APPLICATION OF CRYOGENIC ULTRAVIOLET AND INFRARED SPECTROSCOPY FOR DETERMINATION OF SECONDARY STRUCTURE OF GAS PHASE PEPTIDE IONS

by

John T. Lawler

A Dissertation

Submitted to the Faculty of Purdue University

In Partial Fulfillment of the Requirements for the degree of

Doctor of Philosophy

Department of Chemistry

West Lafayette, Indiana

August 2021

THE PURDUE UNIVERSITY GRADUATE SCHOOL

STATEMENT OF COMMITTEE APPROVAL

Dr. Scott A. McLuckey, Chair

Department of Chemistry

Dr. Timothy S. Zwier

Department of Chemistry

Dr. Garth J. Simpson

Department of Chemistry

Dr. Jonathan Hood

Department of Chemistry

Approved by:

Dr. Christine Hrycyna

Dedicated to my parents, siblings, wife, and children. None of this would be possible without all of you.

ACKNOWLEDGMENTS

The sum of all one's experiences is what makes you who you are therefore, to write this section and name only a few names would feel disingenuous as the impact of any one person cannot be quantified. As I greatly value introspective people, I will instead offer you a thought experiment. If you can look back and say you have helped or aided in my experience, then thank you and I am greatly appreciative.

As the people who have helped me the most are unlikely to read this dissertation, I ask you the reader to please treat those you meet with respect. I have too often observed those with advanced educations treating those that do not as somewhat lesser or incapable and this can not be farther from the truth. The people who were always willing to talk, ask questions, and in general treated me with respect and dignity have been those with little to no formal education. I have the upmost respect for the folks in the Chem Shop, the guys in basement engineering, the janitors, and the ladies in the business/procurement offices.

In an increasingly fragmented society, the need to reach out and talk with others who do not share similar experiences or views as yourself has never been more important. Graduate students spend a lot of time studying the intricacies of the scientific method but rarely apply that same vigor to the outside world. The safety of the academic bubble can not be understated. I would behoove younger students to spend more time out in the community, volunteering and helping people who need it in ways that are meaningful and not merely for social media likes or the approval of peers.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Gas phase studies of the secondary structure of peptides and proteins have become increasingly popular as they offer distinct advantages of small sample usage and experiment time. The mass spectrometer is key to performing these experiments as ions can be manipulated based on their mass to charge ratio. Combining mass spectroscopy with laser spectroscopy birthed a new method for determining gas phase structures, ion spectroscopy. This document begins with an overview of secondary structure analysis using several techniques in solid, liquid, and gas phases. It then describes how ion spectroscopy can also be utilized to obtain detailed fingerprint infrared spectra of ions which are then matched with density functional theory calculations to determine the 3D structure of an ion. After describing the instrumental apparatus, four examples of the use of ion spectroscopy to determine structure are presented. The first study looked to understand the effect of increased flexibility around a proline residue in the diastereomeric pair $YA^{D/L}PGA$ and how a simple switch to glycine can greatly affect beta turn formation in peptides. The next three studies describe an attempt to form a single turn alpha helix in the gas phase using a highly stable tethered peptide motif. Failure to form the single turn helix in the first study led to an interesting examination of the use of computational chemistry to lead the synthesis of peptides where a specific structure is required. After observing the single turn helix attention is then diverted to expanding and controlling its handedness via stereochemistry. In all, this document will guide the reader through the methods and experiments possible with ion spectroscopy.

CHAPTER 1. INTRODUCTION TO STRUCTURAL ANALYSIS AND ION SPECTROSCOPY

1.1 Motivation

The three-dimensional shape of a biomolecule has an extraordinary effect on its function[1-2], so much so that slightly misshapen proteins can become non-functional and even poisonous[3-4] to the host. This structure function relationship is a well understood dogma in biology and effects all the different types of biomolecules from small peptides[5] and megadalton proteins[6] to oligonucleotides[7] and lipids[8]. Due to the importance of this dogma in living beings, studies of the structure and dynamics of biomolecules are of great importance. To this end, a multitude of analytical techniques have been devised to elucidate the geometry of biomolecules in the solid[9-10], liquid[11-12], and gas phases[13-14]. This introduction begins with a short review of a few of the popular techniques used to determine structure in the solid and liquid phases. It then moves on to describe the benefits and pitfalls of gas phase structural elucidation and the techniques used to perform the experiments described in the remainder of this document.

1.2 Structural Elucidation Techniques in the Solution and Solid Phases

An endless amount of acronym laden techniques have been devised to understand the structure of molecules. Typical techniques that one may encounter include, Nuclear Magnetic Resonance[15] (NMR), Infrared[16]/Raman[17] Spectroscopy, Circular Dichroism[18], Mass Spectrometry[19], X-Ray Crystallography[20], and Cryogenic Electron Microscopy[21] (Cryo-EM). While no single form of measurement can truly probe structure in the native environment of the molecule, all of these techniques have proved useful in determining structure in niche circumstances. Solid state approaches like X-Ray Crystallography[22] and ssNMR[23], while not probing native-like structure, have been used heavily by formulation scientists to examine the structures and packing of molecules in pharmaceutical applications. These techniques are not without their pitfalls. Both require a large amount of sample and extensive preparation before data can be collected and they also suffer from a limited ability to observe different conformations in the sample. The solution phase comes with the benefit of being closer to the native state of the biomolecule making measurements taken in solution highly coveted for

understanding how a molecule may be structured. The ubiquity of structural analysis with NMR[24] and Circular Dichroism[25] cannot be denied. Biochemists rely heavily on both techniques to understand not only structure but also the dynamics of how molecules bind to their targets as well. Unfortunately, solution phase techniques share similar pitfalls to those of the solid state. There is a great need for highly pure samples and in large quantities, experiments take large amounts of time, and any detail on the conformations present in the ensemble is lacking.

1.3 Benefits of Studying Cold, Gas Phase Ions

Gas phase approaches to the study of biomolecular structure are becoming increasingly popular as they address several of the issues associated with condensed phase techniques. In particular, mass spectrometry has risen as a powerful tool for structural elucidation as it can utilize smaller, less pure samples with a shorter analysis time than other approaches[26]. The ability to introduce protein molecules into the gas phase in a native-like state via control of solvents, additives, and electrospray parameters has been a boon for protein structural analysis[27]. What followed was a slew of activation techniques that have been developed to analyze these native-like peptides and proteins. Structural determination with ion mobility[28] (IM), ultraviolet photodissociation[29] (UVPD), hydrogen deuterium exchange[30] (HDX), collision induced dissociation[31] (CID), and electron capture/transfer dissociation[32] (ECD/ETD) have all been repeatedly shown to be reliable and quick methods for interrogating the 3D structure of a molecule. These techniques can be broken down into two categories those that dissociate and those that do not. Dissociation techniques (UVPD,CID,ECD,ETD,etc.) work in tandem to create complimentary structural information through destruction of the bonds holding the molecule together. While the remaining methods attempt to maintain the original geometry of the ion as much as possible throughout analysis in an attempt to measure either the body of the molecule (IM) or the accessibility of its components (HDX).

The ability to isolate and manipulate the structure of a biomolecule is a powerful advantage that mass spectrometry maintains over condensed phase approaches. A more subtle and often overlooked advantage is the ability to study a molecule free from interactions with solvent. In the solution phase the geometry of a molecule is dictated by its interactions with the solvent. Intermolecular hydrogen bonding between proteins and water creates a dynamic structure that changes with time, leading to several microstates describing an ensemble of different

structures[33]. This ensemble of structures induces a lot of complexity into spectroscopic determination of structure which limits resolution and utility. All solution phase techniques suffer from this issue to some degree. For example, the broadening of transitions in a solution phase infrared spectrum is in part due to intramolecular interactions. Gas phase spectroscopists use the lack of solvent to their advantage as it allows for the study of purely intermolecular interactions that effect the structure of the ion. This greatly reduces the complexity of the data taken and allows for probing of conformation of smaller biomolecules after cooling.

Gas phase spectroscopic techniques for identifying structure rely heavily on advances in mass spectrometry instrumentation. The ability to isolate and trap an ion of interest greatly enhances resolution of spectral data and increases confidence in results. Pairing mass spectrometry with spectroscopy for determining spectroscopy is not a new endeavor. Infrared multiphoton dissociation (IRMPD) spectroscopy has been used to explore the hydride stretch(XH) and amide I/II regions of several molecules with great success[34-35] when paired with density functional theory (DFT) calculations. The downside to most IRMPD techniques is that they are performed on ions that have high internal temperatures which creates band broadening of spectra, ultimately making structural assignment more difficult. Warm ions can access many different conformational states of the molecules causing spectral congestion. Combine this effect with a room temperature Boltzmann distribution of initial vibrational levels and obtaining narrow resolution infrared spectra becomes near impossible. Cooling ions down to ~5 K can remove much of that internal energy and make spectroscopic interrogation much simpler and easier to interpret. The reasons for this gain in resolution are twofold. Firstly, cooling to cryogenic temperatures (~5-10 K) puts the entire population of ions into the zero-point vibrational levels of the available conformations. Sampling fewer energy levels allows for the visualization of more crisp transitions between said energy levels, making transitions in the ultraviolet or infrared much sharper. Cooling also reduces the number of conformations observed as crashing each population down to its zero-point level means there is no energy available to hop between conformations as observed in warm ions. Studies of cold, gas phase ions offer many advantages compared to condensed phase studies. This method of preparing ions for study by cooling and converting them to the gas phase sets the field for structural determination with one of the most powerful techniques available, ion spectroscopy.

1.4 Ion Spectroscopy

Arguably the most powerful technique for gas phase structural elucidation is ion spectroscopy. The high level of resolution obtainable by this technique and the number of different confirming pieces of data that can be collected are two of its major advantages. As previously discussed, cold, phase molecules lend themselves to high resolution spectra as they are reduced in conformational complexity as well as energy[36]. Obtaining ultraviolet (UV) and conformation specific infrared (IR) spectra to be used as fingerprints for combination with quantum chemical calculations is the bread and butter of ion spectroscopy. IR and UV lasers have also been combined to probe the dynamics of small peptides[37] as well as the effect of hydration on the overall structure of a molecule[38]. Much of the work performed in the ion spectroscopy space has been performed on small peptide systems (<14 residues) as the speed of performing calculations needed to compare to spectra is limited. As quantum computing becomes more of a reality this wall of calculation time and complexity should crumble allowing for spectral assignment of larger molecules. While peptides form the preponderance of studies using ion spectroscopy, the technique has been used to study other biomolecules as well, including sugars[39], steroids[40], and pharmaceuticals[41].

Three major forms of ion spectroscopy have proliferated since its inception. The first, IRMPD[42], is commonly performed on warm ions with the consequences of such a measurement being discussed previously in the chapter. IRMPD benefits from being the quickest and easiest to perform technique of the three as it typically involves a single infrared laser focused into an ion trap and allowed to blast the ion apart. This simplicity in experimental setup gives rise to the least informative spectra of the three. A second form of ion spectroscopy is messenger tagging spectroscopy[43]. For messenger tagging ions are trapped in a quadrupole ion trap at temperatures between 10-40 K which allows for complex formation between the ion of interest and a neutral inert gas, typically helium. The weakness of this complex at higher temperatures is exploited through interrogation with an infrared laser. A typical messenger tagging experiment begins by recording the mass of the ion-tag complex. The complex is then irradiated with a tunable IR laser such that if the laser is resonant with a vibrational transition the energy imparted removes the tag and causes a depletion in the observed signal. Plotting these signal changes versus observed wavenumber of the laser produces an infrared spectrum of the molecule. The benefit to messenger tagging is much better resolution of spectra as the experiment is performed at colder

temperatures. Messenger tagging also allows for the study of a wide range of ions as long as the tag can be adsorbed onto the ion of interest. The last major form of ion spectroscopy, and the one utilized in the studies described in this document, is IR-UV double resonance spectroscopy. IR-UV methods for determining structure were first used to study neutral, gas phase molecules as far back as the 1980s[44]. While the application of this technique to ions is far younger. IR-UV double resonance spectroscopy works in a very similar way to messenger tagging spectroscopy. Instead of the gain/loss of a tag being as a probe for structural change instead a UV chromophore is employed that is irradiated with a UV laser. Absorption of a resonant UV photon imparts a large amount of internal energy into the molecule causing it to fragment. This fragmentation is monitored as a function of wavenumber to produce a UV spectrum. If an IR photon precedes the action of the UV photon an infrared spectrum can be obtained. More details on this technique will be provided below. The benefit to IR-UV techniques is the formation of conformation specific infrared spectra providing a greater level of detail than the messenger tagging technique that is typically not able to probe individual conformers. As one would expect this level of detail comes with the price of a more difficult to perform experiment. While this may be true, the abundance of structural information that can be gleamed from IR-UV techniques is not rivaled by its peers.

1.5 IR-UV Double Resonance Techniques

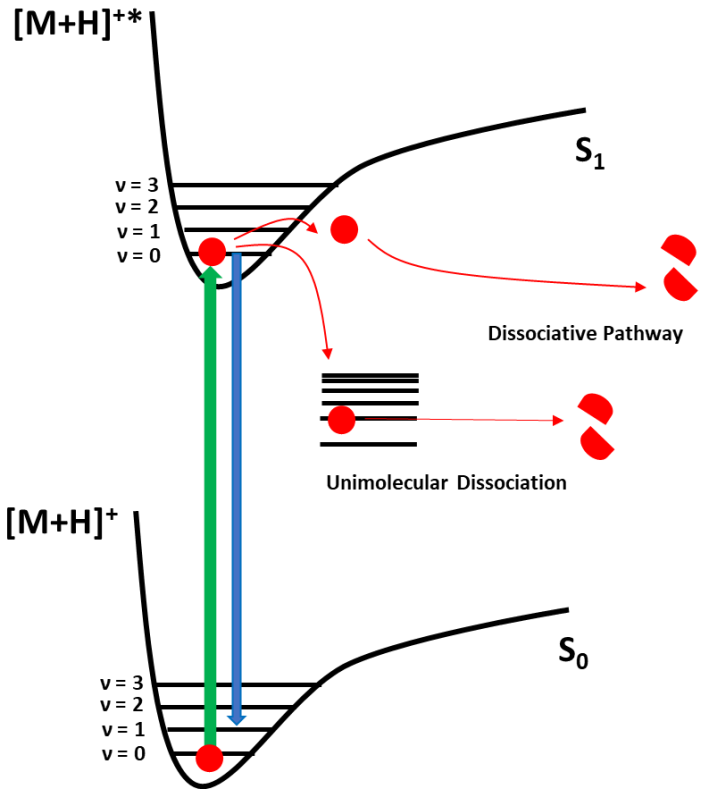

Figure 1. Schematic of the three pathways possible upon UV excitation

1.5.1 Electronic Spectra

As discussed in the previous section IR-UV double resonance spectroscopy relies on an action spectroscopy event, namely fragmentation induced by UV absorption, to monitor the activity of the ion of interest. This section will describe this event in more detail. Figure 1 displays a schematic of the UV excitation event and the consequences of the adsorption. The ion population begins in the v=0 level of the ground electronic state (S_0) due to cooling down to 10 K. As the UV laser is scanned if a photon is absorbed by the chromophore, it can be elevated to the first excited electronic state (S_1). Typical chromophores used for these experiments are tyrosine and phenylalanine which absorb near 280 and 270 nm respectively. Once an ion finds itself in the first excited electronic state there are three pathways in which it may fall. The first pathway is unobservable with UV photofragment techniques but is always happening at some level, namely fluorescence. Fluorescence can be used to probe the vibrational levels of a

molecule and has been widely used with neutral molecules[45] to do so, but as of yet has not been used with cold ions. The remaining two pathways are dissociative processes that occur if the energy imparted by the UV photon is large enough to surpass the dissociation threshold. Vibrational predissociation is the first of these processes and occurs due to internal conversion from the S_1 state to the highly excited vibrational levels of the ground electronic state. Once conversion happens the ion then fragments via a unimolecular dissociation pathway. As the lowest energy bonds dissociate preferably observed fragments are similar to those obtained via collision induced dissociation (CID) in a tandem mass spectrometry experiment. The remaining dissociation method requires a curve crossing from the first excited electronic state to a dissociative electronic potential energy surface. This mode of dissociation creates fragments that tend to be localized to the UV chromophore providing a greater level of detail about the environment in which it lies. Collecting a mass spectrum of these fragments produces a UV photodissociation spectrum which can also be analyzed for structural information. More importantly the electronic spectrum created gives insight into the complexity of the 3D structure of the ion.

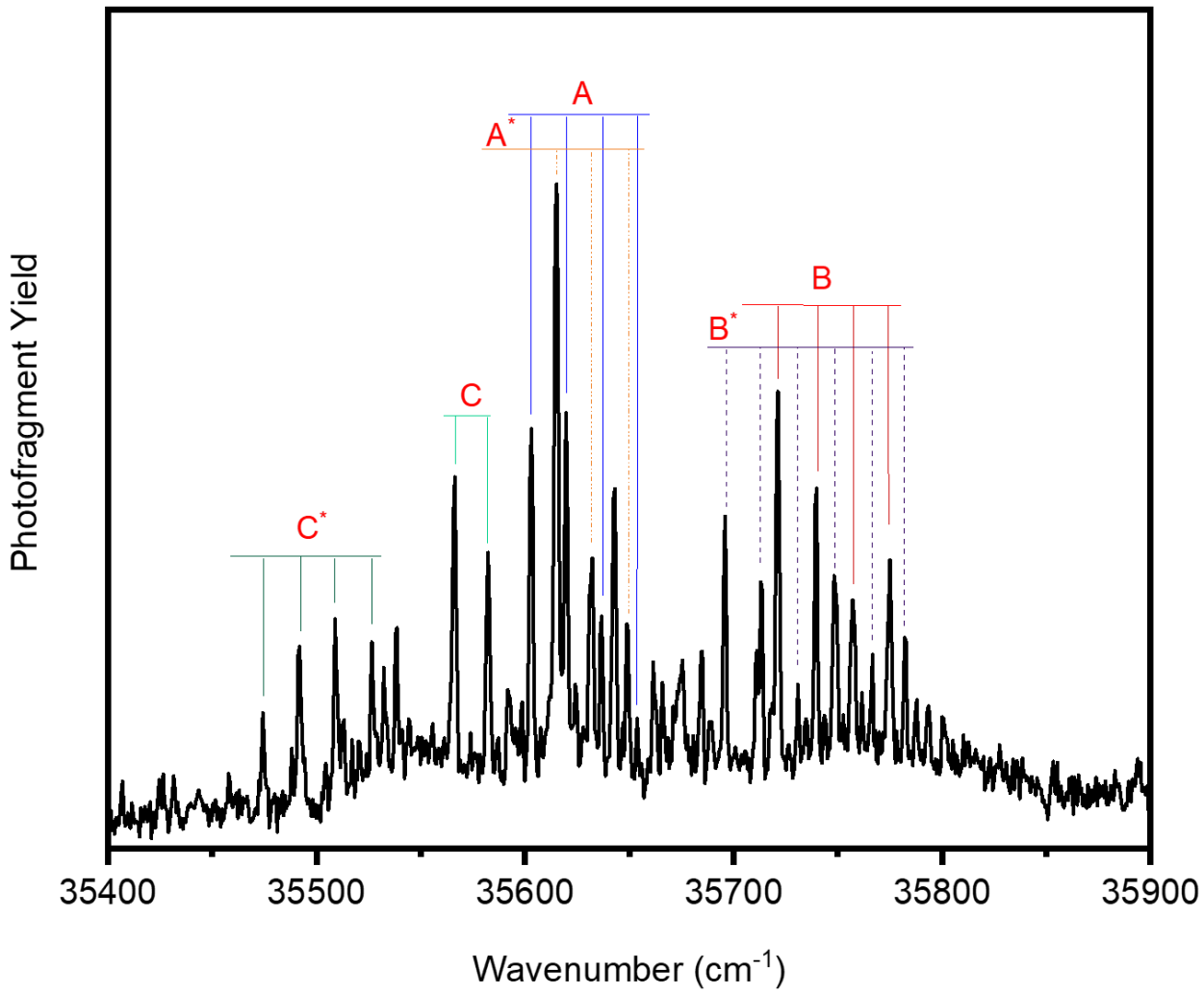

Figure 2. UV spectrum of the ion $[YA^{L}PGA+H]^{+}$

As can be observed in figure 2 several highly resolved transitions can be observed in a typical UV spectrum. For each conformation in an electronic spectrum there exists an electronic origin typically labelled as S_1^0 <- S_0^0 to show the flow of energy between the states without a change in vibrational level. In cases where vibrational level does change with electronic excitation vibronic bands occur. These are the origin of the several transitions that build off of the electronic origin in the UV spectrum and can give some detail as to the vibrational modes of the excited state. Varied intensities of these vibronic bands are observed in figure 1.2. The difference in intensity is determined by the Franck-Condon principle which states that a change from one vibrational level to another is more likely to happen if there is good overlap between the wave functions of the two modes. Therefore, intense Franck-Condon progressions observed

in the UV spectrum of a conformation imply a greater change in geometry upon excitation than when fewer less intense progressions are observed. This effect is readily observed in the example spectrum of YA^{L}PGA provided in figure 2. Conformer C has a short Franck-Condon progression which implies it changes structure less than conformer B which has a very long progression. While the electronic spectrum provides an incredible amount of detail on the structure of the molecule it also allows for probing specific conformations with the addition of an IR laser.

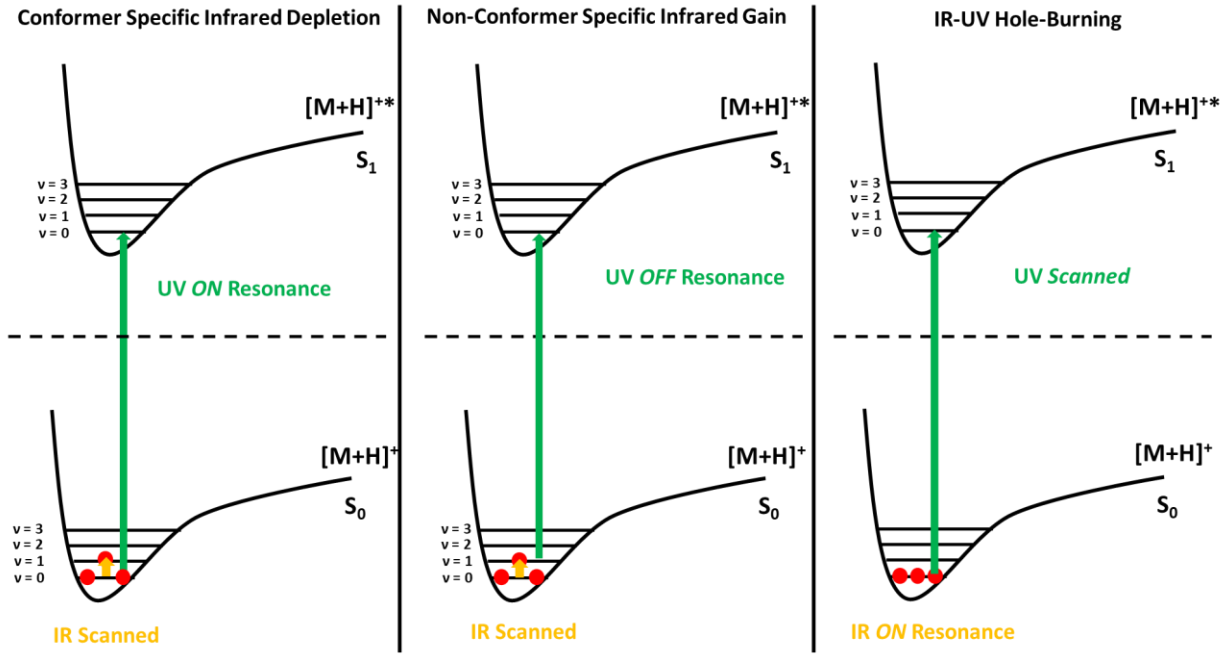

Figure 3. Schematic of different IR-UV double resonance techniques

1.5.2 Conformer Specific Infrared Spectra

Figure 3 displays a schematic of the three most used IR-UV double resonance techniques in the studies described later on in this thesis. The first and most powerful technique is the production of conformer specific infrared spectra. Obtaining these spectra begins by fixing the UV laser on an observed transition in the electronic spectrum which generates a steady stream of photofragments from a specific conformation. Overlapping an IR laser but temporally preceding the UV laser by 100-200 ns allows for the probing of the hydrogen bonding network of the conformer. If a photon from the tuned IR laser is absorbed by the molecule it depopulates the vibrational zero-point level resulting in a smaller observed photofragment signal. This effect can be observed in figure 4a which displays a conformer specific infrared spectrum of

$[YALPGA+H]^+$. The baseline photofragment signal generated by the UV laser is represented by the baseline of the spectrum while the depletions in signal are due to IR interaction with the ion. The produced conformer specific spectrum can then be analyzed in conjunction with DFT calculated spectrum to determine the structure of the molecule. The same method can then be used to record conformer specific infrared spectra of the remaining conformers in the UV spectrum simply by moving the UV laser to the next electronic origin.

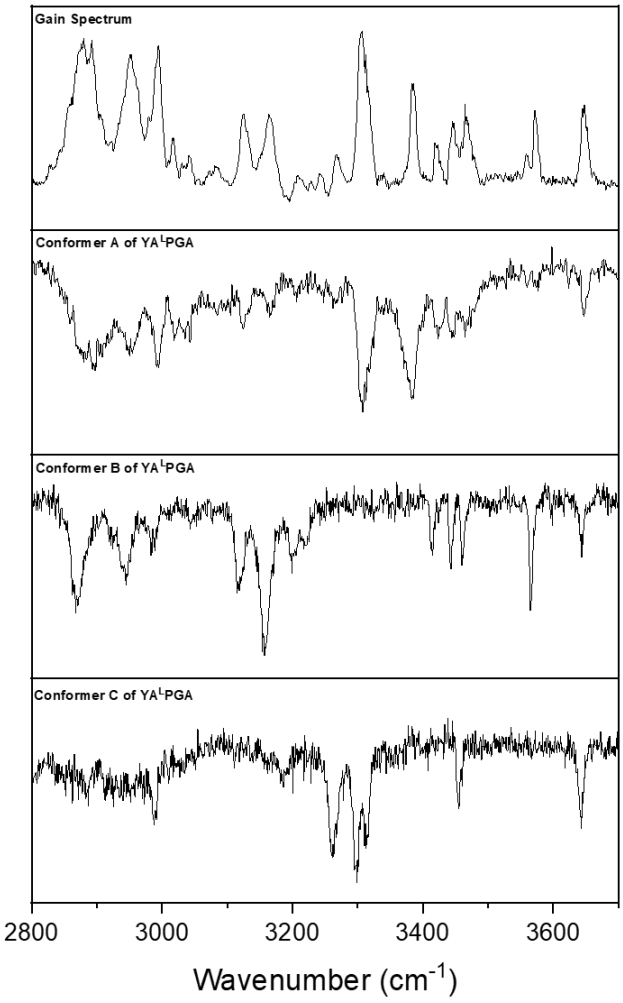

Figure 4. Non-conformer specific gain spectrum compared to conformation specific spectra of $[YA^LPGA+H]^+$

1.5.3 Conformer Non-Specific Gain Spectra

Often the electronic spectrum of an ion is quite complicated containing several transitions that may or may not be electronic origins of different conformers. It is not in the best interest of the experimenter to probe what could be tens of transitions in the electronic spectrum hoping to find all of the origins. Instead, this section and the next will detail two methods through which one can quickly and confidently be sure that all conformers have been studied. The less informative of the two methods requires the production of a non-conformer specific gain spectrum in the hydride stretch region. The process of obtaining this spectrum is in all ways similar to the conformer specific version of the spectrum with one subtle change. To produce a non-conformer specific gain spectrum the UV laser is not placed directly on a transition in the electronic spectrum. Rather, the laser is fixed at a wavenumber a little bit lower than the lowest energy origin in the UV spectrum. At this position only warm ions will be in resonance with the UV laser as the cold ions only produce the observed transitions. When the IR laser is now scanned and comes into resonance with a transition belonging to any of the conformations present photofragments will be produced. Monitoring these photofragments as a function of wavenumber generates the non-conformer specific spectrum. Comparison of the non-conformer specific spectrum to that of a previously recorded conformer specific spectrum can provide insight into whether or not any other conformations remain to be probed. In figure 4 the gain spectrum is compared to the conformer specific spectra. A glance at the two spectra quickly allows the experimenter to see that there are transitions in the gain spectrum that do not align with the conformer specific spectrum.

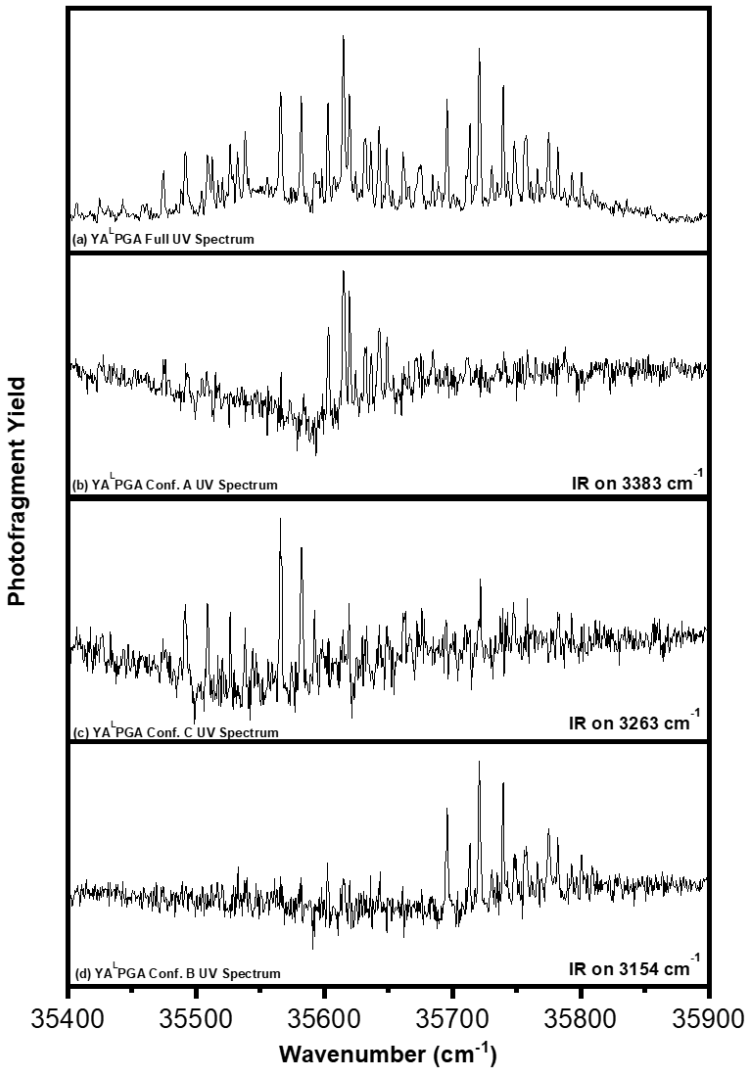

Figure 5. IR-UV Hole-burning spectra of the three conformers of $[YA^{L}PGA+H]^{+}$ compared to the full UV spectrum.

1.5.4 IR-UV Hole-burning Spectra

Once one is confident that all transitions in the infrared are accounted for a second check can be performed by obtaining conformation specific UV spectra of each of the conformers via IR-UV hole-burning. An example set of holeburn spectra is provided in figure 5. These spectra were obtained by reversing the process of obtaining an infrared spectrum. For holeburning, the IR laser is fixed on a resonant transition while the UV laser is tuned. Depopulating the zero-point energy level of ions belonging to the origin of the IR laser leads to a decrease in intensity of the UV transitions belonging to that conformer. Conformers that do not absorb the IR photon are

unaffected and therefore do not modulate in intensity. Subtracting this IR-heated UV spectrum from the non-IR heated spectrum results in a conformer specific UV spectrum. This then allows the experimenter to determine which transitions in the UV spectrum belong to which conformer.

1.6 Computational Chemistry

Key to determining the structure of an ion of interest is performing the calculations necessary to produce candidate spectra. This long process begins with a conformational search designed to cover as much of the conformational landscape as possible. Fortunately, performing the conformational search is a push button process made possible by the use of the Schrodinger software and MacroModel package. Using this software a group of Monte Carlo searches of the structure of the molecule are performed using the Amber*, OPLS3e, and MMFF forcefields. An important note to those using this software is to make sure that the search is being performed in the gas phase (default is solution) and that you widen the energy window for included structures so that more candidate structures can be used to refine the search. In the past only a single forcefield, Amber*, was utilized as it gave good results initially. With more flexible molecules it is worth the time investment to use all three forcefields as the structures produced can wildly differ. The results of the forcefield search will often rank the structures according to a calculated energy value specific to the forcefield. These energy values are highly unreliable and should not be used to influence the selection of candidate structures. Defining convergence for each of the searches is nebulous, but past users go until 1000 structures are obtained per forcefield. With large molecules this has been expanded up to 5000 structures in order to adequately sample the potential energy surface. With a possible pool of up to 15000 structures and no reasonable energy sorting method in the software finding potential candidates can be daunting. The next best step is to cluster the molecules into families as this will reduce the number of structures that must be run through the more time consuming DFT process.

What makes the Macromodel software package so remarkable is the wide range of tools available for analyzing the structures found in the conformational search. The most important of these tools is the clustering analysis program as it will allow for a massive reduction in the number of candidate structures. Best results for clustering have been observed by using backbone RMSD as the comparison parameter but several other parameters exist. The clustering program works by comparing structures to each other and finds backbone structures that are highly like

each other. Choosing a good definition for RMSD is very dependent on the conformations in the pool as they may vary widely form each other or may all be similar. Thankfully, the software calculates the number of clusters that will be produced for a given RMSD value and plots it in what the program calls an R-plot. The best choice in RMSD is where the plot begins to become horizontal as this maximizes differentiation of the conformations without producing too many families with only a handful of structures. After clustering has finished there are typically between 50-100 families that house all the structures from the original pool. The centroid of each family describing the average structure of each family is then selected for DFT calculation.

Running high level DFT on 50-100 structures can take an immense amount of time, especially if the structures are large. It is therefore advisable to perform a DFT energy pre-screen to remove structures that are far too high in energy to be possible. The best way to accomplish this screen is to use the same level of theory as the higher end calculations but negate the polarization correction as it tends to increase calculation time significantly. For example, at the B3LYP-GD3BJ/6-31G*(d) level of theory a single calculation for the tethered peptide Ac-y[kaaad]X_1-NH_2 takes one and a half days to finish. Removal of the polarization parameter, B3LYP-GD3BJ/6-31G(d), reduces calculation time to two hours. The benefits to this method are therefore obvious as screening the whole group of a hundred structures will take less than a day at the lower level of theory while it would take months at the higher level. Relying on a lower level of theory to sort structures can be worrisome as the screen may miss a possible conformation. To increase confidence in the pre-screening method the energies at both levels of theory were plotted against each other for each of the fifteen molecules presented later in this work. Figure 6 shows this plot for the tethered peptide Ac-y[kaaad]X_1-NH_2 where HL stands for high level of theory and LL stands for low level of theory. There is a strong correlation to relative energy that can be observed between the levels of theory which lends credence to the pre-screen being an effective method. This correlation is observed in each of the molecules studied, cementing this technique as an excellent way to reduce computational time.

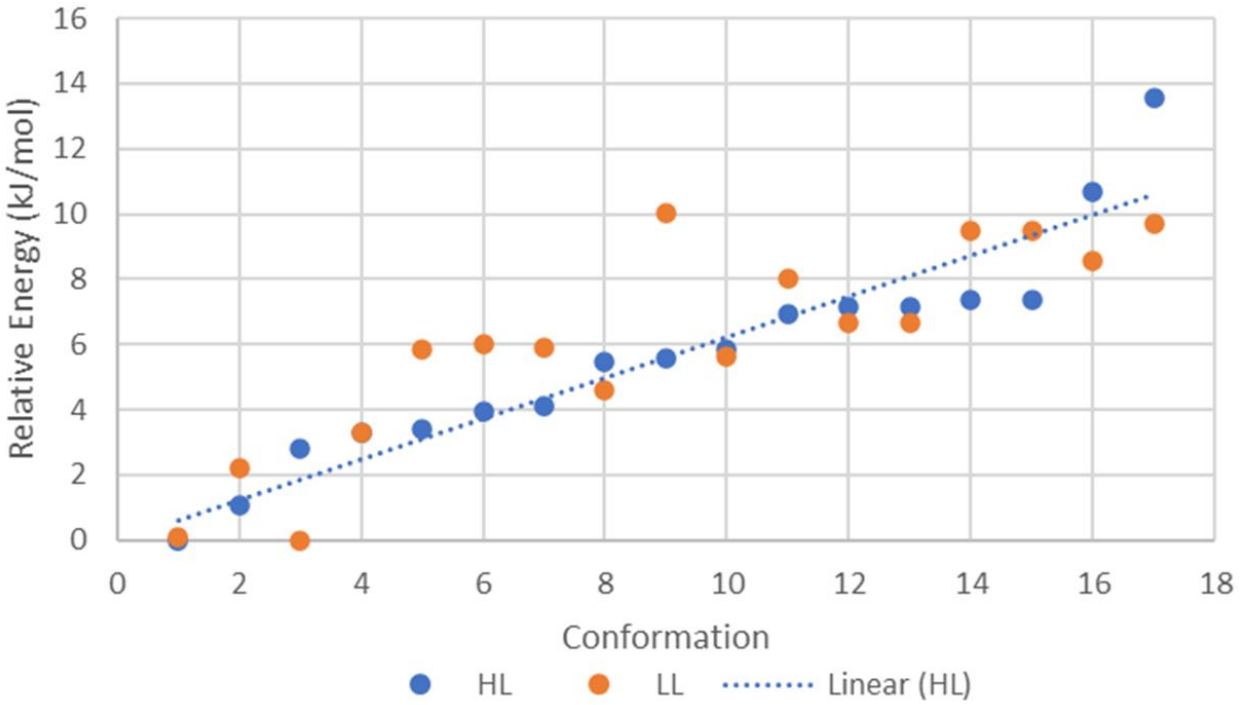

Figure 6 Plot of relative energy versus conformation at a low and high level of theory. A strong correlation between the two plots is observed.

The final step in generating candidate structures is to run the high level of theory calculations on all of the conformations lying below the user defined energy cutoff. For safety this cutoff is usually set to +10 kJ/mol above the global minimum, but if few structures lie within this range, it is best to push it out to +15kJ/mol. When all calculations are completed stick spectra can be generated and compared to the experimental spectrum.

1.7 References

1. Weibel, E. R.; Taylor, C. R.; Hoppeler, H., The concept of symmorphosis: a testable hypothesis of structure-function relationship. *Proceedings of the National Academy of Sciences* **1991,** *88* (22), 10357-10361.

2. Micheletti, C., Comparing proteins by their internal dynamics: Exploring structure–function relationships beyond static structural alignments. *Physics of Life Reviews* **2013,** *10* (1), 1-26.

3. Larrieu, D.; Britton, S.; Demir, M.; Rodriguez, R.; Jackson, S. P., Chemical Inhibition of NAT10 Corrects Defects of Laminopathic Cells. *Science* **2014,** *344* (6183), 527-532.

4. Wang, G.; Wang, F.; Wang, G.; Wang, F.; Zhang, X.; Zhong, M.; Zhang, J.; Lin, D.; Tang, Y.; Xu, Z.; Song, R., Opaque1 Encodes a Myosin XI Motor Protein That Is Required for Endoplasmic Reticulum Motility and Protein Body Formation in Maize Endosperm *The Plant Cell* **2012,** *24* (8), 3447-3462.

5. Fjell, C. D.; Hiss, J. A.; Hancock, R. E. W.; Schneider, G., Designing antimicrobial peptides: form follows function. *Nature Reviews Drug Discovery* **2012,** *11* (1), 37-51.

6. DeMartini, D. G.; Izumi, M.; Weaver, A. T.; Pandolfi, E.; Morse, D. E., Structures, Organization, and Function of Reflectin Proteins in Dynamically Tunable Reflective Cells*. *Journal of Biological Chemistry* **2015,** *290* (24), 15238-15249.

7. Spiegel, J.; Adhikari, S.; Balasubramanian, S., The Structure and Function of DNA G-Quadruplexes. *Trends in Chemistry* **2020,** *2* (2), 123-136.

8. Muro, E.; Atilla-Gokcumen, G. E.; Eggert, U. S., Lipids in cell biology: how can we understand them better? *Molecular Biology of the Cell* **2014,** *25* (12), 1819-1823.

9. Brown, S. P., Applications of high-resolution 1H solid-state NMR. *Solid State Nuclear Magnetic Resonance* **2012,** *41*, 1-27.

10. Hannibal, L.; Smith, C. A.; Jacobsen, D. W., The X-Ray Crystal Structure of Glutathionylcobalamin Revealed. *Inorganic Chemistry* **2010,** *49* (21), 9921-9927.

11. Raman, S.; Huang, Y. J.; Mao, B.; Rossi, P.; Aramini, J. M.; Liu, G.; Montelione, G. T.; Baker, D., Accurate Automated Protein NMR Structure Determination Using Unassigned NOESY Data. *Journal of the American Chemical Society* **2010,** *132* (1), 202-207.

12. Tu, Q.; Chang, C., Diagnostic applications of Raman spectroscopy. *Nanomedicine: Nanotechnology, Biology and Medicine* **2012,** *8* (5), 545-558.

13. Hofmann, J.; Hahm, H. S.; Seeberger, P. H.; Pagel, K., Identification of carbohydrate anomers using ion mobility–mass spectrometry. *Nature* **2015,** *526* (7572), 241-244.

14. Masson, G. R.; Jenkins, M. L.; Burke, J. E., An overview of hydrogen deuterium exchange mass spectrometry (HDX-MS) in drug discovery. *Expert Opinion on Drug Discovery* **2017,** *12* (10), 981-994.

15. Engel, E. A.; Anelli, A.; Hofstetter, A.; Paruzzo, F.; Emsley, L.; Ceriotti, M., A Bayesian approach to NMR crystal structure determination. *Physical Chemistry Chemical Physics* **2019,** *21* (42), 23385-23400.

16. Jamrógiewicz, M., Application of the near-infrared spectroscopy in the pharmaceutical technology. *Journal of Pharmaceutical and Biomedical Analysis* **2012,** *66*, 1-10.

17. Paudel, A.; Raijada, D.; Rantanen, J., Raman spectroscopy in pharmaceutical product design. *Advanced Drug Delivery Reviews* **2015,** *89*, 3-20.

18. Micsonai, A.; Wien, F.; Kernya, L.; Lee, Y.-H.; Goto, Y.; Réfrégiers, M.; Kardos, J., Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy. *Proceedings of the National Academy of Sciences* **2015,** *112* (24), E3095-E3103.

19. Lapthorn, C.; Pullen, F.; Chowdhry, B. Z., Ion mobility spectrometry-mass spectrometry (IMS-MS) of small molecules: Separating and assigning structures to ions. *Mass Spectrometry Reviews* **2013,** *32* (1), 43-71.

20. Shi, Y., A Glimpse of Structural Biology through X-Ray Crystallography. *Cell* **2014,** *159* (5), 995-1014.

21. Scheres, S. H. W., A Bayesian View on Cryo-EM Structure Determination. *Journal of Molecular Biology* **2012,** *415* (2), 406-418.

22. Garman, E. F., Developments in X-ray Crystallographic Structure Determination of Biological Macromolecules. *Science* **2014,** *343* (6175), 1102-1108.

23. Huber, M.; Hiller, S.; Schanda, P.; Ernst, M.; Böckmann, A.; Verel, R.; Meier, B. H., A Proton-Detected 4D Solid-State NMR Experiment for Protein Structure Determination. *ChemPhysChem* **2011,** *12* (5), 915-918.

24. Barnwal, R. P.; Yang, F.; Varani, G., Applications of NMR to structure determination of RNAs large and small. *Archives of Biochemistry and Biophysics* **2017,** *628*, 42-56.

25. Bertucci, C.; Pistolozzi, M.; De Simone, A., Circular dichroism in drug discovery and development: an abridged review. *Analytical and Bioanalytical Chemistry* **2010,** *398* (1), 155-166.

26. Benesch, J. L. P.; Ruotolo, B. T., Mass spectrometry: come of age for structural and dynamical biology. *Current Opinion in Structural Biology* **2011,** *21* (5), 641-649.

27. Leitner, A.; Walzthoeni, T.; Kahraman, A.; Herzog, F.; Rinner, O.; Beck, M.; Aebersold, R., Probing Native Protein Structures by Chemical Cross-linking, Mass Spectrometry, and Bioinformatics*. *Molecular & Cellular Proteomics* **2010,** *9* (8), 1634-1649.

28. Lanucara, F.; Holman, S. W.; Gray, C. J.; Eyers, C. E., The power of ion mobility-mass spectrometry for structural characterization and the study of conformational dynamics. *Nature Chemistry* **2014,** *6* (4), 281-294.

29. Crittenden, C. M.; Novelli, E. T.; Mehaffey, M. R.; Xu, G. N.; Giles, D. H.; Fies, W. A.; Dalby, K. N.; Webb, L. J.; Brodbelt, J. S., Structural Evaluation of Protein/Metal Complexes via Native Electrospray Ultraviolet Photodissociation Mass Spectrometry. *Journal of the American Society for Mass Spectrometry* **2020,** *31* (5), 1140-1150.

