8-2016

Single-conformation spectroscopy of hydrogen bonding networks: Solvation, synthetic foldamers, and neurodegenerative diseases

Patrick S. Walsh
Purdue University

Graduate School Form 30
(Updated 11/20/2014)

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Patrick S. Walsh

Entitled SINGLE-CONFORMATION SPECTROSCOPY OF HYDROGEN BONDING NETWORKS: SOLVATION, SYNTHETIC FOLDAMERS, AND NEURODEGENERATIVE DISEASES

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Timothy Zwier

Lyudmila Slipchenko

Scott A. McLuckey

Libai Huang

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification/Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy on Integrity in Research" and the use of copyrighted material.

Timothy Zwier

Approved by Major Professor(s): _____

_____

Approved by: Timothy Zwier 06/30/2016

Head of the Department Graduate Program Date

SINGLE-CONFORMATION SPECTROSCOPY OF HYDROGEN BONDING NETWORKS: SOLVATION, SYNTHETIC FOLDAMERS, AND NEURODEGENERATIVE DISEASES

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Patrick S. Walsh

In Partial Fulfillment of the

Requirements of the Degree

of

Doctor of Philosophy

August 2016

Purdue University

West Lafayette, Indiana

For my family. I love you.

ACKNOWLEDGMENTS

I would like to begin by thanking Prof. Peter Siska for first introducing me to the strange and beautiful world of quantum mechanics in his freshman general chemistry class. Unfortunately, Dr. Siska passed away before I was able to tell him the impact his class had on my career. I would also like to acknowledge my very first scientific adviser Prof. David Pratt. Dr. Pratt took in a curious junior undergraduate student, and guided me towards being a genuine scientist. Thank you for your passion and genuine interest in my success as a researcher. You provided me with an excellent beginning to my scientific career. I would be remiss for not also mentioning the group of students in Dr. Pratt's lab who helped me: Dr. Phil Morgan, Dr. Leo Alvarez, Dr. Jessica Thomas, and Dr. Ryan Bird. Very special thanks to Dr. Justin Young and Dr. AJ Fleisher for putting up with my never ending questions. I am grateful to count you both as friends. Thanks for "cutting the cord."

I would also like to thank both past and current members of the Zwier research group who made this wild and crazy graduate school experience as fun as it has been: Dr. James Redwine, Dr. Vanesa Vacquero, Dr. Ryoji Kusaka (Kusaka-san!), Dr. Nicole Burke, Dr. Nathan Kidwell, Dr. Deepali Mehta-Hurt, Dr. Joe Korn, Dr. Andrew DeBlase, Dr. Chamara Abeysekera, Dr. Brian Hays, Di Zhang, Dan Hewett, Khadija Jawad, John Hopkins, Karl Blodgett, Alicia Hernandez, Chris Harrilal, Sean Fritz, Piyush Mishra; and

adopted group members Dr. Grant Buckingham (the third), Sebastian Bocklitz (the German), and Dr. Jamie Young (ONE OF US!). Thanks very much to you all for being such good sports putting up with all of my shenanigans, sharing many laughs, and more than just occasionally talking some really great science. Thanks to all the departmental staff that helped solve all my problems, there are too many to name, but thank you all!

Special thanks are in order for several members of the research group: Dr. Evan Buchanan, Dr. Jacob Dean, Dr. Zach Davis, and Joe Gord. Evan and Jacob, your impact on my career is more than you know. Thanks for teaching me the instrument, the science, and for all the fun times we had outside of the lab. Zach, you are a great friend; thanks for always being up for a road trip! Gordo, thanks for going through this process with me. We might bicker like an old married couple, but you made graduate school a great experience.

I would also like to thank Dr. Vas Stavros and his group at Warwick. Thanks for showing this colonist such a great time in the UK. Vielen Dank to my adopted group in Bochum. Christian, you are a tremendous thinker and a great mentor. I am grateful to call you a friend and colleague. Thomas and Sebastian, thanks for showing me a great time in Germany and making me feel at home while I was there (Curley-Wurst is the best).

Tim, I am not sure I can say anything that has not already been said by countless graduate students before me. It has been both an honor and a pleasure to earn a doctoral degree under your tutelage. You exemplify so many outstanding qualities; you are a kind, passionate, genuinely caring person. Your expertise in all things spectroscopy is awe inspiring. You allowed me to grow as a scientist and achieve my full potential, always providing the right amount of encouragement and support. In no uncertain terms I would

not be where I am today without your guidance and support. You are a scholar, a gentleman, and most importantly a friend. For all of this, I can only say thank you!

