* (14), 3374–3390. https://doi.org/10.1002/chem.201705692.

(6) Okumura, M.; Yeh, L. I.; Lee, Y. T. Infrared Spectroscopy of the Cluster Ions $H^+_3 \cdot (H_2)_n$. *The Journal of Chemical Physics* **1988**, *88* (1), 79–91. https://doi.org/10.1063/1.454488.

(7) Okumura, M.; Yeh, L. I.; Lee, Y. T. The Vibrational Predissociation Spectroscopy of Hydrogen Cluster Ions. *The Journal of Chemical Physics* **1985**, *83* (7), 3705–3706. https://doi.org/10.1063/1.449127.

(8) Lisy, J. M. Infrared Studies of Ionic Clusters: The Influence of Yuan T. Lee. *J. Chem. Phys.* **2006**, *125* (13), 132302. https://doi.org/10.1063/1.2338317.

(9) Gorlova, O.; Colvin, S. M.; Brathwaite, A.; Menges, F. S.; Craig, S. M.; Miller, S. J.; Johnson, M. A. Identification and Partial Structural Characterization of Mass Isolated Valsartan and Its Metabolite with Messenger Tagging Vibrational Spectroscopy. *J. Am. Soc. Mass Spectrom.* **2017**, *28* (11), 2414–2422. https://doi.org/10.1007/s13361-017-1767-z.

(10) Khanal, N.; Masellis, C.; Kamrath, M. Z.; Clemmer, D. E.; Rizzo, T. R. Glycosaminoglycan Analysis by Cryogenic Messenger-Tagging IR Spectroscopy Combined with IMS-MS. *Anal. Chem.* **2017**, *89* (14), 7601–7606. https://doi.org/10.1021/acs.analchem.7b01467.

(11) Boyarkin, O. V. Cold Ion Spectroscopy for Structural Identifications of Biomolecules. *International Reviews in Physical Chemistry* **2018**, *37* (3–4), 559–606. https://doi.org/10.1080/0144235X.2018.1547453.

(12) Rizzo, T. R.; Stearns, J. A.; Boyarkin, O. V. Spectroscopic Studies of Cold, Gas-Phase Biomolecular Ions. *International Reviews in Physical Chemistry* **2009**, *28* (3), 481–515. https://doi.org/10.1080/01442350903069931.

(13) Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V. Highly Resolved Spectra of Gas-Phase Gramicidin S: A Benchmark for Peptide Structure Calculations. *J. Am. Chem. Soc.* **2010**, *132* (12), 4040–4041. https://doi.org/10.1021/ja910118j.

(14) Page, R. H.; Shen, Y. R.; Lee, Y. T. Local Modes of Benzene and Benzene Dimer, Studied by Infrared–Ultraviolet Double Resonance in a Supersonic Beam. *J. Chem. Phys.* **1988**, *88* (8), 4621–4636. https://doi.org/10.1063/1.453775.

(15) Voss, J. M.; Kregel, S. J.; Fischer, K. C.; Garand, E. IR-IR Conformation Specific Spectroscopy of Na+(Glucose) Adducts. *J. Am. Soc. Mass Spectrom.* **2018**, *29* (1), 42–50. https://doi.org/10.1007/s13361-017-1813-x.

(16) Leavitt, C. M.; Wolk, A. B.; Fournier, J. A.; Kamrath, M. Z.; Garand, E.; Van Stipdonk, M. J.; Johnson, M. A. Isomer-Specific IR–IR Double Resonance Spectroscopy of D2-Tagged Protonated Dipeptides Prepared in a Cryogenic Ion Trap. *J. Phys. Chem. Lett.* **2012**, *3* (9), 1099–1105. https://doi.org/10.1021/jz3003074.

(17) Yeh, L. I.; Okumura, M.; Myers, J. D.; Price, J. M.; Lee, Y. T. Vibrational Spectroscopy of the Hydrated Hydronium Cluster Ions $H_3O^+ \cdot (H_2O)_n$ ($n$ =1, 2, 3). *The Journal of Chemical Physics* **1989**, *91* (12), 7319–7330. https://doi.org/10.1063/1.457305.

(18) Watson, C. H.; Zimmerman, J. A.; Bruce, J. E.; Eyler, J. R. Resonance-Enhanced Two-Laser Infrared Multiple Photon Dissociation of Gaseous Ions. *The Journal of Physical Chemistry* **1991**, *95* (16), 6081–6086. https://doi.org/10.1021/j100169a010.