30. Masson, G. R.; Burke, J. E.; Ahn, N. G.; Anand, G. S.; Borchers, C.; Brier, S.; Bou-Assaf, G. M.; Engen, J. R.; Englander, S. W.; Faber, J.; Garlish, R.; Griffin, P. R.; Gross, M. L.; Guttman, M.; Hamuro, Y.; Heck, A. J. R.; Houde, D.; Iacob, R. E.; Jørgensen, T. J. D.; Kaltashov, I. A.; Klinman, J. P.; Konermann, L.; Man, P.; Mayne, L.; Pascal, B. D.; Reichmann, D.; Skehel, M.; Snijder, J.; Strutzenberg, T. S.; Underbakke, E. S.; Wagner, C.; Wales, T. E.; Walters, B. T.; Weis, D. D.; Wilson, D. J.; Wintrode, P. L.; Zhang, Z.; Zheng, J.; Schriemer, D. C.; Rand, K. D., Recommendations for performing, interpreting and reporting hydrogen deuterium exchange mass spectrometry (HDX-MS) experiments. *Nature Methods* **2019,** *16* (7), 595-602.

31. Kailemia, M. J.; Li, L.; Ly, M.; Linhardt, R. J.; Amster, I. J., Complete Mass Spectral Characterization of a Synthetic Ultralow-Molecular-Weight Heparin Using Collision-Induced Dissociation. *Analytical Chemistry* **2012,** *84* (13), 5475-5478.

32. Zhang, H.; Cui, W.; Wen, J.; Blankenship, R. E.; Gross, M. L., Native Electrospray and Electron-Capture Dissociation FTICR Mass Spectrometry for Top-Down Studies of Protein Assemblies. *Analytical Chemistry* **2011,** *83* (14), 5598-5606.

33. Bellissent-Funel, M.-C.; Hassanali, A.; Havenith, M.; Henchman, R.; Pohl, P.; Sterpone, F.; van der Spoel, D.; Xu, Y.; Garcia, A. E., Water Determines the Structure and Dynamics of Proteins. *Chemical Reviews* **2016,** *116* (13), 7673-7697.

34. Bythell, B. J.; Dain, R. P.; Curtice, S. S.; Oomens, J.; Steill, J. D.; Groenewold, G. S.; Paizs, B.; Van Stipdonk, M. J., Structure of [M + H − H2O]+ from Protonated Tetraglycine Revealed by Tandem Mass Spectrometry and IRMPD Spectroscopy. *The Journal of Physical Chemistry A* **2010,** *114* (15), 5076-5082.

35. Corinti, D.; De Petris, A.; Coletti, C.; Re, N.; Chiavarino, B.; Crestoni, M. E.; Fornarini, S., Cisplatin Primary Complex with l-Histidine Target Revealed by IR Multiple Photon Dissociation (IRMPD) Spectroscopy. *ChemPhysChem* **2017,** *18* (3), 318-325.

36. Boyarkin, O. V., Cold ion spectroscopy for structural identifications of biomolecules. *International Reviews in Physical Chemistry* **2018,** *37* (3-4), 559-606.

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.

38. Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S., Binding water to a PEG-linked flexible bichromophore: IR spectra of diphenoxyethane-(H2O)n clusters, n = 2-4. *The Journal of Chemical Physics* **2015,** *142* (15), 154303.

39. Simons, J. P.; Jockusch, R. A.; ÇarÇabal, P.; Hünig, I.; Kroemer, R. T.; Macleod, N. A.; Snoek, L. C., Sugars in the gas phase. Spectroscopy, conformation, hydration, co-operativity and selectivity. *International Reviews in Physical Chemistry* **2005,** *24* (3-4), 489-531.

40. He, C. C.; Kimutai, B.; Bao, X.; Hamlow, L.; Zhu, Y.; Strobehn, S. F.; Gao, J.; Berden, G.; Oomens, J.; Chow, C. S.; Rodgers, M. T., Evaluation of Hybrid Theoretical Approaches for Structural Determination of a Glycine-Linked Cisplatin Derivative via Infrared Multiple Photon Dissociation (IRMPD) Action Spectroscopy. *The Journal of Physical Chemistry A* **2015,** *119* (44), 10980-10987.

41. Mucha, E.; González Flórez, A. I.; Marianski, M.; Thomas, D. A.; Hoffmann, W.; Struwe, W. B.; Hahm, H. S.; Gewinner, S.; Schöllkopf, W.; Seeberger, P. H.; von Helden, G.; Pagel, K., Glycan Fingerprinting via Cold-Ion Infrared Spectroscopy. *Angewandte Chemie International Edition* **2017,** *56* (37), 11248-11251.

42. Polfer, N. C., Infrared multiple photon dissociation spectroscopy of trapped ions. *Chem Soc Rev* **2011,** *40* (5), 2211-21.

43. Spieler, S.; Duong, C. H.; Kaiser, A.; Duensing, F.; Geistlinger, K.; Fischer, M.; Yang, N.; Kumar, S. S.; Johnson, M. A.; Wester, R., Vibrational Predissociation Spectroscopy of Cold Protonated Tryptophan with Different Messenger Tags. *The Journal of Physical Chemistry A* **2018,** *122* (40), 8037-8046.

44. Page, R. H.; Shen, Y. R.; Lee, Y. T., Infrared–ultraviolet double resonance studies of benzene molecules in a supersonic beam. *The Journal of Chemical Physics* **1988,** *88* (9), 5362-5376.

45. Mehta-Hurt, D. N.; Korn, J. A.; Navotnaya, P.; Parobek, A. P.; Clayton, R. M.; Zwier, T. S., The spectroscopy and photochemistry of quinioline structural isomers: (E)- and (Z)-phenylvinylnitrile. *The Journal of Chemical Physics* **2015,** *143* (7), 074304.

CHAPTER 2. EXPERIMENTAL APPARATUS

2.1 Instrument for Cold Ion Spectroscopy

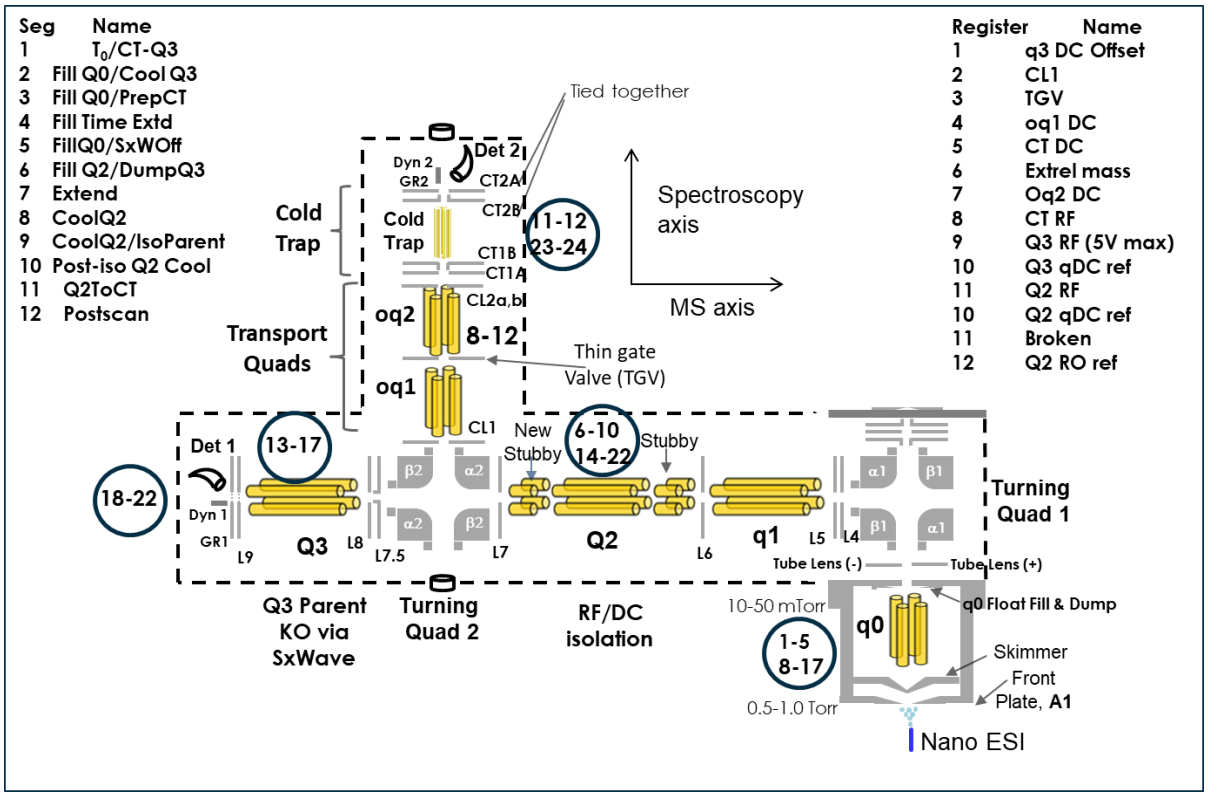

Figure 7 Schematic of the instrument for cold ion spectroscopy with labelled electrical connections

2.2 Experimental Overview

All data described in this thesis were taken on the Instrument for Cold Ion Spectroscopy (ICIS) at Purdue. Figure 7 displays an instrument schematic highlighting both the mass spectrometric and spectroscopic axis's as well as the relevant electrical connections to the Argos controller. What follows is a description of the generation and analysis of ions in the instrument. The design of the instrument is that of a triple quadrupole mass spectrometer in which an orthogonal spectroscopy axis was added between the second (q2) and third (q3) quadrupoles. Ions are generated via nano-ESI and guided into q2 where the ion of interest is isolated via RF/DC isolation. The ion packet is then steered through a turning quadrupole and down the spectroscopy axis where it is trapped in an octupole trap that is cooled to 5K via a closed-cycle

helium cryostat (Sumitomo Heavy Industries, Tokyo, Japan). After cooling to ~10K via collisions with the He buffer gas, the ions are spectroscopically interrogated via a tunable UV laser (ScanMatePro Lambda Physik dye laser frequency doubled by an Inrad Autotracker III). The created photofragments are then extracted back down the spectroscopy axis and turned into a quadrupole for analysis in either "spectroscopy mode" or "mass spectrometry mode". In mass spectrometry mode, a mass spectrum of the photofragments is generated via mass-selective axial ejection[1]. The "spectroscopy mode" entails ejection of remaining precursor ions with SX wave software followed by dumping of all remaining photofragments onto a channeltron detector (4773G, Photonis USA).

UV action spectra are recorded by monitoring the total photofragment signal as a function of laser wavenumber. Upon obtaining a UV spectrum and identifying transitions that may be due to different ion conformations, conformation-specific IR spectra are taken by employing IR-UV double resonance techniques[2]. The UV laser wavelength is fixed on a transition of interest in the UV spectrum creating a steady signal of photofragment ions due to the monitored conformation. A tunable IR laser that is spatially overlapped but temporally precedes the UV laser by 200 ns is scanned in wavelength. Upon absorption of an IR photon by the conformer of interest, a fraction of the population is removed from the ground state, thereby broadening and shifting its UV absorption. This results in a depletion in the total photofragment signal from the single conformation being monitored in the UV. Tuning the IR laser while monitoring this depletion gives a single-conformer IR spectrum.

2.3 Basics of Quadrupole Mass Analysis

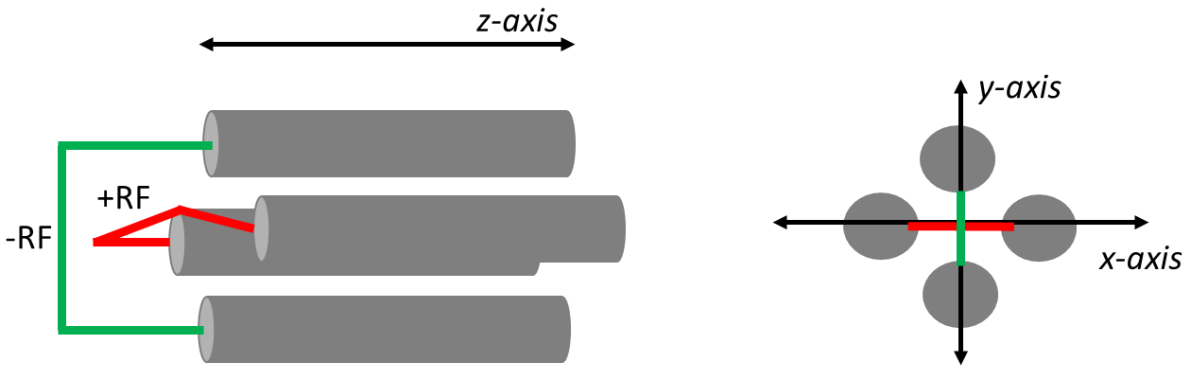

Figure 8 Schematic of a quadrupole rod set showing the RF applied to pairs of rods as well as the coordinate directions of ion confinement

With an overview of the operation of the instrument in hand an explanation of the operation of the most important aspect, the quadrupole, of the setup to be described. Use of the instrument in no way requires an in-depth knowledge of quadrupolar theory, but a scientist who wishes to expand their knowledge of mass spectrometry can find no better place to begin. An excellent review of the basics of quadrupole use and theory was written by D.J. Douglas[3]. This section provides a short summary of the uses of quadrupoles within this instrument. While a quadrupole can be utilized for many purposes its main use in the spectroscopy instrument is trapping, isolation, and analysis. A short explanation of these three functions follows. Trapping of ions in a quadrupole ion trap occurs within the radius inscribed by its two pairs of rods. Figure 8 shows a schematic of a rodset with pairing of the rods and the directions in which the ions must be confined to remain trapped. Upon entering the trap ions remain confined to the z-axis of the rodset due to repulsive voltages applied to lenses on both sides of the trap. Collisions with the helium buffer gas in the trap further removes translational energy from the ions keeping them away from the ends of the rods. Confinement in the x-y dimensions is achieved through the application of a radio frequency (RF) voltage to each of the rods in the set[4-5]. Opposite phases of the RF are applied to adjacent rods (figure 8) to create a quadrupolar field in which trapping occurs. The ions are caught in a push-pull game between the different rods because half of the time they see an attractive potential while the other half of the time they feel a repulsive potential. This pushing and pulling game occurs in both the x and y dimensions which leads to

trapping. The aforementioned review on quadrupoles excellently lays out each of the equations used to define the quadrupolar fields and trapping in each dimension. An interesting effect of the movement of ions in the x-y dimension is that the frequency of that motion is highly dependent on the mass and charge of the ion. The name given to this motion is "secular frequency" and this property of the ion movement can be exploited to maneuver certain ions in the trap in a selective fashion. A small supplemental AC waveform applied to one set of rods allows for this control by tuning to the secular frequency of the ion of interest. This allows the user to mass dependently eject ions from the trap creating a mass spectrum. Where isolation is concerned the secular frequency can be used to remove a certain ion from the trap or conversely remove all but a single ion from the trap.

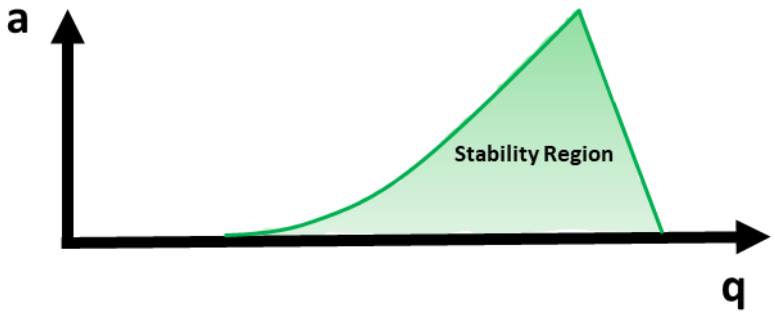

Figure 9 A single solution to the Mathieu Equations showing the stability region of trapped ions.

The stability of ions in the trap can be described by solving the Mathieu equations[6]. Figure 9 shows a graph of a single solution to these equations which describes a stable region for the ions in terms of the dimensionless parameters (a) and (q). Equations governing these parameters are given below, where z is the charge of the ion, e is the elementary charge, m is the mass of the ion, Ω the RF frequency, and r_0 the radius of the trap.

$$(1) \qquad a = \frac{8zeU_{DC}}{m\Omega^2 r_0^2}$$

$$(2) \qquad q = \frac{4zeV_{RF}}{m\Omega^2 r_0^2}$$

The parameters that differ between the two equations relate to variables that can be manipulated to move ions into and out of the stability region. U_{DC} is the DC voltage applied to the rods while, V_{RF} describes the RF voltage applied. One can think of the axes in figure 9 as a function of the

DC and RF applied voltages and therefore ions can be moved into and out of the stability region through manipulation of these parameters. This changing of RF and DC inputs to isolate an ion of interest is performed in the second quadrupole (q2) of the ion spectroscopy instrument accomplished through toggling the waveform generators that control each of the parameters.

2.4 Standard Operating Guidelines for Use of the Updated Ion Spectroscopy Instrument

Basic guidelines for the use of the previous version of the instrument can be found in the theses of its original builders, Nicole Burke and James Redwine. In particular, Nicole's thesis gives a very thorough walkthrough of each of the parts of the instrument. An overview of the updated, and current instrumentation can be found in the thesis of Chris Harrilal. This section will describe some standard operating procedures for the updated instrument with the goal of allowing a new user to be well on their way to collecting data in as short a time possible.

2.4.1 Instrument Startup

Note: Check the pressure in the instrument before beginning to turn it on. If the pressure gauges do not all read "0.0 – 0" in the mass spec axis then there is a leak somewhere.

Begin instrument startup by turning on all laser systems and place the flashlamps in "Internal" mode, wait 30 minutes. Remove the septum from the ion source, screw purge cup in place, and place A1 wire between cup and source flange. After this you may now open the thin gate valve. *Note: If the red line is on the lower end of the scale then the valve is open, if it is on the upper end of the scale it is closed.* Now, the gas lines can be open to the instrument, including the helium lines for both the cold trap and q2. Remember to also open the pinch valve going to the cold trap. Turn the dial for the helium regulator for q2 such that the flow (read on the meter next to the dial) gauge reads roughly 80. *Note: Do not exceed this pressure or you will overload the pumps. Note: The pressure will change over the first 10 minutes of running the instrument as everything equilibrates, be sure to increase the gas flow as needed.* Turn on all electronics except for those associated with the ion source (ESI voltage and top spectrum solutions). Try to maintain a pattern in which the electronics are turned on, this keeps you from forgetting to turn something on. For example, I turn them on in the pattern front, stack, back. Meaning all electronics on the front of the instrument, then those in the stack, then those on the

back of the instrument. *Note: For any piece of electronics with a silver box, the silver box must be turned on before the Kikusui*. Load the "YGRDR – Full Experiment Scan Function" or whichever scan function you are currently using by going to File->Load Scan Function. Press the "Run" arrow on Argos and the instrument is now ready to be operated.

2.4.2 Signal Down the MS Axis

Note: All voltages given as examples herein apply only to YGGFL. If using another sample, voltages on turning quad 2 will change.
Note: A novice user will be tempted to look at the computer screen when tuning for signal, this works but is not optimal. Better signal adjustments are made by watching the scope when tuning.

Begin by obtaining a tip of YGGFL and place the tip into the holder. *Note: The platinum wire in the holder should not exceed the taper of the tip or the tip will shatter when voltage is applied.* Screw the tip holder into the purge cup until the tip is roughly half a centimeter from the orifice. *Note: Ideally, the tip should be in the upper right corner of the orifice. The best spray is achieved here.* Attach the ESI voltage gator clip to the free end of the tip holder then turn on the top spectrum solutions box. Before turning on the ESI voltage, set the voltage to 1000 V. *Note: The higher the voltage when turned on the greater the risk of tip bursting.* Turn on the ESI voltage. If signal appears, tune for maximum signal by starting with the voltages given below, then begin at the front of the instrument and work towards the back of the instrument. If no signal appears, check that all electronics are on, check for ion burn to make sure the tip is spraying, and/or check voltages. There are 8 critical voltages that must be in place to see YGGFL signal down the mass spec axis. They are:

a. L8 : 20.4 V
b. L9 : 22.4 V
c. L7.5 : 12.9 V
d. TQ2 alpha1 : -16.6 V
e. TQ2 alpha2 : 8.0 V
f. TQ2 beta1 : 1.7 V
g. TQ2 beta2 : -4.9 V
h. Argos: Set segments 11 and 12 to -7 from -10
   Note: Be sure to recompile after changing voltages in Argos.

Following these steps should generate signal in the instrument.

2.4.3 Signal Down the Full Experiment

Note: This section assumes you have maximized signal down the triple quad axis first.
With maximum signal going down the triple quad, turn the critical voltages back to their full experiment values:

a. L8 : 8.0 V
b. L9 : 30.4 V
c. L7.5 : -17.7 V
d. TQ2 alpha1 : -16.6 V
e. TQ2 alpha2 : 8.0 V
f. TQ2 beta1 : 20.3 V
g. TQ2 beta2 : -22.5 V
h. Argos: Set segments 11 and 12 to -10 from -7
   Note: Be sure to recompile after changing voltages in Argos.

If signal appears, tune for maximum signal in this order: transport quad voltages, critical voltages, CL1, CL2, and then CT1. If no signal appears first check the pressure in the cold trap. If it is too high (greater than 60 on meter), there is a vacuum problem on the cold side (most likely scroll pump). No trapping will happen with poor vacuum. If the pressure is fine, manipulate the voltage on CT2 (the trapping voltage). If signal is gained by changing CT2 but still poor after tuning it is likely that the trap needs to heat cycle. If signal is near to levels seen going down the triple quad axis then continue with the experiment.

2.4.4 Ion Isolation and Knockout Creation

Note: All previous isolation settings are in an Excel file called "Q2IsoCalibration" these should be used as a guide for isolating new ions. The general rule is the higher the mass the more RF and DC needed to isolate.
Note: The knockout file for YGGFL is an Excel file called "YGGFLknockoutforApexIsolation". Knockout creation relies on the fact that larger ions are placed at lower frequency and smaller ions at higher frequency.

The isolation settings for YGGFL are Q2RF : -6.374 V and Q2DC : 504 V. Use these as a guide when isolating new ions. Isolation should be done in MSAE mode, start by putting the banana plug in the Agilent L9 box from "Dump" to "MSAE". With the Q2 DC off, check to see that the ion of interest is seen in the mass spectrum. Raise the Q2 DC such that a majority of the ion of interest disappears from the mass spectrum. Adjust the Q2 RF such that you maximize the

ion of interest. If no other ions are present in the spectrum, then the isolation is complete and move on to knockout. If other ions do persist then repeat steps the previous steps.

The SxWave software provided by SciEx is used to create the parent knockout in Q3. It works by creating a frequency sweeping waveform that starts at the user input frequency and ends at the mass limit of the system. For example, if you have an ion with an m/z of 634 which correlates to a secular frequency of 62 kHz you would start your knockout at 62 kHz and sweep up in frequency to ensure all high mass ions are ejected. The goal is to ensure that all smaller mass fragments are retained while the much more abundant parent ion is removed. For the knockout, again, use the YGGFL knockout as a guide. Start by underestimating the frequency of your ion in an attempt to overshoot it with the knockout. The goal at first is to not knock anything out with the first attempt. Now, move the frequency value slightly higher and re-check the knockout. If the ion is knocked out then move on to laser alignment, if not then go back to the previous step and repeat until the knockout is fully in place. *Note: The best way to check your knockout is to turn it on, turn the detector up to 1.9 kV, put the instrument into MSAE mode and turn off the isolation. You should see plenty of other ions/fragments in the mass spectrum. The closer they are to your isolated mass the better your knockout.*

2.4.5 Laser Alignment

Note: Rough alignment of the probe laser should be performed with the diode laser as it posses no risk to the trap and is easy to visualize along the full path of the instrument. This being said, alignment of the enhancement laser requires using the IR beam. The UV laser beam is also required for its alignment.

Note: This section will assume that the beam path is for the enhancement experiment and not the old IR-UV setup.

Begin by blocking the UV entrance to the instrument with a card. Using prisms, get the UV laser to the entrance window of the trap. You will want at least two prisms between the laser and the window. Set up the silver wedge at least 1 meter away from the window to the trap. It is important to understand that the enhancement laser will need to go through the open part of the wedge and the probe laser will need to bounce off the angled surface. Turn the enhancement laser to low power and align it through the open “square” part of the wedge. The beam should go straight through and come out the pointed part of the wedge. Use the remaining mirrors to get the

beam to the card blocking the window to the trap. The diode should be setup near the probing IR laser in a position such that the incoming diode beam will intersect the probe beam as near to the probe laser as possible. This makes aligning both beams easier. Use a flipper mount with a silver mirror on it to intersect the path of the probe beam. Align the diode onto the middle of the mirror and follow it down the beam path with the knowledge that the beam must hit the angled portion of the silver wedge at Brewsters angle. Align the red beam from the diode laser on top of the enhancement laser then shutter the enhancement laser. Let the diode through the trap and onto the other side. Use the two mirrors closest to the IR side of the trap to get the laser through. Align the UV laser on top of the diode on the UV side of the trap. With the knockout in place and the pre-amp turned up, tune the UV laser to the origin of YGGFL (560.36). Let the UV laser through the trap and look for photofragments. Maximize photofragment intensity using the second closest prism to the UV window of the trap. *Note: For YGGFL, with 600K isolated ions, at 1.75K Volts on the detector you should have ~250K photofragments. Note: This is the most important step where S/N is concerned. It's easy to get the UV through and get most of the fragments. It takes time to do minor adjustments and get signal that isn't very noisy.*

With photofragment intensity maximized you can now move on to fine alignment of the enhancement laser. Allow the UV laser through the UV side of the trap and block the IR side. Unshutter the enhancement laser. On the IR side of the trap, align the enhancement laser on top of the UV laser using the 2 mirrors closest to the trap. Near alignment should be done with the far mirror and far alignment with the near mirror. When the UV beam is centered on the enhancement beam you can now allow the enhancement beam through the trap. Adjust the second to last mirror on the IR side for enhancement of photofragment signal until it reaches a maximum. *Note: The IR wavelength used for enhancement with YGGFL is 12768.97 cm^{-1}. The extent of enhancement is a function of how much signal you started with.* Enhancement alignment is complete. Now move on to Probe laser alignment.

Block the IR entrance to the cold trap. Turn the waveplate on the enhancement laser down such that you barely see a hot spot on a heat card near the trap entrance. Unshutter the probe laser. Align the probe laser on top of the enhancement laser using the two mirrors before the silver wedge on the probe side of the optics. *Note: Remember that you need to be hitting the angled side of the silver wedge when doing this alignment. This is a hard alignment and will take some time the first few times you do it.* With the probe laser on top of the enhancement laser,

block the probe laser and unblock the IR entrance to the trap. Turn up the power on the enhancement laser till you return to you previously enhanced count of photofragments. Unblock the probe laser. If sitting on a transition in the IR, you should see dip now. Adjust the final mirror on the probe side of the silver wedge for maximum depletion. The lasers are now setup for every experiment.

2.4.6 Recording Spectra

Note: There are two labview programs on the computer for taking spectra. The first is called "UV 3000 Series" and is used for recording UV spectra and the other is called "BioIon ActiveIR 3000 Series" and is used for recording IR spectra. Using these programs is rather straight forward.

Several key points to know before recording a spectrum, the first is to make sure the pre-amplifier is turned up or you will not see fragments. The second is simple but can be overlooked, make sure to unblock the lasers. Where the scope is concerned, make sure the number of averages is set to sixteen, but more importantly make sure the signal on the scope is on the very left side of the screen. The labview programs record what is presented on the screen and do not talk to Argos. So, if you set your recording window to 1 to 6 in the program your signal needs to be in this region on the scope.

Mass spectra are recorded using the Argos software and are saved to the harddisk on the computer that has Argos on it. The first step in recording a mass spectrum is to ensure that the L9 Agilent box is set to MSAE mode by moving the banana plug from "dump" to "MSAE". A mass spectrum should be observed on screen in the Argos software. To record this spectrum first ensure that the number of averages per scan is set to a number of interest. This is usually set to "11" during spectroscopy experiments and "255" during recording of mass spectra. With the averages set the user can now press the camera button on the Argos interface which allows for recording of a spectrum.

A user will likely need to record an IR-UV holeburn spectrum at some point during the analysis of an ion. To do this requires a change in the repetition rate of the infrared laser used to probe the transition of interest. Begin this process by first setting the IR laser to a transition of interest. Then, set both the UV and IR lasers to internal mode, this is an important step as failure to do so will cause the interlock to activate in the next step. The next step is to change the

qswitch timing on the IR laser on both channels C and H on the timing box. Press the appropriate channel button until you see the word “normal” appear on the screen then rotate the toggle until it says “One on/One off”. Press the “run/stop” button to accept the change. This will set the laser in 5Hz mode. The lasers can now be placed into “external” mode and the IR laser timing should be observed. The shot of the laser should be significantly slower than in 10Hz mode, if not then check the settings and try again. Next, go to the scope and ensure that two ion packets are observed on the screen. One packet is IR activated and the second is not so the signal in one packet should be greater than the other. With both packets on screen go to the labview UV software and press the active subtraction button, set up for two gates with the first at the normal time and the second between 101 and 106. Next, change the frequency of collection to 5Hz. A holeburn spectrum should now be able to be recorded. Note, that this spectrum will take twice as long as a normal UV spectrum to collect therefore it is important that stable, long lasting signal is obtained before beginning this experiment.

2.5 Outline of the Thesis

The remainder of the thesis is dedicated to the research that was performed on it. It begins with a study of the peptides $YA^{D/L}PGA$ and $YAAA^{D}PGAAA$ and the secondary structures that they form in the gas phase. The goal of the study is to determine whether the placement of a flexible glycine residue after a proline residue will either inhibit or aid in beta turn formation. This work is a follow on to the studies of $YA^{D/L}PAA$ and $YG^{D/L}PAA$ which were previously performed and guideline the stabilization of beta turns in the gas phase through the use of proline isomers. This work concludes a trilogy of papers on gas phase beta turn formation performed on the ion spectroscopy instrument.

What follows is a new trilogy dedicated to the formation of single turn alpha helices in the gas phase. It begins with a set of tethered peptides containing the motif [KAAAD] which have been shown by David Fairlie and colleagues[7-10] to form lock tight single turn helices in the solution phase. Interrogation of these tethered peptides as ions was a hopeful way of observing the single turn helix and understanding its spectroscopic patterns in the gas phase. What resulted was a mess of structures that largely had no defining secondary structure. This short term let down resulted in a better understanding of the features a small, tethered peptide must contain to possibly observe the helix. The knowledge obtained in the first study was then applied in the

second study which heavily utilized conformational searching algorithms to guide the synthesis of a molecule that would form the single turn helix in the gas phase. Control of the handedness of the helix was also shown using stereochemistry changes in each of the amino acids. The final study described in this thesis details the expansion of the tethered single turn helix to two turns and controlling the handedness of the helices not only to produce pure left- and right-handed versions but to also form mixed left/right-handed helices. Overall, these studies show the power of ion spectroscopy to analyze difficult molecules and aid in their structural determination.

2.6 References

1. Londry, F. A.; Hager, J. W., Mass selective axial ion ejection from a linear quadrupole ion trap. *Journal of the American Society for Mass Spectrometry* **2003,** *14* (10), 1130-1147.

2. 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. *Journal of the American Chemical Society* **2017,** *139* (15), 5481-5493.

3. Douglas, D. J., Linear quadrupoles in mass spectrometry. *Mass Spectrometry Reviews* **2009,** *28* (6), 937-960.

4. Miller, P. E.; Denton, M. B., The quadrupole mass filter: Basic operating concepts. *Journal of Chemical Education* **1986,** *63* (7).

5. Miller, P. E.; Denton, M. B., The transmission properties of an RF-only quadrupole mass filter. *International Journal of Mass Spectrometry and Ion Processes* **1986,** *72* (3), 223-238.

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

7. de Araujo, A. D.; Hoang, H. N.; Kok, W. M.; Diness, F.; Gupta, P.; Hill, T. A.; Driver, R. W.; Price, D. A.; Liras, S.; Fairlie, D. P., Comparative alpha-helicity of cyclic pentapeptides in water. *Angew Chem Int Ed Engl* **2014,** *53* (27), 6965-9.

8. Hill, T. A.; Shepherd, N. E.; Diness, F.; Fairlie, D. P., Constraining cyclic peptides to mimic protein structure motifs. *Angew Chem Int Ed Engl* **2014,** *53* (48), 13020-41.

9. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Left- and Right-Handed Alpha-Helical Turns in Homo- and Hetero-Chiral Helical Scaffolds. *Journal of the American Chemical Society* **2009**.

10. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005**.

CHAPTER 3. STEREOCHEMISTRY, STERIC HINDRANCE, AND PEPTIDE LENGTH EFFECTS ON B-HAIRPIN TURN FORMATION: UNDERSTANDING THE CONFORMATIONAL PREFERENCES OF COLD [YAPGA+H]+ AND [YAAPGAAA+H]+ IONS

3.1 Abstract

Intelligent design of peptide-based therapeutics requires an in depth understanding of the intramolecular forces that shape the secondary structure of the molecule. Inclusion of non-standard amino acids has proven to be a powerful tool in structural control of small peptides. D-proline is widely utilized to form β-hairpin loops in molecules that would otherwise be unstructured. It has been hypothesized that the reduction of steric influence surrounding the proline residue positively influences β-turn formation as well. Herein, conformation specific IR-UV double resonance spectroscopy of the cold (~10K) protonated ^{D}P and ^{L}P diastereomers of YAPGA was performed to elucidate the effect of steric hinderance on β-turn formation in the gas phase. A single conformation was formed from the ^{D}P isomer which presented as a type II β-turn which shows that reducing steric influence does indeed help induce only the β-turn. While the ^{L}P isomer formed three conformations with vastly different structures, none of which was β-hairpin loop. Of the conformations seen with ^{L}P one includes a cis-amide bond between Y_1 and A_2 that forms due to the charge injecting itself into the middle of the peptide. The resulting structures confirm the proclivity of D-Proline to form a β-turn in the gas phase and provide a foundation for understanding the biochemical principles that govern secondary structure formation and the differences between gas and solution phase structures.

3.2 Introduction

The structural advantages to utilizing unnatural D-amino acids in biologics-based materials is a well-researched field.[1-3] Their inclusion in molecules reduces susceptibility to enzymatic degradation[4-6], increases anti-microbial activity[7-9], and aids in self assembly of peptide secondary structures[10-11]. Therapeutic peptides are designed with these properties in mind to enhance the efficacy and selectivity of pharmaceuticals[11-12]. Of all the D-amino acids, the effect D-proline has on secondary structure is one of the most significant. Substitution of L-proline with D-proline is known to greatly stabilize beta turn formation in peptides.[13-18] This finding has been

heavily exploited by synthetic organic chemists to produce scaffolds for drug delivery through the inclusion of a ^{D}PX moiety (X = Ala, or Gly) in peptides.[15-16, 19]

The terminology surrounding the classification of different types of "turn" in biochemistry can be quite complex.[20-21] Turns are first categorized by the separation between residues forming a hydrogen bond, for example, an alpha turn is an *i* to *i+4* hydrogen bond.[22] In simple terms, this means there is a separation between the first and last amino acids forming the turn of four peptide bonds. Whereas a gamma turn is an *i* to *i+2* hydrogen bond, or two peptide bonds between interacting residues.[23-24] Turns are then further classified by the phi and psi bond angles present in the turning region.[20, 22]

Ion	Backbone Structure	Turn Stabilization	Charge Pocket
^{D}PG	Type II' β-hairpins in solution	NA	NA
YGGFL	Type II' β-turn / β-hairpin; NH3+ to turn C=O	C10,C14	C11,C17, π_{NH}
YA^{D}PAA	Conf. B = Type II' β-turn; Conf. A = γ-γ	Conf. B = C10, C14 Conf A. = C7, C7	Conf. B = C11,C17, π_{NH}; Conf. A = C14,C17, π_{NH}
YA^{L}PAA	Charge-stabilizing loop, cis-Pro amide bond	C14, C17	C14,C17, π_{NH}
YG^{L}PAA	Type II' β-turn	C10, C14	C14,C17, π_{NH}
YA^{D}PGA	Type II β-turn / β-hairpin	C10,C14	C14,C17, π_{NH}
YA^{L}PGA	A = cis-Pro β-turn; B=cis-amide charge-stabilizing loop; C = γ-C7	A = C10,C14; B = C11,C17; C = C7,C7	A = C14,C17, π_{NH}; B =C17,C11,C5; C = C14,F_{NH},π_{NH}

Figure 10 Compilation of secondary structure motif, turn stabilization, and charge pocket for the studied set of ions containing β-turns

The beta turn, *i* to *i+3* hydrogen bond, is the most common of all the turns seen in native proteins.[20, 25] Phi and psi angles can be used to group beta turns into different types, of which types I, II, I', and II' are the most prominent.[26] The angles described by these types of turn are compiled in table 1. Of interest are type I' and II' beta turns, as these types are most commonly found in beta hairpins, protein structural elements that cause an inverse in the direction of the protein to form antiparallel beta strands. Studying the origin and conformational inclination of these structures to form is therefore important to understanding and utilizing the larger, antiparallel beta strand.[20, 25-27]

A powerful method to study the conformational preferences of these scaffolds is to analyze them in the gas phase where the lack of solvent interactions shines light on the intramolecular hydrogen bonds that stabilize these structures.[28-31] In the past, our group has leveraged the power of IR-UV double resonance techniques to probe the ^{D}P and ^{L}P diastereomers of $[YAPAA + H]^{+}$

and the $[YG^{L}PAA + H]^{+}$ peptide ions in the gas phase.[32-33] The ^{D}P form of $[YAPAA + H]^{+}$ was found to form two predominant conformations one being a type II' beta turn and the other a gamma turn. The ability of ^{D}P to promote trans amide bond formation was found to be important in relieving steric strain that inhibited formation of the beta turn in the ^{L}P diastereomers. Substitution of the alanine residue in the second position to a glycine reduced steric hinderance enough that a beta turn conformation was seen in $[YG^{L}PAA + H]^{+}$.[32-33]

Two major questions remained after the previous study. The first was what effect the alanine residue that falls c-terminal to the proline has on the conformational landscape of these diastereomers. There is a preponderance of evidence from solution phase studies of natural peptides that the PG motif is rather important for the formation of beta turns. The lack of a side chain on the glycine residue reduces steric hinderance which may be key to forming the type II' beta turn. A goal of this study is to understand the role of the glycine residue in reducing steric hinderance of the PG motif in gas phase peptide ions. To elucidate the effect of glycine flexibility on the formation of type II' beta turns the ^{D}P and ^{L}P diastereomers of $[YAPGA + H]^{+}$ were spectroscopically interrogated.

The second question was what effect the length of the peptide sequence has on the stabilization of the beta-hairpin. Additional hydrogen bonding, brought about through the extension of the peptide, may support the formation of the beta hairpin through the formation of a beta-sheet like structure that is more energetically favorable than its unstructured competitors. Extending the peptide also moves the charge site further away from the turn which may help to elucidate whether charge influences its formation. To answer these questions, we studied the model peptide $[YAA^{D}PGAAA + H]^{+}$ as a gas phase ion via cold ion spectroscopy.

Cold ion spectroscopy continues to be the premier method for 3D structural determination of ions in the gas phase.[30, 34-36] It has been utilized to examine the structure of compounds across the whole spectrum of molecules, from sugars to peptides, and remains a growing field of interest.[30] Cooling ions down to 10K removes vibrational energy in the molecule allowing for specific interrogation of certain vibrational modes using infrared photons. This allows for the mapping of a "fingerprint" infrared spectrum that is unique to every molecule or conformation studied. This spectrum can then be matched to calculated spectra produced by density functional theory calculations to propose a likely 3-dimensional structure. Conformation specific IR-UV double resonance spectroscopic techniques can be applied to determine not only

the 3D structure of an ion but also probe the population and dynamic folding of its conformers.[33, 36] Cold-conformer specific spectroscopic techniques were applied to the peptide diastereomer protonated cations of $YA^{L/D}PGA$ and $YAA^{D}PGAAA$ not only to illuminate their structures, but to help decode the link between stereochemistry and steric hinderance that allows for the formation of beta turns in the gas phase.

3.3 Methods

3.3.1 Experimental

All spectroscopic data were taken on a custom-built instrument for cold ion spectroscopy, previously described.[29, 37] Concisely, the design is that of a triple quadrupole mass spectrometer in which an orthogonal spectroscopy axis was added be-tween the second (q2) and third (q3) quadrupoles. Ions are generated via nano-ESI and guided into q2 where the ion of interest is isolated via RF/DC isolation. The ion packet is then steered around a turning quadrupole and down the spectroscopy axis where it is trapped in the cold quadrupole that is cooled to 5K via a closed-cycle helium cryostat (Sumitomo Heavy Industries, Tokyo, Japan). After cooling to ~10K via collisions with the He buffer gas, the ions are spectroscopically interrogated via IR (Laservision OPO/OPA) and UV (ScanMatePro Lambda Physik frequency doubled by an Inrad Autotracker III and Scanmate Lambda Physik) lasers. The created photofragments are then extracted back down the spectroscopy axis and turned into q3 for analysis in either “spectroscopy mode” or “mass spectrometry mode”. In mass spectrometry mode, a mass spectrum of the photofragments is generated via mass-selective axial ejection.[38] “Spectroscopy mode” entails ejection of remaining precursor ions with SX wave software followed by dumping of the photofragments onto a channeltron detector (4773G, Photonis USA).

UV action spectra are recorded by monitoring the total photofragment signal as a function of laser wavenumber. Upon obtaining a UV spectrum and identifying transitions that may be due to different ion conformations, conformation-specific IR spectra are taken by employing IR-UV double resonance techniques. The UV laser is tuned to a transition of interest in the UV spectrum which creates a steady signal of photofragment ions. Then a tunable IR laser that is spatially overlapped but temporally precedes the UV laser by 200 ns is scanned in wavelength. Upon absorption of an IR photon by the conformer of interest, a fraction of the

population is removed from the ground state which results in a depletion in the total photofragment signal.

3.3.2 Solution Preparation

Solutions of YA^{D}PGA, YA^{L}PGA , and YAA^{D}PGAAA (Genscript) were prepared as 50:50 Methanol:Water to a concentration of 200 μM and electrosprayed to create the corresponding protonated ion.

3.3.3 Computational

Conformational searches were performed for $[YA^{D}PGA + H]^{+}$, $[YA^{L}PGA+ H]^{+}$ and $[YAA^{D}PGAAA + H]^{+}$ structures using the Monte Carlo multiple minimum method with the AMBER* and OPLS3 forcefields inside the MACROMODEL software package.[39] Approximately 1000 structures were found between the two forcefields for each of the molecules. To narrow down the potential structures further, the 1000 structure pool was clustered based on backbone RMSD. The centroid of each of the clusters was then optimized at the B3LYP/6-31G*-GD3BJ level of theory using Gaussian16 and harmonic vibrational frequencies/intensities were calculated to obtain predicted infra-red spectra to compare with experiment.[40] To correct for anharmonicity the calculated spectra were scaled by 0.958 for the NH, CH, and OH stretching regions, 0.973 for the free OH transitions, 0.981 for Amide I, and 0.970 for Amide II. Adjusting the calculated transitions by these values puts them into agreement with previously calculated and fitted spectra.[29, 32-33] "Stick" spectra were generated for the centroid of each cluster. These calculated spectra were then matched to the experimental spectrum and the best fit centroids were found. All clusters of structures with best fit centroids that lie within 10 kJ/mol of the global minimum were then run through DFT at the same level of theory used for the centroids. After completing all calculations, the structure that best fit the experimental spectrum and was lowest in energy was called as the most likely structure of the ion.

3.4 Results and Analysis

3.4.1 UV Action Spectra

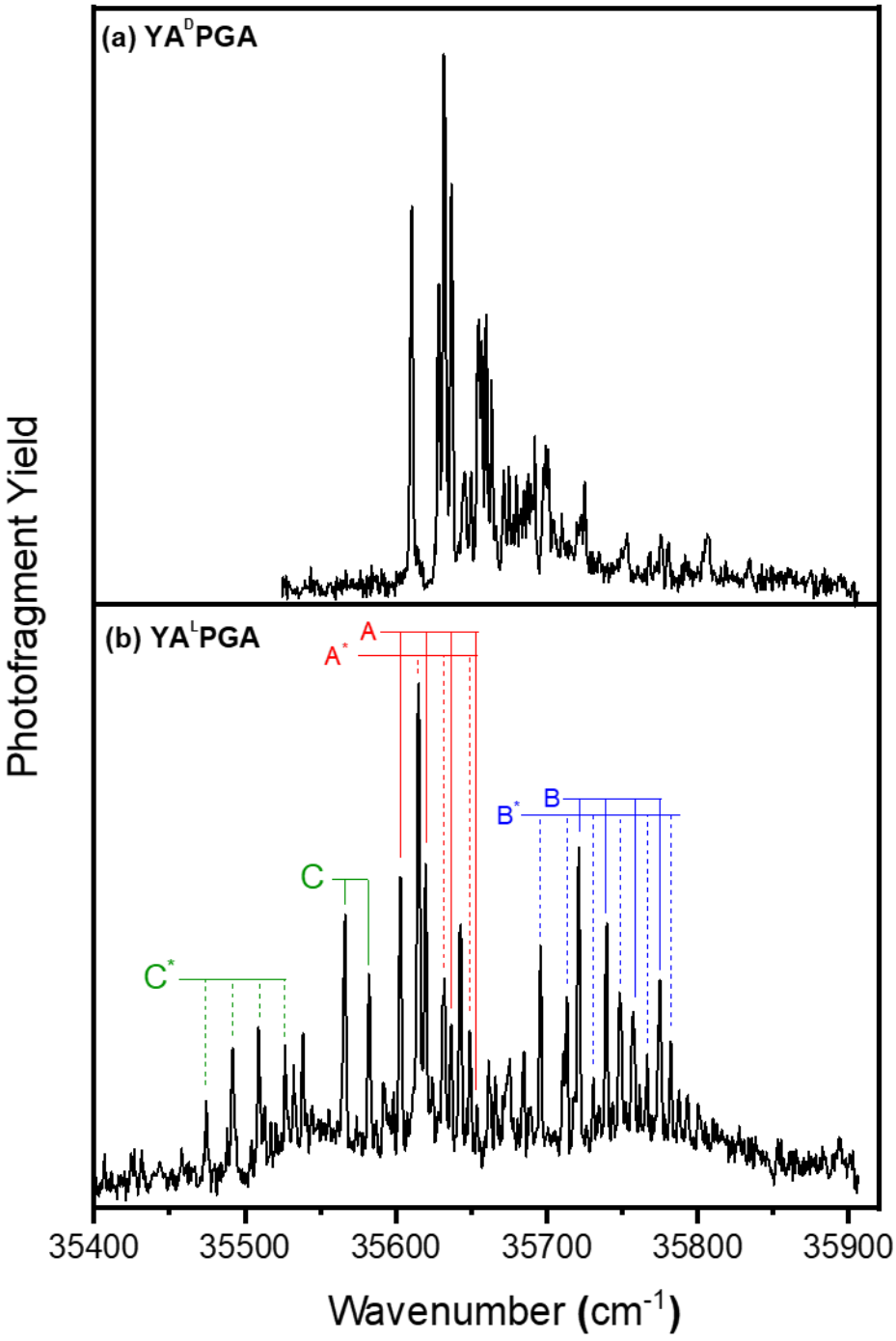

Figure 11 UV action spectra of (a) $[YA^{D}PGA+H]^{+}$ and (b) $[YA^{L}PGA+H]^{+}$. The origins and Franck-Condon progressions of the different conformers are labelled A, B, and C with the asterisk(*) representing tyrosine rotamers of the conformer.