Last, but certainly not least, I would like to thank my family. To all of my extended family who constantly supported me; thanks for the love and the care packages and the little notes of support. To my brother Michael, thanks for always being there to talk and urge me forward, and for always being there to give me a good laugh. Your friendship and support is unyielding. To my mom and dad, words cannot express how much I owe you both. From my earliest memories you both taught me to put my best foot forward and never give up. Thank you for nurturing my curious nature and pushing me to be a better person. The ways in which you have supported me during this journey are beyond measure. Thanks for it all, without you I am not here today. I love you more than you know.

TABLE OF CONTENTS

Page

Page

Page

Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

Figure Page

ABSTRACT

Walsh, Patrick S. Ph.D., Purdue University, August 2016. Single-Conformation Spectroscopy of Hydrogen Bonding Networks: Solvation, Synthetic Foldamers, and Neurodegenerative Diseases. Major Professor: Timothy S. Zwier.

The hydrogen bond is one of the most important interactions in natural processes ranging from protein folding to chemical reactions. Two complementary methodologies are applied to understanding this important interaction: top-down and bottom-up. Top-down methods use large molecules, such as proteins, revealing secondary structure information. Bottom-up experiments are performed on small molecules, utilizing high-resolution spectroscopy to reveal underlying quantum mechanical effects. The complexity gap is formed between these two experimental regimes; between large and small molecules; between bulk and individual solvent molecules; between classical mechanics calculations and quantum chemical calculations. This dissertation will focus on the application of gas phase, single-conformation ultraviolet (UV) and infrared (IR) spectroscopies to the study of molecules and clusters in the size and solvation complexity gap, with the goal of bridging the gulf between the two experimental approaches.

Single-conformation spectroscopy is perfectly suited to study solvation. Solvent molecules, in many instances water, can be frozen onto the solute in a stepwise manner. Here, we solvate a prototypical flexible bichromophore—1,2-diphenoxyethane

(DPOE)—by stepwise addition of water molecules. Single-conformation spectroscopy reveals both the structural perturbations associated with water molecule addition and their effect on DPOE's pair of closely spaced excited electronic states. These experimental studies provide excellent insight into solvent effects on vibronic and excitonic coupling, and can be used to further develop the models used to describe such processes.

Similarly, single-conformation spectroscopy can reveal the effects of conformational flexibility on the innate conformational preferences and hydrogen bonding motifs in peptides. Results obtained from a study of a cyclically constrained γ peptide, γ_{ACHC}, reveal that increased conformational flexibility can be controlled by synthetic chemists in order to direct folding into pre-programmed secondary structures and that these structures are stabilized with intramolecular rather than intermolecular hydrogen bonds. This theme of conformational flexibility is continued in studies of glutamine containing peptides. Glutamine—with its flexible, hydrogen bond forming sidechain—is intimately involved with neurodegenerative diseases such as Huntington's disease. Single-conformation studies help reveal the delicate interplay between three different types of hydrogen bonds within the molecule: backbone-backbone, sidechain-backbone, and sidechain-sidechain hydrogen bonds. The importance of these competing hydrogen bonds on the conformational preferences will be discussed both locally and within the larger context of disease pathogenesis.

CHAPTER 1. INTRODUCTION

1.1 Motivation: Hydrogen Bonding Networks

'Structure dictates function' is the mantra of various fields of research ranging from chemical physics to chemical biology. In much the same way that it is possible to know the function of a key in a lock by understanding the shape of the key, it is possible to discern the function of a molecule by understanding its structure. Realizing that structure is so vital, it is critical to recognize the interactions that stabilize structural motifs—chief among them the hydrogen bond.

The pervasive nature of the hydrogen bond implicates its importance in most, if not all, biological and chemical processes. As such, there are several active fields of research aimed at discovering the nature of hydrogen bond networks in both protein and aqueous environments. Historically, X-ray crystallographic and NMR methodologies have been used to elucidate the structure of proteins and peptide oligomers. These methods are still the gold standard for structural determination of large molecules. Recently, pioneering work by Eisenberg and coworkers have used X-ray crystallography to reveal, with atomistic resolution, the structure of several peptide oligomers intimately related to neurodegenerative diseases.[1,2] Tycko *et al.* have used solid state NMR to characterize the structure of Alzheimer's disease Amyloid-β fibrils.[3,4] These methods excel at revealing the atomistic detail of three-dimensional structure, but provide no

spectroscopic signatures for these interactions and structures; this deficiency looms large with the rise of new optical spectroscopy methods to interrogate similar molecules.