(19) Peiris, D. M.; Cheeseman, M. A.; Ramanathan, R.; Eyler, J. R. Infrared Multiple Photon Dissociation Spectra of Gaseous Ions. *The Journal of Physical Chemistry* **1993**, *97* (30), 7839–7843. https://doi.org/10.1021/j100132a009.

(20) Sinha, R. K.; Nicol, E.; Steinmetz, V.; Maître, P. Gas Phase Structure of Micro-Hydrated [Mn(ClO4)]+ and [Mn2(ClO4)3]+ Ions Probed by Infrared Spectroscopy. *Journal of the American Society for Mass Spectrometry* **2010**, *21* (5), 758–772. https://doi.org/10.1016/j.jasms.2010.02.014.

(21) Altinay, G.; Metz, R. B. Comparison of IRMPD, Ar-Tagging and IRLAPS for Vibrational Spectroscopy of Ag+(CH3OH). *International Journal of Mass Spectrometry* **2010**, *297* (1), 41–45. https://doi.org/10.1016/j.ijms.2010.05.016.

(22) Sinha, R. K.; Erlekam, U.; Bythell, B. J.; Paizs, B.; Maître, P. Diagnosing the Protonation Site of b 2 Peptide Fragment Ions Using IRMPD in the X–H (X = O, N, and C) Stretching Region. *Journal of The American Society for Mass Spectrometry* **2011**, *22* (9), 1645–1650. https://doi.org/10.1007/s13361-011-0173-1.

(23) Durand, S.; Rossa, M.; Hernandez, O.; Paizs, B.; Maître, P. IR Spectroscopy of B4 Fragment Ions of Protonated Pentapeptides in the X–H (X = C, N, O) Region. *J. Phys. Chem. A* **2013**, *117* (12), 2508–2516. https://doi.org/10.1021/jp400634t.

(24) Nosenko, Y.; Menges, F.; Riehn, C.; Niedner-Schatteburg, G. Investigation by Two-Color IR Dissociation Spectroscopy of Hoogsteen-Type Binding in a Metalated Nucleobase Pair Mimic. *Phys. Chem. Chem. Phys.* **2013**, *15* (21), 8171–8178. https://doi.org/10.1039/C3CP44283G.

(25) Duong, C. H.; Yang, N.; Kelleher, P. J.; Johnson, M. A.; DiRisio, R. J.; McCoy, A. B.; Yu, Q.; Bowman, J. M.; Henderson, B. V.; Jordan, K. D. Tag-Free and Isotopomer-Selective Vibrational Spectroscopy of the Cryogenically Cooled H9O4+ Cation with Two-Color, IR-IR Double-Resonance Photoexcitation: Isolating the Spectral Signature of a Single OH Group in the Hydronium Ion Core. *J Phys Chem A* **2018**, *122* (48), 9275–9284. https://doi.org/10.1021/acs.jpca.8b08507.

(26) Duong, C. H.; Gorlova, O.; Yang, N.; Kelleher, P. J.; Johnson, M. A.; McCoy, A. B.; Yu, Q.; Bowman, J. M. Disentangling the Complex Vibrational Spectrum of the Protonated Water Trimer, H+(H2O)3, with Two-Color IR-IR Photodissociation of the Bare Ion and Anharmonic VSCF/VCI Theory. *J. Phys. Chem. Lett.* **2017**, *8* (16), 3782–3789. https://doi.org/10.1021/acs.jpclett.7b01599.

(27) Yang, N.; Duong, C. H.; Kelleher, P. J.; Johnson, M. A.; McCoy, A. B. Isolation of Site-Specific Anharmonicities of Individual Water Molecules in the I−·(H2O)2 Complex Using Tag-Free, Isotopomer Selective IR-IR Double Resonance. *Chemical Physics Letters* **2017**, *690*, 159–171. https://doi.org/10.1016/j.cplett.2017.09.042.

(28) Lang, J.; Gaffga, M.; Menges, F.; Niedner-Schatteburg, G. Two-Color Delay Dependent IR Probing of Torsional Isomerization in a [AgL1L2]+ Complex. *Phys. Chem. Chem. Phys.* **2014**, *16* (33), 17417–17421. https://doi.org/10.1039/C4CP02045F.