Cold UV action spectra of $[YA^{D}PGA + H]^{+}$ and $[YA^{L}PGA+ H]^{+}$ were recorded by monitoring their complete photofragmentation signal as a function of laser wavelength (fig. 9). The striking difference in the complexity of the two UV spectra is evident even upon a cursory examination. $[YA^{D}PGA + H]^{+}$ produces a simple UV spectrum with a single major Franck-Condon progression whereas, the UV spectrum of $[YA^{L}PGA+ H]^{+}$ is far more complex containing multiple unique transitions that point to multiple conformations.

To ascertain the identity of the conformations present in these spectra, IR-UV double resonance techniques were utilized. Assignment of the transitions in the UV spectra were completed by first obtaining conformation specific infrared spectra of the most intense peaks in the spectra. Conformation specific UV spectra were then obtained by fixing the IR laser on a transition unique to each conformer and scanning the UV laser while concurrently taking a spectrum with the IR laser off. These spectra are then subtracted from one another to produce a conformation specific UV spectrum. The resulting spectra are provided in the supporting information.

3.4.2 Conformation-Specific Infrared Spectrum of YA^{D}PGA

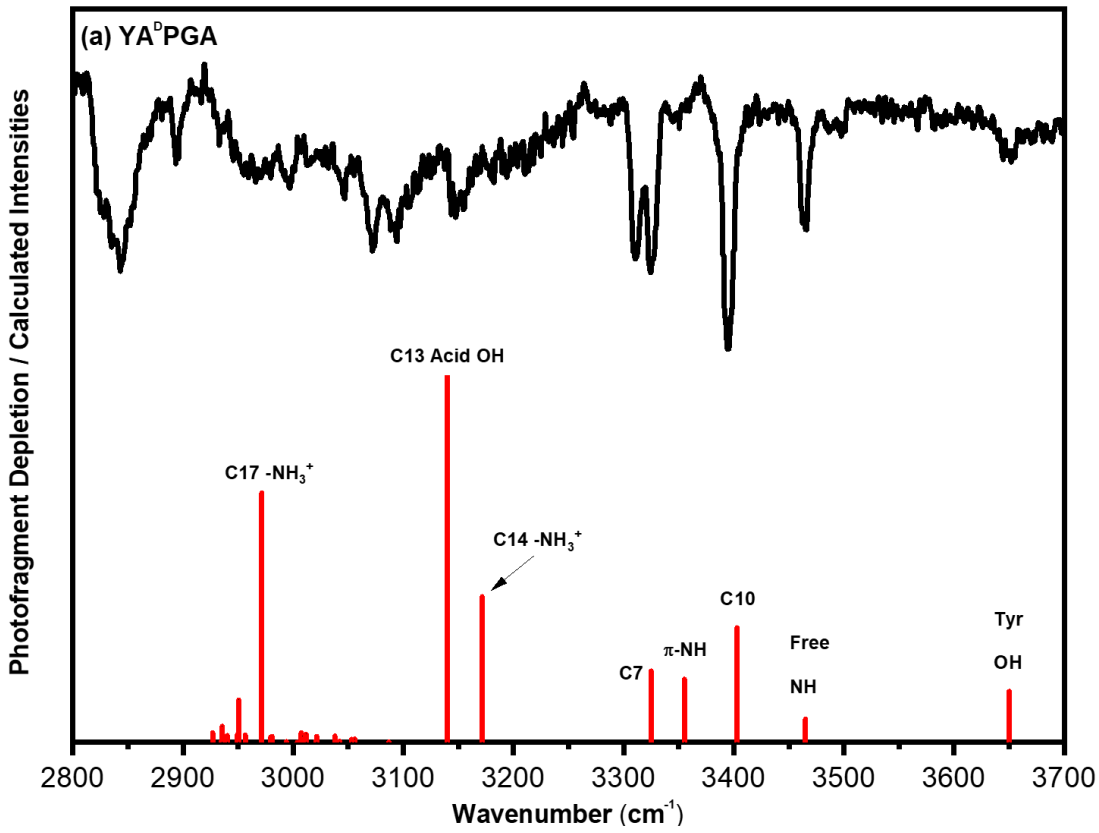

Figure 12 Conformation specific infrared spectrum of the single conformer of [YA^{D}PGA + H]$^{+}$ (black) in the hydride stretch region, compared to the calculated vibrational frequencies and infrared intensities of the best-fit conformation, which is also the global minimum.

Figure 12 presents the conformation-specific infrared spectrum of the single witnessed conformer of [YA^{D}PGA + H]$^{+}$. Below the experimental spectrum is the calculated lowest energy conformer which duplicates the fingerprint pattern of the experimental spectra exemplary. The NH/OH stretching region (3200-3700 cm^{-1}) holds an abundance of information about the hydrogen bonding pattern of a molecule. With [YA^{D}PGA + H]$^{+}$, the OH stretching region shows a characteristic Tyr-free OH stretch near 3648 cm^{-1} and lacks an Acid-free OH stretch at 3571 cm^{-1}. The presence of the Tyr-free OH is expected as the chromophore is too bulky/sterically hindering to interact with a peptide of this size. More informative is the lack of the Acid-free OH stretch at 3571 cm^{-1}, this points to the carboxyl terminus of the molecule being involved in a strong hydrogen bond which shifts the stretch to a position lower in energy in the spectrum.

Upon examination of the calculated lowest energy structure, the Tyr-free OH in the calculated spectrum aligns perfectly with the experimental and the Acid-free OH stretch is absent.

The free NH stretching region (3400-3500 cm^{-1}) of the experimental spectrum contains one strong peak at 3465 cm^{-1}, while the bound stretching region contains three intense peaks. The first and most intense at 3394 cm^{-1} followed by two closely spaced peaks at 3324 and 3310 cm^{-1}. This conformer must therefore contain one free NH stretch and three hydrogen bound NH stretches. In the calculated structure the free NH is associated with the backbone amide of the alanine residue between tyrosine and proline. The amide carbonyl oxygen of this alanine is held in a tight C10 hydrogen bond with the glycine residue opposite it, this transition appears at 3394 cm^{-1}. Of the remaining hydrogen bound NH modes are an NH-π interaction at 3324 cm^{-1} and at 3310 cm^{-1} a C7 hydrogen bond between the amide carbonyl of the glycine residue and the amide NH of the last alanine residue.

Lower in energy than the NH/OH stretching region is the CH stretching region which extends from 2800 – 3200 cm^{-1}. The presence or absence of broad bands in this region has been previously associated with strong hydrogen bonds of the acid OH and charged N-terminal amine. The conformer specific infrared spectrum of $[YA^{D}PGA + H]^{+}$ displays very broad transitions in this region implying the existence of hydrogen bonds to either the Acid OH, charged amine, or both. Analysis of the calculated spectrum reveals a C14 hydrogen bond between the acid OH and the proline amide carbonyl oxygen. The broad distribution also contains two hydrogen bonds to the charged N-terminus. The highest energy of which forms a C14 bond with the amide carbonyl of the c-terminal alanine while the lower energy transition forms a C17 hydrogen bond with carbonyl oxygen of the carboxylic acid on the c-terminus.

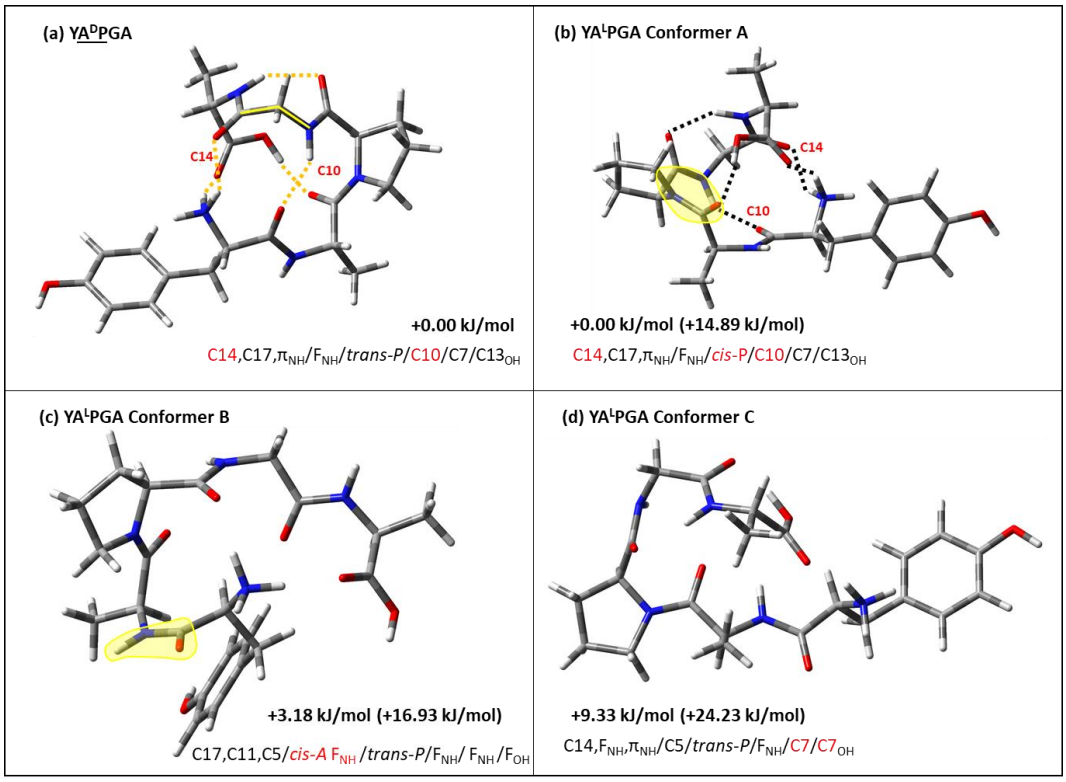

Figure 13 Assigned conformations of (a) $[YA^{D}PGA+H]^{+}$, and (b-d) the three assigned conformers of [YALPGA+H]+. A summary of the H-bonding architecture in terms of the size H-bonded rings formed, starting from the N-terminus NH3+. Relative energies are given comparing within a given diastereomer, with the comparison across diastereomers given in parentheses.

3.4.3 Conformation-Specific Infrared Spectra of YA^{L}PGA

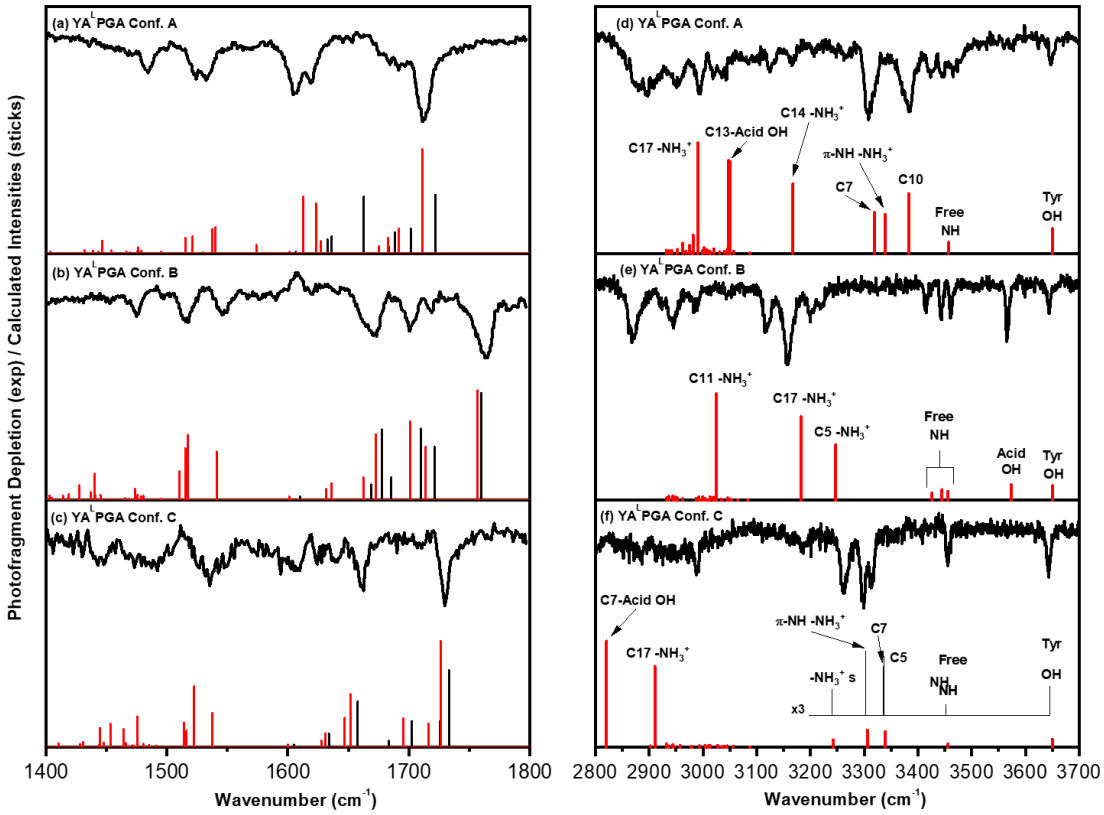

Figure 14 (a-c): Corresponding experimental and calculated spectra in the 1400-1800 cm^{-1} region. The black stick spectra are the results of a harmonic treatment using a single scale factor for all vibrations in the 1600-1800 cm-1 region, while the red stick spectra those resulting from the local mode Hamiltonian model. (d-f): Conformation specific infrared spectrum of the three observed conformers of [YA^{L}PGA + H]$^{+}$ (black) in the hydride stretch region, compared to the calculated vibrational frequencies and infrared intensities of the best-fit conformations.

Conformer A

The infrared spectrum of conformer A of [YA^{L}PGA+ H]$^{+}$ strikingly resembles that of [YA^{D}PGA + H]$^{+}$ (fig. 12). Comparison of the lowest energy calculated structure for [YA^{L}PGA+ H]$^{+}$ to the experimental spectrum magnifies this similarity. The acid OH is shifted to lower energy, as in [YA^{D}PGA + H]$^{+}$, down to around 3000 cm^{-1} where it is seen as making a C13 hydrogen bond with the carbonyl oxygen of the n-terminal alanine residue. Also, akin to YA^{D}PGA is the NH stretching region where the established pattern is echoed. A free NH stretch

at 3445 cm^{-1} belonging to the n-terminal alanine is followed by a set of three transitions. A C10 hydrogen bond between the glycine NH and the tyrosine carbonyl oxygen appears at 3385 cm^{-1}. Near this transition is both a C7 hydrogen bond (3307 cm^{-1}) between the proline carbonyl and the NH of the c-terminal alanine group and an NH-π interaction (3320 cm^{-1}) between the tyrosine ring and the charged amine. The protonated amine also forms two more hydrogen bonds. The first, a C14 hydrogen bond, with the glycine carbonyl group and the second forms a C17 with the carbonyl of the c-terminus.

Conformer B

Appearing in stark contrast to the complex hydrogen bonding network of conformer A is the lack of hydrogen bonding seen in conformer B. This is first observed in the OH stretching region where along with the expected free-Tyr OH stretch (3848 cm^{-1}) is an unexpected free acid OH stretch at 3566 cm^{-1}. More stunning is a trio of free NH stretches (3460,3443, and 3416 cm^{-1}) corresponding to every backbone amide NH in the molecule being uninvolved in hydrogen bonding. The only stabilization through hydrogen bonding observed with this conformer is due to the charged amine burying itself within the c-terminal half of the molecule. A weak C5 hydrogen bond between the charged amine and the adjacent carbonyl group on the tyrosine residue is the highest energy of the three charge involved hydrogen bonds. The other two hydrogen bonds are similar to what was noticed with both conformer A and $YA^{D}PGA$. A higher energy C17 hydrogen bond to the c-terminal carbonyl and a C11 hydrogen bond to the proline carbonyl. The eccentricity of this structure is hypothesized to be due to the presence of a cis-amide bond between the tyrosine and alanine residues adjacent to the proline. This motif has previously been observed in solution phase structures of small peptides and occurs in increased frequency in amide bonds adjacent to proline residues. To our knowledge this is the first non-proline cis-amide conformation discovered by cold ion spectroscopy.

Conformer C

Conformer C stands as a middle ground in hydrogen bond complexity between conformer A and C. The acid OH is bound as in conformer A but appears far lower in energy. A single free NH (3454 cm^{-1}) appears in the spectrum, where unlike conformer A, this free NH belongs to the

glycine NH. The NH stretching region contains the ever-present NH-π interaction (3320 cm^{-1}), but also incorporates three transitions not observed in the previous conformers. A set of closely spaced transitions belonging to C7 and C5 hydrogen bonds are recognized adjacent to the NH-π transition. The C7 cycle is formed from the c-terminal alanine NH and the proline carbonyl, whereas the C5 is formed from the n-terminal alanine NH interacting with its carbonyl oxygen. A pseudo-C5 hydrogen bond is formed by the symmetric stretch of the charged amine and the adjacent carbonyl on the tyrosine residue. The only charged amine, hydrogen bound macrocycle is a C17 hydrogen bond to the carbonyl on the c-terminus. This exact motif is present in all three conformers of $YA^{L}PGA$.

3.5 Discussion

3.5.1 Formation of a Gas Phase Type II β-Turn in $[YA^{D}PGA + H]^{+}$: Comparison to $[YA^{D}PAA + H]^{+}$ and $[YG^{L}PAA + H]^{+}$

In a previous study it was hypothesized that the removal of bulky, sterically hindering groups around the proline residue would increase flexibility and therefore better induce type II' beta turn formation. The presence of C10 and C14 hydrogen bonds and a trans amide bond about the proline residue are strong pieces of evidence for turn formation in $[YA^{D}PGA + H]^{+}$ as these motifs are expected to induce β-turn formation. The type of turn can only be determined through examination of the phi and psi angles forming the turn. Inspection of the Ramachandran angles points to a type II beta turn and not the expected type II' turn. This result is inconsistent with the notion that the ^{D}PX moiety locks in a type II' beta turn and therefore induces β-hairpin formation. To better elucidate a possible explanation for type II beta turn formation it is helpful to compare the structure and spectrum of $[YA^{D}PGA + H]^{+}$ to previously reported spectra of $[YA^{D}PAA + H]^{+}$ and $[YG^{L}PAA + H]^{+}$ both of which formed type II' beta turns(fig 4).[32]

As the type II' beta turn is the inverse turn of the type II turn, the spectroscopic signals of each should be similar. This is evident when comparing conformer B of $[YA^{D}PAA + H]^{+}$ to $[YA^{D}PGA + H]^{+}$. Both contain the expected spectroscopic signal of a turn, C10 and C14 hydrogen bonds, but the arrangement of these bonds both structurally and spectroscopically differ. The first "rung" of the β-hairpin is the C10 hydrogen bond that forms the β-turn. Placement of the proline residue in the turn that forms this hydrogen bond differs between the two molecules. In $[YA^{D}PAA + H]^{+}$ the proline residue lies in the classical i+1 position, putting it

near the front of the turn where it anchors the formation of the turn. Whereas, in $[YA^{D}PGA + H]^{+}$ the proline finds itself in the i+2 position near the end of the turn. The relative position of the A_2 methyl group in relation to the P_3 methylene was postulated to influence turn formation as the molecule contorts itself to minimize the interaction of these two groups. Configuring the proline residue into the i+2 position of the turn curtails this steric effect as the distance between the groups is maximized while still maintaining the secondary structure.

Minimization of this steric strain in $[YA^{D}PAA + H]^{+}$ required a "kink" in the peptide chain as this was the only way to minimize the effects of the A_2 and A_4 methyl groups while keeping the turn. This kink was stabilized by a C11 hydrogen bond between the protonated amine and the proline carbonyl oxygen. Lacking the A_2 methyl group allowed $[YG^{L}PAA + H]^{+}$ to form this kink around the A_4 methyl group effectively minimizing its influence on the structure. This required an interaction between the same proline carbonyl oxygen, but not to the amine, instead it formed a C13 hydrogen bond with the carboxylic acid OH. As $[YA^{D}PGA + H]^{+}$ retains the A_2 methyl group but gains the flexibility of a G_4 hydrogen atom, one would expect this kink to form about the A_2 methyl to reduce its interaction with the rest of the molecule. Instead, the kink unexpectedly forms a C13 hydrogen bond about the G_4 residue much like it did in the case of A_4 for $[YG^{L}PAA + H]^{+}$. Having proline in the i+2 position reduces the need for the molecule to contort the A_2 methyl group away from the proline residue meaning the kink seen in $[YA^{D}PAA + H]^{+}$ is unneeded in $[YA^{D}PGA + H]^{+}$. Instead the c-terminal end of the peptide kinks towards the middle of the structure allowing it to hydrogen bond with both the proline carbonyl oxygen as well as the charged amine.

With the infrared spectra of $[YA^{D}PGA + H]^{+}$ and $[YG^{L}PAA + H]^{+}$ being nearly identical the expectation would be that the structures are also nearly identical. This is surprisingly true across the greater portion of both molecules with the hydrogen bonding pattern of each roughly indistinguishable. Bond angles of $[YG^{L}PAA + H]^{+}$ align it with a type II' beta turn, exactly the opposite of the type II turn formed by $[YA^{D}PGA + H]^{+}$. Both molecules fly in the face of the expected result that the ^{D}PX motif will give type II' turns while ^{L}PX should give a type II turn. The difference in the two molecules being the proline isomerization and the location of the glycine residue, it may be that the presence, not location, of the glycine has a significant effect on β-turn formation. Allowing the molecule to adjust itself such that the steric issues on one end of the turn are relieved by the other end appears necessary to preferentially forming a type II/II'

turn. To better understand why $[YA^{D}PGA + H]^{+}$ formed a type II turn and not the expected type II' turn, its diastereomer pair, $[YA^{L}PGA + H]^{+}$, was spectroscopically interrogated with the expectation of finding the inverse turn(fig 5).

3.5.2 Conformational Complexity of [YALPGA + H]+

Conformer A: The "almost" β-turn

The infrared spectrum of conformer A of $[YA^{L}PGA + H]^{+}$ is decidedly analogous to that of $[YA^{D}PGA + H]^{+}$, this spectral similarity is reflected in the structure of the molecules. Conformer A retains the proline in the i+2 position of the "turn" as a means of reducing steric strain against the A_2 methyl group. While also maintaining the kink in the chain about G_4 to form the C13 hydrogen bond between the carboxylic acid OH and the proline carbonyl. The charge site anchors the back end of the molecule in two ways. First, it forms a C14 hydrogen bond with the G_4 carbonyl which composes the second rung of what would be a β-hairpin. The charged amine also forms a strong C17 hydrogen bond with the A_5 carboxylic acid carbonyl. This hydrogen bonding pattern keeps the back end of the molecule compact, keeping the charge site solvated and leaving the portion of the peptide involved in the turn free from its direct influence.

Conformer A presented the best opportunity at forming the type II' beta turn analogue to the type II turn formed by $[YA^{D}PGA + H]^{+}$. The C10 bond formed between the Y_1 carbonyl and the X_4 amide NH has been a stable indicator of type II/II' β-turn formation in our past spectroscopic experiments involving turns. Conformer A is the first evidence of this C10 hydrogen bond that does not result in the type II turn or its inverse. In trying to understand how the same hydrogen bonding pattern could lead to a different structure it is important to probe the effect of the stereochemistry change of the proline. The difference in these structures lie in the contortion of the backbone that occurs around $^{L}P_3$, in conformer A, which twists in such a way as to destroy the bond angles necessary to classify the turn as a type II β-turn. The bond angles instead point to a type IV β-turn, a catch all term that contains all turns that have the hydrogen bonding pattern of a β-turn but lack the specific angles of the other numbered turns. The dissimilarity to $[YA^{D}PGA + H]^{+}$ in this region is expected as the difference in stereochemistry between the proline residues forces the molecule to contort itself to maintain the same hydrogen bonding pattern.

This difference is best visualized by looking down the long axis of both molecules and focusing on the bonds projecting out of the proline residue(fig 6). In $[YA^{D}PGA + H]^{+}$ the arms of the molecule project out of the same side of the proline ring, the effect of this positioning is to better align the backbone amide groups to make tighter, more stable hydrogen bonds. To achieve this same hydrogen bonding stability in conformer A the arms of the molecule cross over the long axis of the molecule, twisting over the proline. Projecting the arms straight out of the proline ring, as in $[YA^{D}PGA + H]^{+}$, would still allow for the formation of the C10 and C14 hydrogen bonds of the turn. The net benefit of this twist is the ability to form the C13 hydrogen bond between the carboxylic acid OH and the proline carbonyl.

Conformer C: A γ-turn in a β-turn world

The real complexity in the conformational landscape of $[YA^{L}PGA + H]^{+}$ begins to unfold with the structure of conformer C. This conformer is the first to present without C10 and C14 hydrogen bonds which immediately disqualifies it from the possibility of having a β-turn. Instead, the bulk of the hydrogen bonding occurs on the c-terminal end of the proline residue with two C7 hydrogen bonds. The first C7 hydrogen bond forms between the P_3 carbonyl and the A_5 amide NH. The dihedral angles of the amide bond about this hydrogen bond, $\varphi = -79$ $\psi = 72$, point to a γ-turn. A second C7 cycle is made by the carboxy terminus orienting itself in a trans configuration which grants the acid OH group bonding access to G_4 carbonyl group. The trans contour of the c-terminus is stabilized by a C17 hydrogen bond between its carbonyl and the charged amine. This C17 macrocycle is interesting as it is a familiar motif seen in all of the other conformers but presents the only hydrogen bond between the charged amine and the rest of the molecule for this conformer. In conformer A of $[YA^{L}PGA + H]^{+}$, as well as with the single conformation of $[YA^{D}PGA + H]^{+}$, the c-terminus of the molecule was formed in the cis position. This allowed the acid OH to form a C13 hydrogen bond to the A_2 carbonyl which twists the c-terminal end of the peptide. This twist exposes the G_4 carbonyl to the charged amine leading to the formation of the C14 hydrogen bond marking the second rung of the β-turn. The trans acid seen in conformer C disrupts this bonding pattern instead leading to the aforementioned C7 hydrogen bonding motif.

Contrary to the complexity of the c-terminal end of the proline residue, the n-terminal end is very linear and exhibits few interactions with the opposite end of the molecule. Two

hydrogen bonds are observed on this part of the molecule. The first is a partial C5 bond built upon an intra-amino acid interaction between the A_2 NH and carbonyl oxygen. The previously mentioned C17 between the charged amine and the c-terminal carbonyl being the other hydrogen bond. Typically, the charge site has a large effect, positive or negative, on the structure of small molecules. So far, the charge site has shown a positive effect by anchoring the ends of the molecule together and forming a C14 hydrogen bond reminiscent of a β-turn. In this conformer the charge is positioned away from the bulk of the molecule reducing its ability to interact with and form the hydrogen bonding pattern that helped lock in the turn. A NH-π interaction between the charged amine and the tyrosine phenol ring acts to stabilize the charge.

Conformer B: Formation of a cis-amide bond not involving Proline

The conformer specific infrared spectrum of conformer B proved quite complex in its simplicity. It presented with several transitions in the free NH and OH portions of the spectrum which indicated very little hydrogen bonding occurring in the 3D structure. This seemed odd as both conformers A and C as well as the single conformer of $[YA^{D}PGA + H]^{+}$ exhibited several hydrogen bound transitions. Also noteworthy was the lack of a broad background in the CH stretching region which has historically been associated with hydrogen bonding between the charge site and the rest of the molecule. Instead, sharp transitions with little baseline noise are observed. The complicated infrared spectrum is mirrored in the complexity of the structure. Conformer B has an intricate structure that is formed via the very strong impact the charge has on the molecule. The effect of the charge is so strong as to overcome any influence the proline may have on locking in a β-turn. Instead, a weak C7 and strong C11 hydrogen bonds forming about the proline residue act to turn the molecule about itself. Facilitating the position of the charge in the center of the molecule is the presence of a rarely observed cis amide bond between Y_1 and A_2. The aggregate of these unique molecular substructures is a conformation that defies expectation.

All three conformers of $[YA^{L}PGA + H]^{+}$ have their structures in some way mediated by the charge site. In conformer A the charge maintains a positive influence by helping form the last rung of the turn. While in conformer C, the charge is relatively isolated and impacts the structure minimally. Conformer B shows what negative influence charge can have on the molecule by burying into a nest of carbonyl groups and forcing itself into the heart of the structure. The

charged amine hydrogen bonds with the rest of the molecule at sites all along the backbone with the effect of it wrapping the structure around itself. It starts with a small C5 hydrogen bond to its neighboring carbonyl on the Y_1 residue. This bonding motif is common of these small peptides as the tyrosine carbonyl is typically not involved in a hydrogen bond with the rest of the molecule. Instead, it is far more influenced by the strong pulling effect of the nearby charge site and the energetic reduction in removing the steric influence of the tyrosine phenol ring from the rest of the molecule. Where it is found hydrogen bound, as in conformer A, this C5 motif remains as a testament to the energetic favorability of this motif.

Common to all of the conformers of $[YA^{L}PGA + H]^{+}$ is a C17 hydrogen bond between the charged amine and the c-terminal carbonyl of the A_5 residue. This pattern is once again present in conformer B. Unlike the other conformers, the acid OH on the opposite end of the C17 is not found in a hydrogen bond. To accommodate the interaction of the charge with the carbonyl of the c-terminus the acid OH is set apart form the rest of the molecule. The charged amine also forms a C11 macrocycle with the proline carbonyl. In the other conformers of $[YA^{L}PGA + H]^{+}$ as well as the single conformer of $[YA^{D}PGA + H]^{+}$ this carbonyl is found in a C7 hydrogen bond with the A_5 NH group. The C11 hydrogen bond forms what would be the second rung of a γ-turn originating from a very weak C7 hydrogen bond about the proline residue. As this ultra-tight turn is typically not centered around proline, the strain of this interaction leads to dihedral angles that are far from the typical γ-turn.

This unique charge dominated hydrogen bonding pattern would not be energetically likely without the presence of a cis amide bond between the Y_1 and A_2 residues. To test this conclusion the exact same structure was built in Gaussian with the exception of turning the amide bond in question from a cis to a trans configuration. After optimization and energy calculations were completed the trans amide structure was found to be 34.7 kJ/mol higher in energy than its cis counterpart. This large relative energy makes it unlikely that this conformation is populated in the reduced energy environment of the cold trap. The question that then remains is why the cis amide structure is so low in energy, +3.18 kJ/mol above global minimum, compared to the trans isomer. Both structures are very similar on the c-terminal end of the proline residue maintaining both the C11 and C17 bonds to the charged amine. Hydrogen bond distances are smaller in the cis amide isomer, implying stronger hydrogen bonds. The structures significantly differ around the proline residue where flipping the Y_1-A_2 amide bond to

the trans formation led to an isomerization of the proline amide bond to the cis configuration. Changing the orientation of the proline amide bond spun the A_2 carbonyl group away from the G_4 NH removing the weak C7 hydrogen bond which formed the γ-turn. What is gained by isomerization to a trans Y_1-A_2 amide bond is a weak C11 hydrogen bond from the A_2 NH to the G_4 carbonyl. The small amount of stabilization afforded by this bond is dwarfed by the energetic gains of tighter hydrogen bonds to the charge site, a C7 γ-turn, and a trans proline amide bond.

3.5.3 Energy Level Diagrams for $[YA^{L}PGA + H]^{+}$

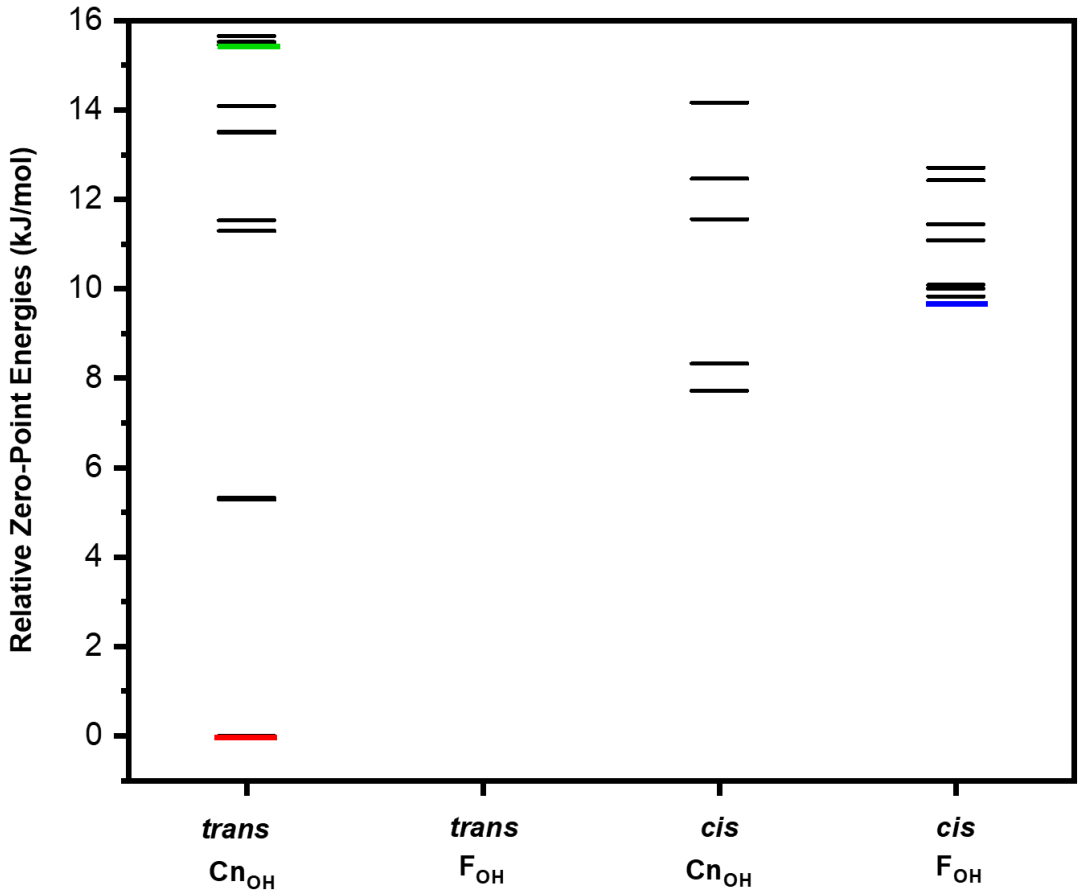

Figure 15 Relative energies at 0 K, including zero-point energy corrections for the conformers of $[YA^{L}PGA + H]^{+}$. The x-axis divides the structures into groups based upon the presence or absence of a *cis* amide bond in the molecule (*cis/trans*) and whether the acid OH is in a hydrogen bond (Cn_{OH}) or not (F_{OH}). Each line in the figure represents the energy of a calculated structure while the energies of the three conformers that have best fits to the observed conformers are shown in red (conformers A), blue (conformer B), and green (conformer C).

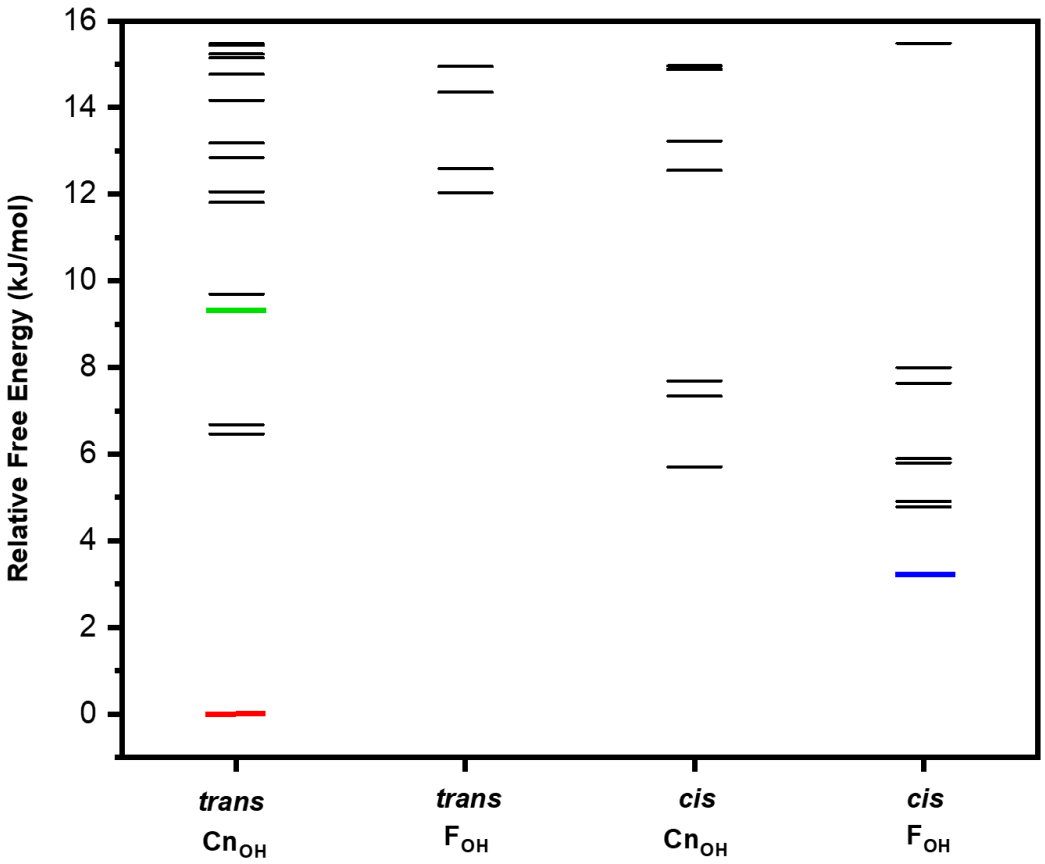

Figure 16 Gibbs Free Energy(ΔG 298K) level diagram for $[YA^{L}PGA + H]^{+}$ with relative energy on the y-axis. The x-axis breaks the structures down into groups based upon the presence or absence of a cis amide bond in the molecule (*cis/trans*) and whether the acid OH is in a hydrogen bond (Cn_{OH}) or not (F_{OH}). Each line in the figure represents the energy of a calculated structure, with those labeled in red (conformer A), blue (conformer B), and green (conformer C) mark the energies of the best fit structures to each infrared spectrum.

Conformers B and C proved to be troublesome structures to obtain by our normal structural screening methods. The presence of a free acid OH transition appearing at 3566 cm^{-1} along with three transitions in the free NH region of the IR spectrum provided quite a challenge for determining the structure of conformer B of $[YA^{L}PGA + H]^{+}$ as a low energy structure with several non-hydrogen bound groups was unexpected. It was only after expanding the conformational search algorithm to rotate around the full circumference of the amide bond dihedral angles that conformer B could accurately be assigned. Conformer C has a higher relative energy compared with the other two conformers and has six conformers between it and conformer B. To better understand this energy difference and the complex conformational landscape of $[YA^{L}PGA + H]^{+}$ energy level diagrams were constructed.

Figure 16 shows a plot of the Gibbs free energies (ΔG, 298K) against the set of components that best explains the energy differences between these structures. A corresponding energy level diagram at 0 K which was built using zero-point corrected energies can be found in the Supporting Information. As with our previous studies of $[YA^{D}PAA + H]^{+}$ and $[YG^{L}PAA + H]^{+}$ better agreement was found between calculated and experimental spectra of structures based on thermal free energy at 298 K than the zero-point corrected energies at 0K. The first and major component is the presence of a cis or trans amide bond in the molecule. Inclusion of structures containing cis amide bonds into the conformational search was not expected to affect the low energy conformations as this configuration is considered rare in nature. In reality, cis amide containing structures dominate the low energy portion of the conformational landscape of $[YA^{L}PGA + H]^{+}$. Other than the global minimum, conformer A, the next six of seven low energy structures all contain a cis amide bond between Y_1-A_2. The lowest energy of the cis amide containing structures was confidently assigned as conformer B. The second component used to differentiate these structures was the presence or absence of a hydrogen bond with the acid OH group. In the previous study of $[YA^{D/L}PAA + H]^{+}$ conformers containing a trans amide bond about the proline residue and having a hydrogen bound acid OH group dominated the conformational scene. This trend is also observed in $[YA^{L}PGA + H]^{+}$ with the majority of conformers being present in this state. Cis amide containing structures interestingly do not follow this general guideline. Instead, both free and hydrogen bound populations are relatively equally distributed energetically.

The prevalence of cis amide containing structures partially explains why conformer C has several conformers between it and the global minimum. Common to all of the Y_1-A_2 cis amide structures is the presence of several free NH stretches. The infrared spectrum of conformer C has only one transition firmly in the free NH region of the spectrum, making it unlikely that this structure belongs to any of the cis amide containing compounds. Several of the cis amide incorporated structures also have a free acid OH stretch that is absent in the experimental spectrum of conformer C which removes these structures from consideration. What remains between the assigned structure and the global minimum is a single structure with its Y_1-A_2 amide bond in the trans configuration. This molecule has all of its NH groups bound and has no calculated intensities near the free NH transition in the experimental spectrum at 3454 cm^{-1}.

3.5.4 Peptide length and β-turn formation: The study of $[YAA^{D}PGAAA+H]^{+}$

Facilitating the energetics of the β-turn motif is the indirect role the charge site plays in stabilizing the structure. Where the proline anchors the turn end of the molecule, the charged amine keeps the ends together and plays a vital role in forming the C14 hydrogen bond underpinning the second rung of the turn. The single conformation spectroscopy of $[YA^{L}PGA + H]^{+}$ showed that charge can either stabilize the formation of the classical β-turn or project itself into the center of the molecule and destabilize the compact secondary structure. By extending the peptide chain on either side of the proline residue we hope to elucidate the effect of charge and chain length on the stabilization of the turn. To this end the single charged peptide $[YAA^{D}PGAAA+H]^{+}$ was subjected to IR-UV double resonance spectroscopy.

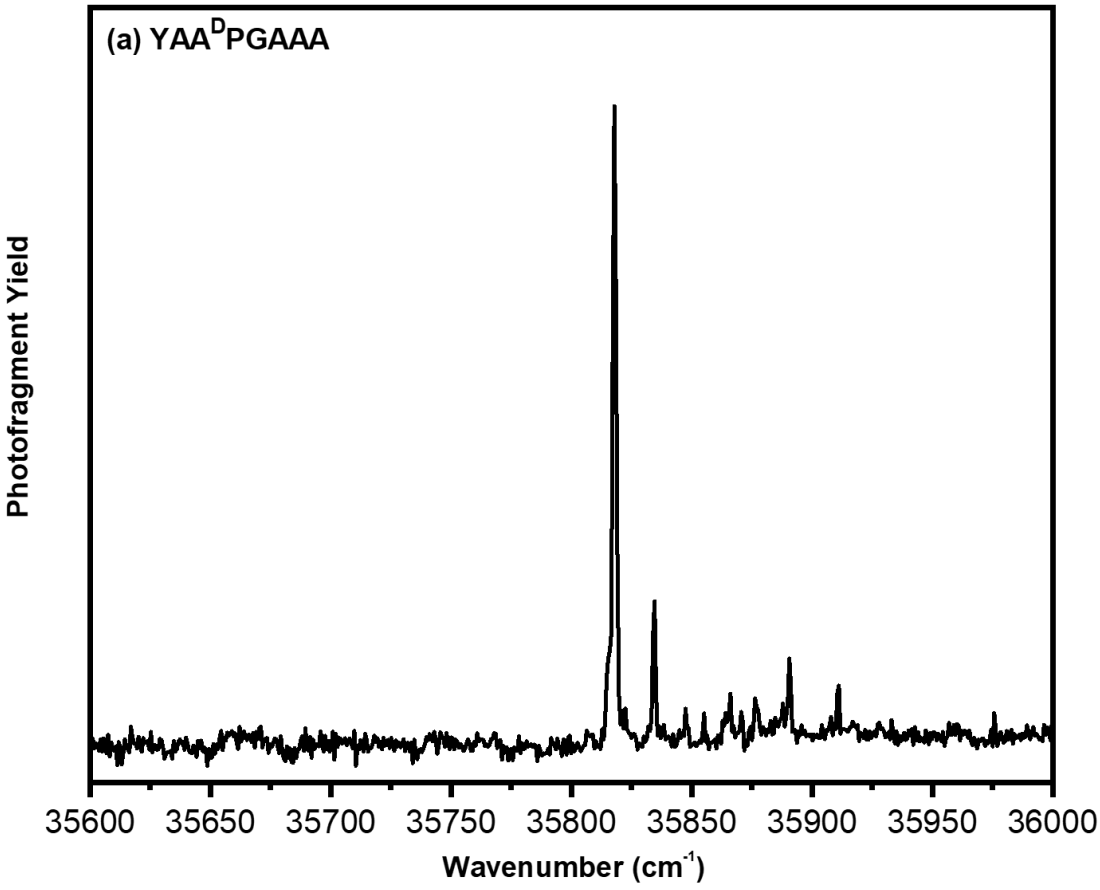

Figure 17 UV photofragment spectrum of cryo-cooled $[YAA^{D}PGAAA+H]^{+}$ in the S_0-S_1 origin region.