These go to methods for structure identification also lack, in most instances, any temporal resolution. Biological and chemical processes are dynamic; viewing these processes with temporal resolution will undoubtedly provide deeper comprehension of such processes. These traditional approaches might not provide the best view into a topic such as protein folding and misfolding, where the structure is important, but so is knowing how it reached that structure

Linear optical spectroscopy in condensed phase environments, although timeless, is replete with pitfalls making it less and less useful for analyzing macromolecular structure and hydrogen bonding. In the infrared (IR), interesting spectral features are often obscured by the absorptions arising from the bulk solution—this should be clear to any person who has taken an IR spectrum of a molecule that is contaminated with water while in undergraduate organic class. Some new and exciting techniques have been developed to ameliorate this problem, but their application and scope are still somewhat limited.[5,6] Most solution phase spectroscopy measurements—in both the infrared and ultraviolet regions—are also greatly affected by spectral broadening and increased spectral complexity. The increased spectral complexity comes from the plethora of thermally populated quantum states, while the broadening arises from collisions and other interactions between the solute and bulk solvent or other solute molecules. These factors result in spectral lines that are orders of magnitude broader than their natural lifetime linewidths, hindering the identification of interesting spectral features, and their origin.

With the advent of ultrafast laser sources, multidimensional optical spectroscopy methods have been developed to study both the structure and dynamics of large molecules in solution. These multidimensional spectroscopy methods should be thought of as optical analogues of their radiofrequency counterparts widely used in NMR. In these methods the "on-diagonal" spectrum is roughly analogous to the linear absorption spectrum, while the "off-diagonal" features arise from the coupling between two sets of transitions (e.g. nuclear spins in NMR and vibrational modes in IR). Two- and three-dimensional IR spectroscopy have been developed to overcome some of the challenges inherent to linear solution phase spectroscopy, to study the structure and solvation of macromolecules. These multidimensional methods provide more spectral features which can be used to reveal the structure and dynamics. 2D-IR has been used to elucidate the structure of several important protein aggregates related to Huntinton's disease and Alzheimer's disease, while 3D-IR has been utilized to explore the structure and dynamics of solvent molecules around a solute.[7-10] 2D-electronic spectroscopy has been used to better understand the fate of electronic excitations in complex multichromophore arrays, such as those found in photosynthetic light harvesting proteins. These 2D-electronic spectroscopy experiments have pointed towards coherent energy transfer mechanism between chromophores, opening the door on a new field of quantum biology.[11,12]

The aforementioned methods are generally considered "top-down" methods of research because they tackle research problems using large macromolecular systems, focusing on large scale structure rather than an atomistic comprehension of the hydrogen bonding networks that stabilize the structure. These top-down approaches face numerous challenges. The sizes of the systems under investigation are typically so large that

comprehensive, first-principles quantum mechanical calculations are impossible. As such, theoretical models based on frequency and coupling maps for small molecules are developed to interpret the results obtained from such experiments. These models would benefit from rigorous experimental tests using complementary methods.[13-19]

The opposite approach is termed "bottom-up" and focuses on small systems that can be rigorously tested with high-level quantum chemical calculations. This approach is often used by chemical physicists to expose the underlying chemical and quantum phenomena. Many of the experimental difficulties present in solution phase experiments are avoided altogether by performing these experiments on gas phase samples. Many of the complexities introduced by collisional broadening and the high number of populated quantum states simply vanish when experiments are performed in the cold, collision free environment of the molecular beam. This allows for spectral lines to be recorded that are much closer to their natural linewidth. Although experimental sources of broadening will still exist (e.g. laser linewidth, Doppler broadening, etc.), these perturbations are generally small and still allow for the underlying quantum mechanical phenomena to be observed without much trouble (e.g. Fermi-resonances, Herzberg-Teller coupling, the Jahn-Teller effect, Renner-Teller coupling, etc.). With careful experiment design, many such problems can be greatly reduced, or accounted for, providing exceptional experimental resolution.[20-23] Such bottom-up studies are routinely used to explore chemical reactions and isomerization in, and between, small molecules.[24-26] The biggest drawback of these bottom-up approaches is extrapolating results on small molecules to systems as large and complex as a solvated protein. In fact many times such an extrapolation is impossible.