(29) Redwine, J. G.; Davis, Z. A.; Burke, N. L.; Oglesbee, R. A.; McLuckey, S. A.; Zwier, T. S. A Novel Ion Trap Based Tandem Mass Spectrometer for the Spectroscopic Study of Cold Gas Phase Polyatomic Ions. *International Journal of Mass Spectrometry* **2013**, *348*, 9–14. https://doi.org/10.1016/j.ijms.2013.04.002.

(30) Schwarz, H.; Asmis, K. R. Identification of Active Sites and Structural Characterization of Reactive Ionic Intermediates by Cryogenic Ion Trap Vibrational Spectroscopy. *Chemistry – A European Journal* **2019**, *25* (9), 2112–2126. https://doi.org/10.1002/chem.201805836.

(31) Burke, N. L.; DeBlase, A. F.; Redwine, J. G.; Hopkins, J. R.; McLuckey, S. A.; Zwier, T. S. Gas-Phase Folding of a Prototypical Protonated Pentapeptide: Spectroscopic Evidence for Formation of a Charge-Stabilized β-Hairpin. *J. Am. Chem. Soc.* **2016**, *138* (8), 2849–2857. https://doi.org/10.1021/jacs.6b00093.

(32) Burke, N. L.; Redwine, J. G.; Dean, J. C.; McLuckey, S. A.; Zwier, T. S. UV and IR Spectroscopy of Cold Protonated Leucine Enkephalin. *International Journal of Mass Spectrometry* **2015**, *378*, 196–205. https://doi.org/10.1016/j.ijms.2014.08.012.

(33) Leavitt, C. M.; DeBlase, A. F.; Johnson, C. J.; van Stipdonk, M.; McCoy, A. B.; Johnson, M. A. Hiding in Plain Sight: Unmasking the Diffuse Spectral Signatures of the Protonated N-Terminus in Isolated Dipeptides Cooled in a Cryogenic Ion Trap. *J. Phys. Chem. Lett.* **2013**, *4* (20), 3450–3457. https://doi.org/10.1021/jz401681y.

(34) Schinle, F.; Jacob, C. R.; Wolk, A. B.; Greisch, J.-F.; Vonderach, M.; Weis, P.; Hampe, O.; Johnson, M. A.; Kappes, M. M. Ion Mobility Spectrometry, Infrared Dissociation Spectroscopy, and Ab Initio Computations toward Structural Characterization of the Deprotonated Leucine-Enkephalin Peptide Anion in the Gas Phase. *J. Phys. Chem. A* **2014**, *118* (37), 8453–8463. https://doi.org/10.1021/jp501772d.

(35) Bouchet, A.; Klyne, J.; Ishiuchi, S.; Fujii, M.; Dopfer, O. Conformation of Protonated Glutamic Acid at Room and Cryogenic Temperatures. *Phys. Chem. Chem. Phys.* **2017**, *19* (17), 10767–10776. https://doi.org/10.1039/C6CP08553A.

(36) Craig, S. M.; Menges, F. S.; Duong, C. H.; Denton, J. K.; Madison, L. R.; McCoy, A. B.; Johnson, M. A. Hidden Role of Intermolecular Proton Transfer in the Anomalously Diffuse Vibrational Spectrum of a Trapped Hydronium Ion. *PNAS* **2017**, *114* (24), E4706–E4713. https://doi.org/10.1073/pnas.1705089114.

(37) 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* (8), 2096–2107. https://doi.org/10.1021/acs.jpca.7b12751.

(38) Wang, X.-B.; Wang, L.-S. Development of a Low-Temperature Photoelectron Spectroscopy Instrument Using an Electrospray Ion Source and a Cryogenically Controlled Ion Trap. *Review of Scientific Instruments* **2008**, *79* (7), 073108. https://doi.org/10.1063/1.2957610.

(39) Pereverzev, A. Y.; Kopysov, V.; Boyarkin, O. V. High Susceptibility of Histidine to Charge Solvation Revealed by Cold Ion Spectroscopy. *Angewandte Chemie International Edition* **2017**, *56* (49), 15639–15643. https://doi.org/10.1002/anie.201709437.

(40) Voronina, L.; R. Rizzo, T. Spectroscopic Studies of Kinetically Trapped Conformations in the Gas Phase: The Case of Triply Protonated Bradykinin. *Physical Chemistry Chemical Physics* **2015**, *17* (39), 25828–25836. https://doi.org/10.1039/C5CP01651G.