The UV photofragmentation spectrum of the tyrosine chromophore of $[YAA^{D}PGAAA+H]^{+}$ is remarkably simple with a single UV origin at 35818 cm^{-1} followed by smaller vibronic bands (fig 8). The simplicity of this spectrum was expected as the shorter chained $[YA^{D}PGA+H]^{+}$ also presented with a simple UV spectrum and a single conformer. The origin of $[YAA^{D}PGAAA+H]^{+}$ is shifted ~ 200 cm^{-1} higher in energy than its smaller chain counterpart reflecting greater interaction between the phenol ring of tyrosine and the structure.

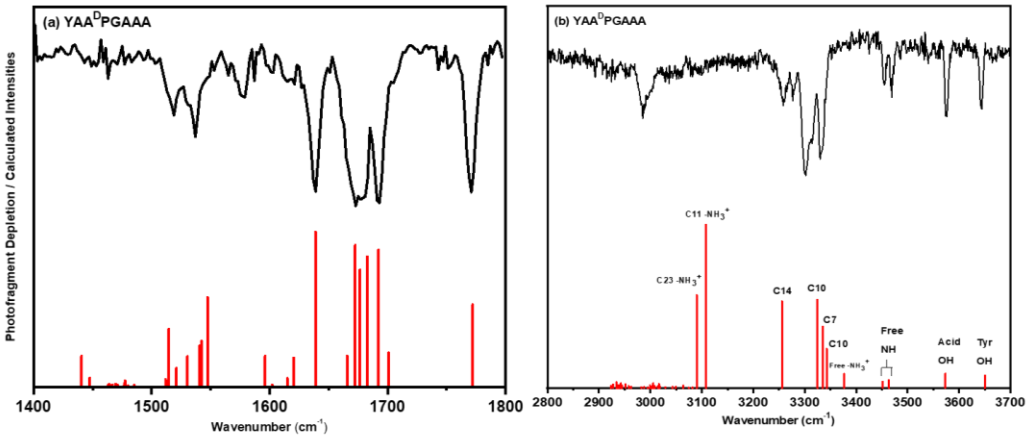

Figure 18 (b) Conformation specific infrared spectrum of the single observed conformer of $[YA^{L}PGA + H]^{+}$ (black) in the hydride stretch region, compared to the calculated vibrational frequencies and infrared intensities of the best-fit conformation, which is also the global minimum. (a): Corresponding experimental and calculated spectra in the 1400-1800 cm^{-1} region using the local mode Hamiltonian model.

The single conformer infrared spectrum is extraordinarily simple given the size of the molecule and fits well to the global minimum (ΔG, 298K) calculated experimental spectrum of the type II β-turn like the turn motif observed with $[YA^{D}PGA+H]^{+}$ (fig 9,supp 4). A few noticeable differences between the two structures are present in the experimental spectrum. Beside the expected free tyrosine OH is an unexpected acid OH stretch at 3576 cm^{-1}. In its smaller counterpart $[YA^{D}PGA+H]^{+}$, the acid OH was found in a C13 bond to the A_2 carbonyl oxygen. Instead this carbonyl is locked in a C11 hydrogen bond with the charge site. Keeping this carbonyl in a hydrogen bond with the charge orients it away from the G_5 backbone NH which effects the structure in two ways. Firstly, it prevents a C7 hydrogen bond that would force a γ-turn like that seen with conformer C of $[YA^{L}PGA+H]^{+}$. It also helps aid in the formation of

the type II β-turn by making it more likely that the G_4 NH will form a C10 hydrogen bond with the A_2 carbonyl.

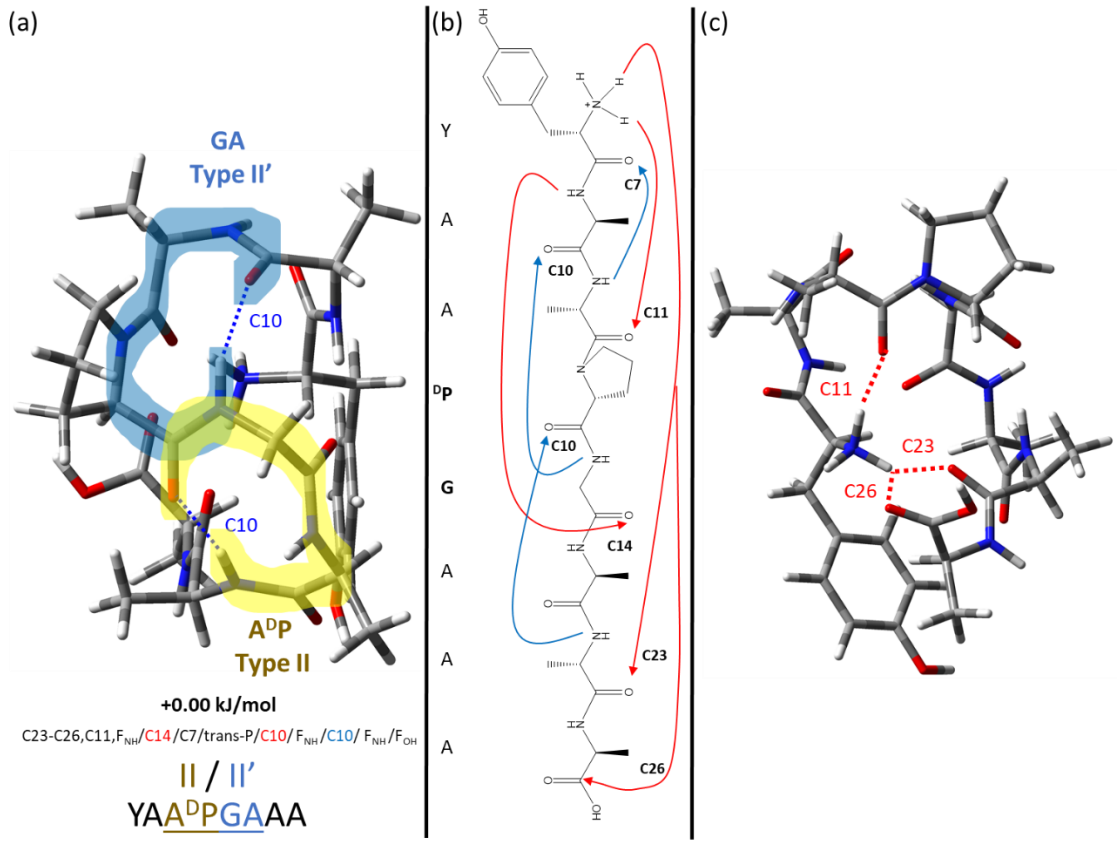

Figure 19 (a) Calculated lowest energy structure of $[YAA^{D}PGAAA + H]^{+}$ showing the sequential β-turns from a side view. The individual turns are highlighted in blue and yellow. (b) Stick structure with highlighted hydrogen bonding pattern. A red colored arrow represents a N-C terminal hydrogen bond while a blue arrow is indicative of the opposite. (c) Top-down view of the structure highlighting the binding pocket surrounding the NH_3^+.

Important to answering the question of whether or not the turn is stabilized by increasing the chain length is the presence of an additional rung to the β-sheet like structure that was not seen in the smaller molecule. The first rung of the molecule is the C10 hydrogen bond formed about the P_3 residue that forms the turn itself. Rung 2 is composed of a C14 macrocycle between the A_2 NH and G_5 carbonyl. In $[YA^{D}PGA+H]^{+}$, the G_5 carbonyl was locked into a similar hydrogen bond with the charged amine. As there is no additional carbonyl on the n-terminal side of the peptide to form the traditional third rung of the sheet, the charged amine is the only group remaining capable of forming a pseudo-third rung. It is expected then that this macrocycle, a

C20, would form between the charge and the A_6 carbonyl. Instead, a C23 is built between the charged amine and the A_7 carbonyl because the A_6 carbonyl is faced away from the charge site due to its amide bound partner, the A_7 NH, being in a hydrogen bond. This C10 hydrogen bond is composed of the aforementioned A_7 NH and the P_4 carbonyl. Recall that in $[YA^{D}PGA+H]^{+}$ the P_3 carbonyl was stabilized by a C7 hydrogen bond with the A_5 NH group. The C14 between the A_2 NH and the G_5 carbonyl pivots the G_5-A_6 amide linkage in such a way that forming a C7 bond is unlikely. This leaves the proline carbonyl open to forming the next biggest stable macrocycle, a C10, with the nearby A_7 NH. The overarching effect of these additional hydrogen bonds is a structure that is stabilized about the turn.

The remaining question to probe is the effect of the charge and whether its influence on the formation of the turn been negated by increasing the peptide chain length. Its enforcement of the turn via formation of a C11 hydrogen bond with the A_3 carbonyl and its construction of the third rung in the sheet, effects that have been previously discussed, provide ample evidence that the charge site does indeed still play a role in stabilizing the turn. This leads to a new question of how many amino acid residues need to be between the charge site and the turn for its effect to be adequately negated. Given a similar hydrogen bonding motif, it may be that the introduction of another alanine residue between Y_1 and A_2 would be all that is necessary to remove the charge. This would put it out of hydrogen bonding range of the A_3 carbonyl and may allow the more traditional third rung of the sheet to form.

3.6 Conclusion

Cold ion spectroscopy was used to probe the conformational proclivity of ^{L}P and ^{D}P diastereomers with a flexible glycine residue directly c-terminal to the proline. With $[YA^{D}PGA+H]^{+}$, an unexpected type II β-turn was the single conformer found. This structure is the inverse of the type II' β-turn assigned previously to one of the conformations of $[YA^{D}PAA+H]^{+}$. $[YA^{L}PGA+H]^{+}$ presented with three very different conformers starting with a structure very near to its diastereomer analogue in conformer A. The two remaining conformers were the γ-turn of conformer C and the globular structure containing a non-proline cis amide bond in conformer B. The flexibility provided by the G_4 residue and the influence of the charge site are the expected culprits in the formation of these interesting structures. Lastly, with $[YAA^{D}PGAAA+H]^{+}$ the charged amine was moved further from the site of the turn to reduce the

scale of its effect on the stabilization of the structure. While it still maintained some effect on the structure, early signs of β-sheet formation were witnessed.

3.7 References

1. Li, J.; Kuang, Y.; Gao, Y.; Du, X.; Shi, J.; Xu, B., D-amino acids boost the selectivity and confer supramolecular hydrogels of a nonsteroidal anti-inflammatory drug (NSAID). *J Am Chem Soc* **2013,** *135* (2), 542-5.

2. Feng, Z.; Xu, B., Inspiration from the mirror: D-amino acid containing peptides in biomedical approaches. *Biomol Concepts* **2016,** *7* (3), 179-87.

3. Grishin, D. V.; Zhdanov, D. D.; Pokrovskaya, M. V.; Sokolov, N. N., D-amino acids in nature, agriculture and biomedicine. *Frontiers in Life Science* **2019,** *13* (1), 11-22.

4. Novick, P. A.; Lopes, D. H.; Branson, K. M.; Esteras-Chopo, A.; Graef, I. A.; Bitan, G.; Pande, V. S., Design of beta-amyloid aggregation inhibitors from a predicted structural motif. *J Med Chem* **2012,** *55* (7), 3002-10.

5. Ohide, H.; Miyoshi, Y.; Maruyama, R.; Hamase, K.; Konno, R., D-Amino acid metabolism in mammals: biosynthesis, degradation and analytical aspects of the metabolic study. *J Chromatogr B Analyt Technol Biomed Life Sci* **2011,** *879* (29), 3162-8.

6. Shi, J.; Du, X.; Yuan, D.; Zhou, J.; Zhou, N.; Huang, Y.; Xu, B., D-amino acids modulate the cellular response of enzymatic-instructed supramolecular nanofibers of small peptides. *Biomacromolecules* **2014,** *15* (10), 3559-68.

7. Lee, J.; Lee, D. G., Structure-antimicrobial activity relationship between pleurocidin and its enantiomer. *Exp Mol Med* **2008,** *40* (4), 370-6.

8. Wallace, B. A., Common structural features in gramicidin and other ion channels. *BioEssays* **2000,** *22* (3), 227-234.

9. Oliva, R.; Chino, M.; Pane, K.; Pistorio, V.; De Santis, A.; Pizzo, E.; D'Errico, G.; Pavone, V.; Lombardi, A.; Del Vecchio, P.; Notomista, E.; Nastri, F.; Petraccone, L., Exploring the role of unnatural amino acids in antimicrobial peptides. *Sci Rep* **2018,** *8* (1), 8888.

10. Luo, Z.; Zhao, X.; Zhang, S., Structural dynamic of a self-assembling peptide d-EAK16 made of only D-amino acids. *PLoS One* **2008,** *3* (5), e2364.

11. Garton, M.; Nim, S.; Stone, T. A.; Wang, K. E.; Deber, C. M.; Kim, P. M., Method to generate highly stable D-amino acid analogs of bioactive helical peptides using a mirror image of the entire PDB. *Proc Natl Acad Sci U S A* **2018,** *115* (7), 1505-1510.

12. Martin, E. B.; Williams, A.; Richey, T.; Wooliver, C.; Stuckey, A.; Foster, J. S.; Kennel, S. J.; Wall, J. S., Evaluation of the effect of D-amino acid incorporation into amyloid-reactive peptides. *J Transl Med* **2017,** *15* (1), 247.

13. Espinosa, J. F.; Gellman, S. H., A Designedβ-Hairpin Containing a Natural Hydrophobic Cluster. *Angewandte Chemie International Edition* **2000,** *39* (13), 2330-2333.

14. Fisk, J. D.; Gellman, S. H., A parallel beta-sheet model system that folds in water. *J Am Chem Soc* **2001,** *123* (2), 343-4.

15. Fisk, J. D.; Powell, D. R.; Gellman, S. H., Control of Hairpin Formation via Proline Configuration in Parallel β-Sheet Model Systems. *Journal of the American Chemical Society* **2000,** *122* (23), 5443-5447.

16. Stanger, H. E.; Gellman, S. H., Rules for Antiparallel β-Sheet Design: d-Pro-Gly Is Superior tol-Asn-Gly for β-Hairpin Nucleation1. *Journal of the American Chemical Society* **1998,** *120* (17), 4236-4237.

17. Fisk, J. D.; Schmitt, M. A.; Gellman, S. H., Thermodynamic analysis of autonomous parallel beta-sheet formation in water. *J Am Chem Soc* **2006,** *128* (22), 7148-9.

18. Walsh, P. S.; Blodgett, K. N.; McBurney, C.; Gellman, S. H.; Zwier, T. S., Inherent Conformational Preferences of Ac-Gln-Gln-NHBn: Sidechain Hydrogen Bonding Supports a beta-Turn in the Gas Phase. *Angew Chem Int Ed Engl* **2016,** *55* (47), 14618-14622.

19. Fu, H.; Grimsley, G. R.; Razvi, A.; Scholtz, J. M.; Pace, C. N., Increasing protein stability by improving beta-turns. *Proteins* **2009,** *77* (3), 491-8.

20. de Brevern, A. G., Extension of the classical classification of beta-turns. *Sci Rep* **2016,** *6*, 33191.

21. Kabsch, W.; Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983,** *22* (12), 2577-637.

22. Richardson, J. S., The anatomy and taxonomy of protein structure. *Adv Protein Chem* **1981,** *34*, 167-339.

23. Baruah, P. K.; Sreedevi, N. K.; Gonnade, R.; Ravindranathan, S.; Damodaran, K.; Hofmann, H. J.; Sanjayan, G. J., Enforcing periodic secondary structures in hybrid peptides: a novel hybrid foldamer containing periodic gamma-turn motifs. *J Org Chem* **2007,** *72* (2), 636-9.

24. Némethy, G.; Printz, M. P., The γ Turn, a Possible Folded Conformation of the Polypeptide Chain. Comparison with the β Turn. *Macromolecules* **1972,** *5* (6), 755-758.

25. Gibbs, A. C.; Bjorndahl, T. C.; Hodges, R. S.; Wishart, D. S., Probing the structural determinants of type II' beta-turn formation in peptides and proteins. *J Am Chem Soc* **2002,** *124* (7), 1203-13.

26. Kopple, K. D.; Go, A.; Pilipauskas, D. R., Studies of peptide conformation. Evidence for beta structures in solutions of linear tetrapeptides containing proline. *J Am Chem Soc* **1975,** *97* (23), 6830-8.

27. Trevino, S. R.; Schaefer, S.; Scholtz, J. M.; Pace, C. N., Increasing protein conformational stability by optimizing beta-turn sequence. *J Mol Biol* **2007,** *373* (1), 211-8.

28. 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 beta-Hairpin. *J Am Chem Soc* **2016,** *138* (8), 2849-57.

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.

30. Boyarkin, O. V., Cold ion spectroscopy for structural identifications of biomolecules. *International Reviews in Physical Chemistry* **2018,** *37* (3-4), 559-606.

31. Garand, E.; Kamrath, M. Z.; Jordan, P. A.; Wolk, A. B.; Leavitt, C. M.; McCoy, A. B.; Miller, S. J.; Johnson, M. A., Determination of noncovalent docking by infrared spectroscopy of cold gas-phase complexes. *Science* **2012,** *335* (6069), 694-8.

32. 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 beta-Hairpin Turn. *J Am Chem Soc* **2017,** *139* (15), 5481-5493.

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.

34. Dean, J. C.; Burke, N. L.; Hopkins, J. R.; Redwine, J. G.; Ramachandran, P. V.; McLuckey, S. A.; Zwier, T. S., UV photofragmentation and IR spectroscopy of cold, G-type beta-O-4 and beta-beta dilignol-alkali metal complexes: structure and linkage-dependent photofragmentation. *J Phys Chem A* **2015,** *119* (10), 1917-32.

35. Pereverzev, A. Y.; Boyarkin, O. V., Exploring the relevance of gas-phase structures to biology: cold ion spectroscopy of the decapeptide neurokinin A. *Phys Chem Chem Phys* **2017,** *19* (5), 3468-3472.

36. 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* (1), 014304.

37. 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.

38. Londry, F. A.; Hager, J. W., Mass selective axial ion ejection from a linear quadrupole ion trap. *Journal of the American Society for Mass Spectrometry* **2003,** *14* (10), 1130-1147.

39. Mohamadi, F.; Richards, N. G. J.; 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. *Journal of Computational Chemistry* **1990,** *11* (4), 440-467.

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

CHAPTER 4. CHARACTERIZING TETHERED PEPTIDE ALPHA HELICES VIA COLD ION SPECTROSCOPY OF MODEL COMPOUNDS

4.1 Abstract

Model tethered peptides that form highly stable single turn alpha helices in the solution phase were interrogated by IR-UV double resonance as gas phase ions. The stability of these structures in the solution phase did not transfer to the gas phase as no single turn helix was observed. Protonation of the n-terminus on $[Y[KAAAD]\text{-}NH_2 + H]^+$ resulted in three similar conformations where large macrocycles predominated. Acetylation of the n-terminus and insertion of an arginine residue as a charge site proved fruitless in $[Ac\text{-}YR[KAAAD]\text{-}NH_2 + H]^+$ as the charged guanidinium group wove itself into the heart of the molecule and displaced the hydrogen bonds needed to form the helix. Lastly, repositioning the charged residue to the c-terminal end of the molecule in $[Ac\text{-}Y\ [KAAAD]R\text{-}NH_2 + H]^+$ showed promise as a portion of the molecule presented as a 3_{10} helix, alas the helix unwound itself in the second half of the molecule as the charge site reoriented the hydrogen bonding framework towards itself.

4.2 Introduction

Biological means of controlling disease have come into greater focus over the last decade as scientists obtain a better understanding of the structure-function relationship of proteins[1-2]. Key to protein interaction in ailments such as Creutzfeldt-Jakob and Human Papilloma Virus (HPV) infection is the presence of binding sites that recognize small alpha helical segments acting as switches that turn on/off enzyme activity[3-5]. Inhibition of protein interactions by small molecule mimics is seen as a promising avenue for establishing a cure. To this end, synthetic chemists have attempted to form small (<4 residues) alpha helices in solution that act as antagonists to the larger disease-causing proteins. This pathway has been met with limited success due in part to the entropic cost of forming a single turn alpha helix in solution[6-7]. Hydrogen bonding to surrounding water molecules being far more favorable than the formation of intramolecular hydrogen bonds.

One of the most effective ways of producing these small, single-turn alpha helices has been to form a lactam tether between the side chains of lysine and aspartic acid residues, and the

addition of a C-terminal primary amine in place of the carboxylic acid[6, 8]. Fairlie and coworkers have been able to synthesize these highly stable alpha helices and in NMR studies have shown their strong propensity for alpha helicity[6]. The durability of these peptides was tested along a range of different temperatures, pH, solvent types, and denaturants and the stability of the helices remained[8]. The incredible stability of these helices in solution begs the question as to whether that stability will be transferred into the gas phase as ions where the lack of intermolecular interactions with solvent may lead to different folding pathways.

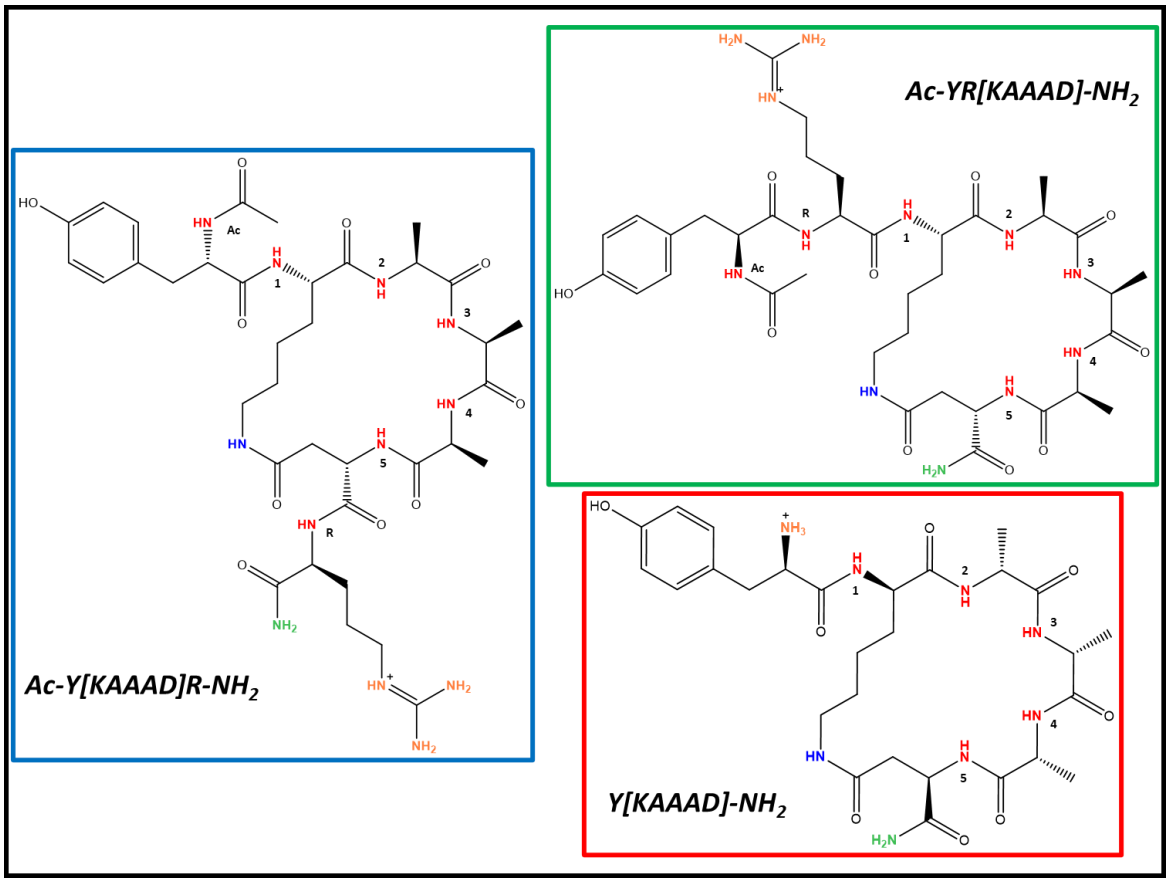

Figure 20 Stick structures of the Y[KAAAD]-NH_2, YR[KAAAD]-NH_2, and Y[KAAAD]R-NH_2 Backbone NH oscillators are labelled in red. "Ac" refers to the NH oscillator of the acetylated n-terminus, numbers above the backbone NH groups correlate to transitions in the infrared spectra, "R" refers to the NH oscillator of the backbone NH belonging to the arginine residue. Charge site NH groups are labelled in orange while the tether NH is in blue and the c-terminal NH_2 cap is labelled in green. This color coordination is maintained when analyzing the DFT produced stick spectra below.

The propensity of non-tethered peptides, both long and short, to form gas phase alpha helices has been previously investigated[9-10]. Jarrold used Ion Mobility Mass Spectrometry (IM-MS) to elucidate the folding preferences of $[Ala_{19}+H]^+$ and $[Ac\text{-}Ala_{19}\text{-}Lys+H]^+$ [9]. He concluded that if the charge site was located on the N-terminus a globular structure resulted. By contrast, the presence of a C-terminal charged lysine residue induced formation of an alpha helix due in part to the positive charge interacting with the macrodipole of the helix and the stabilizing effect of its hydrogen bonding to three nearby carbonyl groups. Russell and colleagues used IM-MS to study larger peptides (12-30 residues) with multiple charge carrying sites in an attempt to understand the factors that form helices[10]. They found that peptides consisting of poly-(AAKAA) sequences correlated to structures with partial helices whereas, peptides with poly-(AEAAKA) sequences formed only charge solvated globules.

Spectroscopic signatures of poly-alanine alpha helices containing a C-terminal lysine residue were identified by Rizzo and coworkers[11]. Utilizing IR-UV double resonance techniques they were able to identify two conformers for both $[Ac\text{-}Phe\text{-}(Ala)_5\text{-}Lys+H]^+$ and $[Ac\text{-}Phe\text{-}(Ala)_{10}\text{-}Lys+H]^+$. The conformer specific infrared spectra for each of these compounds led to a few conclusions regarding the IR spectra of helices. Transitions above 3380 cm-1 are generally narrow and weaker which is indicative of free or weakly H-bonded N-H bonds. While transitions below 3380 cm-1, being more intense and broad, are associated with strong hydrogen bonding of varying strength to N-H groups.

The aim of this study is first to elucidate whether single-turn alpha helices can be transferred into the gas phase as ions. Through the utilization of conformer specific IR-UV double resonance techniques the unique structures of these ions can be deduced and the intramolecular forces holding them together can be identified. The second goal is to identify the infrared spectral signatures that single turn helices produce, both in the XH stretching region as well as the Amide I and II regions. Three molecules that form single turn alpha helices in the solution phase were studied as protonated ions in the gas phase, $Y[KAAAD]\text{-}NH_2$, $Ac\text{-}YR[KAAAD]\text{-}NH_2$, and $Y[KAAAD]R\text{-}NH_2$, as a means to achieving these goals.

4.3 Methods

4.3.1 Experimental

All spectroscopic data were taken on a custom-built instrument for cold ion spectroscopy, previously described[12]. In brief, the design of the instrument comprises a triple quadrupole mass spectrometer in which an orthogonal spectroscopy axis was added between the second (q2) and third (q3) quadrupoles. Ions are formed via nano-electrospray ionization(nESI) and transported into q2 where the ion of interest is isolated via RF/DC isolation. The ion packet is then steered around a turning quadrupole and down the spectroscopy axis where it is trapped in the cold quadrupole that is cooled to 5K via a closed-cycle helium cryostat (Sumitomo Heavy Industries, Tokyo, Japan). After cooling to ~10K via collisions with the He buffer gas, the ions are spectroscopically interrogated via infrared (Laservision OPO/OPA) and ultraviolet (ScanMate-Pro Lambda Physik frequency doubled by an Inrad Autotracker III and Scanmate Lambda Physik) lasers. The created photofragments are then extracted back down the spectroscopy axis and trapped in q3 for analysis in either “spectroscopy mode” or “mass spectrometry mode”. In mass spectrometry mode, a mass spectrum of the photofragments is generated via mass-selective axial ejection[13]. Spectroscopy mode entails ejection of the remaining precursor ions with SX wave[14] software followed by propelling of the photofragments onto a channeltron detector (4773G, Photonis USA).

UV action spectra are recorded by monitoring the total photo-fragment signal as a function of laser wavenumber. Upon obtaining a UV spectrum and identifying transitions that may be due to different ion conformations, conformation-specific IR spectra are taken by employing IR-UV double resonance techniques. The UV laser is tuned to a transition of interest in the UV spectrum which creates a steady signal of photofragment ions. Then a tunable IR laser that is spatially overlapped but temporally precedes the UV laser by 200 ns is scanned in wavelength. Upon absorption of an IR photon by the conformer of interest, a fraction of the population is removed from the ground state which results in a depletion in the total photofragment signal.

4.3.2 Solution Preparation

Solutions of Y[KAAAD]-NH2, Ac-YR[KAAAD]-NH2, and Ac-Y[KAAAD]R-NH2 were prepared as 50:50 Methanol:Water to a concentration of 200 μM and electrosprayed to create the corresponding protonated ion.

4.3.3 Computational

Conformational searches were performed for $[Y[KAAAD]\text{-}NH_2 + H]^+$,$[Ac\text{-}Y[KAAAD]R\text{-}NH_2 + H]^+$, and $[Ac\text{-}YR[KAAAD]\text{-}NH_2 + H]^+$ structures using the Monte Carlo multiple minimum method with the AMBER* and OPLS3 forcefields inside the MACROMODEL[15] software package. Approximately 1000 structures were found between the two forcefields for each of the molecules. To narrow down the potential structures further, the 1000 structure pool was clustered based on backbone RMSD. The centroid of each of the clusters was then optimized at the B3LYP/6-31G*-GD3BJ level of theory using Gaussian16[16] and harmonic vibrational frequencies/intensities were calculated to obtain predicted infrared spectra to compare with experiment. To correct for anharmonicity the calculated spectra were scaled by 0.958 for the NH, CH, and OH stretching regions, 0.973 for the free OH transitions, 0.981 for Amide I, and 0.970 for Amide II. Adjusting the calculated transitions by these values puts them into agreement with previously calculated and fitted spectra. “Stick” spectra were generated for the centroid of each cluster. These calculated spectra were then matched to the experimental spectrum and the best fit centroids were found. All clusters of structures with best fit centroids that lie within 10 kJ/mol of the global minimum were then run through DFT at the same level of theory used for the centroids. After completing all calculations, the structure that best fit the experimental spectrum and was lowest in energy was determined to be the most likely structure of the ion.

4.4 Results and Analysis

4.4.1 UV Action Spectra

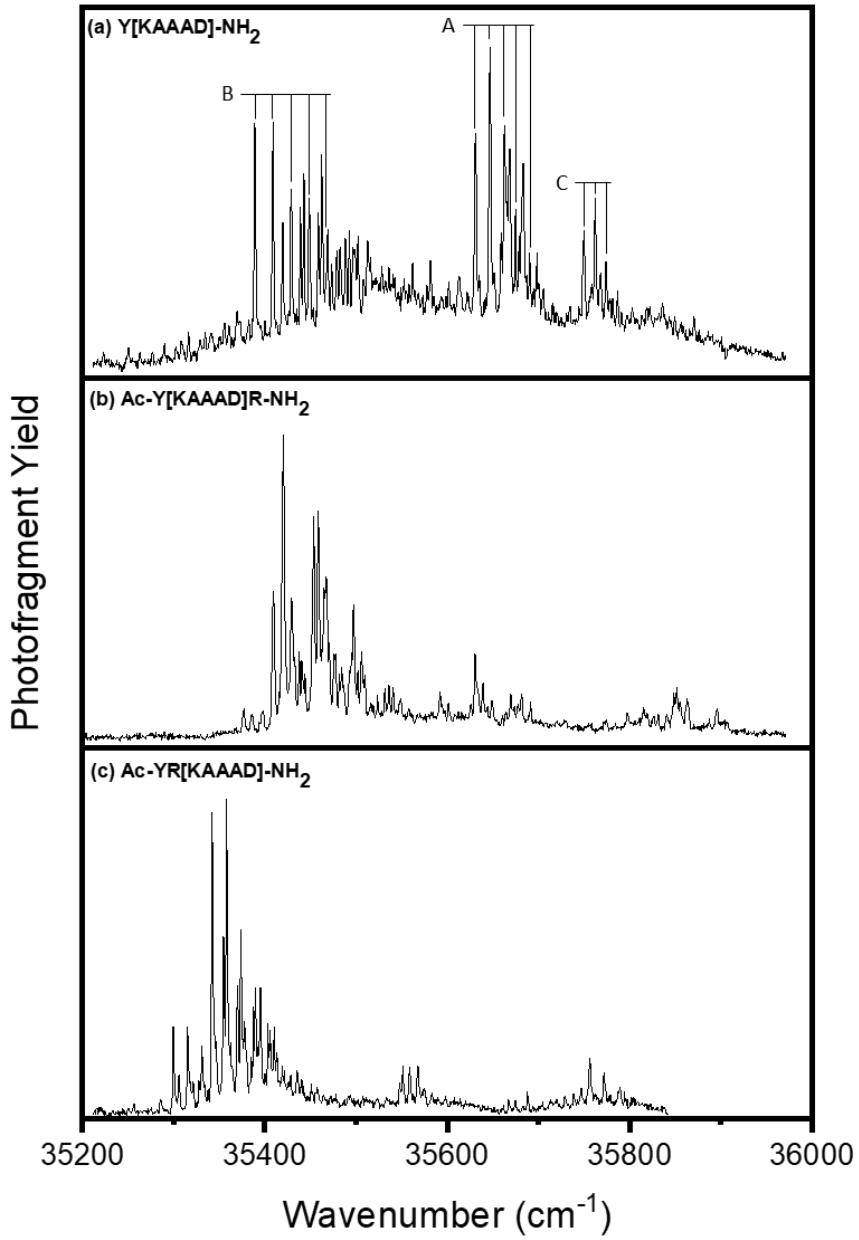

Figure 21 UV spectra of (a) $[Y[KAAAD]\text{-}NH_2 + H]^+$, (b) $[Ac\text{-}Y[KAAAD]R\text{-}NH_2 + H]^+$, and (c) $[Ac\text{-}YR[KAAAD]\text{-}NH_2 + H]^+$. The origins and Franck-Condon progressions of the different conformers are labelled A, B, and C in the UV spectrum of $Y[KAAAD]\text{-}NH_2$ (a).

Cold UV action spectra of $[Y[KAAAD]\text{-}NH_2]^+$ and $[Ac\text{-}Y[KAAAD]R\text{-}NH_2]^+$ were taken by recording total photofragment signal as a function of laser wavenumber. These are shown in Figures 21a and b, respectively. The sharp peaks and limited background noise indicate that the ions were fully cooled down to temperatures near 10 K before interrogation by the laser. The spectrum of $[Y[KAAAD]\text{-}NH_2]^+$ (Fig. 1a) reveals several transitions (marked A-C) with several unique Franck-Condon progressions, hinting at the potential presence of multiple conformers. Alternatively, the UV photofragment spectrum of $[Ac\text{-}Y[KAAAD]R\text{-}NH_2]^+$ each contains one central progression and some vibronic activity. To unravel the complexity of these spectra we utilize IR-UV double resonance techniques which allow for single conformers to be selected for and analyzed.

4.4.2 Conformation-Specific Infrared Spectra of $[Y[KAAAD]\text{-}NH_2 + H]^+$

The conformation specific infrared spectra of conformers A, B, and C of $[Y[KAAAD]\text{-}NH_2 + H]^+$ are presented in Figure 22. Immediately recognizable in all three conformers is the free Tyr-OH transition at 3648 cm^{-1}, as seen in the spectra of other ions containing a tyrosine residue.

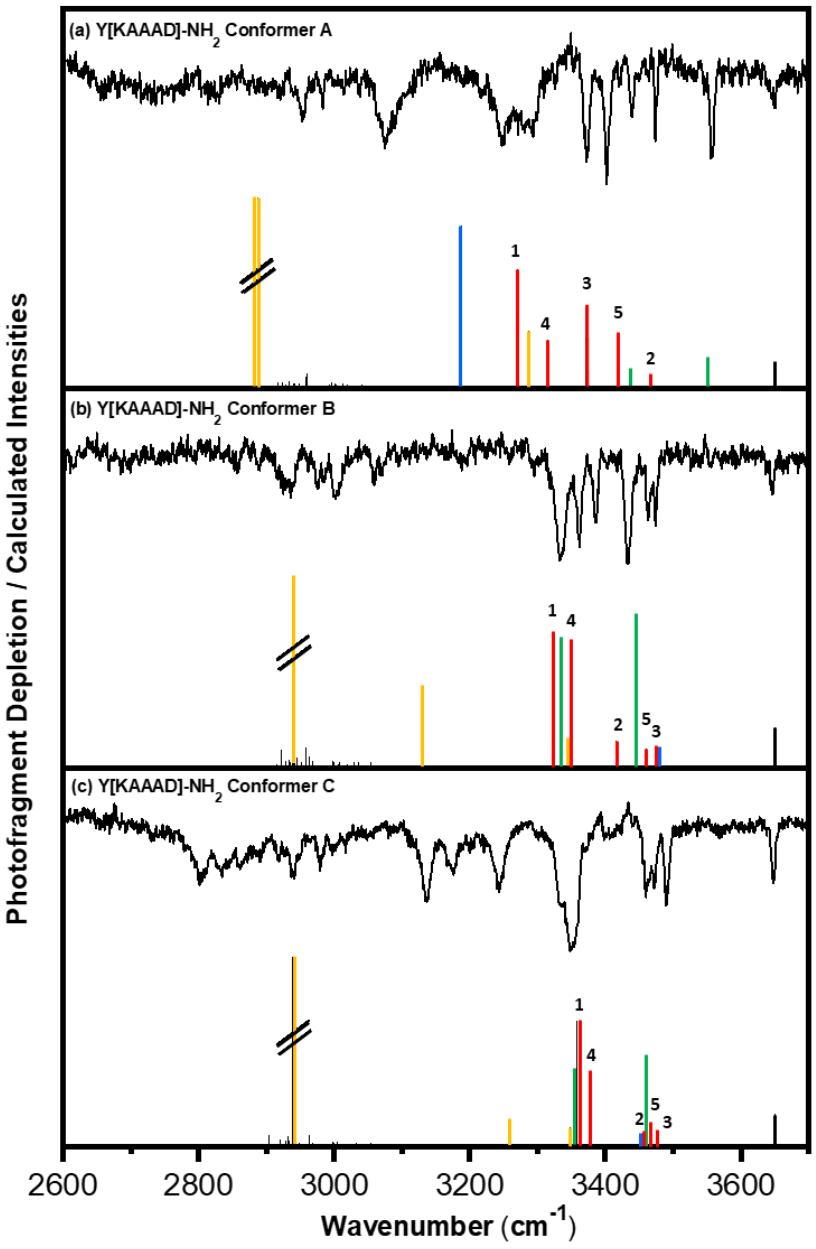

Figure 22 Conformation specific infrared spectra of the three conformations of [Y[KAAAD]-NH_2 + H]$^+$ (black) and best fit calculated stick spectra. Colors of each of the sticks in the calculated spectra correlate to the color scheme provided in figure 23.

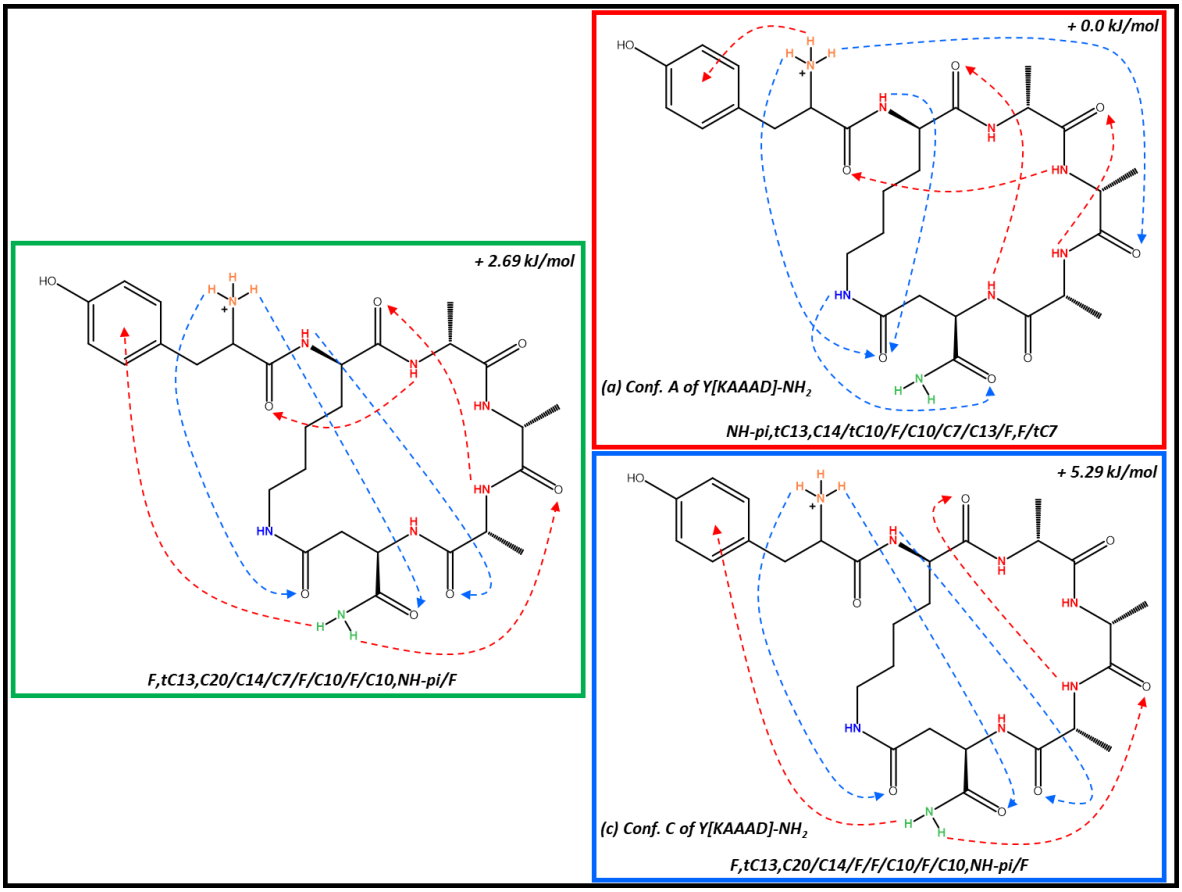

Figure 23 Stick structures of Y[KAAAD]-NH_2 conformers A, B, and C showing hydrogen bonds

Dashed arrows indicate hydrogen bonds with the arrow originating from the donor and terminating on the acceptor. Blue coloration denotes an N->C terminal hydrogen bond while red indicates a C->N terminal bond. Shorthand notation describing the presence or absence of hydrogen bonding at NH groups starting from the N-terminus is provided utilizing Cn notation where n equals the number of atoms contained by the macrocycle.

Conformer A

The experimental infrared spectrum of Conformer A reveals three free NH stretches in the region of 3400 to 3500 cm^{-1} while the 3200 to 3400 cm^{-1} region is packed with bound NH stretches of varying strength. Assigned to this spectrum is the global minimum of the observed conformational landscape, a structure that is filled with large cyclic hydrogen bonds (C10, C13) and does not resemble a helix. Highest in energy of the free NH stretches is the asymmetric stretch of the C-terminal NH_2 cap at 3556 cm^{-1} its partner symmetric stretch appearing near 3436 cm^{-1}. Remaining in the free NH stretch region is a sole transition (3476 cm^{-1}) belonging to the first alanine residue in the molecule, labelled as 2 in the spectrum.

The hydrogen bound NH stretch region provides a great depth of insight into the hydrogen bonding patterns that describe the secondary structure of the molecule. Analyze this region of the spectrum of conformer A and two distinct microregions emerge. Higher energy hydrogen bonds containing sharp transitions appear at 3372 and 3400 cm^{-1} which is ascribed to C10 and C13 hydrogen bonds of the A_2 and D_6 residues respectively.

The lower energy hydrogen bound region encompasses a clump of three broad transitions. Broadness in this region of the spectrum has been previously associated with the protonated amine. With the protonation site of this molecule containing a protonated amine it is therefore not surprising to see broad features in this region. The highest energy transition in this region belongs to a C7 hydrogen bond between the A_3 NH, labelled as 4, and the A_1 carbonyl. Lowest in energy of this clump of transitions is a C10 hydrogen bond from the NH of K_1 to the carbonyl oxygen of the tether. Between these two backbone transitions is an NH-π bond from the charge to the tyrosine ring which, as previously stated, is suspected to be the reason for the local broadening of the region. The protonated amine injects itself into the center of the molecule through the formation of a C13 hydrogen bond with the tether carbonyl and a C14 hydrogen bond with the A_2 carbonyl. The remaining peak in the spectrum, labelled in blue, coincides with a strong C7 hydrogen bond between the tether NH and the D_6 carbonyl.