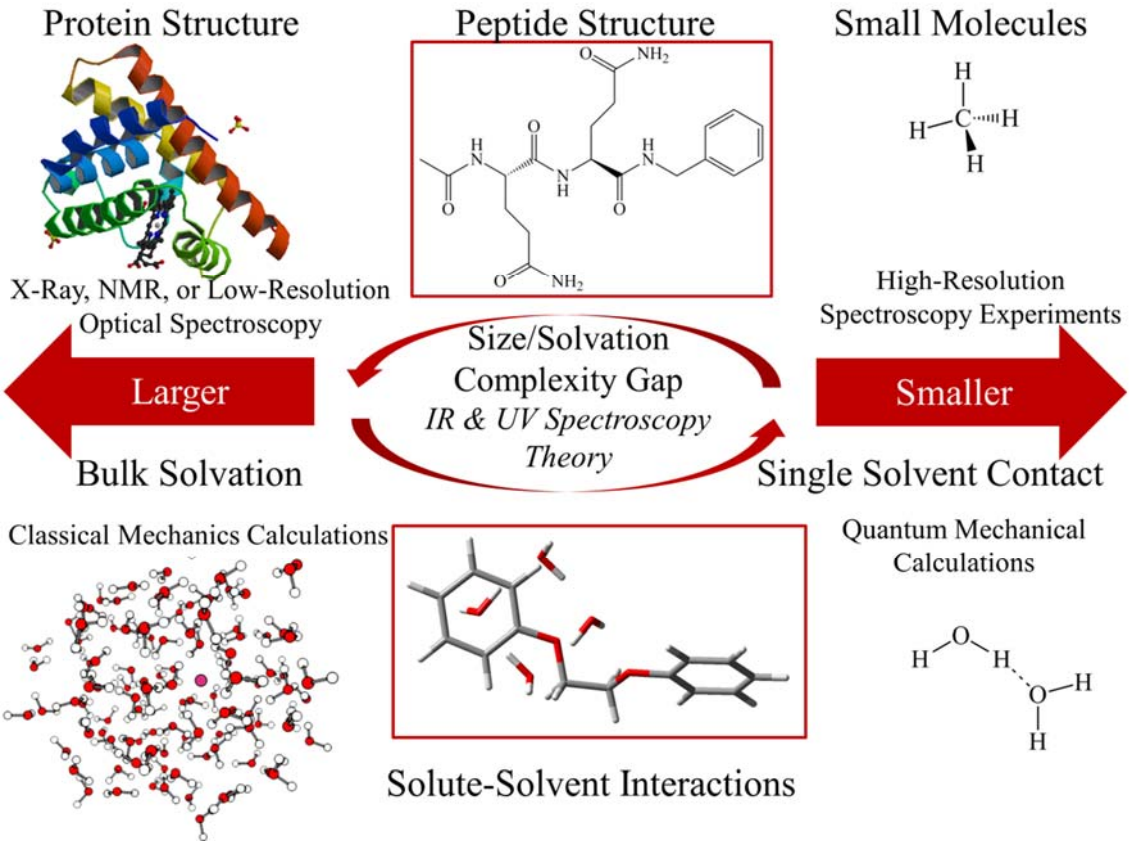


Figure 1.1. Visual demonstration of both the size, and solvation complexity gaps. Single-conformation spectroscopy provides an excellent method for bridging these gaps, providing data to test and extend theoretical modeling of both solvation and protein environments. Protein structure is Myoglobin published in ref. [27].

Between these two experimental approaches there exists a "Complexity Gap" in both size and degree of solvation (Figure 1.1).[28-30] The focus of this dissertation will be on studying molecules and molecular clusters that reside in the complexity gap with single-conformation UV and IR spectroscopy, providing a link between the two structural extremes. The molecules studied here push the limits of cutting edge quantum calculations to larger and more complex systems. Molecules in this size regime can also

be studied with high enough spectral resolution, and are of a size large enough to adopt natural secondary structures, that then can provide incisive tests for the models used to comprehend results obtained in top-down measurements.[13-19] Single-conformation studies provide both atomistic level detail and unique spectral signatures of hydrogen bonding networks, results which help further the understanding of such networks in large proteins. The following sections will lay out the important role single-conformation spectroscopy has in bridging the complexity gap.