(41) Polfer, N. C.; Oomens, J.; Suhai, S.; Paizs, B. Infrared Spectroscopy and Theoretical Studies on Gas-Phase Protonated Leu-Enkephalin and Its Fragments: Direct Experimental Evidence for the Mobile Proton. *J. Am. Chem. Soc.* **2007**, *129* (18), 5887–5897. https://doi.org/10.1021/ja068014d.

(42) Yoon, S. H.; Chamot-Rooke, J.; Perkins, B. R.; Hilderbrand, A. E.; Poutsma, J. C.; Wysocki, V. H. IRMPD Spectroscopy Shows That AGG Forms an Oxazolone B2+ Ion. *J. Am. Chem. Soc.* **2008**, *130* (52), 17644–17645. https://doi.org/10.1021/ja8067929.

(43) Oomens, J.; Young, S.; Molesworth, S.; Stipdonk, M. Spectroscopic Evidence for an Oxazolone Structure of the B2 Fragment Ion from Protonated Tri-Alanine. *Journal of the American Society for Mass Spectrometry* **2009**, *20* (2), 334–339. https://doi.org/10.1016/j.jasms.2008.10.012.

(44) Oomens, J.; Roij, A. J. A. van; Meijer, G.; Helden, G. von. Gas-Phase Infrared Photodissociation Spectroscopy of Cationic Polyaromatic Hydrocarbons. *ApJ* **2000**, *542* (1), 404–410. https://doi.org/10.1086/309545.

(45) Yacovitch, T. I.; Heine, N.; Brieger, C.; Wende, T.; Hock, C.; Neumark, D. M.; Asmis, K. R. Vibrational Spectroscopy of Bisulfate/Sulfuric Acid/Water Clusters: Structure, Stability, and Infrared Multiple-Photon Dissociation Intensities. *J. Phys. Chem. A* **2013**, *117* (32), 7081–7090. https://doi.org/10.1021/jp400154v.

(46) DeBlase, A. F.; Harrilal, C. P.; Lawler, J. T.; Burke, N. L.; McLuckey, S. A.; Zwier, T. S. Conformation-Specific Infrared and Ultraviolet Spectroscopy of Cold [YAPAA+H]+ and [YGPAA+H]+ Ions: A Stereochemical "Twist" on the β-Hairpin Turn. *J. Am. Chem. Soc.* **2017**, *139* (15), 5481–5493. https://doi.org/10.1021/jacs.7b01315.

(47) Laskin, J. Energetics and Dynamics of Fragmentation of Protonated Leucine Enkephalin from Time- and Energy-Resolved Surface-Induced Dissociation Studies. *J. Phys. Chem. A* **2006**, *110* (27), 8554–8562. https://doi.org/10.1021/jp057229r.

(48) Han, H.; Londry, F. A.; Erickson, D. E.; McLuckey, S. A. Tailored-Waveform Collisional Activation of Peptide Ion Electron Transfer Survivor Ions in Cation Transmission Mode Ion/Ion Reaction Experiments. *Analyst* **2009**, *134* (4), 681–689. https://doi.org/10.1039/b821348h.

# VITA

Christopher Harrilal was born on February 18th, 1993 in Queens, NY to Annette Dipierro and Gewan Harrilal. He spent the next 10 years growing up in Far Rockaway, NY before moving to Edison, NJ. He spent the next 8 years in NJ and moved to Florida after graduating from J. P. Stevens High School. Chris spent the next 4 fours attending Florida International University (FIU) where he received his B.S in Chemistry. During this time, he was advised by Dr. Francisco Lima and worked on trapped ion mobility. When not in school Chris worked in a gym as a personal trainer. During the 4-year span Chris spent the majority of his time in the gym, at lab, or by the beach. After graduating FIU in 2015 he joined the PhD program at Purdue University and entered into a joint collaboration between Dr. McLuckey and Dr. Zwier performing ion spectroscopy. During this time, he gained a great deal of knowledge in the realm of mass spectrometry, laser spectroscopy, and chemical kinetics. He also met his wonderful girlfriend, Sydney Chow who took him around the country within the span of a few months. He received his PhD in 2019 and moved to Washington State as a post-doc at PNNL under Yehia Ibrahim and Dick Smith. This was a difficult decision but felt it was the best.