Conformer B

Conformer B lies 2.69 kJ/mol higher in energy than the global minimum conformer A. Three small free NH stretches appear between 3450 and 3500 cm^{-1}. A doublet of free NH stretches on the higher energy end of the spectrum belong to the tether and the second alanine residue with the third and lowest energy NH stretch is possessed by the aspartic acid. The asymmetric stretch of the C-terminal NH_2 cap is shifted lower in wavenumber (3434 cm^{-1}) than its free counterpart in conformer A. It is positioned between the tyrosine ring and the A_2 carbonyl forming both NH-π and a C10 hydrogen bonds. Flanking the free stretch region is a weak C7 hydrogen bond between the NH of the first alanine residue and the carbonyl of the tyrosine. Several intense transitions appearing in the 3300-3400 cm^{-1} region signify strong hydrogen bonds. Highest in energy of these transitions is a C10 bond involving the A_3 NH and the lysine carbonyl. Lost in the forest of strong transitions in this region is a weakly intense free NH stretch coming from the protonated amine (3345 cm^{-1}). This small peak is encompassed by

the two remaining strong transitions in the is region. The first of these peaks being the symmetric stretch of the C-terminal cap which like its asymmetric version is shifted lower in wavenumber and the second belonging to a C14 hydrogen bond incorporating the lysine NH and the A_3 carbonyl. Of the two NH stretches remaining in the molecule both, expectedly, constitute the remaining NH stretches of the charged amine. The higher energy transition of the pair being in a C20 hydrogen bond to the aspartic acid carbonyl and the massively intense, lower energy member forming a C21 hydrogen bond to the carbonyl of the tether amide.

Conformer C

The infrared spectrum of Conformer C resembles that of Conformer B with its two sets of clustered transitions in the free and bound stretching regions. Congestion in the free stretching region of conformer C is greater than any of the previously seen conformations with five mostly weak transitions appearing within 20 cm^{-1} of each other. Three of backbone stretches take up residence in this region, in order of high to low energy, these include the NH group of A_2, D_5, and A_1. Shifted lower in energy than its free form, the asymmetric stretch of the C-terminal NH_2 cap is found at a similar wavenumber as that of conformer B. Interestingly, the cap is also in an analogous position as conformer B with one arm pointed towards the tyrosine ring forming an NH-π bond and the other forming a C10 hydrogen bond with A_2. A free NH stretch of the tether amide group rounds out the free stretching region. Structural similarities to conformer B continue in the hydrogen bonding region starting with the highest energy bond being a C10 macrocycle formed by the A_3 NH and the lysine carbonyl. A C14 hydrogen bond between the NH group of the lysine residue and the carbonyl group of A_3 lies 20 cm^{-1} lower in energy than the C10 cycle and is the lowest energy of the backbone NH transitions. Near the backbone C10 transition is the symmetric stretch of the c-terminal cap and the free NH stretch of the charged amine. The final two transitions, belonging to the charged amine, form both a C20 macrocycle to the aspartic acid carbonyl and a C21 cyclic hydrogen bond to the tether carbonyl.

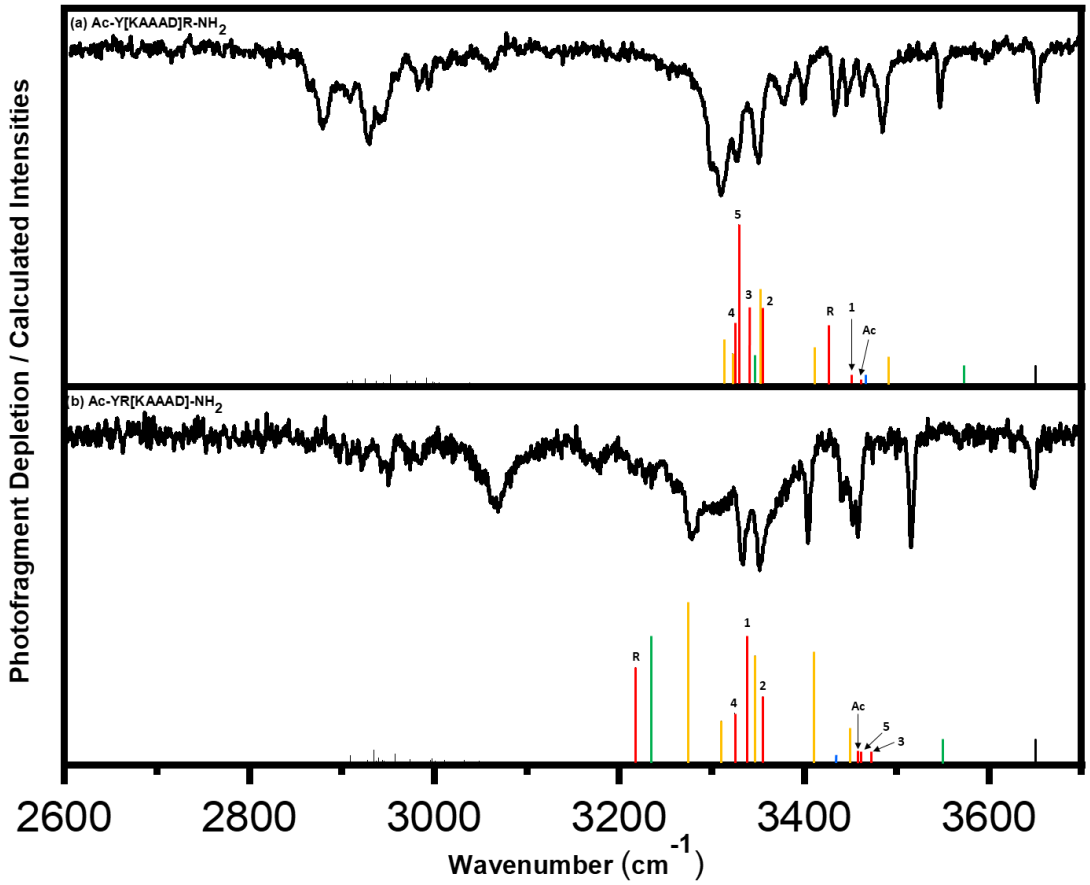

Figure 24 IR spectra of Y[KAAAD]R-NH_2 and YR[KAAAD]-NH_2 with calculated sticks. Conformation specific infrared spectra of [Y[KAAAD]R-NH_2 + H]$^+$ (a) and [YR[KAAAD] - NH_2 + H]$^+$ (b) with their best fit calculated stick spectra. Colors and names of each of the sticks in the calculated spectra correlate to the color scheme provided in figure 23.

4.4.3 Conformation-Specific Infrared Spectra of [Ac-Y[KAAAD]R-NH_2 + H]$^+$

The addition of an arginine residue to the peptide scaffold immediately complicates the infrared spectrum through the addition of several NH stretches including two symmetric and two asymmetric NH_2 stretches as well as a single NH stretch. Free and bound asymmetric NH_2 stretches are expected to appear in the same region as the C-terminal NH_2 cap. The arginine residue in Y[KAAAD]R-NH_2 has one NH_2 group that is fully hydrogen bound and another that is singly bound. The doubly hydrogen bound NH_2 forms both C14 and C17 macrocycles with the A_4 and A_5 carbonyl oxygens. Its asymmetric stretch appears, expectedly, at a lower wavenumber (3411 cm^{-1}) than its singly bound partner which has an asymmetric stretch at 3492 cm^{-1}. This asymmetric stretch provides the base for a C20 hydrogen bond with the carbonyl of A_3. Both of these asymmetric NH_2 stretches are far lower in energy than the C-terminal NH_2 cap that is in a

C7 hydrogen bond with the D_6 carbonyl group. Both symmetric NH_2 stretches of the arginine group unfortunately add to the congestion of the hydrogen bound NH stretch region making spectral assignments of closely spaced transitions difficult. The symmetric stretch of the doubly hydrogen bound NH_2 at 3322 cm^{-1} is positioned slightly higher in energy than the singly bound species at 3314 cm^{-1}. Higher in energy, 3347 cm^{-1}, is the symmetric stretch of the C-terminal NH_2 cap.

The addition of arginine and the subsequent acetylation of the N-terminus adds two additional NH stretches to the "backbone" of the molecule that aren't present in $[Y[KAAAD]\text{-}NH_2 + H]^+$. In the schematic (figure 23) as well as the infrared spectra these positions have been labelled "R" for arginine and "Ac" for acetylation so as not to perturb the numbering order defined for the smaller molecule. Of the seven backbone NH groups, the first two N-terminal NH oscillators are free. They include the acetylated N-terminus at 3460 cm^{-1} and the NH group of the lysine residue at 3451 cm^{-1}. Also in this region is the free NH of the tether at 3466 cm^{-1}. The last remaining backbone oscillator in the region correlates to a weak C16 hydrogen bond between the NH group of the arginine residue and the carbonyl of the lysine. Four backbone NH groups are found in hydrogen bonds. Of the four, the highest in energy is a C10 from the A_1 NH to the carbonyl of the n-terminal cap. A second, lower energy C10 hydrogen bond is composed of the A_2 NH and the tyrosine carbonyl group. The largest transition in the cluster belongs to the C6 macrocycle found between the aspartic acid NH and the tether carbonyl. Remaining in the bound backbone set of transitions is a third C10 hydrogen bond this time between the NH group of A_3 and the carbonyl belonging to lysine.

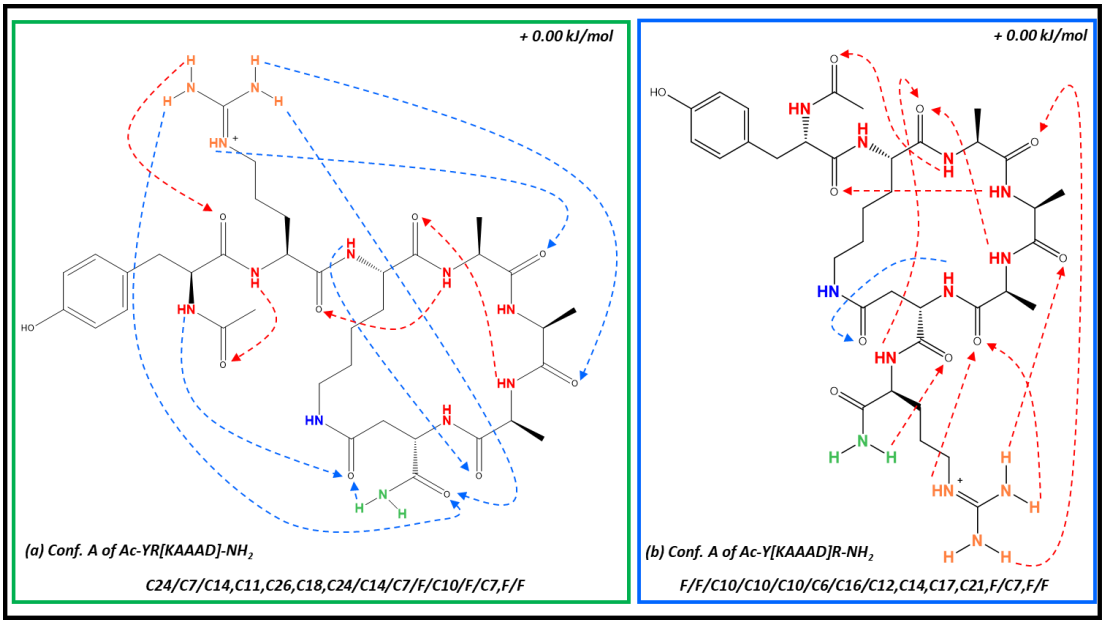

Figure 25 Stick structures of YR[KAAAD]-NH_2 (a) and Y [KAAAD]R-NH_2 (b) showing hydrogen bonds. Dashed arrows indicate hydrogen bonds with the arrow originating from the donor and terminating on the acceptor. Blue coloration denotes an N->C terminal hydrogen bond while red indicates a C->N terminal bond. . Shorthand notation describing the presence or absence of hydrogen bonding at NH groups starting from the N-terminus is provided utilizing Cn notation where n equals the number of atoms contained by the macrocycle.

4.4.4 Conformation-Specific Infrared Spectra of [Ac-YR[KAAAD]-NH_2 + H]$^+$

Placing the arginine residue n-terminal to the macrocyclic region of the molecule allowed the charge site to expand its reach and hydrogen bond to multiple different sites on the molecule. This had the effect of spreading the transitions associated with arginine throughout the spectrum standing in stark contrast to the more compacted transitions seen when the charged residue is placed on the c-terminus. Near 3440 cm^{-1} is the asymmetric stretch of one of the pair of NH_2 groups of the arginine with its corresponding symmetric stretch found at 3277 cm^{-1}. This set of transitions is involved in both C26 and C11 hydrogen bonds. The C26 formed with the aspartic acid carbonyl and the C11 with the tyrosine carbonyl. The sister asymmetric stretch belonging to the second NH_2 group of the arginine is shifted 36 cm^{-1} lower in energy than its counterpart. Its symmetric stretch is also shifted similarly and is found at ~3310 cm^{-1}. The shift to lower energy of these oscillators corresponds to a tighter hydrogen bonding pair. The first macrocyle is a C19 hydrogen bond to the A_2 carbonyl and the second is a second C25 to the same aspartic acid carbonyl as the other NH_2. At 3516 cm^{-1} the asymmetric stretch of the c-terminal NH_2 cap is

shifted lower in energy than where it is found in [Ac-Y[KAAAD]R-NH_2 + H]$^+$, its complementary symmetric stretch is also shifted much further down in wavenumber than it has been seen as well. A massive shift lower in wavenumber is indicative of the strong C7 hydrogen bond between one of the NH oscillators of this group to the tether carbonyl.

Lending to the large differences in structure between the arginine containing structures, the backbone hydrogen bonding pattern of [Ac-YR[KAAAD]-NH_2 + H]$^+$ is highly different than its complementary sister structure. Free NH oscillators are concentrated more towards the c-terminal end of the molecule with the A_2, aspartic acid, and acetalyated n-terminal NH groups forming a small cluster in this region. The highest in energy of the bound NH oscillators is the A_1 NH which is found in a C7 hydrogen bond with the carbonyl group of the arginine residue. Establishing a C14 macrocycle is the NH of the lysine and the carbonyl of A_2 which buts against a C10 hydrogen bond between A_2 and the lysine carbonyl. An extremely strong C7 hydrogen bond between the amide NH of the arginine residue and the carbonyl of the acetylated n-terminus.

4.5 Discussion

4.5.1 Fit and Structure Analysis for $[Y[KAAAD]\text{-}NH_2 + H]^+$

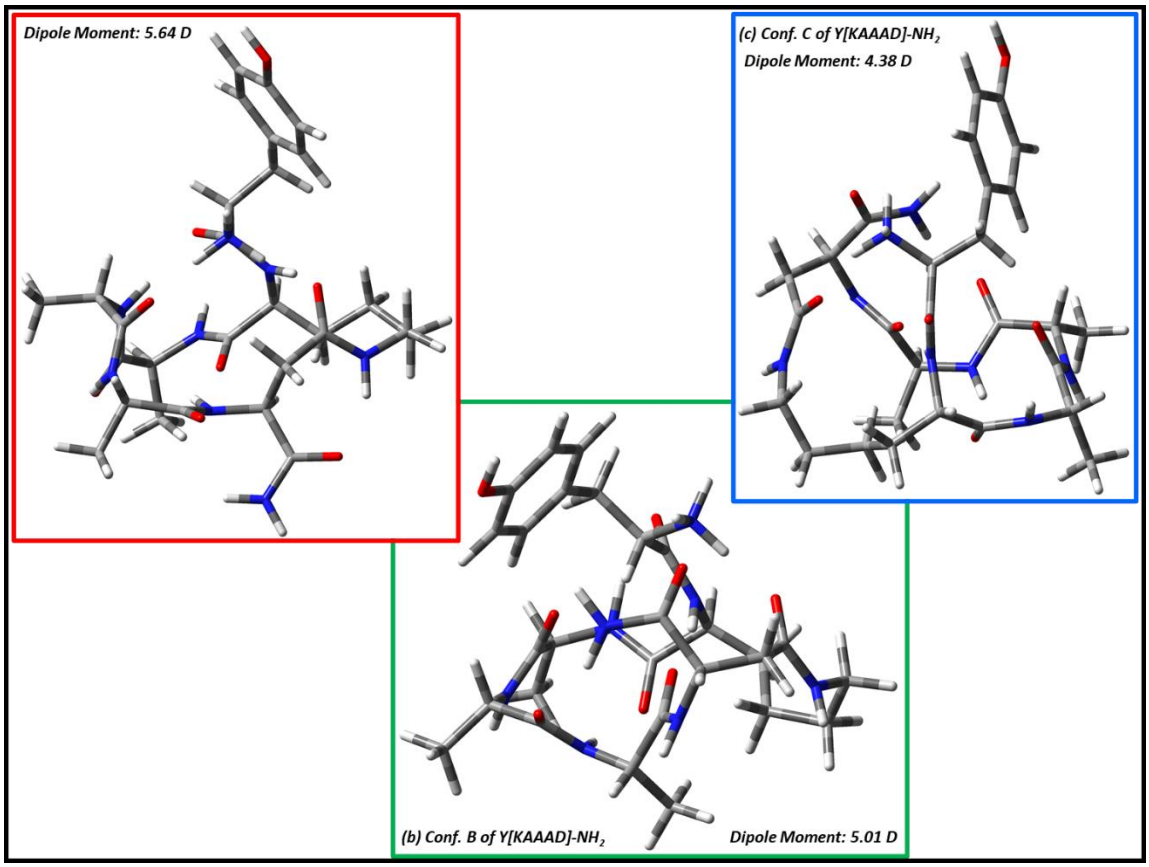

Figure 26 Assigned structures for each of the conformations of $[Y[KAAAD]\text{-}NH_2 + H]^+$ are shown.

The expectation for these molecules was to detect a single conformation comprised of the C13 hydrogen bonds that are the staple of the single turn alpha helix. $[Y[KAAAD]\text{-}NH_2 + H]^+$ instead appears to have three conformations. Two of which are very similar, differentiated only by the existence of a single hydrogen bond, and another that finds itself in a completely different hydrogen bonding pattern than the others. A structural feature that all three have in common is an expanded ovular macrocycle of a backbone that is formed, in part, by the carbonyl oxygen atoms orienting themselves toward the charged amine in the center of the structure. The backbone NH groups of the three conformers showcase the major difference in structure between conformer A and B/C. Conformer A has all but one of its NH (belonging to the first alanine residue) oscillators oriented in large hydrogen bonding cycles. Beginning with lysine, the

shorthand hydrogen bonding pattern for the backbone of conformer A is *tC10/F/C10/C7/C13/tC7* where the "t" denotes a hydrogen bond to the tether. As can be seen in figure 26, most of the hydrogen bonds in this conformer are large and bound across the molecule forming a ladder framework that holds the macrocycle together. This network of bonds stands in contrast to the backbone of conformers B (*C14/C7/F/C10/F/F*) and C (*C14/F/F/C10/F/F*) which exhibit far fewer hydrogen bonds. Conformer B has three free NH groups: A_2, D, and the tether while conformer C has four: A_1, A_2, D, and the tether. Both conformers share a C14 hydrogen bond from the lysine NH to the A_3 carbonyl and a C10 bond between the A_3 NH and lysine carbonyl.

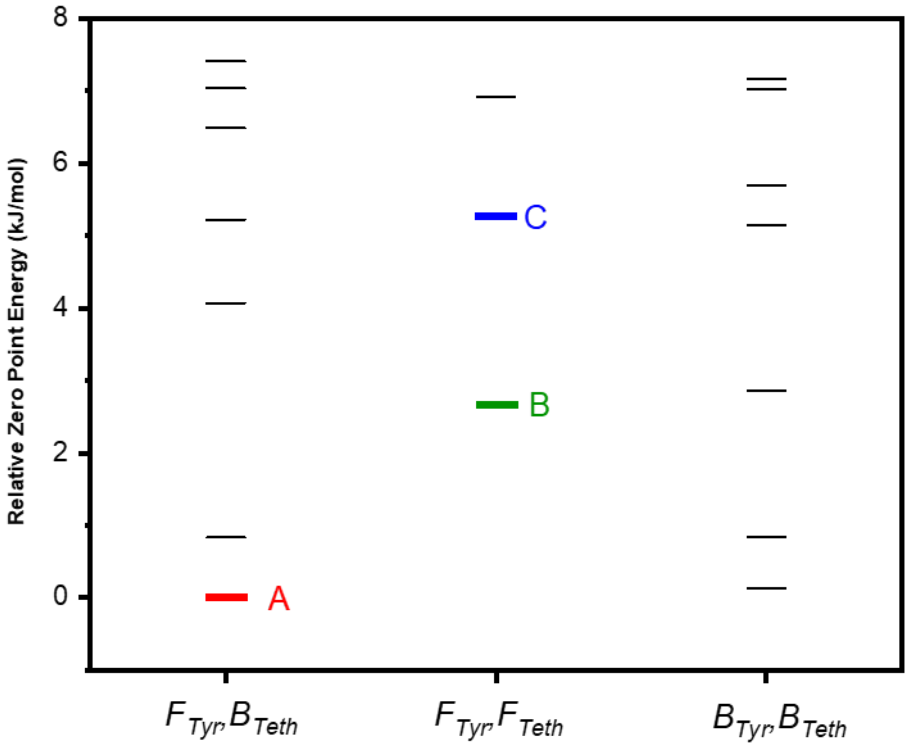

Figure 27 Zero-Point Energy (0 K) level diagram for [Y[KAAAD]-NH_2 + H]$^+$ with relative energy on the y-axis. The x-axis breaks the structures down into groups based upon the presence or absence of a bound tether NH group in the molecule and whether the tyrosine OH is in a hydrogen bond (B_{Tyr}) or not (F_{Tyr}). Each line in the figure represents the energy of a calculated structure while the red (conformer A), green (conformer B), and blue (conformer C) sticks represent the best fit structures to each infrared spectrum. The DFT produced alpha helix lies 150 kJ/mol above the global minimum.

The turning regions in large macrocyclic peptides are typically associated with common turning motifs such as the β-turns observed in sunflower trypsin inhibitor (SFTI)[17] or Gramacidin S[18]. Conformer A has a C7 hydrogen bond between the A_1 carbonyl and the A_3 NH that forms one turn while the second has a bifurcated C10/C13 bond from the carbonyl of the tether to the lysine NH and charged amine. Neither of these regions exhibit the expected bond angles or hydrogen bonding patterns of common turning motifs. Both conformers B and C share a C10/C14 hydrogen bonding structure and β-turn on one end of the macrocyle. The second turn in both conformers has the inverse angles of the first but lacks hydrogen bonds as this turn is comprised of the tethered region of the molecule. In all, the general lack of common motifs in all three conformers is not unexpected as the structures as the presence of the charged amine looms large over the whole molecule. Large differences in backbone hydrogen bonding between conformers A and B/C would insinuate vastly different structures. Conformer A with its numerous hydrogen bonds would expectedly have a more compact structure than B/C which are freer. Reality is that all three structures remain compact due to a complex interplay between the tether, which keeps the structure tight, and the charged amine/c-terminal cap that orient themselves in a way that maximizes their "solvation". Our discussion of the effects of these structural components begins with the charged amine.

4.5.2 Impact of the Charged Amine on the Conformations of $[Y[KAAAD]\text{-}NH_2 + H]^+$

Charge and its effect on the structure of gas phase ions has been widely reported and debated[19-22]. For $[Y[KAAAD]\text{-}NH_2 + H]^+$, the influence of the charge can not be understated, warping the molecule from the expected single turn helix and instead producing three conformations with no pronounced secondary structure. The impact of the charge is not equal in all three conformations. In conformer A it orients itself on top of the macrocycle where it reaches out and "pinches" the molecule together through the formation of hydrogen bonds with the A_2 (C14) and tether (tC13) carbonyls. The third NH group of the charged amine points itself toward the tyrosine ring, an orientation that is shared, in part, by the other two conformers but this is the extent of the similarity in charge placement and impact. Conformers B and C share the same placement and bonding structure of the charged amine with each other, as they do with most of the general features of the molecule. Unlike conformer A, in these structures the charge is not oriented over the top center of the molecule. Instead it still finds itself on top of the macrocyle

but oriented towards the tether side of the molecule. This may be due to the presence of the NH_2 cap which orients itself near the charge site where it also interacts with the tyrosine ring. Positioning of the cap near the tyrosine keeps the charge from forming the same hydrogen pattern that is observed in conformer A. Further, different from conformer A is the presence of a free NH group in the charged amine in both conformers B and C. What remains are two NH groups that anchor themselves to the tether side of the molecule by forming hydrogen bonds with the tether carbonyl (tC13) and the aspartic acid carbonyl (C20). Common between all three conformers is having the charge in an area that is not conducive to helix formation. Previously reported gas phase studies of helical ions show that placement of the charge is key to secondary structure formation[10-11]. In linear, non-cyclic peptides a charged amine on the N-terminus produces mostly globular structures like what is observed in this study. Ultimately, the influence of the charged amine dwarfs that of the tether and cap combination when attempting to observe gas phase helices.

4.5.3 Tether and Cap for $[Y[KAAAD]\text{-}NH_2 + H]^+$

The amide tether linkage that forms the macrocycle as well as the c-terminal NH_2 are key structural components of these molecules that are intended to seed the single turn helix. It is therefore of interest to examine their position and effect on the conformations observed. With conformer A emerging as the global minimum of the conformational landscape of protonated $Y[KAAAD]\text{-}NH_2$ it was expected that we would see these components in there most stable formation. Conformer A contains several unique features that are not observed in the two other found conformations. The most noticeable of these features is the free NH_2 cap of the c-terminus and the bound NH group of the tether. In the original solution-phase NMR studies completed on the uncharged prototype of this molecule, Fairlie and colleagues utilize the c-terminal cap as a foundation for seeding the hydrogen bonding pattern needed to form the single turn helix[6, 8]. In our protonated analogue it is therefore likely that this same hydrogen bond needs to be detected if the single turn helix is also to be observed. The symmetric NH_2 stretch found in the free position at 3556 cm^{-1} is the first major piece of evidence that points to the presence of an unexpected secondary structure for this conformer. Complementing the uniqueness of the free NH_2 cap is the amide tether which is in an uncharacteristic hydrogen bond unlike its paternal uncharged prototype where the tether NH group is found unbound to any adjacent carbonyl. This

strong C7 hydrogen bond links the tether amide to the aspartic acid carbonyl. Compared to the peculiar positioning of the tether and cap in conformer A, conformers B and C have these components in their expected condition (bound cap, free tether) but in unexpected locations.

As conformers B and C differ only in the appearance/lack thereof of a C7 hydrogen bond unrelated to either the tether or the amide cap it is expected that they have similar positioning of these motifs. Indeed, this premonition turns out to be true as the c-terminal NH_2 cap in both conformers is in a C10 hydrogen bond with the carbonyl of the second alanine residue. There is also a significant interaction between the cap and the π-cloud of the tyrosine residue. This binding motif expectedly shifts the observed (symmetric and asymmetric stretches) transitions in the infrared spectrum lower in energy than their free counterparts observed in conformer A. The asymmetric NH_2 stretch appears at 3445 cm^{-1} which comprises a near 100 cm^{-1} shift from its free position found in conformer A. That same 100 cm^{-1} shift is also observed for the symmetric stretch which moves out of its free position at 3435 cm^{-1} down to 3335 cm^{-1} in both conformers B and C.

4.5.4 Structural Analysis of [Ac-YR[KAAAD]-NH_2 + H]$^+$ and [Ac-Y[KAAAD]R-NH_2 + H]$^+$

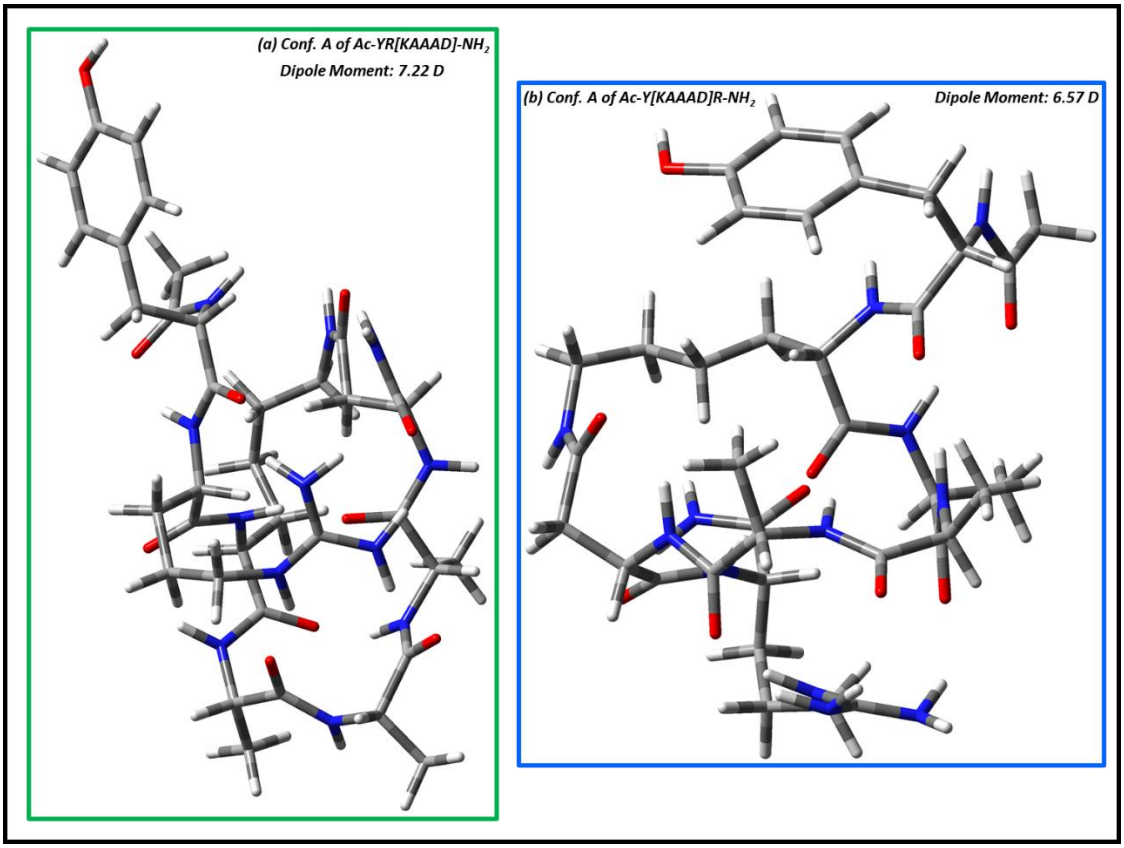

Figure 28 DFT calculated structures of YR[KAAAD]-NH_2 and Y[KAAAD]R-NH_2 Assigned structures for each of the conformations of [YR[KAAAD]-NH_2 + H]$^+$ (a) and [Y[KAAAD]R-NH_2 + H]$^+$ (b) are shown.

To negate the affect of the protonated amine two molecules where synthesized with arginine as the charge site. Arginine was chosen because the Fairlie group had previously reported that the molecule *Ac-R[KAAAD]-NH_2* forms a single turn helix in the solution phase[6]. Therefore, converting this prototype into its gas phase analogue presented another opportunity to observe the gas phase single turn helix. Utilization of our gas phase ion spectroscopy techniques required the addition of a tyrosine residue as was done with *[Y[KAAAD]-NH_2 + H]$^+$*. To ensure that the arginine residue is only charge site observed we kept the acetylation on the N-terminus from the prototype molecule. The analogous synthesized gas phase peptide ion is thus *[Ac-YR[KAAAD]-NH_2 + H]$^+$*. A second molecule was synthesized with the arginine residue on the opposite end of the tethered region. This produced the ion *[Ac-Y[KAAAD]R-NH_2 + H]$^+$* where it

was hypothesized that having the charge on the C-terminal end of the molecule would help induce the helix as was observed in previous gas phase studies.

Effect of Charge and Cap on [Ac-YR[KAAAD]-NH_2 + H]$^+$

Immediately apparent in the single conformation of *[Ac-YR[KAAAD]-NH_2 + H]$^+$* is the absence of any helical motif and the presence of the large macrocycle observed with all three conformations of *[Y[KAAAD]-NH_2 + H]$^+$*. Not unexpectedly, the formation of this structure is seeded by five hydrogen bonds, figure 28, between the arginine guanidinium group and the backbone. The charge having complete control of the backbone carbonyls disallows the hydrogen bonds necessary to form the helix. Addition of arginine as the charged residue presented a roadblock when examining the infrared spectrum as two additional NH oscillators complicate assignment. This complication does afford the opportunity to observe the location of three pairs of NH_2 vibrations (symmetric and asymmetric) in possibly different environments. The first of these pairs belong to the NH_2 cap that was placed on the C-terminus. A sharp, intense peak near 3500 cm^{-1} is indicative of the asymmetric stretch of the partially hydrogen bound NH_2 cap while its symmetric pair appears at ~3350 cm^{-1}. This partially hydrogen bound version of the cap provides the last piece information needed to identify the effect hydrogen bonding has on the shifting of the transition in the infrared spectrum. Conformer A of *[Y[KAAAD]-NH_2 + H]$^+$* with its completely asymmetric stretch of the cap at 3556 cm^{-1} and the fully bound cap with its asymmetric stretch at ~3450 cm^{-1} for conformers B and C complete the set. The 50 cm^{-1} shift to lower energy with each additional hydrogen bond has been previously observed[23-24]. Both residual pairs of NH_2 transitions belong to the charged guanidinium group. Higher in energy between the pairs is a set of oscillators forming C11 and C26 hydrogen bound macrocycles to the carbonyl of the tyrosine and the aspartic acid, respectively. Weakness of this hydrogen bonding pair compared to its neighbor can be explained by the indirect, nearly perpendicular, bond to the aspartic acid carbonyl. Whereas the second NH_2 pair forms direct, traditional hydrogen bonds, of a much shorter distance, with the same aspartic acid carbonyl as well as the A_2 carbonyl.

Structural Effects of the Tether and Backbone Hydrogen Bonding on [Ac-YR[KAAAD]-NH_2 + H]$^+$

The charged guanidinium site on the arginine residue plays a major role in the three-dimensional structure of this molecule. Interaction of the charge with the backbone directly shapes the hydrogen bonding patterns seen in the remainder of the backbone. Disruption of the expected hydrogen bonding patterns by the charge site leaves the backbone in the distinct ring shape that is observed with the other molecules. This ring-like backbone takes on two distinct structural motifs on its respective loop ends. One end exhibits both C10 and C14 hydrogen bonds forming a classical β-turn. The other loop end of this "ring" is far less compact and does not have the formal structure of the β-turn end. It is dominated by the tether with its long alkyl chain forming one section and the head of the tether, bound in a C7 hydrogen bond with the c-terminal cap, on the other. β-turn configurations of small peptides, especially cyclic peptides, is ubiquitous in nature[18, 25] and have been well observed in the gas phase[12]. With the seeding hydrogen bonds and formation of a macrodipole, induced by the charge site, that are necessary to observe the single turn helix being absent a more easily formed secondary structure, the β-turn, is frequently observed in its place.

Effect of Charge and Cap on [Ac-Y[KAAAD]R-NH_2 + H]$^+$

At first glance, [Ac-Y[KAAAD]R-NH_2 + H]$^+$ looked to be a promising structure as it is almost helical. Three C10 hydrogen bonds on the n-terminal side give the appearance of the formation of a 3_{10} helix. Upon close examination this "almost" helix lacks the coordinated hydrogen bonding pattern at its c-terminal end needed to produce the tight secondary structure. The presence of the arginine residue on the c-terminal end of the molecule plays the deciding factor in this motif. As has been observed by Rizzo and colleagues[11], specific placement of the charged residue is integral for helix formation in medium length, untethered peptides. In gas phase [Ac-Phe-$(Ala)_{10}$-Lys-H^+] the protonated lysine residue formed hydrogen bonds to the three nearest backbone carbonyl oxygen atoms. This anchors the back end of the molecule and orients the backbone into the first turn of the helix. A similar pattern is realized in [Ac-Y[KAAAD]R-NH_2 + H]$^+$ where the arginine residue also forms strong hydrogen bonds to the three nearest backbone carbonyl oxygens. The major difference being lysine is a more compact charged species making the range of possible hydrogen bonding distances quite small. Whereas the

arginine residue has a much larger charged surface area making it more capable of a large variety of hydrogen bonding distances and motifs. The ability to produce a more varied state of hydrogen bonds reduces the arginine residues likelihood of organizing the nearby carbonyls into the tight formation required to observe the common alpha helix.

With this structure nearing what would be a 3_{10} helix the NH_2 cap should be expected to play a major role as it does in the solution phase. In actuality, a similar hydrogen bonding scheme as [Ac-YR[KAAAD]-NH_2 + H]$^+$, a C7 macrocycle to the aspartic acid backbone carbonyl and a free NH stretch. This continues the theme observed with these gas phase ions where the cap does not form the hydrogen bond necessary to help stabilize the helix seen in the solution phase. The charge, whether N-terminus or guanidinium, serves to diminish any effect the cap might have on the molecule.

Structural Effects of the Tether and Backbone Hydrogen Bonding on [Ac-Y[KAAAD]R-NH_2 + H]$^+$

In its simplest definition an $i \rightarrow i+3$ hydrogen bonding pattern defines the 3_{10} helix[26-27]. While the c-terminal end of [Ac-Y[KAAAD]R-NH_2 + H]$^+$ is a soup of hydrogen bonds dominated by the protonated guanidinium, the n-terminal end presents with two "pure" C10 hydrogen bonds followed by one bifurcated C10/C16 hydrogen bond. The presence of these C10 bonds is necessary for characterization of the 3_{10} helix, but not sufficient. For example, back-to-back β-turns in proteins and peptides occur at a high frequency[28]. These double turns are characterized by a pair of C10 hydrogen bonds occurring one after the other. It becomes quite clear then that complex secondary structures can not be defined by the presence of certain hydrogen bonding patterns alone; a stricter definition is required. For a structure to represent a 3_{10} helix each amino acid should lead to a 120º turn in the structure and the dihedral angles (φ,ψ) should approximate (-49º, -26º)[26]. The 120º turn criteria is met by the Tyr and Lys residues as the C10 hydrogen bond that anchors these two residues lead to ~112º turns in the structure. The second C10 hydrogen bond between the Tyr carbonyl and the Ala_2 NH does not create the angles turn necessary to wrap the structure around in the helix. Instead, the remainder of the molecule flattens out, forming a near straight line for three residues (Ala_2,Ala_3, Asp). With the turn criteria for a 3_{10} helix only partially satisfied an analysis of the dihedral angles about the residues in question may provide more insight. Observation of the *phi, psi* angles of Tyr(-54º,-35º), Lys(-

65º,-23º), and Ala_1(-56º,-25º) lends credence to the notion that this region of the molecule is indeed helical. Overall, this structure follows the same pattern of both its arginine positional isomer and the charged N-terminus analogues where the charged site ultimately dismantles any opportunity for a cohesive secondary structure.

4.5.5 Charged N-terminus vs. Protonated Guanidinium

The propensity of these tethered peptides to form single-turn α-helices in the solution phase was a possible boon for seeing helices in the gas phase outside of well-known motifs motivated our study of them in the gas phase, to spectroscopically characterize them and concatenate them with one another to form longer helical segments. Unfortunately, the only peptide of the group to form a partial helix was $[Ac\text{-}Y[KAAAD]R\text{-}NH2]^+$. Formation of the 3_{10} helix was not surprising as Jarrold had previously shown that a C-terminal charged residue will induce helix formation. Of note is the reduction of conformers concurrent with the addition of the arginine residue. The presence of the arginine seems to lock the molecule into a stable helix, consequently reducing the number of conformers. This may be due to the arginine residue's propensity to form hydrogen bonds with many nearby carbonyl groups. This activity may force the molecule into a particular way of folding.

Without the stabilizing effect of Arginine, $[Y[KAAAD]\text{-}NH2]^+$ ballooned into a macrocycle stabilized by internal hydrogen bonding. Without arginine to orient the dipole of the molecule, more flexibility opened paths to different structures. What resulted is an increased amount of conformations with no defined secondary structure. The large C10+ hydrogen bonds needed to form the helix in $[Ac\text{-}Y[KAAAD]R\text{-}NH2]^+$ are replaced with shorter C5 and C7 hydrogen bonds which tighten small segments of the molecule. This localized tightening retards helix formation by reorienting the carbonyls of each amino acid so that no macrodipole is formed.

The location and nature of the charge site on an ion can have a great impact on its structure. A protonated N-terminus on a peptide has the potential to perturb structure away from its solution phase structure in ways that aren't expected with a charged amino acid side chain, such as with arginine. In $[Y[KAAAD]\text{-}NH2]^+$, the charged site does not work its way into the center of the molecule. Instead, it remains on the outside of the molecule forming a C5 hydrogen bond with the carbonyl of the tyrosine group. While this bond may seem innocuous, when one

examines the acetylated N-terminus of $[Ac\text{-}Y[KAAAD]R\text{-}NH2]^+$ the detriment that the protonated N-terminus has becomes clear. The N-terminal acetylation seeds a C10 hydrogen bond to the amide N-H of the first alanine residue. The combination of this helix forming hydrogen bond and the c-terminal arginine seeds provides nucleation sites for a helical structure into at both ends of the peptide. This strongly encourages the n-terminal end of the peptide to fall into place to form the 3_{10} helix.

4.6 Conclusions

Cold ion spectroscopy and IR-UV double resonance techniques were used to determine the conformations and structures of $[Y[KAAAD]\text{-}NH2]^+$ and $[Ac\text{-}Y[KAAAD]R\text{-}NH2]^+$. $[Y[KAAAD]\text{-}NH2]^+$ presents with three conformations none of which form single turn alpha helices. To the contrary $[Ac\text{-}Y[KAAAD]R\text{-}NH2]^+$, almost forms a 3_{10} helix in the gas phase, but the charge ultimately interferes with its completion. We have shown that the nature of the charge site and its placement matter greatly for nucleating helix formation. Placing an arginine residue near at the c-terminus of the peptide helps to induce helix formation through hydrogen bond formation with several carbonyl groups near the end of the molecule. Acetylation of the N-terminus also helps form the needed C10 hydrogen bonds to seed the helix. This work has elucidated important hydrogen bonding patterns that need to be in place to see single turn alpha helices in the gas phase.

4.7 References

1. Manning, M. C.; Chou, D. K.; Murphy, B. M.; Payne, R. W.; Katayama, D. S., Stability of Protein Pharmaceuticals: An Update. *Pharmaceutical Research* **2010,** *27* (4), 544-575.

2. Fowler, S. A.; Blackwell, H. E., Structure–function relationships in peptoids: Recent advances toward deciphering the structural requirements for biological function. *Organic & Biomolecular Chemistry* **2009,** *7* (8), 1508-1524.

3. Gonzalez, M. W.; Kann, M. G., Chapter 4: Protein Interactions and Disease. *PLOS Computational Biology* **2012,** *8* (12), e1002819.

4. Ryan, D. P.; Matthews, J. M., Protein–protein interactions in human disease. *Current Opinion in Structural Biology* **2005,** *15* (4), 441-446.

5. Lage, K., Protein–protein interactions and genetic diseases: The interactome. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease* **2014,** *1842* (10), 1971-1980.

6. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005**.

7. Harrison, R. S.; Shepherd, N. E.; Hoang, H. N.; Ruiz-Gómez, G.; Hill, T. A.; Driver, R. W.; Desai, V. S.; Young, P. R.; Abbenante, G.; Fairlie, D. P., Downsizing human, bacterial, and viral proteins to short water-stable alpha helices that maintain biological potency. *Proceedings of the National Academy of Sciences* **2010,** *107* (26), 11686.

8. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Left- and Right-Handed Alpha-Helical Turns in Homo- and Hetero-Chiral Helical Scaffolds. *Journal of the American Chemical Society* **2009**.

9. Robert R. Hudgins, M. A. R., and Martin F. Jarrold, Design of Helices That Are Stable in Vacuo. *Journal of the American Chemical Society* **1998**.

10. McLean, J. R.; McLean, J. A.; Wu, Z.; Becker, C.; Pérez, L. M.; Pace, C. N.; Scholtz, J. M.; Russell, D. H., Factors That Influence Helical Preferences for Singly Charged Gas-Phase Peptide Ions: The Effects of Multiple Potential Charge-Carrying Sites. *The Journal of Physical Chemistry B* **2010,** *114* (2), 809-816.

11. Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R., Spectroscopic signatures of gas-phase helices: Ac-Phe-(Ala)5-Lys-H+ and Ac-Phe-(Ala)10-Lys-H+. *J Am Chem Soc* **2007,** *129* (45), 13820-1.

12. 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. *Journal of the American Chemical Society* **2017,** *139* (15), 5481-5493.

13. Londry, F. A.; Hager, J. W., Mass selective axial ion ejection from a linear quadrupole ion trap. *Journal of the American Society for Mass Spectrometry* **2003,** *14* (10), 1130-1147.

14. 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.

15. Mohamadi, F.; Richards, N. G. J.; 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. *Journal of Computational Chemistry* **1990,** *11* (4), 440-467.

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

17. Mylne, J. S.; Colgrave, M. L.; Daly, N. L.; Chanson, A. H.; Elliott, A. G.; McCallum, E. J.; Jones, A.; Craik, D. J., Albumins and their processing machinery are hijacked for cyclic peptides in sunflower. *Nature Chemical Biology* **2011,** *7* (5), 257-259.

18. Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V., Highly Resolved Spectra of Gas-Phase Gramicidin S: A Benchmark for Peptide Structure Calculations. *Journal of the American Chemical Society* **2010,** *132* (12), 4040-4041.

19. Jhingree, J. R.; Bellina, B.; Pacholarz, K. J.; Barran, P. E., Charge Mediated Compaction and Rearrangement of Gas-Phase Proteins: A Case Study Considering Two Proteins at Opposing Ends of the Structure-Disorder Continuum. *Journal of the American Society for Mass Spectrometry* **2017,** *28* (7), 1450-1461.