1.2 Organization

1.2.1 Experimental Approaches to Spectroscopy in the Complexity Gap

Ever since the pioneering studies by Levy and co-workers on the amino acid tryptophan, laser spectroscopy of important biologically relevant molecules entrained in a molecular beam has become a mainstay experiment for physical chemists.[31-33] Since these early days several improvements have occurred that allow for single-conformation spectra to be recorded in both the IR and UV in a mass-resolved manner. Chapter 2 will focus on the experimental techniques used throughout this dissertation. The heart of every experiment is the supersonic expansion which allows us to prepare the gas phase samples in a cold, collision free environment. A newly incorporated laser desorption source has opened the door to new and challenging projects that were previously impossible due to the limitations inherent in thermal vaporization. The time-of-flight mass spectrometer and hardware required for acquisition of the signals generated by our laser spectroscopy methods will be detailed in this chapter. These laser methods include

resonant two-photon ionization (R2PI), single-conformation UV spectroscopy (IR-UV holeburning and UV-UV holeburning), and resonant ion-dip infrared (RIDIR) spectroscopy for recording single conformation IR spectra. The computational methods used to generate the assigned conformations will also be discussed.

1.2.2 Understanding Solvation Effects in the Complexity Gap

As water is nature's solvent, understanding its effects on a solute are critically important to comprehending processes from chemical reactions to protein folding. The process of solvating a molecule is extremely complex, so gaining knowledge of the hydrogen bonding networks at atomistic level detail is extremely difficult. The notion of exploring solvation by step-wise addition of individual water molecules to the solute molecule, allowing for the detailed investigation of the structural changes and other perturbations caused by each added water molecule, is heuristically pleasing. Such heuristic experiments are perfectly suited for the mass-resolved, conformation-specific spectroscopy methods used in this dissertation. Solvent molecules can be "frozen" onto the solute molecule of interest and studied one at a time.[28,29,33,34]

Chapters 3 and 4 will explore the aqueous solvation of a prototypical bichromophore 1,2-diphenoxyethane (DPOE). These chapters expand on initial studies performed in our group on the structure and excited electronic state behavior of DPOE monomer.[35-37] Chapter 3 focuses primarily on the ground electronic state conformations of the DPOE-$(H_2O)_n$ (n=2-4) clusters.[38] Using a combination of both single-conformation IR spectroscopy and quantum chemical calculations, we will discuss the structural perturbations associated with stepwise solvation of an aromatic rich solute. In this study

we focus on both the solvent perturbations induced on the solute, and the perturbations of the solute on the solvent structure. The most striking result found in this work is that with addition of only three water molecules a conformational isomerization occurs in the solute, causing a structure to be adopted that is not observed in the absence of water. This structure also allows water to adopt a trimer chain geometry—a geometry which is not found in a free water trimer.

Beyond structural influences, solvent interactions can, and do, have profound impacts on the solute's excited states. With the discovery of quantum coherences in photosynthetic multichromophore arrays, a need has been generated for understanding the behavior of close-lying and strongly coupled electronic states, and the role vibrational modes and solvation have on the coupling of these excited states.[12,39] The gas phase spectroscopy community has been extremely active studying the quantum mechanical manifestations of such coupled electronic states through the study of excitonic splitting in so called flexible bichromophores.[36,40-42] The results of such studies are being used to test and develop cutting edge theories of excitonic and vibronic coupling.[40,43-47]

DPOE is unique in that its first two excited electronic states are weakly coupled and are split by ~1 cm^{-1}.[36] This splitting is greatly increased in DPOE-H_2O where water binds asymmetrically to the two chromophores, thereby causing the excited states to be split by ~190 cm^{-1}.[37] In Chapter 4 we focus on the effects of increasing solvation on the excited state behavior of the DPOE solute in DPOE-$(H_2O)_n$ (n=2-4).[48] Through novel application of excited state RIDIR spectroscopy, we were able to use the water molecules' excited state IR spectra to report on the electronic excitations. These excited

state IR spectra allowed us to determine which chromophore is responsible for which excitation, and the degree of excited state mixing as solvation is increased.