20. Gadzuk-Shea, M. M.; Bush, M. F., Effects of Charge State on the Structures of Serum Albumin Ions in the Gas Phase: Insights from Cation-to-Anion Proton-Transfer Reactions, Ion Mobility, and Mass Spectrometry. *The Journal of Physical Chemistry B* **2018,** *122* (43), 9947-9955.

21. Allen, S. J.; Schwartz, A. M.; Bush, M. F., Effects of Polarity on the Structures and Charge States of Native-Like Proteins and Protein Complexes in the Gas Phase. *Analytical Chemistry* **2013,** *85* (24), 12055-12061.

22. Munshi, M. U.; Berden, G.; Martens, J.; Oomens, J., Gas-phase vibrational spectroscopy of triphenylamine: the effect of charge on structure and spectra. *Physical Chemistry Chemical Physics* **2017,** *19* (30), 19881-19889.

23. Walsh, P. S.; Blodgett, K. N.; McBurney, C.; Gellman, S. H.; Zwier, T. S., Inherent Conformational Preferences of Ac-Gln-Gln-NHBn: Sidechain Hydrogen Bonding Supports a beta-Turn in the Gas Phase. *Angew Chem Int Ed Engl* **2016,** *55* (47), 14618-14622.

24. Blodgett, K. N.; Fischer, J. L.; Lee, J.; Choi, S. H.; Zwier, T. S., Conformation-Specific Spectroscopy of Asparagine-Containing Peptides: Influence of Single and Adjacent Asn Residues on Inherent Conformational Preferences. *J Phys Chem A* **2018,** *122* (44), 8762-8775.

25. Korsinczky, M. L.; Schirra, H. J.; Rosengren, K. J.; West, J.; Condie, B. A.; Otvos, L.; Anderson, M. A.; Craik, D. J., Solution structures by 1H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant. *J Mol Biol* **2001,** *311* (3), 579-91.

26. Barlow, D. J.; Thornton, J. M., Helix geometry in proteins. *Journal of Molecular Biology* **1988,** *201* (3), 601-619.

27. Zhang, L.; Hermans, J., 310 Helix Versus .alpha.-Helix: A Molecular Dynamics Study of Conformational Preferences of Aib and Alanine. *Journal of the American Chemical Society* **1994,** *116* (26), 11915-11921.

28. Shapovalov, M.; Vucetic, S.; Dunbrack, R. L., Jr., A new clustering and nomenclature for beta turns derived from high-resolution protein structures. *PLOS Computational Biology* **2019,** *15* (3), e1006844.

CHAPTER 5. COMPUTATIONAL CHEMISTRY MEDIATED DESIGN OF A GAS PHASE SINGLE TURN ALPHA HELIX

5.1 Abstract

Computational chemistry methods were used to isolate chemical structures that provided a high likelihood of forming gas phase single turn alpha helices. Cold ion spectroscopy and IR-UV double resonance techniques were used to confirm the presence of a single helical conformer for the tethered peptides, Ac-Y[KAAAD]X_1-NH_2 and Ac-Y [KAAAD]X_2-NH_2. The presence of the helix confirms the power of searching the conformational landscape in finding and synthesizing a specific chemical moiety. Control of the handedness of the helix was accomplished with changes in the stereochemistry of the amino acids forming the molecule. The relationship between the tether and charge in controlling conformational preferences was also probed.

5.2 Introduction

The organic syntheses of possible drug targets has historically been an arduous process involving the production of tens of molecules with the hope that one binds its target[1]. Reducing the time spent producing non-functional molecules stands to save a great deal of money and better help those in need. With the advent of powerful supercomputers and functional GUIs computational chemistry is quickly finding its way into the design and analysis of compounds ranging from small molecules to large megadalton proteins[2]. Being able to search the conformational landscape of a molecule to determine the structure most likely to be observed after synthesis is a powerful tool for chemists[3-5]. Key to increasing the use and efficacy of computations in chemical synthesis is providing powerful examples of its ability to solve complex problems.

In a previous study, an attempt was made to observe a single turn alpha helix in the gas phase as an ion using tethered, capped peptides. These molecules were chosen because they produced extremely stable solution phase helices that were not easily denatured[6-8]. When no single turn alpha helix was detected an in-depth analysis of the structural motifs causing the perturbation of the structure was completed. It was found that the nature and location of the charge site in the ion holds great influence over the observed secondary structures as solvation of the charge takes

precedence. While not observing the single turn helix was a setback, the search for a molecule that would form the helix provided an opportunity to show off the powerful combination of conformational searching and density functional theory(DFT) calculations.

This study utilizes both conformational searching as well as DFT calculations to produce target structures for synthesis that give the highest likelihood of forming single turn alpha helices when ionized in the gas phase. Multiple charge configurations and locations were used to produce a library of potential structures(>10,000) that were then clustered and sorted energetically to find two target molecules that had the greatest propensity to form gas phase helices. These structures, Ac-Y[KAAAD]X_1-NH_2 and Ac-Y [KAAAD]X_2-NH_2, where the brackets denote the macrocycle and X denotes the charged residue, were expected to form solid helical structures when interrogated in the gas phase.

To determine whether the results of the computationally driven synthesis of these molecules was accurate, cold ion spectroscopy was used to probe the gas phase structures. IR-UV double resonance techniques have been a proven method for structural determination of small molecules and peptides in the gas phase for several years[9-11]. Single conformations in the UV spectrum were probed and conformer specific infrared spectra produced for each molecule. The combination of experimental spectra and DFT calculated spectra was used to determine the secondary structure of these molecules.

The observation that charge has more influence than the tether in the formation of the structure of the ion was hypothesized based on our previous work with similar peptides. While this was assumed to be true no study of both a tethered and non-tethered set of peptides was performed. To explore the relationship between the tether and the secondary structure a third molecule was synthesized, Ac-YKAAANX_2, that mimicked the primary structure of the other peptides but did not include the tether. An analysis of the structures that this ion forms in the gas phase provided deeper insight into the charge/tether relationship.

5.3 Experimental

5.3.1 Instrumentation

All spectroscopic data were taken on a custom-built instrument for cold ion spectroscopy, previously described.[12-13] In brief, the design of the instrument comprises a triple quadrupole

mass spectrometer in which an orthogonal spectroscopy axis was added between the second (q2) and third (q3) quadrupoles. Ions are formed via nano-electrospray ionization(nESI) and transported into q2 where the ion of interest is isolated via RF/DC isolation. The ion packet is then steered around a turning quadrupole and down the spectroscopy axis where it is trapped in the cold quadrupole that is cooled to 5K via a closed-cycle helium cryostat (Sumitomo Heavy Industries, Tokyo, Japan). After cooling to ~10K via collisions with the He buffer gas, the ions are spectroscopically interrogated via infrared (Laservision OPO/OPA) and ultraviolet (ScanMate-Pro Lambda Physik frequency doubled by an Inrad Autotracker III and Scanmate Lambda Physik) lasers. The created photofragments are then extracted back down the spectroscopy axis and trapped in q3 for analysis in either "spectroscopy mode" or "mass spectrometry mode". In mass spectrometry mode, a mass spectrum of the photofragments is generated via mass-selective axial ejection.[14] Spectroscopy mode entails ejection of the remaining precursor ions with SX wave software followed by propelling of the photofragments onto a channeltron detector (4773G, Photonis USA).

UV action spectra are recorded by monitoring the total photo-fragment signal as a function of laser wavenumber. Upon obtaining a UV spectrum and identifying transitions that may be due to different ion conformations, conformation-specific IR spectra are taken by employing IR-UV double resonance techniques.[9] The UV laser is tuned to a transition of interest in the UV spectrum which creates a steady signal of photofragment ions. Then a tunable IR laser that is spatially overlapped but temporally precedes the UV laser by 200 ns is scanned in wavelength. Upon absorption of an IR photon by the conformer of interest, a fraction of the population is removed from the ground state which results in a depletion in the total photofragment signal.

5.3.2 Solution Preparation

Solutions of Ac-Y[KAAAD]X_1-NH_2, Ac-Y [KAAAD]X_2-NH_2, Ac-y[kaaad]X_1-NH_2, and Ac-YKAAANX_2 were prepared as 50:50 Methanol:Water to a concentration of 200 μM and electrosprayed to create the corresponding protonated ion.

5.3.3 Computational

Conformational searches were performed for [Ac-Y[KAAAD]X_1-NH_2 + H]$^+$,[Ac-Y[KAAAD]X_2-NH_2 + H]$^+$, [Ac- y[kaaad]X_1-NH_2 + H]$^+$, and Ac-YKAAANX_2 structures using the Monte Carlo multiple minimum method with the AMBER* and OPLS3 forcefields inside the MACROMODEL software package.[15] Approximately 10000 structures were found between the two forcefields for each of the molecules. To narrow down the potential structures further, the 10000-structure pool was clustered based on backbone RMSD. The centroid of each of the clusters was then optimized at the B3LYP/6-31+G(d)-GD3BJ level of theory[16-18] using Gaussian16 and harmonic vibrational frequencies/intensities were calculated to obtain predicted infrared spectra to compare with experiment. To correct for anharmonicity the calculated spectra were scaled by 0.958 for the NH, CH, and OH stretching regions, 0.973 for the free OH transitions, 0.981 for Amide I, and 0.970 for Amide II. Adjusting the calculated transitions by these values puts them into agreement with previously calculated and fitted spectra.[9-11] "Stick" spectra were generated for the centroid of each cluster. These calculated spectra were then matched to the experimental spectrum and the best fit centroids were found. All clusters of structures with best fit centroids that lie within 10 kJ/mol of the global minimum were then run through DFT at the same level of theory used for the centroids. After completing all calculations, the structure that best fit the experimental spectrum and was lowest in energy was determined to be the most likely structure of the ion. A further analysis of the forcefield searching method is detailed in section 3.1.

5.4 Results and Analysis

5.4.1 Use of Forcefield Searching to Guide Synthesis of a Gas Phase Single Turn Alpha Helix

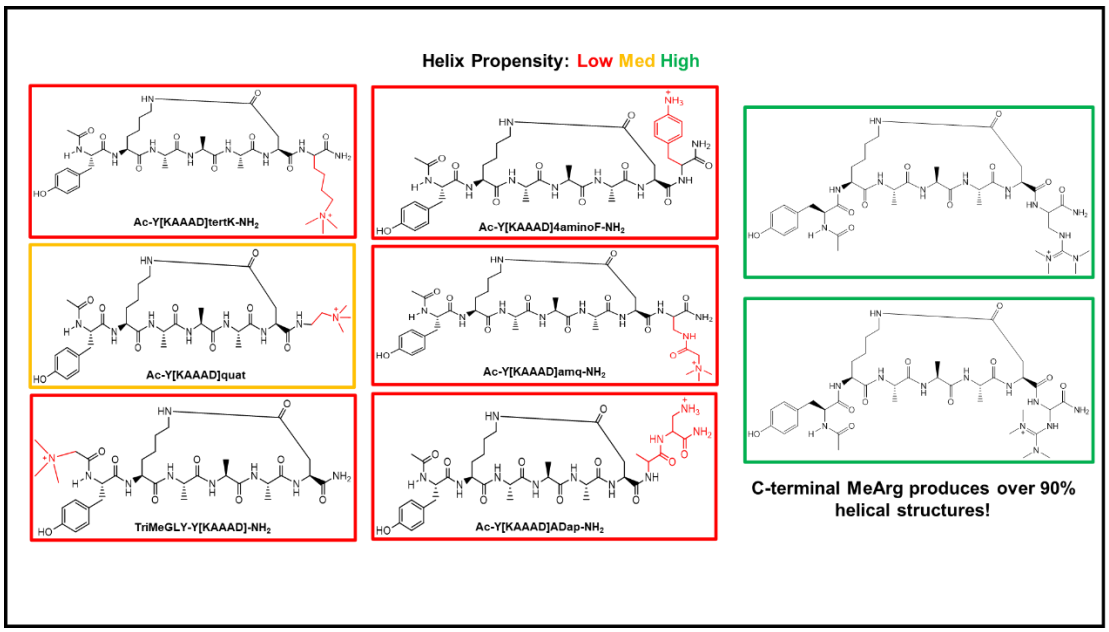

Figure 29 A small selection of the ion types studied computationally are provided with the charge on the molecule highlighted in red. Results of the conformational searching showed that only two charged groups had a significant potential to form single turn helices (green box), while the majority of molecules showed little (orange box) to no (red box) propensity for helix formation.

Our previous attempt to bring a solution phase single turn alpha helix into the gas phase resulted with several confirmations that in no way retained their solution phase structure. The main takeaway from that study was the nature of the charge matters when attempting to maintain structural integrity during the transition from the solution to the gas phase. This investigation began with a search for a charged site that was small enough as not to perturb the structure and contained little to no N-H oscillators as to not complicate the infrared spectrum. Several examples of charged moieties that fit these criteria exist, figure 29 presents a few poignant examples. Synthesis and analysis of a large number of molecules to determine which would produce the helix was unfeasible. Instead, it was decided that a computational approach to selection and synthesis of a few target molecules would produce better results in a shorter amount of time.

The use of forcefield searches to generate a considerable number of potential structures is one of the key components of analysis in ion spectroscopy[9, 19-20]. With increases in computational power and resources, large scale studies of the potential energy surface of a molecule of interest can be completed quickly and accurately[21-22]. A diverse set of tools to achieve this task exists, from existing software packages that make conformational searches easy to accomplish[15] to academic written programs that allow for highly detailed searches[23]. For this application, the MacroModel software package was used as it not only makes searching for structures simple, but it also makes analysis of the generated structures an easy task. The OPLS3 and Amber* forcefields have proven track records in handling peptide and small protein molecules[24-25]. Each forcefield produced roughly five thousand potential structures for each of the molecules studied, for a total of ten thousand structures to analyze per molecule. With 10+ molecules the number of potential structures to sift through was enormous and any energy sorting via DFT methods would take a lifetime. Remedying this problem began by combining the structures from both the OPLS3 and Amber* forcefield searches into a single data set and employing a clustering analysis to group the structures together[26].

Clustering was accomplished through comparison of backbone root mean square deviation (RMSD). Focusing on the backbone of the peptide structure allowed for quick comparison and visualization of the structures to confirm the presence of the helix. Clustering of the structures resulted in between 50-100 clusters per molecule. After clustering a centroid of each cluster was chosen that exemplified the backbone structure of the cluster. This procedure made simple determining which charge site produced the most helical structures. Removal of most of the candidate structures occurred at this point. Interestingly, several quaternary amines were eliminated during this step as they showed a very low propensity for formation of the helix in the conformational search. This result goes against our initial intuition that making the charge as small and near to a point as possible would be the best way to go. Even more surprising was the charged moieties that produced the largest number of helical structures. In our previous study the molecule that got closest to forming the single turn helix contained a charged arginine residue on the c-terminus of the molecule. Results of the conformational search suggested that if the normal guanidinium group on the arginine residue was replaced with a methylated version that we can be highly confident that a helical structure would appear. In fact, over 90% of the conformations observed with the forcefield search were helical for this charge site. Figure 29

shows the structures of the two charged residues that we have chosen for this study. The power of forcefield searching saved a great deal of time and resources that would have gone into synthesizing potential structures.

Forcefield analysis reduced the amount of DFT calculations required to sort the structures greatly but an evaluation using a high level of theory was still an insurmountable task given the number of parent molecules and limited computational resources. An energy sorting via a low level of DFT theory (B3LYP/6-31G(d)) was employed to further reduce the number of candidates. Every centroid was optimized at this level of theory and the energies of all structures sorted with a cut off +15 kJ/mol above the global minimum used to further reduce the number of structures down to 20-30 per molecule. These 20-30 conformations were then submitted to a higher level of DFT theory (B3LYP/6-31+G(d)) for final sorting and production of the calculated infrared spectra.

5.4.2 Analysis of UV Spectra of Ac-Y[KAAAD]X_1-NH_2, Ac-Y[KAAAD]X_2-NH_2, Ac-y[kaaad]X_1-NH_2, and Ac-YKAAANX$_1$

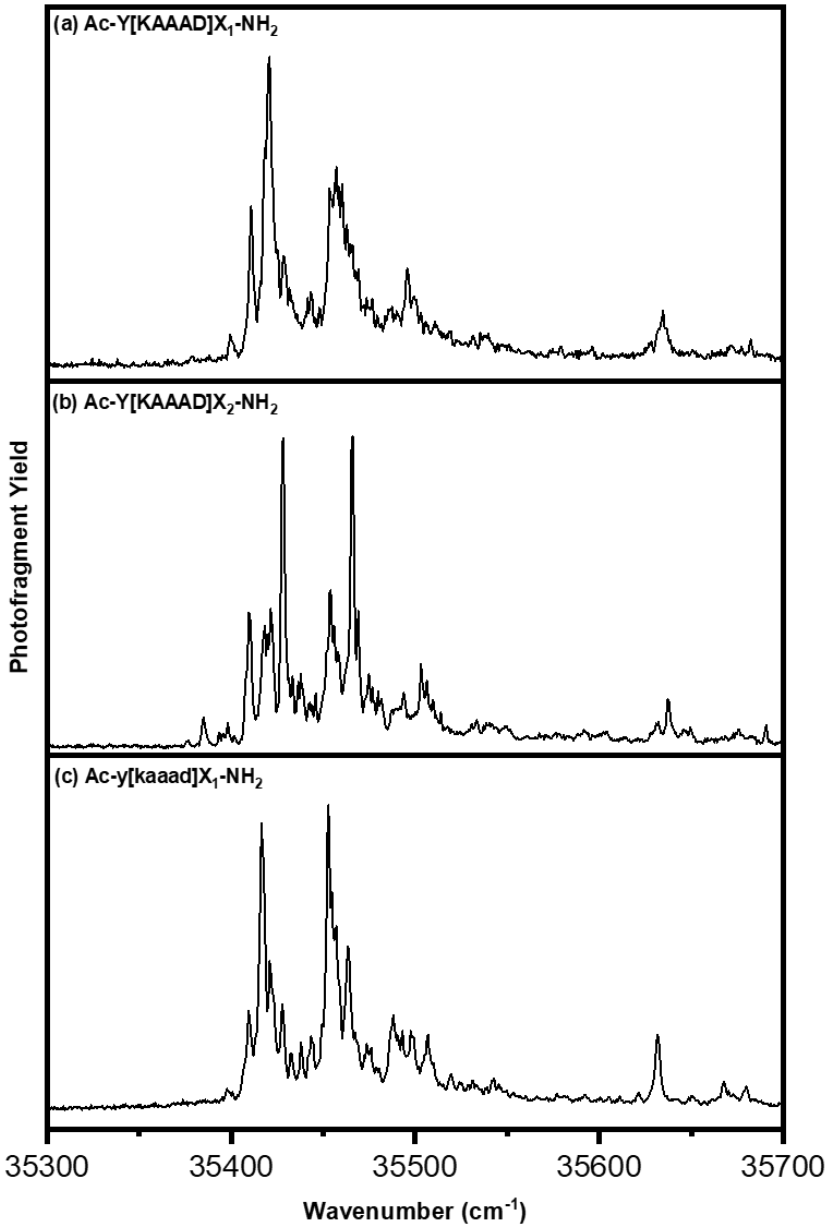

Figure 30 UV action spectra of (a) [Ac-Y[KAAAD]X_1-NH_2+H]$^+$, (b) [Ac-Y[KAAAD]X_2-NH_2+H]+, and (c) [Ac-y[kaaad]X_1-NH_2+H]$^+$.

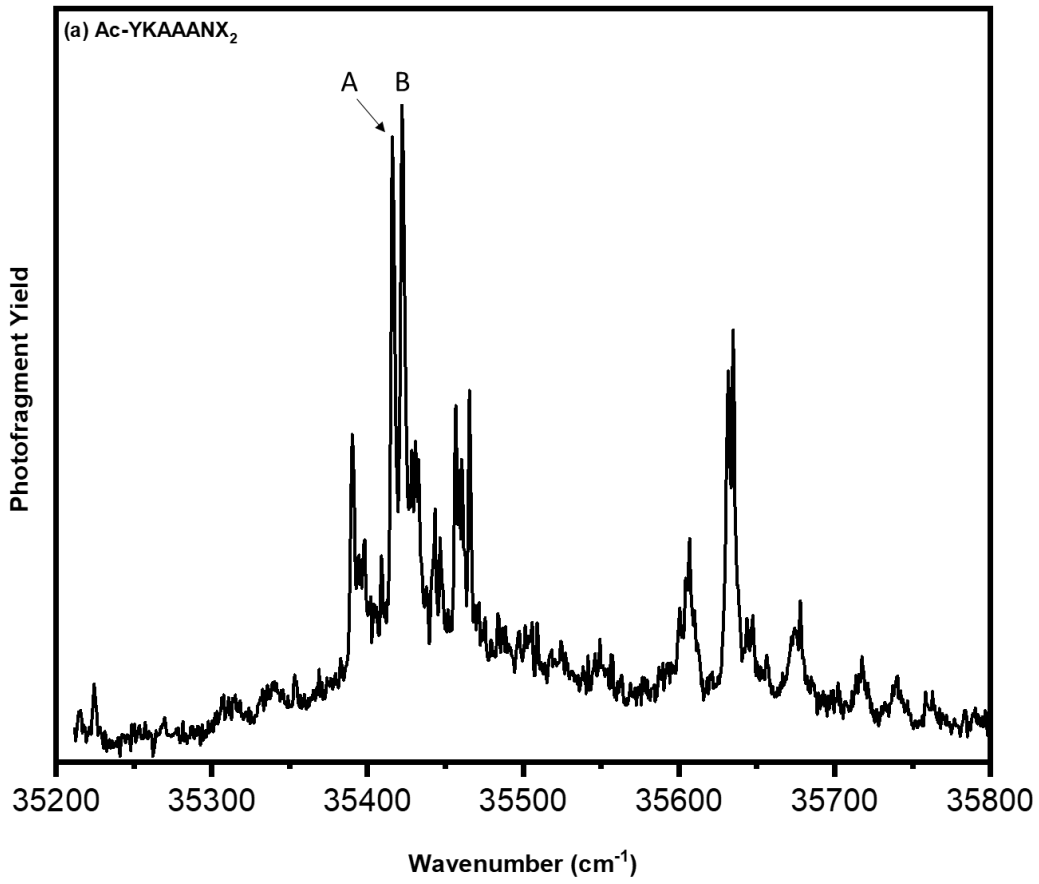

Figure 31 UV spectra of Ac-YKAAANX$_1$. The origins and Franck-Condon progressions of the different conformers are labelled A and B.

Structural analysis of all synthesized peptides began with acquiring a cold UV action spectrum in the tyrosine absorption region. UV spectra for all ions were produced by monitoring total photofragment signal across all loss channels versus laser wavenumber. Figure 30 contains the UV spectra of the cyclic peptides Ac-Y[KAAAD]X$_1$-NH$_2$, Ac-Y[KAAAD]X$_2$-NH$_2$, and Ac-y[kaaad]X$_1$-NH$_2$, while figure 31 displays the UV spectrum of the linear peptide Ac-YKAAANX$_1$. For all four peptides a definite electronic origin is found at 35400 cm^{-1}, blue shifted from that of protonated tyrosine (35081 cm^{-1}) by almost 400 cm^{-1}, indicating some interaction between the tyrosine ring and the rest of the peptide. Strong vibronic activity and a long Franck-Condon progression dominates all four spectra. The striking similarity between the cyclic peptide spectra was positive evidence that the location of the tyrosine residue, and maybe the structures themselves, were also very similar. Simplicity of the spectra also pointed to a single conformation for each of the cyclic peptides. After obtaining IR-UV infrared spectra for

each of the molecule's IR-UV holeburning was performed to confirm the number of conformations in each spectrum (Figures 31a-c). Results of the holeburning spectra indicate a single conformation present for the three cyclic peptides while the linear peptide has two distinct conformers. A more in-depth structural analysis of the secondary structure of the peptides required the recording of conformation specific infrared spectra utilizing IR-UV double resonance techniques.

5.4.3 Infrared Spectral and Structural Analysis of Ac-Y[KAAAD]X_1-NH_2, Ac-Y[KAAAD]X_2-NH_2, Ac-y[kaaad]X_1-NH_2, and Ac-YKAAANX$_1$

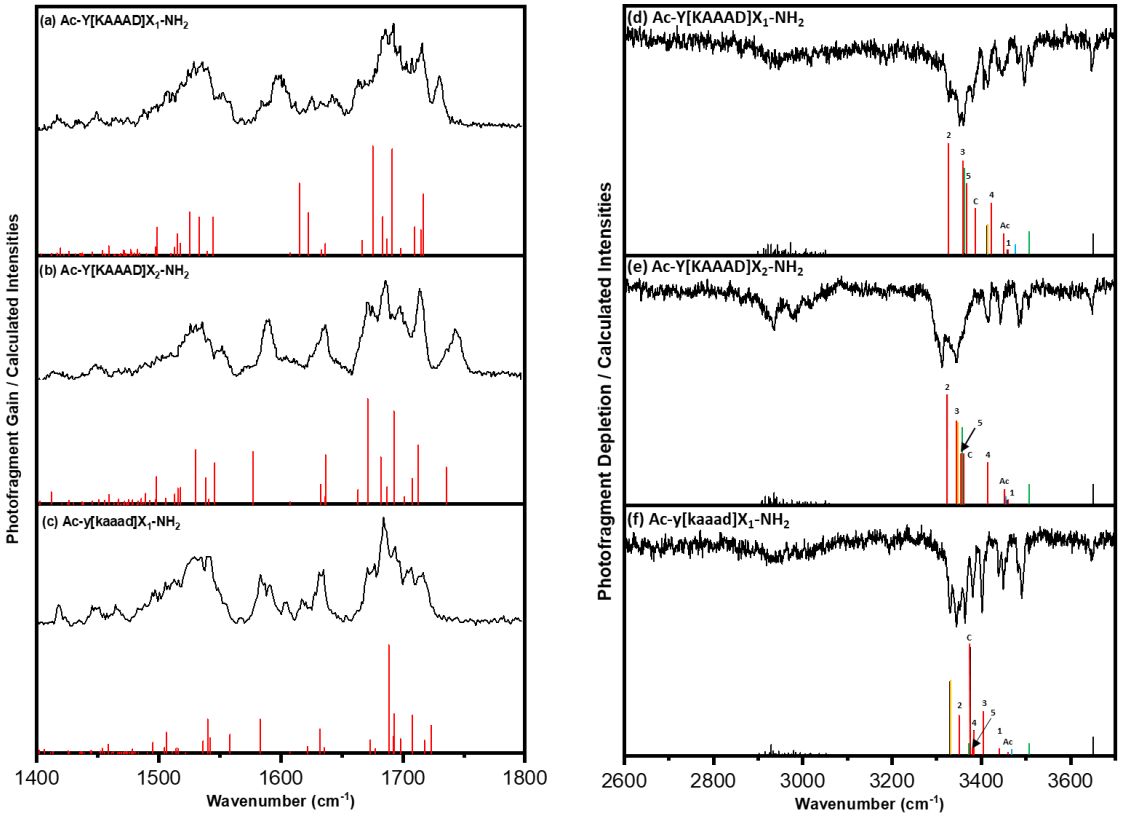

Figure 32 Conformation Specific Infrared Spectra of Ac-Y[KAAAD]X_1-NH_2, Ac-Y[KAAAD]X_2-NH_2, and Ac-y[kaaad]X_1-NH_2. (a-c): Corresponding experimental and calculated spectra in the 1400-1800 cm^{-1} region. (d-f): Conformation specific infrared spectrum in the hydride stretch region, compared to the calculated vibrational frequencies and infrared intensities of the best-fit conformations.

Conformation Specific Infrared Spectrum of Ac-Y[KAAAD]X_1-NH_2

The conformation specific infrared spectrum of Ac-Y[KAAAD]X_1-NH_2 is provided in figure 32, while a backbone schematic with labelled X-H oscillators is shown in figure 33. All but one of the transitions in the infrared spectrum are between 3300 and 3500 cm^{-1} with the bulk of the transitions centered around 3350 cm^{-1}. The pile up of transitions in this region is indicative of multiple strong hydrogen bonds, a positive sign that the secondary structure of the molecule is tight albeit not evidence alone of a helix. A lone transition outside of this X-H stretching region (~3650 cm^{-1}) is known to belong to the tyrosine OH moiety suggesting that it is free and non-interacting with the rest of the molecule. Deciphering of specific transitions in the infrared spectrum required the use of DFT calculations to generate infrared spectra that could be matched to the experimental spectrum. Figure 32 displays the calculated stick spectra matched to the experimental spectrum. Denoted as X_1, the charged group placed on this molecule is a methylated guanidinium group attached via a single methylene to the backbone chain. A single N-H oscillator is present on the charge, an intentional choice as the guanidinium group on the arginine residue can confound spectral analysis. This single N-H group appears in the infrared spectrum at 3413 cm^{-1} and is found in a weak hydrogen bond with the nearby carbonyl belonging to the charged moiety itself. Expectations for the charge site were that it would help induce the helix and through formation of a macrodipole and not inhibit helix formation via interactions with the backbone. This was achieved wonderfully as the charge only interacts with a local carbonyl and does not interfere with secondary structure formation by the rest of the molecule.

Vital to the formation of the solution phase helix was the c-terminal NH_2 cap that seeded the structure. The lack of a charged group in the solution phase to force a dipole meant that the burden of helix formation was placed upon this cap. In the gas phase the cap loses its relevance as the aforementioned charge takes priority. In Ac-Y[KAAAD]X_1-NH_2 one of the two N-H oscillators associated with the cap is found in a free position (3510 cm^{-1}) while the other is strongly hydrogen bound (3361 cm^{-1}) to the carbonyl of the tether forming a C10 hydrogen bond. As the carbonyl of the tether is in no position to interact with the remainder of the molecule an innocuous C10 hydrogen bond to the NH_2 cap does neither perturbs nor enhances the construction of the helix. The N-H group of the tether is not hydrogen bound (3476 cm^{-1}) and therefore plays no direct role in the secondary structure of the molecule. Most important to the visualization of the helix is the nature of the backbone N-H stretches of the ion.

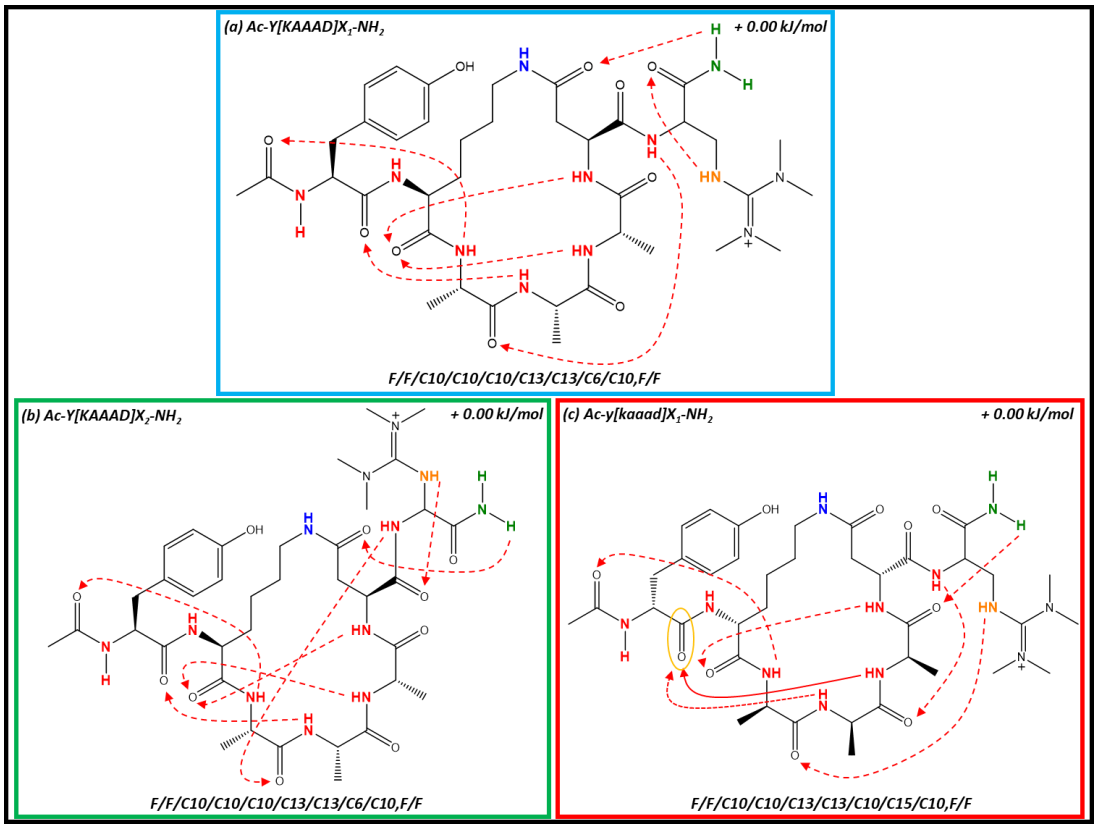

Figure 33 Stick structures of Ac-Y[KAAAD]X_1-NH_2, Ac-Y[KAAAD]X_2-NH_2, and Ac-y[kaaad]X_1-NH_2. Stick structure with highlighted hydrogen bonding pattern. A red colored arrow represents a N-C terminal hydrogen bond while a blue arrow is indicative of the opposite. Shorthand notation describing the presence or absence of hydrogen bonding at NH groups starting from the N-terminus is provided utilizing Cn notation where n equals the number of atoms contained by the macrocycle.

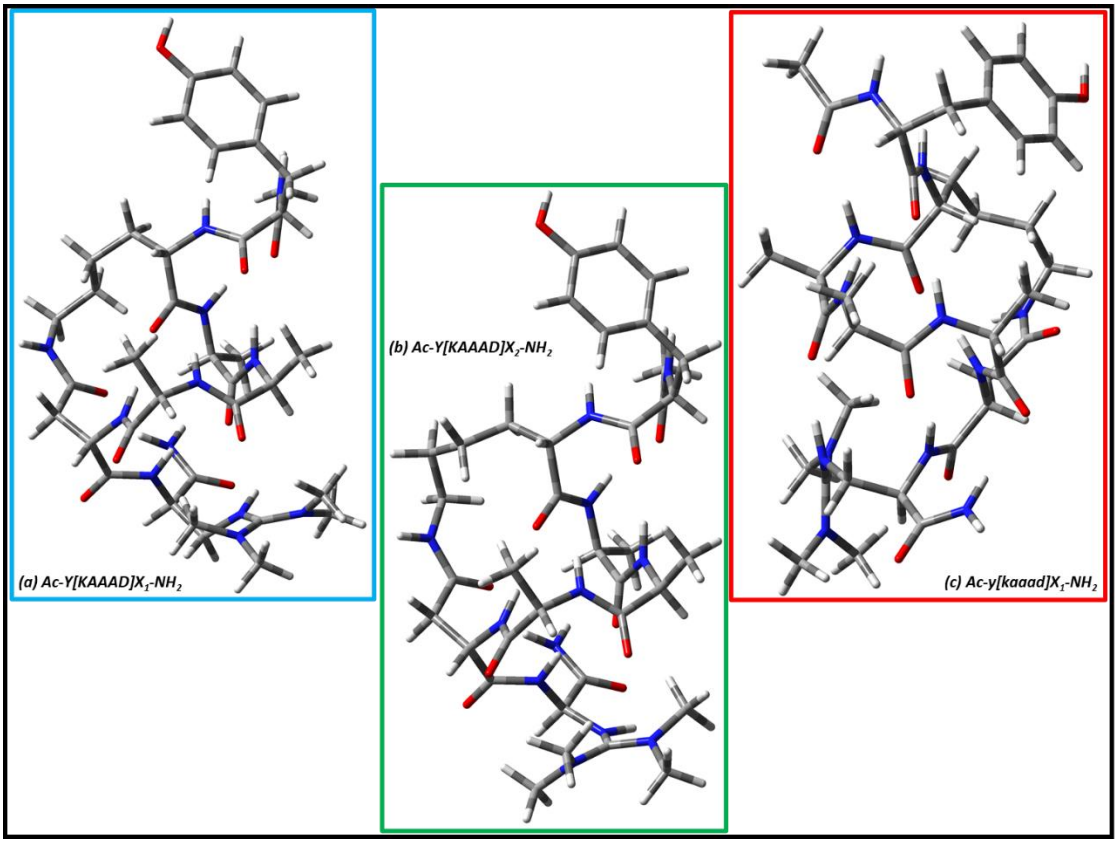

Figure 34 DFT calculated, best fit structures for Ac-Y[KAAAD]X_1-NH_2, Ac-Y[KAAAD]X_2-NH_2, and Ac-y[kaaad]X_1-NH_2

Figure 31 shows the naming scheme used to decipher the location of each of the backbone stretches. In brief, numbers 1-5 refer to the amide N-H oscillators along the traditional backbone of the structure. The term "Ac" indicates the acetylated n-terminus while "C" describes the N-H group on the backbone amide near the charge site. Coincidentally, the first two N-H oscillators in the series (Ac and 1) are the free oscillators of the group (3450 and 3458 cm^{-1}). Both are oriented towards the tyrosine side of the molecule following the direction of the remaining N-H groups along the macrodipole of the molecule. Their positioning in this manner meant that they were unable to form a hydrogen bond with any of the nearby carbonyl moieties, another structural motif that is necessary for formation of the helix. Crucial to the induction of the single turn helix is the hydrogen bonding patten observed with N-H groups 2-5. Observation of an alpha helix requires these groups to be in C13 hydrogen bonds to nearby carbonyls while the oft observed 3_{10} helix requires them to form C10 macrocycles. The infrared spectrum and DFT calculations reveal a mixed set of hydrogen bonds with the more n-terminal trio (2,3,4)

composing strict C10 hydrogen bonds to the acetyl carbonyl (2, 3326 cm^{-1}), tyrosine carbonyl (3, 3358 cm^{-1}), and lysine carbonyl (4, 3421 cm^{-1}). An abundance of C10 bonds early in the structure is opposed by the preponderance of C13 bonds formed at the opposite end of the ion. The remaining N-H oscillators (5, C) twist themselves into these C13 macrocycles through formation of hydrogen bonds with the lysine carbonyl (5, 3366 cm^{-1}) and the carbonyl of the first alanine (C, 3386 cm^{-1}). Resulting from this mixed set of hydrogen bonds is a molecule with a secondary structure that is neither pure alpha nor 3_{10} helix. Instead, a mixed C10/C13 helix appears to predominate given the hydrogen bonding lattice. While construction of the pure alpha (C13) helix was expected, observation of a helix of any type was a boon as it confirms the ability of the forcefield/DFT calculation approach to predict the expected conformations observed in the gas phase.

Conformation Specific Infrared Spectrum of Ac-Y[KAAAD]X_2-NH_2

The second charge site that produced mostly helical structures in the force field search was remarkably like the charge used in the previous molecule. This new charge site labelled X_2 in all figures contains the methylated guanidium group and the single N-H oscillator but is connected directly to the backbone of the molecule and does not have the methylene linkage that X_1 has. As X_2 is a less flexible charge site it was postulated that it would be less likely to form unwanted side structures. After obtaining the conformer specific infrared spectrum of this ion it was surprising, but not unexpected, to observe that the infrared spectrum nearly matched that of its sister molecule Ac-Y[KAAAD]X_1-NH_2. The largest difference between the two spectra being that Ac-Y[KAAAD]X_2-NH_2 has a larger, more convoluted set of transitions in the strongly bound X-H stretching region (3300-3400 cm^{-1}). As expected, the transition in the infrared spectrum that moved belong to the charge site, see yellow stick in figure 32. With the methylene linker absent the shorter chain forms a shorter C5 hydrogen bond with the nearby carbonyl of the backbone. This closer, stronger bond shifts the transition in the infrared spectrum to a lower wavenumber. In fact, several of the transitions in the spectrum of this molecule are shifted slightly lower in wavenumber than Ac-Y[KAAAD]X_1-NH_2 indicating a tighter stronger set of hydrogen bonds. Comparison of the stick spectra between the molecules highlights the similarity not only in hydrogen bonding pattern but of placement of the transitions in the spectra. The majority of transitions, except for the aforementioned charge site N-H, retain their order between

the spectra. Retention of the C10/C13 mixed helix with the introduction of a slightly different charge site was not unexpected.

Conformation Specific Infrared Spectrum of Ac-y[kaaad]X_1-NH_2

With the success of the of the methylated guanidinium group in seeding formation of the helix in Ac-Y[KAAAD]X_1-NH_2 and Ac-Y[KAAAD]X_2-NH_2 the next step was determining what control could be had over helix formation. A staple of the solution phase Fairlie helix is the ability to control handedness of the structure via the stereochemistry of the amino acids[7]. Switching the amino acid configuration to the D- form should correspond to a flip in the handedness of the helix. To test whether this switch in handedness would survive the transition to the gas phase a replica of the X_1 charge site containing peptide was synthesized with D-amino acids in place of the L-form for all residues except the charge site. Unfortunately, a D- version of the charged site was not readily available. To decipher between the two peptides a lower-case lettering scheme is used to represent the D- amino acids while upper case is maintained for the L- forms. This new peptide, dubbed Ac-y[kaaad]X_1-NH_2, was then interrogated with IR-UV double resonance techniques to determine its structure.

A single conformation was observed for Ac-y[kaaad]X_1-NH_2 which followed its L-form sister structure. Even more striking was the general similarity in infrared spectra with a large group of transitions again in the 3300-3400 cm^{-1} region of the spectrum and few outside of this region. More important was the differences between the spectra as the transitions in Ac-y[kaaad]X_1-NH_2 are shifted to a higher wavenumber than its sister structure, implying weaker hydrogen bonding. The transitions in the 3300-3400 cm^{-1} region are also better resolved as two of the peaks appear shifted lower in wavenumber clearing the congestion in the region. Further interpretation of the infrared spectrum required a comparison to the DFT generated, best fit stick spectrum (figure 32).

With the charge retaining the L- stereochemistry there was an opportunity for the structure to be perturbed greatly from the helix found in the all L- form of the peptide. The innocuous binding of the charge site N-H group to its nearby carbonyl in Ac-Y[KAAAD]X_1-NH_2 allowed the charge to seed the macrodipole without interfering with the structure. The charge in Ac-y[kaaad]X_1-NH_2 finds itself in a strong (3330 cm^{-1}) C16 hydrogen bond with the backbone carbonyl group of the first alanine residue. While this hydrogen bond does not affect

the formation of the helix it is still involved in an interesting motif as it forms the second rung of a beta-turn nestled into the c-terminal end of the peptide. The first rung of this beta turn is a C10 macrocycle formed by the backbone amine of the charge site (C, 3374 cm^{-1}). This set of hydrogen bonds represents the largest deviation from Ac-Y[KAAAD]X_1-NH_2 observed. The similarities between the peptides begin at the charged end of the peptide where a C10 hydrogen bond between a single N-H of the NH_2 cap and the carbonyl of the third alanine residue (3372 cm^{-1}). Its sister stretch is free (3507 cm^{-1}) along with the tether (3468 cm^{-1}) as was seen with the L- analogue. The backbone N-H groups also mirror that of its sister molecule with its first two being free (Ac, 1) followed by a set of C10 macrocycles (2,3). While the C10 formed by the amine group of the first alanine is of similar strength to its analogue (2, 3350 cm^{-1}) the second C10 arising from the second alanine is much weaker (3, 3404 cm^{-1}). A pair of C13 hydrogen bonds completing the mixed C10/C13 helix follows. The first between the N-H group of the third alanine residue and the tyrosine carbonyl (4, 3383 cm^{-1}) and the second from the aspartic acid amine to the lysine carbonyl (5, 3379 cm^{-1}). Maintenance of the mixed helix upon stereochemical inversion of the amino acids was a boon to understanding how these structures form in the gas phase.

Conformation Specific Infrared Spectrum of Ac-YKAAANX$_1$-NH$_2$

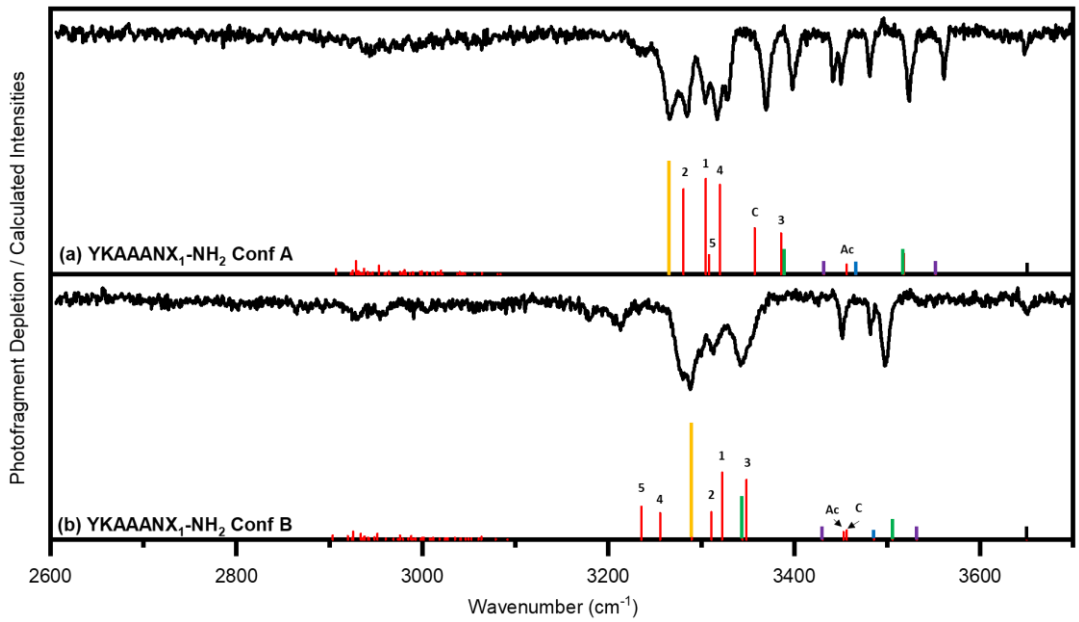

Figure 35 Conformation Specific Infrared Spectra of Conformers A (a) and B (b) of YKAAANX$_1$-NH$_2$

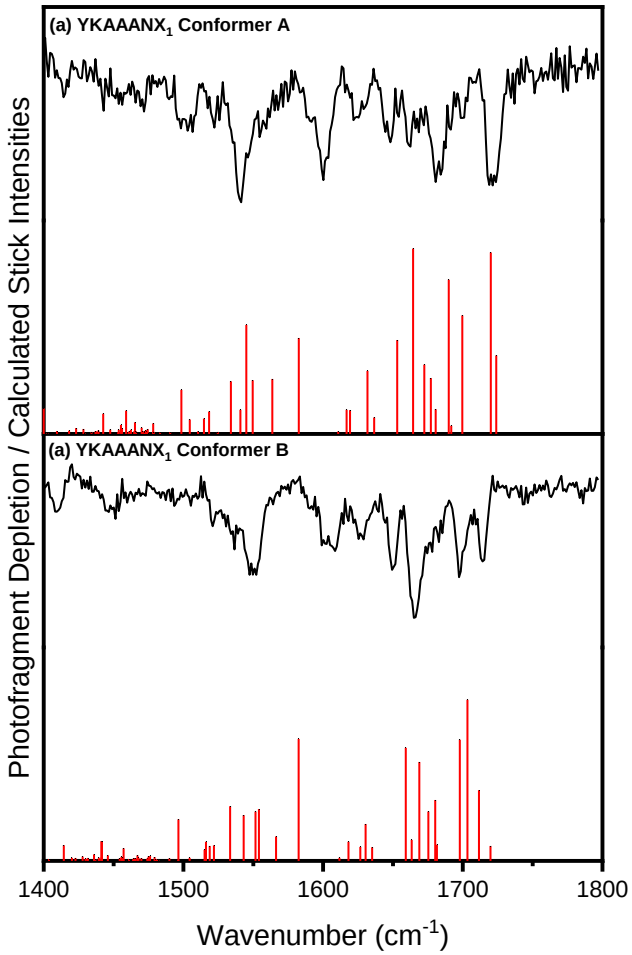

Figure 36 Amide I and II spectra of Conformers A (a) and B (b) of $YKAAANX_1$-NH_2

In the solution phase the interplay between the tether and the c-terminal NH_2 cap overcomes the ability of water to fully solvate the molecule which allows for the single turn helix to predominate. Where in the gas phase there is no water to disturb the structure instead the charge site plays the predominate role in structural formation. Of the tethered peptides studied for this article the terminal NH_2 cap predominately formed innocuous hydrogen bonds to nearby carbonyl groups which largely had no effect on helix formation. What is left to be understood is the affect of the tether on the structures of these peptides in the gas phase. As the tether is not directly engaging in hydrogen bonding in the majority of structures studied, it is difficult to discern its role in producing the helix. To better understand the role of the tether in formation of the single turn helix the linear, untethered peptide, Ac-$YKAAANX_1$-NH_2 was synthesized.

The cold UV spectrum of Ac-$YKAAANX_1$-NH_2 pointed to two possible conformations. IR-UV double resonance methods were used to acquire conformer specific infrared spectra of

each of the conformers (figure 35). Common to both infrared spectra is the large number of transitions observed in the 3300-3400 cm^{-1} region. This congestion of peaks having been observed with the tethered peptides was a sign of possible helix formation. While the similarities between the spectra are clear, starker are the differences. Conformer A produced several strong, sharp transitions in the region above 3350 cm^{-1} indicating several free N-H stretches in the molecule, whereas conformer B is littered with smaller, weaker transitions in the same region. These differences pointed to two different conformations with wildly dissimilar structures.

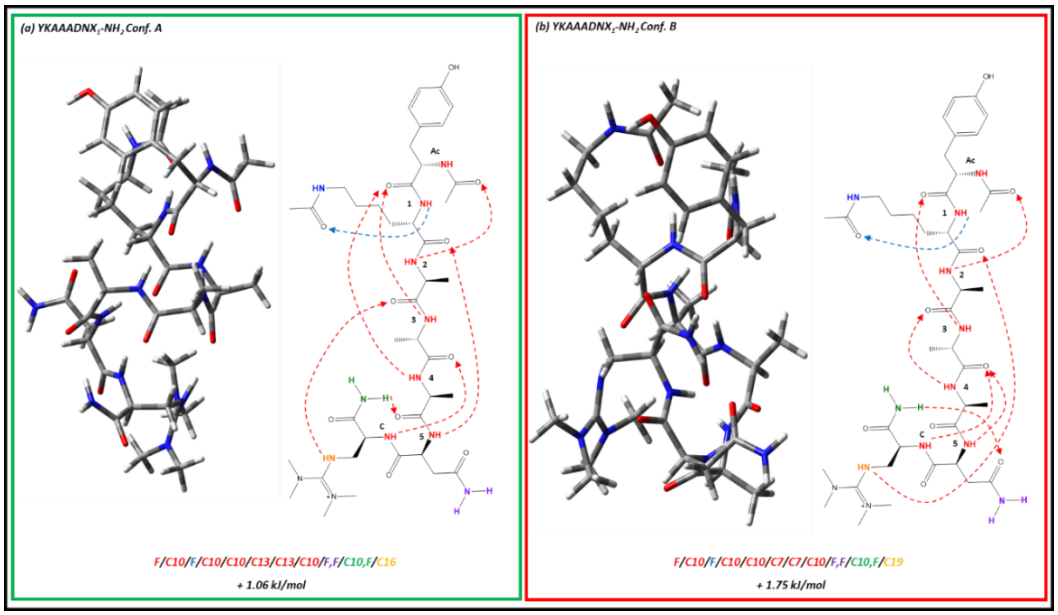

Figure 37 Calculated structures and stick spectra of Conformers A (a) and B (b) of $YKAAANX_1$-NH_2. Stick structure with highlighted hydrogen bonding pattern. A red colored arrow represents a N-C terminal hydrogen bond while a blue arrow is indicative of the opposite. Shorthand notation describing the presence or absence of hydrogen bonding at NH groups starting from the N-terminus is provided utilizing Cn notation where n equals the number of atoms contained by the macrocycle.

DFT calculated structures and their calculated stick spectra are provided underneath the experimental spectra in figure 37. Color and numerical coordination of the sticks in the calculated spectra follow the same guidelines as used with the previous molecules. Figure 37 displays the 3D structures as well as the hydrogen bonding pattern for both conformers. Conformer A presents as the mixed C10/C13 helix that was observed with the tethered peptides. Expectedly, the hydrogen bonding pattern of the backbone matches closely that of Ac-$Y[KAAAD]X_1$-NH_2 and Ac-$Y[KAAAD]X_2$-NH_2. A major difference between the tethered

species and this molecule is the bond formed by the amine group of the charge site. In the tethered species a small C6 macrocycle was observed originating with the charge N-H and terminating at the aspartic acid carbonyl. With conformer A the N-H group of the charge forms a C16 hydrogen with the carbonyl of the first alanine residue. This produces the second rung of a beta-turn that originates with the last backbone amine and ends at the carbonyl of the second alanine. Observation of the beta turn is unique as only the L- stereoisomer version of the tethered peptide produced the beta turn at the same location. Where conformer A follows the recognizable secondary structure that has been observed with the tethered helices, conformer B presents a weaving ball of hydrogen bonds that has no defined secondary structure. The n-terminal end of the peptide follows the similar pattern observed with the rest of the molecules, a free N-H stretch (Ac) followed by a series of C10 macrocycles (1,2,3). Key to construction of the helix are the hydrogen bonds in the center of the molecule formed by the second and third alanine residues. The structures that formed the helix all had a core pair of C13 hydrogen bonds in this region that defined the structure. Conformer B has a pair of C7 bonds in this area that sharply turn the backbone of the molecule back upon itself leading to a globular structure. From analysis of these structures, it can be concluded that while the tether is not necessary for formation of the single turn helix it does have a beneficial affect as it reduces the number of conformations observed.

5.4.4 Energy Level Diagrams

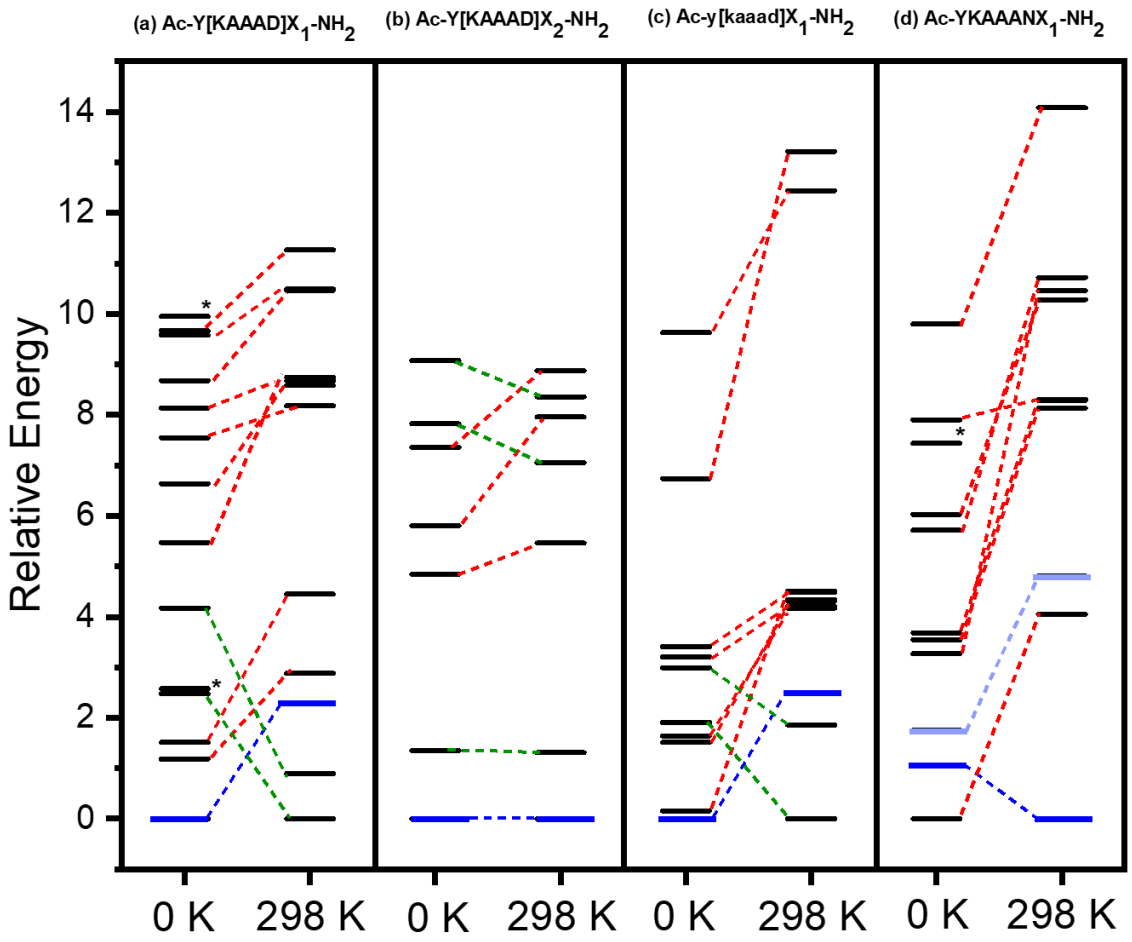

Figure 38 Energy level diagram for the four molecules studied containing both Zero Point Corrected Energies (0K) as well as Relative Free Energies (298K). Assigned conformations are highlighted in dark and light blue all other structures have black bars. An asterisk(*) denotes a change in energy from 0K to 298K that goes off the scale (above +15 kJ/mol). Red dashed lines indicate an increase in energy from 0 to 298K whereas a green dashed line indicates the inverse.

As the size and number of possible hydrogen bonding pairs of a molecule increases so does the complexity of the infrared spectrum. A common motif observed with all four molecules studied was a pile up of transitions in the 3300-3400 cm^{-1} region of the spectrum. This spectral complexity makes a firm call of the structure from calculated spectra much more difficult as several potential fits can exist. Numerous forms of analysis can be used to induce more confidence in the spectral fit. TD-DFT calculated S_o-> S_1 transition energies can be calculated to determine the origin of each conformer which can be matched to the experimental UV spectrum

(figure 34). Another excellent method for confirming spectral mapping is to use a different region of the infrared spectrum (Amide I and II) which is less congested and more amenable to matching (figure 34). A definitive call on the structure of a molecule can be made not only by producing calculated stick spectra but also through comparison of lowest energy structures produced for each molecule. To this end the use of energy level diagrams describing the energy of differing conformations and their structural differences goes a long way.

Figure 38 displays the energy level diagrams for each of the molecules studied. Energies provided are both zero-point corrected (0K) and free energies at 298K. Of note is that all conformations below +10 kJ/mol above the global minimum are helices for all the ions studied. In the region between +5-10 kJ/mol tyrosine bound helices dominated the structural pool. The presence of a free tyrosine stretch in all experimental spectra removed these structures as possibilities. A more in-depth analysis of the structures below +5 kJ/mol was needed to discern the final structure of each ion. Highlighted in blue is the assigned structure in each column with a blue dashed tie line showing the change in energy between 0 and 298K. Dashed red lines indicate an increase in energy between 0 and 298 K, while green lines indicate a decrease in energy.

At 0K the assigned structure of Ac-Y[KAAAD]X_1-NH_2 is the global minimum structure (figure 38a). At 298K this conformation rises in energy by 2.1 kJ/mol while two other structures that were much higher in energy at 0K (+2.3,+4.1 kJ/mol) fall beneath it in energy. The 298K global minimum structure is easily discounted as a possible conformation as it presents with a strongly bound NH_2 cap and a highly red shifted NH stretch belonging to the charged group. As neither of these features are present in the experimental spectrum this conformation could not possibly be the observed structure. In between the observed conformation and the minimum at 298 K is a structure with a highly similar amide I/II spectrum but in the stretching region deviates due to the presence of a very intense band belonging to the A_1 NH that appears shifted 60 cm^{-1} higher than where it is observed in the global minimum at 0K. The presence of this strong transition changes the profile of the calculated spectrum such that it in no way matches the experimental.

Of the four molecules studied, only Ac-Y[KAAAD]X_2-NH_2 remained the global minimum at both 0 and 298K (figure 38b). The exceptional stability of this molecule can be seen in just how low in energy it is compared to the other similar structures found for this ion. Only

two feasible conformations appear below 0 and +5 kJ/mol with the global minimum being the assigned structure. The 1.2 kJ/mol difference in energy between the minimum and this structure is difficult to explain by comparison of the calculated spectra as the difference between the two is only a small (2 cm^{-1}) change in the location and intensity of the NH stretch associated with the NH group on the charge. As the resolution of our experimental spectrum was not adequate to resolve this region confidence in the calling of one structure over the other is reduced. As both structures are very nearly identical it can be concluded that the observed conformation is likely the global minimum structure as this structure should predominate at ~10K.

With the experimental spectrum of Ac-y[kaaad]X_1-NH_2 following closely to its all L-amino acid version the expectation was that the energetics of the calculated structures would follow. This expectation became reality with the assigned global minimum structure at 0K rising in energy by nearly the same amount (+2.5 kJ/mol) at 298K (figure 38c). A major difference between this molecule and the L-amino acid version is the presence of a calculated conformation close in energy to the minimum (+0.515 kJ/mol). This conformer can be omitted as a potential structure as the calculated spectrum omits the pair of transitions observed near 3400 cm^{-1} in the experimental spectrum. At 298K the 0K global minimum is supplanted by two structures neither of which replicate the experimental spectrum. Both the 298K minimum and the structure falling at +2.1 kJ/mol fail to reproduce the congested region of the experimental spectrum between 3325 and 3375 cm^{-1}, specifically an intense transition at 3370 cm^{-1}. This leaves the assigned structure as the only low energy structure to accurately mirror the experimental spectrum.

With two conformers and the similarity in structures observed form the conformational search, Ac-YKAAANX$_1$-NH_2 was expected to be difficult to match. Fortunately, the increase in flexibility afforded to the molecule by the removal of the tether meant that many of the calculated low energy structures contained a bound tyrosine OH moiety. With both experimental spectra presenting with a free tyrosine OH spectral matching to calculated structures was simplified. Of the six structures below +5kJ/mol at 0K (figure 38d) four are tyrosine bound helices that do not match up to the experimental spectra (figure 38). The remaining two calculated spectra lie +1.06 and +1.75 kJ/mol above the minimum. Comparison of the experimental spectra of conformers A and B to these calculated spectra produces good alignment making them the most likely choices for the structure of the ions. At 298 K the assigned conformer A becomes the new global minimum with conformer B raising in energy by almost 3

kJ/mol. Only one structure lies lower in energy than conformer B at 298K, namely the aforementioned global minimum at 0K.

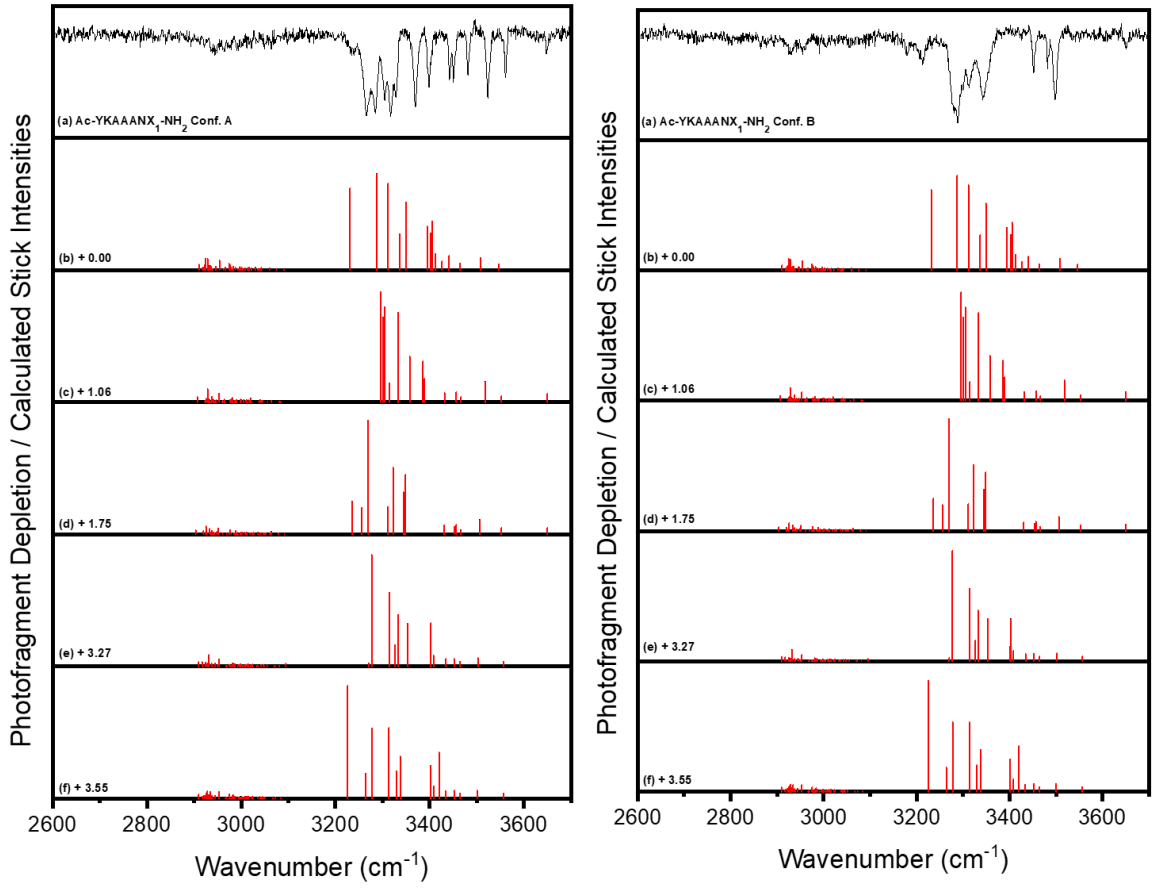

Figure 39 Calculated stick spectra of the five lowest energy conformers of $YKAAANX_1$ compared to the experimental spectra of each conformer

5.5 Conclusions

Chemical synthesis via directed computational design was used to predict the formation of a gas phase alpha helix. Cold ion spectroscopy confirmed not only the presence of strong helical structures in both Ac-Y[KAAAD]X_1-NH_2 and Ac-Y[KAAAD]X_2-NH_2 but confirmed the formation of the opposite turn helix with Ac-y[kaaad]X_1-NH_2. The ability of amide tether to control the conformational landscape of the molecule was affirmed through the study of Ac-$YKAAANX_1$-NH_2 which generated multiple conformations in the gas phase unlike its tethered analogues. Formation and control of the helix opens the door to further studies on how much control over the secondary structure is possible with this motif.

5.6 References

1. Nicolaou, K. C., Organic synthesis: the art and science of replicating the molecules of living nature and creating others like them in the laboratory. *Proc Math Phys Eng Sci* **2014,** *470* (2163), 20130690-20130690.

2. Elkin, M.; Newhouse, T. R., Computational chemistry strategies in natural product synthesis. *Chemical Society Reviews* **2018,** *47* (21), 7830-7844.

3. Law, R.; Barker, O.; Barker, J. J.; Hesterkamp, T.; Godemann, R.; Andersen, O.; Fryatt, T.; Courtney, S.; Hallett, D.; Whittaker, M., The multiple roles of computational chemistry in fragment-based drug design. *Journal of Computer-Aided Molecular Design* **2009,** *23* (8), 459-473.

4. Mandal, S.; Moudgil, M. n.; Mandal, S. K., Rational drug design. *European Journal of Pharmacology* **2009,** *625* (1), 90-100.

5. Korshunova, M.; Ginsburg, B.; Tropsha, A.; Isayev, O., OpenChem: A Deep Learning Toolkit for Computational Chemistry and Drug Design. *Journal of Chemical Information and Modeling* **2021,** *61* (1), 7-13.

6. de Araujo, A. D.; Hoang, H. N.; Kok, W. M.; Diness, F.; Gupta, P.; Hill, T. A.; Driver, R. W.; Price, D. A.; Liras, S.; Fairlie, D. P., Comparative alpha-helicity of cyclic pentapeptides in water. *Angew Chem Int Ed Engl* **2014,** *53* (27), 6965-9.

7. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Left- and Right-Handed Alpha-Helical Turns in Homo- and Hetero-Chiral Helical Scaffolds. *Journal of the American Chemical Society* **2009**.

8. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005**.

9. 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 beta-Hairpin Turn. *J Am Chem Soc* **2017,** *139* (15), 5481-5493.

10. 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 beta-Hairpin. *J Am Chem Soc* **2016,** *138* (8), 2849-57.

11. 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.

12. 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.

13. 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.

14. Londry, F. A.; Hager, J. W., Mass selective axial ion ejection from a linear quadrupole ion trap. *Journal of the American Society for Mass Spectrometry* **2003,** *14* (10), 1130-1147.

15. Mohamadi, F.; Richards, N. G. J.; 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. *Journal of Computational Chemistry* **1990,** *11* (4), 440-467.

16. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993,** *98* (7), 5648-5652.

17. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011,** *32* (7), 1456-1465.

18. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H., Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr. *Chemical Physics Letters* **1989,** *157* (3), 200-206.

19. Kranenburg, R. F.; van Geenen, F. A. M. G.; Berden, G.; Oomens, J.; Martens, J.; van Asten, A. C., Mass-Spectrometry-Based Identification of Synthetic Drug Isomers Using Infrared Ion Spectroscopy. *Analytical Chemistry* **2020,** *92* (10), 7282-7288.

20. 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* (7), 1374-8.

21. Sliwoski, G.; Kothiwale, S.; Meiler, J.; Lowe, E. W., Computational Methods in Drug Discovery. *Pharmacological Reviews* **2014,** *66* (1), 334-395.

22. Ou-Yang, S.-s.; Lu, J.-y.; Kong, X.-q.; Liang, Z.-j.; Luo, C.; Jiang, H., Computational drug discovery. *Acta Pharmacologica Sinica* **2012,** *33* (9), 1131-1140.

23. Rackers, J. A.; Wang, Z.; Lu, C.; Laury, M. L.; Lagardère, L.; Schnieders, M. J.; Piquemal, J.-P.; Ren, P.; Ponder, J. W., Tinker 8: Software Tools for Molecular Design. *Journal of Chemical Theory and Computation* **2018,** *14* (10), 5273-5289.

24. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A., OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation* **2016,** *12* (1), 281-296.
25. Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A., Development and testing of a general amber force field. *Journal of Computational Chemistry* **2004,** *25* (9), 1157-1174.

26. Downs, G. M.; Barnard, J. M., Clustering Methods and Their Uses in Computational Chemistry. In *Reviews in Computational Chemistry*, 2002; pp 1-40.

CHAPTER 6. STEREOCHEMICAL CONTROL OF THE STRUCTURAL PROPERTIES OF GAS PHASE SINGLE TURN HELICES

6.1 Abstract

Cold ion spectroscopy was used to interrogate a series of tethered peptides that utilize differences in stereochemistry to form left- and right-handed alpha helices. A pair of homochiral molecules, Ac-Y[KAAAD][KAAAD]X_1-NH_2 and Ac-y[kaaad][kaaad]X_1-NH_2 were found to form two turn alpha helices with right and left handedness, respectively. Heterochiral versions of these peptides were studied to determine whether mixed helices could be found in the gas phase. One molecule, Ac-Y[KAAAD][kaaad]X_1-NH_2, was found to form a left-handed helix similar to the homochiral molecules. While another heterochiral molecule, Ac-y[kaaad][KAAAD]X_1-NH_2, loses almost all of its helicity. Structural analysis of the molecule, Ac-YAAA[KAAAD]X_1-NH_2 provides insight into the necessity of the tether in continuing the helical structure of the peptide.

6.2 Introduction

Control of the structural properties of synthetic molecules is key to their function as medicines[1], pesticides[2], or polymers[3]. Synthetic chemists achieve these structural changes most often through the manipulation of functional groups attached to a molecule. For example, a simple addition of an isopropyl group to a small drug molecule can greatly affect its binding ability[4]. Another typical way to control the structure of small molecules is through selective chirality changes[5]. Biological systems governing metabolism and excretion of small molecules are greatly affected by chiral interactions[6]. Therefore, control of the chirality of a synthetic molecule is a vital step in organic syntheses. Where peptide therapeutics are concerned adjusting the chirality of certain amino acids can stabilize different secondary structural motifs such as beta turns[7] and alpha helices[8] which can then be used as drug scaffolds.

Where control of small, single turn alpha helices is concerned, Fairlie and colleagues have devised an ingenious strategy of tethering small peptide molecules in ways that induce helix formation in the solution phase[8-10]. They have also found a way to control the handedness of these helices through control of the stereochemistry of different amino acid residues[8]. In our search for a single turn helix in the gas phase this chemical control mechanism was also utilized

to switch between left- and right-handed helices. This study proposes taking this method of control one step further through the formation of double turn helices of mixed handedness in the gas phase.

To test whether double turn helices of mixed handedness can be observed in the gas phase as ions synthesized peptides utilizing the Fairlie tethering motif were produced. Two homochiral peptides, Ac-Y[KAAAD][KAAAD]X_1-NH_2 and Ac-y[kaaad][kaaad]X_1-NH_2 are first studied to ensure that transfer of these species to the gas phase does not perturb their secondary structure found in the solution phase. Brackets denote the location of the tethered regions of each peptide while X_1 stands for the charged synthetic amino acid, a methylated guanidinium group, that is used to form the ion. Capitalized letters refer to L-amino acids and lower-case letters refer to D-amino acids. Observation of a right-handed helix for L-amino acid containing tethered peptides and left-handed helices for D-amino acid containing rings was expected. The heterochiral peptides, Ac-y[kaaad][KAAAD]X_1-NH_2 and Ac-Y[KAAAD][kaaad]X_1-NH_2 where then prepared and analyzed with the expectation of observing both left- and right- handed helical sections inside the molecule. Lastly, a molecule with only one tethered ring was synthesized, Ac-YAAA[KAAAD]X_1-NH_2, to examine the effect of a single tethered region to seed helix formation in an untethered section of the molecule. Structural analysis of these gas phase ions was performed via cold ion spectroscopy utilizing IR-UV double resonance spectroscopic techniques. In the past this method of structural determination has proven invaluable in the study of beta turn formation[11], gas phase population studies[12], and formation of single turn alpha helices. The focus of this study is to understand the spectroscopic characteristics of gas phase double turn helices.

6.3 Experimental

6.3.1 Instrumentation

All spectroscopic data were taken on a custom-built instrument for cold ion spectroscopy, previously described.[13-14] In brief, the design of the instrument comprises a triple quadrupole mass spectrometer in which an orthogonal spectroscopy axis was added between the second (q2) and third (q3) quadrupoles. Ions are formed via nano-electrospray ionization(nESI) and transported into q2 where the ion of interest is isolated via RF/DC isolation. The ion packet is

then steered around a turning quadrupole and down the spectroscopy axis where it is trapped in the cold quadrupole that is cooled to 5K via a closed-cycle helium cryostat (Sumitomo Heavy Industries, Tokyo, Japan). After cooling to ~10K via collisions with the He buffer gas, the ions are spectroscopically interrogated via infrared (Laservision OPO/OPA) and ultraviolet (ScanMate-Pro Lambda Physik frequency doubled by an Inrad Autotracker III and Scanmate Lambda Physik) lasers. The created photofragments are then extracted back down the spectroscopy axis and trapped in q3 for analysis in either "spectroscopy mode" or "mass spectrometry mode". In mass spectrometry mode, a mass spectrum of the photofragments is generated via mass-selective axial ejection.[15] Spectroscopy mode entails ejection of the remaining precursor ions with SX wave software followed by propelling of the photofragments onto a channeltron detector (4773G, Photonis USA).

UV action spectra are recorded by monitoring the total photo-fragment signal as a function of laser wavenumber. Upon obtaining a UV spectrum and identifying transitions that may be due to different ion conformations, conformation-specific IR spectra are taken by employing IR-UV double resonance techniques.[11] The UV laser is tuned to a transition of interest in the UV spectrum which creates a steady signal of photofragment ions. Then a tunable IR laser that is spatially overlapped but temporally precedes the UV laser by 200 ns is scanned in wavelength. Upon absorption of an IR photon by the conformer of interest, a fraction of the population is removed from the ground state which results in a depletion in the total photofragment signal.

6.3.2 Solution Preparation

Solutions of Ac-Y[KAAAD][KAAAD]X_1-NH_2, Ac-y[kaaad][KAAAD]X_1-NH_2, Ac-Y[KAAAD][kaaad]X_1-NH_2, Ac-y[kaaad][kaaad]X_1-NH_2, and Ac-YAAA[KAAAD]X_1-NH_2 were prepared as 50:50 Methanol:Water to a concentration of 200 μM and electrosprayed to create the corresponding protonated ion.

6.3.3 Computational

Conformational searches were performed for Ac-Y[KAAAD][KAAAD]X_1-NH_2, Ac-y[kaaad][KAAAD]X_1-NH_2, Ac-Y[KAAAD][kaaad]X_1-NH_2, Ac-y[kaaad][kaaad]X_1-NH_2, and

Ac-YAAA[KAAAD]X_1-NH_2 structures using the Monte Carlo multiple minimum method with the AMBER* and OPLS3 forcefields inside the MACROMODEL software package.[16] Approximately 10000 structures were found between the two forcefields for each of the molecules. To narrow down the potential structures further, the 10000-structure pool was clustered based on backbone RMSD. The centroid of each of the clusters was then optimized at the B3LYP/6-31+G(d)-GD3BJ level of theory[17-19] using Gaussian16 and harmonic vibrational frequencies/intensities were calculated to obtain predicted infrared spectra to compare with experiment. To correct for anharmonicity the calculated spectra were scaled by 0.958 for the NH, CH, and OH stretching regions, 0.973 for the free OH transitions, 0.981 for Amide I, and 0.970 for Amide II. Adjusting the calculated transitions by these values puts them into agreement with previously calculated and fitted spectra.[11-12, 20] "Stick" spectra were generated for the centroid of each cluster. These calculated spectra were then matched to the experimental spectrum and the best fit centroids were found. All clusters of structures with best fit centroids that lie within 10 kJ/mol of the global minimum were then run through DFT at the same level of theory used for the centroids. After completing all calculations, the structure that best fit the experimental spectrum and was lowest in energy was determined to be the most likely structure of the ion. A further analysis of the forcefield searching method is detailed in section 3.1.

6.4 Results and Analysis

6.4.1 Analysis of UV Spectra of Ac-Y[KAAAD][KAAAD]X_1-NH_2, Ac-y[kaaad][KAAAD]X_1-NH_2, Ac-Y[KAAAD][kaaad]X_1-NH_2, Ac-y[kaaad][kaaad]X_1-NH_2, and Ac-YAAA[KAAAD]X_1-NH_2

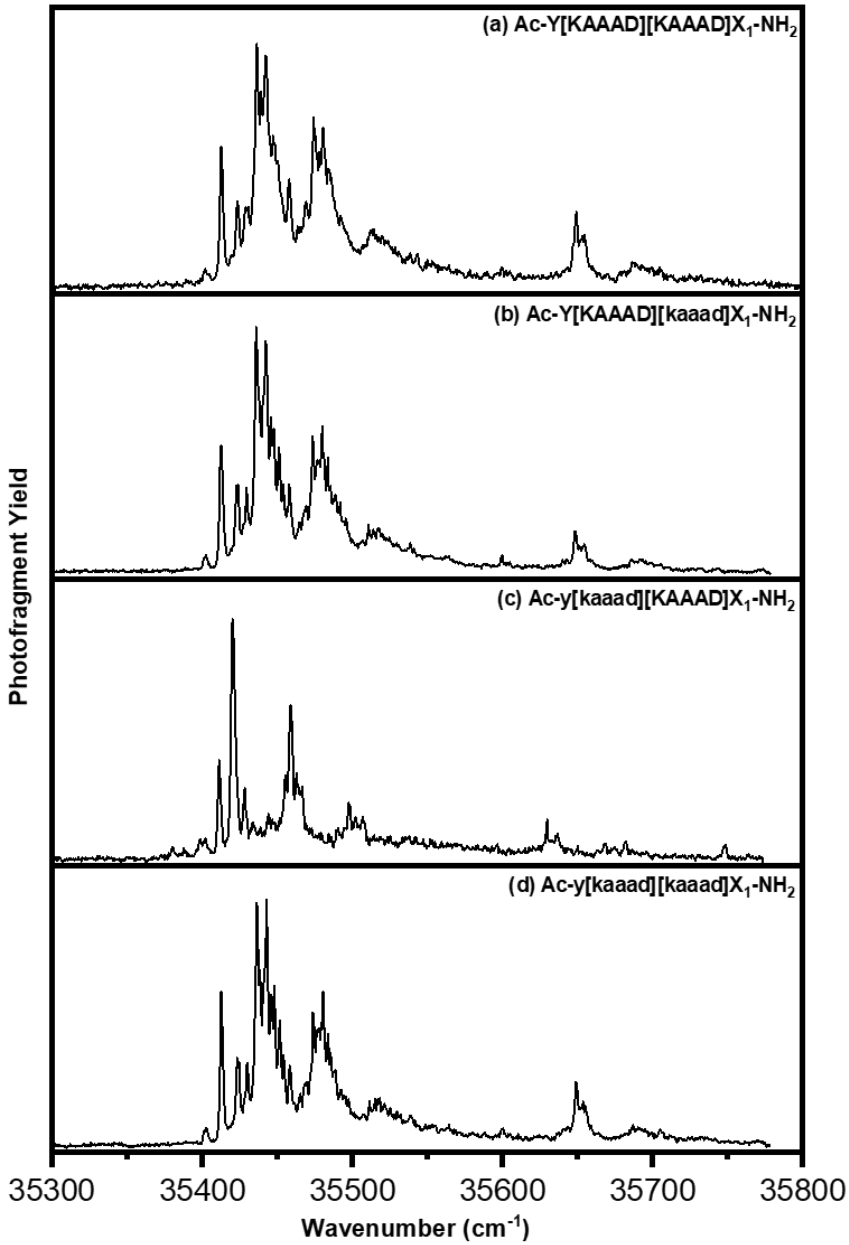

Figure 40 UV spectra of (a) Ac-Y[KAAAD][KAAAD]X_1-NH_2, (b) Ac-Y[KAAAD][kaaad]X_1-NH_2, (c) Ac-y[kaaad][KAAAD]X_1-NH_2,and (d) Ac-y[kaaad][kaaad]X_1-NH_2

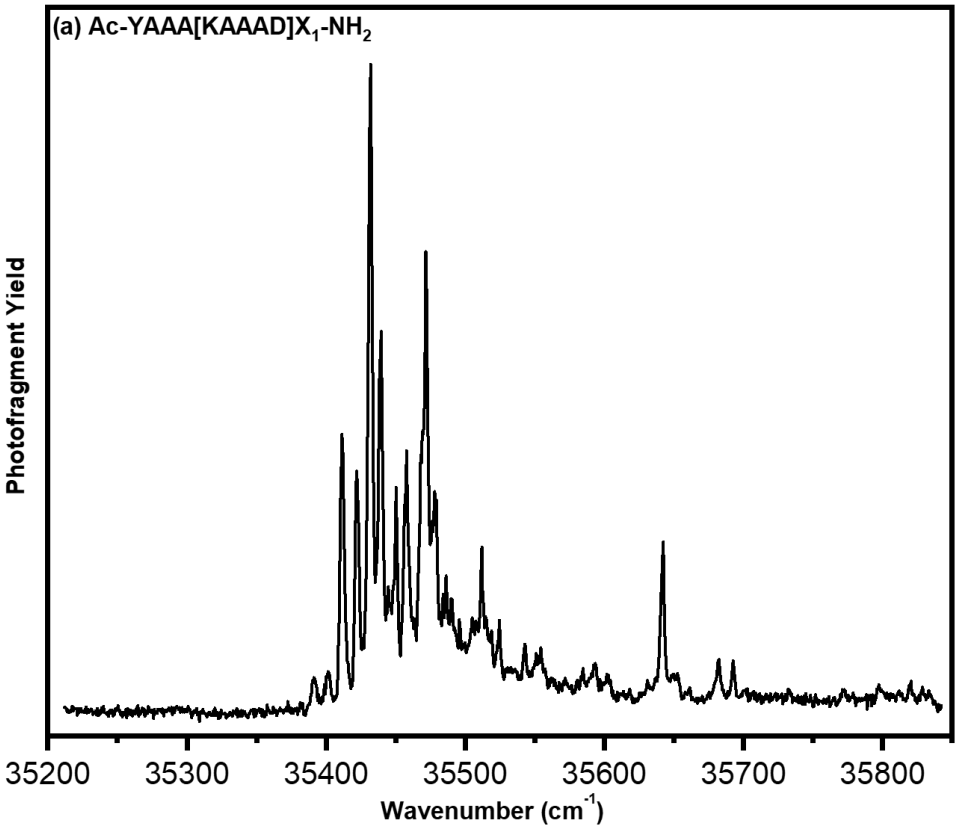

Figure 41 UV spectrum of (a) Ac-YAAA[KAAAD]X_1-NH_2

Initial analysis of the set of tethered peptides began with acquiring a cold UV action spectrum in the tyrosine absorption region. UV spectra for all ions were produced by monitoring total photofragment signal across all loss channels against laser wavenumber. Figure 41 contains the UV spectra of the tethered peptides Ac-Y[KAAAD][KAAAD]X_1-NH_2, Ac-y[kaaad][KAAAD]X_1-NH_2, Ac-Y[KAAAD][kaaad]X_1-NH_2, Ac-y[kaaad][kaaad]X_1-NH_2, while figure 42 displays the UV spectrum of the partially linear peptide Ac-YAAA[KAAAD]X_1-NH_2. All five peptides display a definite electronic origin at 35400 cm^{-1}, blue shifted from that of protonated tyrosine (35081 cm^{-1}) by almost 400 cm^{-1}, indicating some interaction between the tyrosine ring and the rest of the peptide. Strong vibronic activity and a long Franck-Condon progression dominates all five spectra. The striking similarity between three of the four tethered peptide spectra was positive evidence that the location of the tyrosine residue, and maybe the structures themselves, were also very similar. Simplicity of the spectra also pointed to a single conformation for each of the cyclic peptides. After obtaining IR-UV infrared spectra for each of the molecule's non-conformation specific IR gain spectra were recorded to confirm the number

of conformations in each spectrum (Figures 42a-d). Comparison of the gain and conformer specific spectra indicate a single conformation present for all five of the tethered peptides studied. A more in-depth structural analysis of the secondary structure of the peptides required the recording of conformation specific infrared spectra utilizing IR-UV double resonance techniques.

6.4.2 Infrared Spectral and Structural Analysis of Ac-Y[KAAAD][KAAAD]X_1-NH_2, Ac-y[kaaad][KAAAD]X_1-NH_2, Ac-Y[KAAAD][kaaad]X_1-NH_2, Ac-y[kaaad][kaaad]X_1-NH_2, and Ac-YAAA[KAAAD]X_1-NH_2

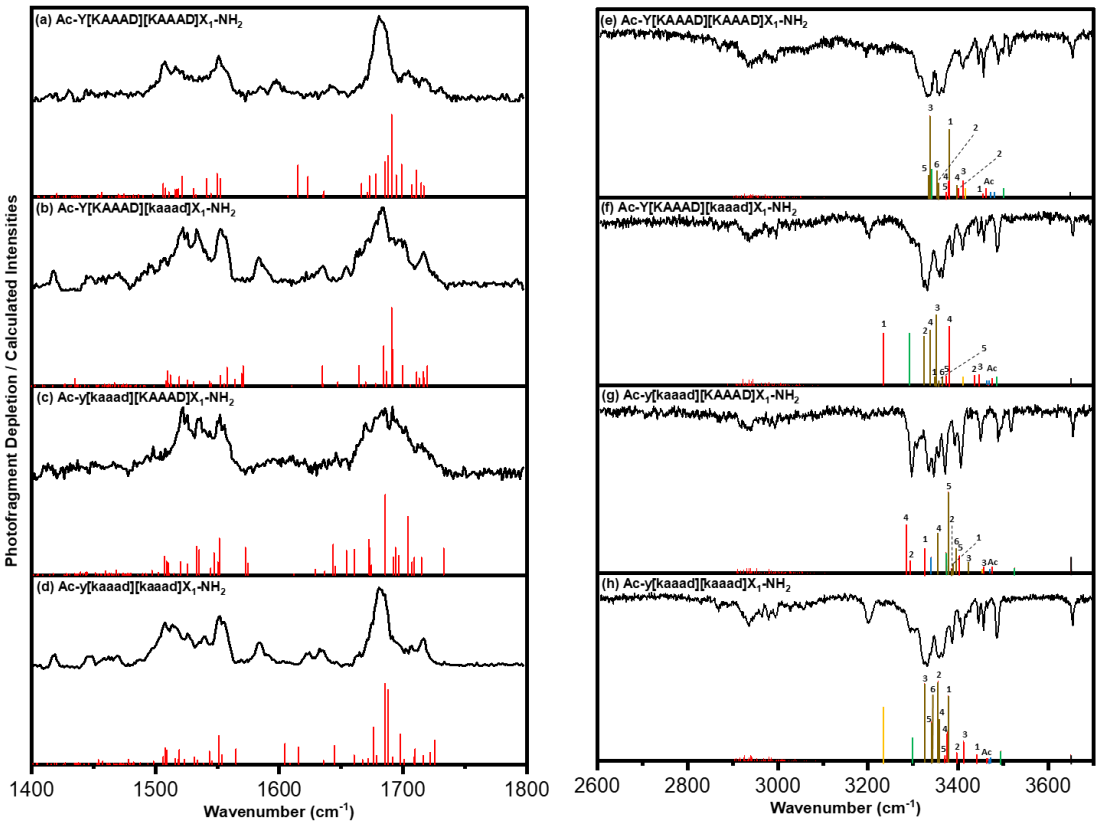

Figure 42 IR spectra of (a) Ac-Y[KAAAD][KAAAD]X_1-NH_2, (b) Ac-Y[KAAAD][kaaad]X_1-NH_2, (c) Ac-y[kaaad][KAAAD]X_1-NH_2,and (d) Ac-y[kaaad][kaaad]X_1-NH_2 in the Amide I/II region and the stretching region (e-h) Conformation specific infrared spectra of all ions are shown in black and best fit calculated stick spectra in color below. Colors of each of the sticks in the calculated spectra correlate to the color scheme provided in figure 40.

Conformation Specific Infrared Spectrum of Ac-Y[KAAAD][KAAAD]X_1-NH_2

The conformer specific infrared spectrum, taken in the stretching and amide regions, for Ac-Y[KAAAD][KAAAD]X_1-NH_2 is provided in figure 42e. With 17 XH oscillators present in the molecule the expectation, and reality, was a congested experimental spectrum. Two features stand out in the spectrum. The first is the free tyrosine OH stretch at 3650 cm^{-1} and the second is the pile up of transitions in the 3300-3400 cm^{-1} region. A free tyrosine OH is an important indicator of structure as the flexibility of these large molecules increases the probability of the molecule to fold over on itself. The tyrosine being free indicates that this likely did not occur. In our previous studies of the shorter versions of these peptides the double mountains of transitions in the 3300-3400 cm^{-1} region was a clear indicator of the formation of a helix as the NH oscillators form similar strength bonds. Further understanding of the experimental spectrum required the calculation and fitting of possible structures and their spectra.

The best fit calculated spectrum to the experimental can be seen in figure 42e while the structure and hydrogen bonding pattern that fit this spectrum are provided in figure 42a. With the complexity of the experimental spectrum and the large pool of possible low energy structures bearing similar calculated spectra confidence in the fit must be increased by examining multiple pieces of evidence. Examining the data through the window of the energy level diagram given in figure 6a will drive a path to the best fit structure. DFT obtained energies of a large pool of possible conformations were obtained and used to filter structures that were far too high in energy to be achieved. Using zero-point corrected energies there are a total of seven structures in the range of +10 kJ/mol above the global minimum. Two of these seven structures have bound tyrosine OH groups which immediately removes them from consideration. The remaining five structures include the global minimum and a cluster of conformations between +3-4.5 kJ/mol. Three of the five structures can be omitted as they contain a strongly bound NH group associated with the charge site (labelled orange in stick structure) that appears in the calculated spectrum at 3250 cm^{-1} which does not appear in the experimental spectrum. Differentiation between the remaining two structures can be observed through comparison of the infrared spectra in the amide I/II regions (figure 3a) where the assigned structure is an unmistakably better fit. With greater confidence in the global minimum as the assigned structure a last step of conformation was obtained by calculating the S0-S1 transition energy (figure S3a). The calculated origin of the

molecule fits very well to the origin observed in the experimental UV spectrum providing greater assurance that the assigned structure was indeed the correct fit.

With certainty of the assignment in hand analysis of the observed structure can be performed. Based on our previous studies of smaller single turn versions of these molecules a helical structure was expected and indeed the assigned global minimum is a two-turn alpha helix where the tethered regions anchor each of the turns. Formation of the alpha helix requires a successive chain of C13 hydrogen bonds that begins with the second alanine residue (labelled in red, number 3 on stick structure) and works its way down the backbone. At the c-terminal end of the backbone the charged residue anchors the formation of the helix through stabilization of the macrodipole. The utility of the two tethers to reduce the flexibility of the molecule and therefore increase the likelihood of the helix find themselves unbound and non-interacting with the rest of the molecule. While the c-terminal NH_2 cap that anchors the helix in the solution phase[10] forms only a single C10 hydrogen bond to the nearby carbonyl of the tether thereby limiting its effect on the greater structure of the molecule.

These common features of the helix result in a congested spectrum as the large number of similar hydrogen bonds creates a pile up of transitions. Comparison of the calculated stick spectrum to the experimental spectrum does provide some interesting results. Of note are the differences in strength and intensity of the hydrogen bonds that form the backbone C13 bonds of the molecule. The backbone NH groups of the first ring (labelled in red) appear less intense and shifted higher in wavenumber than those of the second ring (labelled in brown). This implies that the hydrogen bonds formed by the second ring are much stronger and tighter than those of the first ring. With the second ring being bookended by the first ring as well as the charge this result is not surprising. The charge site is a strong anchor for the positioning of the second ring and the inflexibility of the first ring means that the second has little room for forming multiple structures. Contrast this to the first ring which has the free tyrosine residue on one end and the second ring on the other. The free n-terminal end, with no nearby carbonyl groups to form strong hydrogen bonds, is more flexible and able to affect the structure of the first ring. Analysis of the structures produced by the conformational search lend to this idea as many of the molecules found themselves with a helical second ring and non-helical first ring. In all, the expectation for this molecule was that the two turn helix would be observed. This ion served as a control molecule

with which to compare the attempts at stereochemical control and observe any spectroscopic differences which may result.

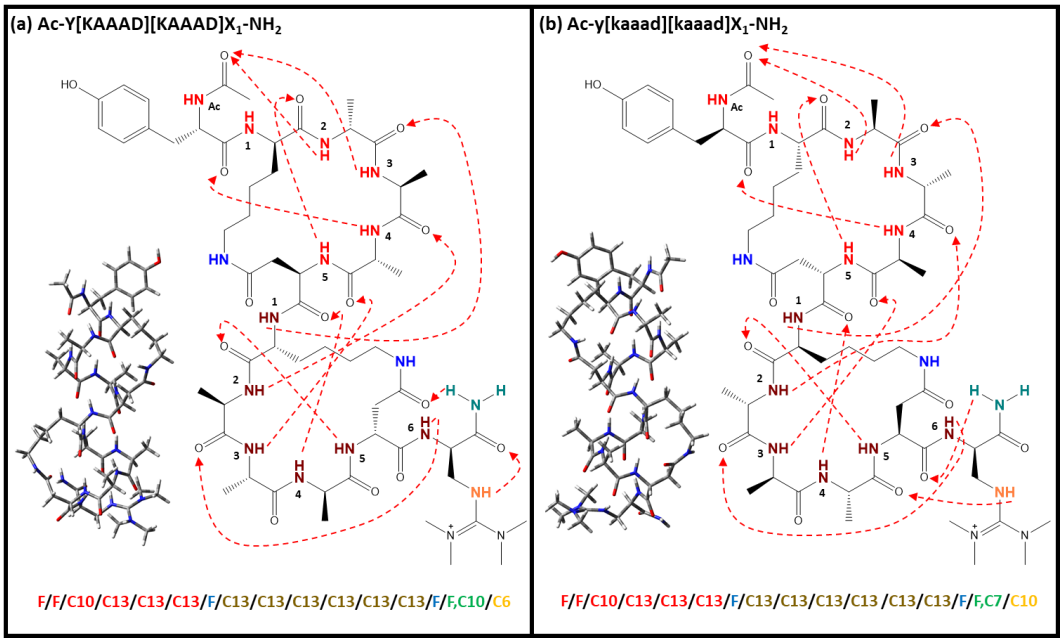

Figure 43 Stick structures of (a) Ac-Y[KAAAD][KAAAD]X_1-NH_2 and (b) Ac-y[kaaad][kaaad]X_1-NH_2 Dashed arrows indicate hydrogen bonds with the arrow originating from the donor and terminating on the acceptor. Blue coloration denotes an N->C terminal hydrogen bond while red indicates a C->N terminal bond. Shorthand notation describing the presence or absence of hydrogen bonding at NH groups starting from the N-terminus is provided utilizing Cn notation where n equals the number of atoms contained by the macrocycle.

Conformation Specific Infrared Spectrum of Ac-y[kaaad][kaaad]X_1-NH_2

With the assignment of the all L-amino acid version of the molecule secured the first attempt at stereochemical control of the molecule involved switching to all D-amino acids. The expectation of this switch would be that an opposite handed helix would form that mirrors its sister structurally and spectroscopically. Figure 42h displays the conformation specific infrared spectrum of this molecule in both the amide I/II (d) and XH stretching (h) regions. Immediately apparent is the similarity to the opposite stereochemistry spectrum. The set of large transitions between 3300 and 3400 cm^{-1} remain as do the small transitions higher in wavenumber, including the free tyrosine OH stretch. One difference stands out, a transition at 3200 cm^{-1} indicating a strong hydrogen bond. To determine what this spectroscopic feature may indicate the best fit structure first had to be ascertained.

Eight structures populate the lowest 10 kJ/mol above the global minimum in zero-point corrected energy (figure 46b). Of the eight structures only two display a transition in the region of 3200 cm^{-1} that could possibly match the same transition in the experimental spectrum. Differentiation between the two remaining structures required analysis of the amide I/II spectra given in figure 42d. Only the assigned global minimum structure adequately described the doublet at 1589/1597 cm^{-1}. TD-DFT calculated S0->S1 transition energy was calculated and used to determine the electronic origin of the assigned conformer. This origin closely matches that of the experimental UV spectrum (figure 42d) greatly increasing confidence in this structure being the correct match.

Expectations for this molecule were that of a mirror turn to the all L- amino acid form and this is precisely what was observed. A striking similarity between the bonding pattern of both rings as well as the n-terminal region of the molecule is observed. The C13 hydrogen bonds that hallmark the alpha helix are present throughout the chain. The genesis of the mystery transition in the experimental spectrum could now be uncovered. Due to the inability to locate a D-form of the charge residue for synthesis of this peptide the charge site remains in the canonical L-form. While this change had no significant effect on the overall secondary structure of the molecule it did change the local hydrogen bonding environment around itself. The effect of this change is a different set of hydrogen bonds formed at the c-terminal end of the molecule involving the charge NH and the NH_2 cap. Firstly, the NH_2 cap forms a C7 macrocycle with the aspartic acid residue of the second ring. Secondly, the charge site NH group finds itself in a strong C10 hydrogen bond to the carbonyl of the third alanine group in the second ring (brown, 5). This bond is the origin of the mystery transition at 3200 cm^{-1}.

Spectroscopic similarities between the two enantiomers are predictable as the structures so closely resemble each other. The second ring dominates the spectrum with stronger and more intense bonds that shift the cluster of transitions closer to 3300 cm^{-1}. Weaker bonds are observed in the first ring these transitions centering around 3400 cm^{-1}. Difference in strength of the hydrogen bonds of the rings and the separation in the transitions that result begins to form a fingerprint for what a double turn helix with these molecules should look like in the infrared spectrum. At 3301 cm^{-1} the NH_2 cap symmetric stretch is blue shifted about 30 cm^{-1} from its observed position in the control molecule indicative of a strong hydrogen bond.

Conformation Specific Infrared Spectrum of Ac-Y[KAAAD][kaaad]X₁-NH₂

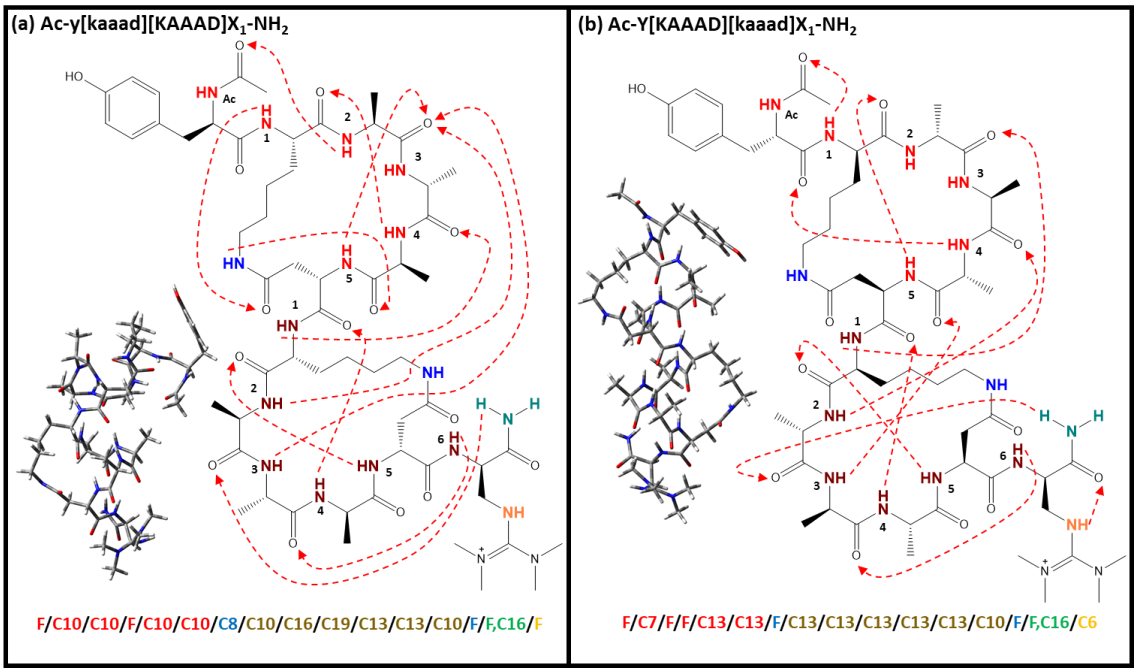

Figure 44 Stick structures of (a) Ac-y[kaaad][KAAAD]X_1-NH_2 and (b) Ac-Y[KAAAD][kaaad]X_1-NH_2 Dashed arrows indicate hydrogen bonds with the arrow originating from the donor and terminating on the acceptor. Blue coloration denotes an N->C terminal hydrogen bond while red indicates a C->N terminal bond. Shorthand notation describing the presence or absence of hydrogen bonding at NH groups starting from the N-terminus is provided utilizing Cn notation where n equals the number of atoms contained by the macrocycle.

With the all L/D- amino acid forms of the molecule behaving as expected, attention was then devoted to the possibility of forming double turn helices with mixed turns. Changing the chirality of the amino acids in one of the rings was expected to allow for the formation of the mixed helix without perturbing the structure greatly. Upon obtaining the conformation specific infrared spectrum of Ac-Y[KAAAD][kaaad]X_1-NH_2 (figure 42b and 42f) the similar spectra to the mono-helices was an excellent sign of possible helix formation. The common motifs for these helices were again observed including the large pile up of transitions between 3300 and 3400 cm^{-1}, the smaller transitions at higher wavenumber, and the single transition at 3200 cm^{-1}. Determination of the best fit structure to the experimental spectrum would provide more insight into the nature of this structure.

The energy level diagram describing the conformational landscape of this molecule is provided in figure 45c. Only six structures populate the lowest 10 kJ/mol above the global

minimum. As with the other molecules structural similarity abounds in the low energy conformers of this ion. Use of the transition at 3200 cm^{-1} as a delineator of spectral similarity quickly removes all but two conformations from consideration. Of the two remaining conformations the global minimum forms a left-handed helix throughout the majority of the structure that begins to unravel at the n-terminus. The conformer +0.98 kJ/mol higher in energy forms a similar helix in the ring two section of the molecule but breaks down completely into a macrocycle in ring one. Analysis of the calculated spectra of each of the structures easily places the global minimum as the best fit structure as it replicates the double mountain of transitions near 3300 cm^{-1}, but also has a single transition at 3200 cm^{-1} matching the experimental spectrum. The higher energy conformation does not replicate the double mountain motif nor does it have a single transition in the 3200 cm^{-1} region, instead it has three.

The goal in inverting the stereochemistry between the rings was to create a helix that is partially right-handed (ring 1) and partially left-handed (ring 2). Instead, what is observed is a pure left-handed helix that mimics the structure of the all D- amino acid configuration with subtle differences near the ends of the structure. At the n-terminus Ac-Y[KAAAD][kaaad]X_1-NH_2 has the most free NH oscillators of any of the molecules studied. The C7 bond formed by the lysine NH and the acetylation site carbonyl twists this end of the molecule in such a way that the third NH group (red,3) does not form the typical C10 observed in all the other structures. Added to this free amine are the free NH groups of the acetylation site and the second amine. On the c-terminal end of the molecule the NH_2 cap, which is innocuously bound in the purely homochiral helices, now finds itself in a C16 macrocycle with the carbonyl of the first alanine of ring 2. The closeness of the cap to this carbonyl may be what displaces the formation of a C13 bond to the sixth NH of the second ring and this carbonyl as was observed in the pure helices. The tether and hydrogen bound NH_2 cap appear to flank the helical backbone of the structure locking it in place. Backbone hydrogen bonding follows the pattern of a prototypical alpha helix with C13 hydrogen bonds originating from the aspartic acid amine of the second ring (brown,5) and extending up the structure to the third alanine residue of the first ring (red,4).

More insight into the structural properties of this tethered peptide can be gleamed through understanding how the experimental spectrum changes with structure. The largest structure in the experimental spectrum is the double mountain of transitions between 3300 and 3400 cm^{-1}. In the homochiral structures these peaks are associated with the NH stretches along the backbone for

each ring of the molecule. This holds true with Ac-Y[KAAAD][kaaad]X_1-NH_2 as the amine stretches associated with the second ring populate the mountain closer to 3300 cm^{-1} and those of the first ring the mountain near 3400 cm^{-1}. Of interest is the peak close to 3200 cm^{-1} which was associated with the charge site in the D-amino acid version of the tethered peptide. In this molecule that transition is in line with a strong C7 hydrogen bond between the first amine of ring one and the carbonyl of the acetylation site. The bonding of this amine is contrary to how it was found in the homochiral molecules. When bound it orients the second and third amines in such a way as they are not in a position to hydrogen bond. This change is observed in the spectrum where a doublet appears at 3444 and 3456 cm^{-1} that is associated with the free stretching of those two amines.

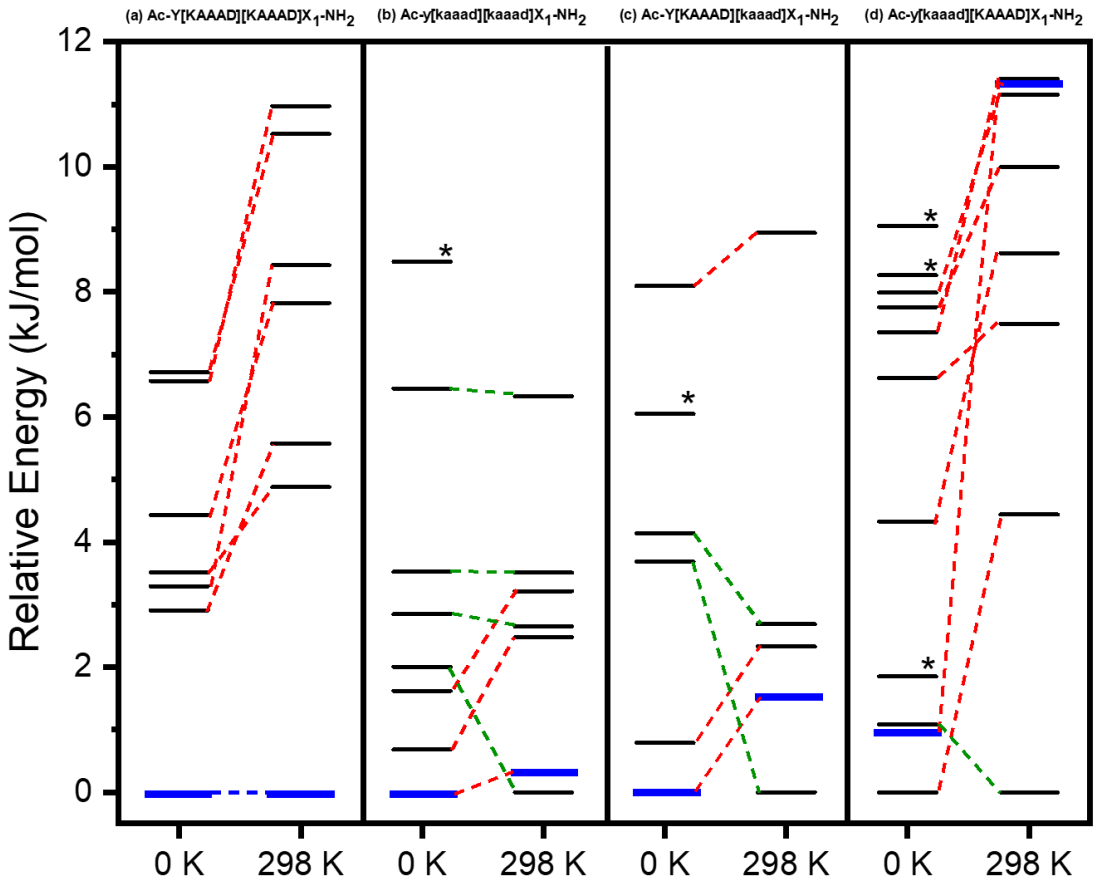

Figure 45 Energy level diagram for the four molecules studied containing both Zero Point Corrected Energies (0K) as well as Relative Free Energies (298K). Assigned conformations are highlighted in dark blue all other structures have black bars. An asterisk(*) denotes a change in energy from 0K to 298K that goes off the scale (above +15 kJ/mol). Red dashed lines indicate an increase in energy from 0 to 298K whereas a green dashed line indicates the inverse.

Conformation Specific Infrared Spectrum of Ac-y[kaaad][KAAAD]X_1-NH_2

The experimental infrared spectrum of Ac-y[kaaad][KAAAD]X_1-NH_2 was initially jarring as it did not follow the established pattern of the other double turn helices in this study. The oft mentioned double mountain motif between 3300 and 3400 cm^{-1} is replaced by several well resolved, sharp transitions between 3250 and 3400 cm^{-1}. As the typical infrared pattern was not associated with a set of mixed turn helices in the previous molecule studied the next thought was that this may be the infrared fingerprint of the mixed helices. To determine the validity of this hypothesis the experimental spectrum was fit to the calculated best fit spectrum and the structure that that spectrum belonged to assigned as the conformer.

With eleven structures populating the region +10 kJ/mol above the global minimum, this molecule had the most potential conformations of any studied. A diagram depicting the relative energies of these structures at 0K and 298K is provided in figure 45d. Of note is that this is the only molecule where the assigned conformation was not the global minimum, instead this structure lies +0.95 kJ/mol above the global minimum. The global minimum, as well as four other structures, were omitted from the pool as it contained a bound tyrosine OH group that was not observed in the experimental spectrum. Only two structures below +7 kJ/mol exhibited a free tyrosine OH, the assigned structure (+0.95 kJ/mol) and the other at (+1.09 kJ/mol). This second structure was a similar fit to the assigned but did not match critical transitions in the spectrum as well as the former, including the large transition at 3291 cm^{-1}. The same holds true for the bulk of structures above +7 kJ/mol where the transitions in the calculated spectrum were generally shifted to the higher wavenumber region of the spectrum resulting in misalignment with the experimental. Indeed, the best fit of all the structure was the assigned structure at +0.95 kJ/mol above the minimum. It not only replicated the XH stretch region of the spectrum excellently but reproduced the amide I/II regions (figure 42c) as well was the TDDFT calculated origin matched the UV spectrum.

With the fit of the experimental spectrum complete one can begin to understand what about the structure of the molecule makes the infrared spectrum so unique. As with its sister molecule, the goal of Ac-y[kaaad][KAAAD]X_1-NH_2 was to produce a structure containing both left- and right-handed helices. This proved to be a difficult task as the assigned structure presents with only the second ring as partially helical while the first ring is an unstructured macrocycle. In all the molecules studied so far, the second ring has been the most stable helix of the two with its stronger hydrogen bonds and little room for flexibility this is hypothesized to be due to the proximity of the charge. While the first ring completely breaks down in this structure the second ring manages to retain some of its helical structure even while losing the majority of its C13 hydrogen bonds. Amine groups 1-3 in the second ring do not form the C13 macrocycles observed in all the other structures. Instead, they form C10,C16, and C19 bonds to the less structured first ring. The helical secondary structure of the second ring is maintained by the bonds formed by amines 4 and 5, keeping the C13 hydrogen bonds needed to observe the alpha helix in the second ring. While the second ring attempts to maintain the helix the first ring abandons it entirely. A series of C10 hydrogen bonds that turn the backbone upon itself defines

the first ring of the molecule. The most interesting aspect of the first ring is a single carbonyl group, belonging to the first alanine residue, that is the donor for three different hydrogen bonds. Two of these bonds originate with the second ring, C16 and C19 bonds from the first and second alanine residues. While a C10 to the same carbonyl also arises from the aspartic acid amine of the first ring. The first ring also has the only hydrogen bound tether of all the molecules studied, forming a C8 bond to the carbonyl of the third alanine residue. Vast differences in hydrogen bonding pattern and strength serve to separate transitions in the infrared spectrum and gives it the unique appearance.

Comparison of the experimental spectrum to the calculated spectrum reveals some unique features. Unlike the other molecules the transitions belonging to the second ring are less intense and shifted higher in wavenumber, a move that explains the more relaxed structure of the second ring. Meanwhile, the transitions associated with the first ring are shifted lower in wavenumber implying stronger bonds. The bound tether of the first ring appears at 3340 cm^{-1} shifted far down from its typical unbound position near 3450 cm^{-1}.

Conformation Specific Infrared Spectrum of Ac-YAAA[KAAAD]X$_1$-NH$_2$

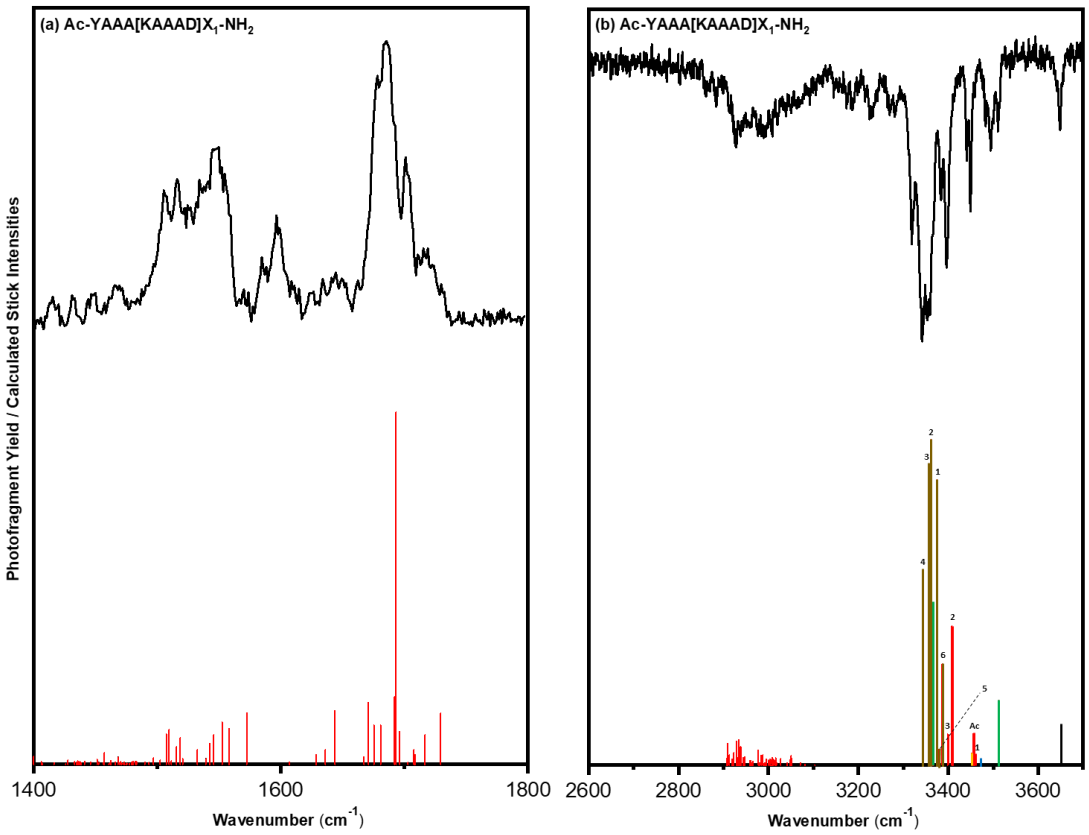

Figure 46 IR spectra of Ac-YAAA[KAAAD]X$_1$-NH$_2$. Conformation specific infrared spectra of the ion is shown in black and the best fit calculated stick spectrum in color below. Colors of each of the sticks in the calculated spectra correlate to the color scheme.

The finding that the homochiral molecules formed stable helices in both rings was excellent in confirming the power of these motifs to form tight secondary structures. Observing that result begged the questions as to whether a single tethered ring in combination with the charge site was enough to seed helix formation in an untethered section. To examine whether this would occur or not the molecule Ac-YAAA[KAAAD]X$_1$-NH$_2$ was synthesized. Placing the tether in the second ring was thought to provide the best opportunity to observe a continuous helical structure throughout the molecule. The similarity of the experimental spectrum to those of the homochiral molecules provided hope that the helix would indeed be observed, but to confirm DFT calculated structures had to first be obtained and sorted.

The increased flexibility afforded to the n-terminal end of this molecule due to removal of the tether meant that far more variable structures appeared in the conformational search. Focusing on the region below +10 kJ/mol above the global minimum, ten conformations appear as possible candidates for the structure of the molecule. Three of the higher energy structures can quickly be eliminated from contention as they contain strong transitions in the calculated spectrum below 3300 cm^{-1} that do not appear in the experimental spectrum. Another has a bound tyrosine OH which also excludes it as a possibility. The remaining six structures had very similar calculated spectra in the XH stretching region. Only an analysis of the amide I/II regions resulted in differentiation of the structures. Four structures were immediately removed as they contain a strong transition at 1730 cm^{-1} that does not appear in the experimental spectrum. The remaining two molecules had similar calculated spectra in both regions. Delineation between the two was made possible by examining which of the two calculated spectra best fit the doublet at 1581/1598 cm^{-1}. Only the assigned global minimum fit these peaks well enough to be considered the correct structure.

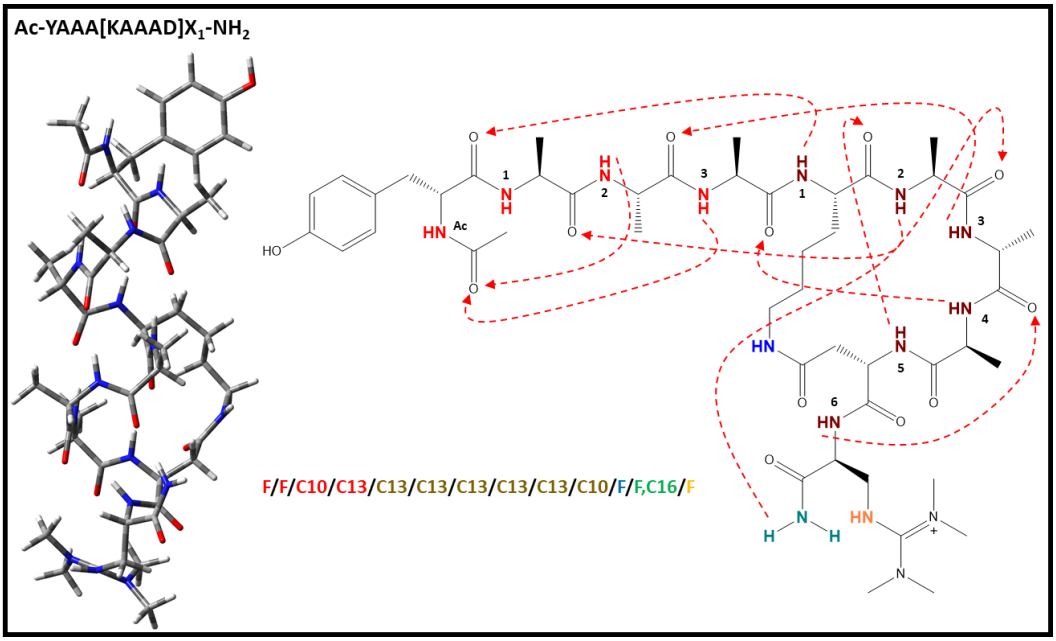

Figure 47 Stick structure of Ac-YAAA[KAAAD]X_1-NH_2. Dashed arrows indicate hydrogen bonds with the arrow originating from the donor and terminating on the acceptor. Blue coloration denotes an N->C terminal hydrogen bond while red indicates a C->N terminal bond. Shorthand notation describing the presence or absence of hydrogen bonding at NH groups starting from the N-terminus is provided utilizing Cn notation where n equals the number of atoms contained by the macrocycle.

Removal of the tether on the first ring was a test of the ability of the charge and second ring to stabilize the helix throughout the whole molecule. In the assigned structure the tethered ring forms the same C13 hydrogen bound alpha helix observed in the other molecules. A C16 macrocycle between the NH_2 cap and the carbonyl of the first alanine residue is the only hydrogen bond found c-terminal to the ring as the charge NH group is free. The residues outside of the ring took on several forms in the structural search with many conformations having a random jumble of hydrogen bonds in this region. However, these conformations were higher in energy than the global minimum helix structure. The common motif of this region in the other molecules studied is retained here with the first two amine groups being free, followed by a C10/C13 bifurcated hydrogen bond from the second and third alanine to the acetyl carbonyl. The stability afforded by the formation of the helix makes its observation here not at all surprising.

While the structure follows closely the components of the other homochiral helices as does the infrared spectrum. The double mountain motif between 3300 and 3400 cm^{-1} is present but muted due to the reduced number of NH groups on the n-terminal end of the molecule in compared to the others. As with the other molecules the transitions associated with the second ring are higher in intensity and lower in wavenumber. This difference is afforded to the stronger hydrogen bonds formed by the ring amine groups in comparison to untethered region. The free stretching region is populated by a stack of closely spaced transitions belonging to the charge, acetyl, first alanine, and tether amine groups. In all the spectrum closely aligns with the expectation for the helix showing several sections that match the other helix spectra closely.

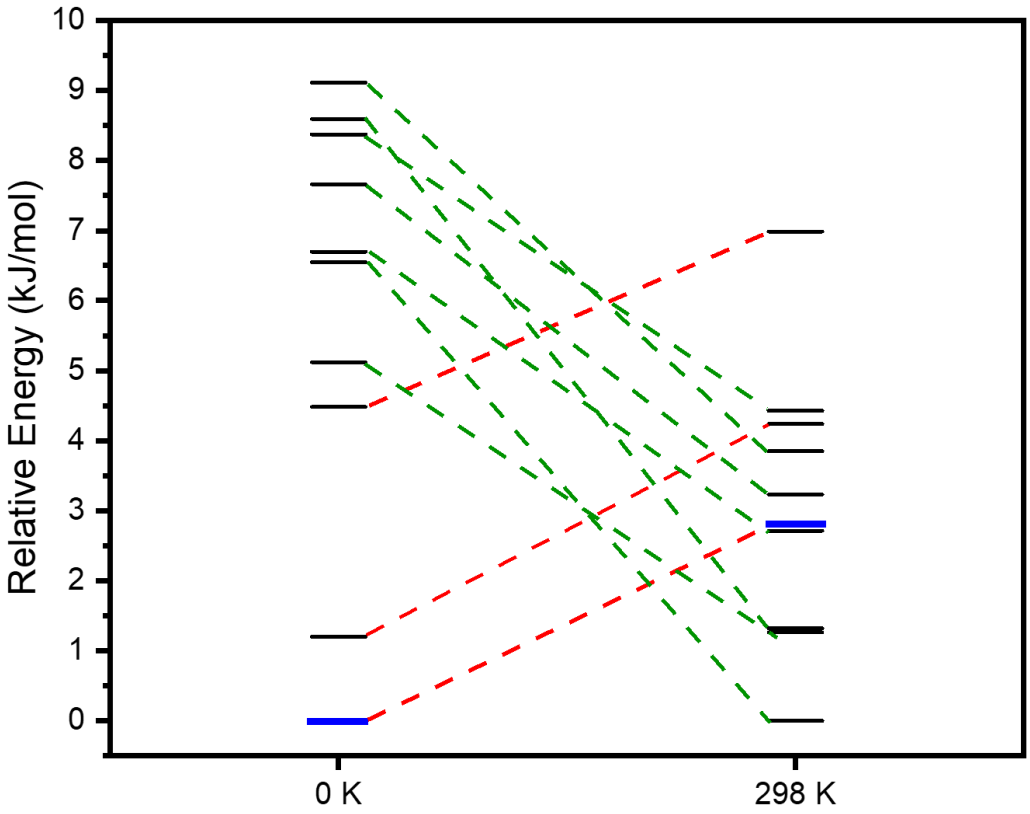

Figure 48 Energy level diagram for Ac-YAAA[KAAAD]X_1-NH_2 containing both Zero Point Corrected Energies (0K) as well as Relative Free Energies (298K). Assigned conformations are highlighted in dark blue all other structures have black bars. Red dashed lines indicate an increase in energy from 0 to 298K whereas a green dashed line indicates the inverse.

6.4.3 Comparison of the Helices

Determination of how much control over the handedness of these molecules could be observed began with a study of the homochiral species (L- vs. D- amino acids). There was great certainty that Ac-Y[KAAAD][KAAAD]X_1-NH_2 would form the two-turn helix as a previous study of the single turn helix form of the same molecule was successful. The simplicity of the infrared spectrum lent to this confidence as the large number of backbone NH stretches would only form the uncomplicated double mountain motif if they were of similar strength and intensity. As observed with Ac-y[kaaad][KAAAD]X_1-NH_2 a breakdown in the helicity of part of the structure has a great affect on the infrared spectrum. Transition pileup in the double mountain motif dissolves and better resolution is obtained due to the change in secondary structure. This lends credence to the observation that the spectral similarity observed in all of the other molecules is a major sign of helix formation. When the oppositely homochiral molecule, Ac-

y[kaaad][kaaad]X_1-NH_2 was studied the similar infrared spectrum was a sign of relief. In synthesizing the molecule, the inability to obtain a D- version of the charged residue could have resulted in deformation of the structure away from the helix. The lack of flexibility in the charge site combined with the dearth of NH groups for hydrogen bonding worked in favor of helix formation even with the change in chirality of the residue. In fact, the L-amino acid version of the charge in this molecule may have been a boon as it allowed for a different hydrogen bonding scheme in the charge region which was observed in the infrared spectrum. The only difference between the homochiral spectra was the single transition near 3200 cm^{-1} in the D-amino acid molecule that was assigned to a strong hydrogen bond between the charge amine and the carbonyl of the third alanine residue in the second ring. Without this singular difference in bonding there would be no way to decipher a difference between the two molecules in the infrared spectrum. This small, unintended difference could theoretically be used to decipher the ratio of one enantiomer to another in a mixture of enantiomers as it belongs solely to the D-amino acid form.

A more difficult test of stereochemical control was the attempt to form mixed helices using heterochiral molecules. A different set of molecules was used to investigate this possibility in the solution phase where NMR spectroscopy was used to visualize the handedness change between the two helical sections[8]. Important to the formation of the mixed helices in solution was the behavior of the section of the molecule that linked the two rings together. In order to form the mixed helix there had to be a transition from one motif to the other. This produced a more open and unwound section of the molecule that allowed for the transition to occur. A major roadblock to observation of the mixed helices in the gas phase was whether or not this transitional region would be allowed to form as it is most likely energetically unfavorable in comparison to a more compact structure in this region. In fact, this was an issue that led to no mixed helix being observed with either heterochiral molecule studied. However, the more interesting result was the large difference between the two molecules.

Ac-Y[KAAAD][kaaad]X_1-NH_2 followed suit spectrally and structurally with the homochiral molecules through the formation of same handed alpha helices in both rings. This was an odd result as it seems the propensity of the stereochemical change to form opposite handed helices was overridden by the stabilization of forming the same helix. The more likely result of the heterochirality of these structures was that seen with Ac-y[kaaad][KAAAD]X_1-NH_2.

Only partial helicity was maintained in this structure, as it unwound forming a more random pattern of hydrogen bonds. As the location and nature of the charge site is thought to seed the formation of the alpha helix in the gas phase it is not surprising that the c-terminal end of the second ring maintains its C13 hydrogen bonds. What is surprising is how the n-terminal portion of the second ring is unable to maintain the pattern throughout the rest of the molecule. The lack of structure in the first ring works to undo the second as the backbone amine groups in the n-terminal portion of the second ring no longer have their needed hydrogen bonding partners for forming the helix. The finding of a bound tether in Ac-y[kaaad][KAAAD]X_1-NH_2 may have a role in why the structure was so perturbed as it interacts with the transition region and breaks up the formation of the needed C13 macrocycle. This hypothesis could be tested by formation and analysis of a molecule that had an untethered first ring but maintained the tether and charge in the second ring. Such a molecule, Ac-YAAA[KAAAD]X_1-NH_2 was synthesized and studied. The increased flexibility on the n-terminal end of Ac-YAAA[KAAAD]X_1-NH_2 was anticipated to be the death knell of any helix formation in the molecule. Instead, alpha helix formation is maintained throughout the body of the structure. This result lies in direct opposition to the tethered heterochiral molecules that struggled to maintain helicity in their n-terminal section. The hypothesis that the hydrogen bond formed by the tether may destabilize the formation of the helix in Ac-y[kaaad][KAAAD]X_1-NH_2 holds credence given the structure seen with Ac-YAAA[KAAAD]X_1-NH_2. The tethered ring also maintained its C13 hydrogen bonds as the n-terminal end of the molecule oriented itself in the helical structure.

6.5 Conclusions

An attempt at stereochemical control of the handedness of tethered helix forming peptides was performed. Tethered peptides formed through the use of oppositely chiral amino acids were synthesized and analyzed via cold IR-UV double resonance spectroscopy in the gas phase. Homochiral versions of the peptides formed stable two turn alpha helices with a change in handedness observed between the two molecules. The same could not be said for the heterochiral molecules which formed vastly different structures to each other, one helical and the other not. Untethering of the first ring produced a molecule that was able to form a helix throughout the entirety of its structure. Insight into the chemical synthesis and control of alpha helix handedness in the gas phase is increased by this study. The subtle differences in structure that lead to very

different conformations between the solution and gas phase will help inform further structural analysis between the two states of matter. Future studies would examine the placement of the tether in both rings and its effect on helix formation as well as modification of the linker section between the rings with the intent to observe mixed helices in the heterochiral molecules.

6.6 References

1. Campos, K. R.; Coleman, P. J.; Alvarez, J. C.; Dreher, S. D.; Garbaccio, R. M.; Terrett, N. K.; Tillyer, R. D.; Truppo, M. D.; Parmee, E. R., The importance of synthetic chemistry in the pharmaceutical industry. *Science* **2019,** *363* (6424), eaat0805.

2. Aliferis, K. A.; Jabaji, S., Metabolomics – A robust bioanalytical approach for the discovery of the modes-of-action of pesticides: A review. *Pesticide Biochemistry and Physiology* **2011,** *100* (2), 105-117.

3. Lutz, J.-F.; Ouchi, M.; Liu, D. R.; Sawamoto, M., Sequence-Controlled Polymers. *Science* **2013,** *341* (6146), 1238149.

4. Bissantz, C.; Kuhn, B.; Stahl, M., A Medicinal Chemist's Guide to Molecular Interactions. *Journal of Medicinal Chemistry* **2010,** *53* (14), 5061-5084.

5. Simeonov, S. P.; Nunes, J. P. M.; Guerra, K.; Kurteva, V. B.; Afonso, C. A. M., Synthesis of Chiral Cyclopentenones. *Chemical Reviews* **2016,** *116* (10), 5744-5893.

6. Kasprzyk-Hordern, B., Pharmacologically active compounds in the environment and their chirality. *Chemical Society Reviews* **2010,** *39* (11), 4466-4503.

7. Fisk, J. D.; Powell, D. R.; Gellman, S. H., Control of Hairpin Formation via Proline Configuration in Parallel β-Sheet Model Systems. *Journal of the American Chemical Society* **2000,** *122* (23), 5443-5447.

8. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Left- and Right-Handed Alpha-Helical Turns in Homo- and Hetero-Chiral Helical Scaffolds. *Journal of the American Chemical Society* **2009**.

9. de Araujo, A. D.; Hoang, H. N.; Kok, W. M.; Diness, F.; Gupta, P.; Hill, T. A.; Driver, R. W.; Price, D. A.; Liras, S.; Fairlie, D. P., Comparative alpha-helicity of cyclic pentapeptides in water. *Angew Chem Int Ed Engl* **2014,** *53* (27), 6965-9.

10. Nicholas E. Shepherd, H. N. H., Giovanni Abbenante, and David P. Fairlie, Single Turn Peptide Alpha Helices with Exceptional Stability in Water. *Journal of the American Chemical Society* **2005**.

11. 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 beta-Hairpin Turn. *J Am Chem Soc* **2017,** *139* (15), 5481-5493.

12. 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.

13. 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.

14. 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.

15. Londry, F. A.; Hager, J. W., Mass selective axial ion ejection from a linear quadrupole ion trap. *Journal of the American Society for Mass Spectrometry* **2003,** *14* (10), 1130-1147.

16. Mohamadi, F.; Richards, N. G. J.; 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. *Journal of Computational Chemistry* **1990,** *11* (4), 440-467.

17. Becke, A. D., Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics* **1993,** *98* (7), 5648-5652.

18. Grimme, S.; Ehrlich, S.; Goerigk, L., Effect of the damping function in dispersion corrected density functional theory. *Journal of Computational Chemistry* **2011,** *32* (7), 1456-1465.

19. Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H., Results obtained with the correlation energy density functionals of becke and Lee, Yang and Parr. *Chemical Physics Letters* **1989,** *157* (3), 200-206.

20. 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 beta-Hairpin. *J Am Chem Soc* **2016,** *138* (8), 2849-57.

VITA

John Thomas Lawler was born on September 3rd, 1986 in Chicago, IL to John Christopher Lawler and Mirka Lawler. He is the older of three children, with a younger brother, Patrick, and sister, Victoria. John attended Wheaton Academy high school, graduating in 2004. After high school John began his undergraduate career at Purdue University majoring in Biology and Philosophy. After becoming frustrated with the lack of career prospects and slow pacing of courses, John left Purdue and went back to day trading stocks and options which became a lucrative but mind-numbing endeavor. In 2014, he decided to give academics another try and enrolled at Northern Illinois University to study biochemistry. It is here that he met and befriended Dr. Michael Lesslie and Dr. Victor Ryzhov who allowed John to work in the lab all he desired. The freedom to create and productivity in the research performed at NIU convinced John that graduate school might be a worthwhile endeavor. Returning to Purdue in June 2016, he began his graduate studies working in the laboratories of Professors Scott McLuckey and Timothy Zwier studying the secondary structure of small peptide ions. During this time John met his wife, Khadija Jawad and had two children, John and Munira. John defended his Ph.D. in July 2021 and is still looking for work.

LIST OF PUBLICATIONS

Conformation-Specific Infrared and Ultraviolet Spectroscopy of Cold [YAPAA+H]$^+$ and [YGPAA+H]$^+$ Ions: A Stereochemical "Twist" on the β-Hairpin Turn
(Journal of the American Chemical Society, 10.1021/jacs.7b01315)

Near-UV Water Splitting by Cu, Ni, and Co Complexes in the Gas Phase
(Journal of Physical Chemistry A, 10.1021/acs.jpca.7b12445)

Infrared Population Transfer Spectroscopy of Cryo-Cooled Ions: Quantitative Tests of the Effects of Collisional Cooling on the Room Temperature Conformer Populations
(Journal of Physical Chemistry A, 10.1021/acs.jpca.7b12751)

Gold (I) Cationization Promotes Ring Opening in Lysine Containing Cyclic Peptides
(Journal of the American Society for Mass Spectrometry, 10.1007/s13361-019-02247-x)

Gas-Phase Sequencing of Cyclotides: Introduction of Selective Ring Opening at Dehydroalanine via Ion/Ion Reaction
(Analytical Chemistry, 10.1021/acs.analchem.9b03671)