1.2.3 Conformational Flexibility and Peptide Structure in the Complexity Gap

Understanding the effects of conformational flexibility and hydrogen bonding in peptides is critically important for developing a more complete comprehension of the way proteins fold into their desired structure. Gas phase single-conformation spectroscopy is a powerful tool for studying the inherent conformational preferences of peptides in the complexity gap. Recent advances in laser desorption vaporization sources and electrospray instruments has increased the size of molecules which can be introduced into the gas phase, pushing us further into the complexity gap.[28,29,49-57] Studies of this sort are well positioned to push the limits of existing quantum mechanical and density functional methods, and provide an incisive test of force fields and spectroscopic models being developed to study large peptides and proteins.[13-19]

Conformational flexibility is an important concept in naturally occurring amino acids and peptides, but it is leveraged in an especially fruitful way by the synthetic foldamer community. By carefully controlling the conformational flexibility of the foldamer subunits (e.g., unnatural amino acids) it is possible to promote folding propensities that lead to new secondary structures that either mimic or build on nature's tool set.[58] Numerous studies by our group have sought to understand the influences of increased conformational flexibility on the inherent conformational preferences and hydrogen bonding patterns of such synthetic foldamers.[55,56,59]

Chapter 5 explores the effects of controlling the conformational flexibility through constraining groups which direct the folding of γ-peptides into pre-programed conformations.[60] In this work an otherwise extremely flexible γ-peptide (which has two additional carbon atom spacers in the amide backbone) is cyclically constrained to induce a well-defined bend into the peptide backbone.[61,62] These foldamer oligomers are designed to form a 2.6_{14}-helix—a helix with 2.6 residues per turn and held together with 14-membered hydrogen bonded rings. Using conformation-specific IR and UV spectroscopy we studied the conformational preferences of a capped amino acid and dipeptide containing the γ-amino acid γ_{ACHC}. These studies reveal that the 2.6_{14}-helix was formed at the earliest possible stage in the peptide.

Conformational flexibility and its influence on the hydrogen bonding network of a peptide is also the major theme in the work discussed in Chapters 6 and 7. These chapters focus on the amino acid glutamine which features a long flexible sidechain terminating in an amide group. Glutamine is an important amino acid as it relates to neurodegenerative diseases, specifically CAG codon repeat diseases, of which Huntington's disease (HD) is the most widely known.[63] Knowing the structure and hydrogen bonding propensities of glutamine is critically important to gaining deeper knowledge of disease pathogenesis and potential therapeutic approaches. Once again single-conformation spectroscopy is uniquely suited to explore these questions.

In Chapter 6 a series of glutamine containing capped amino acids and a capped dipeptide are studied using conformation specific spectroscopy in the UV and IR.[64] Experimental spectra were compared to density functional theory calculations and conformational isomers were assigned, revealing the unique hydrogen bonding

interactions that are available to glutamine with its flexible hydrogen bonding sidechain. Detailed insight into the interplay between backbone-to-backbone and sidechain-to-backbone hydrogen bonds was gained. These results highlighted glutamine's ability to form strong hydrogen bonds which "lock-in" particular secondary structures including an inverse γ-turn and a type-I β-turn.

Chapter 7 represents the first experimental exploration of a model "polyQ" segment using conformation-specific spectroscopy. In the work, Ac-Gln-Gln-NHBn is studied in the cold, isolated environment of the molecular beam with R2PI and RIDIR spectroscopy. In this case three competing hydrogen bonding types are available, enabling the study of the delicate interplay between them. This marks the first instance where sidechain-to-sidechain H-bonds can be formed in the model molecule, along with the aforementioned backbone-to-backbone and sidechain-to-backbone hydrogen bonds. Sidechain-sidechain hydrogen bonds have been long implicated in holding the β-sheets together in the fibril structures so prevalent in neurodegenerative diseases.[1,2] The results of this study are striking, revealing that Ac-Gln-Gln-NHBn prefers to fold into a type-I β-turn that is locked in and stabilized by two successive sidechain-to-backbone hydrogen bonds. This result is all the more important when it is placed in context: the toxic species in HD is a compact β-turn/β-haripin structure that evolves from the misfolding of a peptide in a random coil.[7,65-67] Results from this study demonstrate that glutamine containing compounds can readily fold into the needed turn geometry, which is stabilized in part by hydrogen bonds formed by the flexible sidechain. The unique nature of these single-conformation experiments provides once again both atomistic insight and unique

spectral signatures, of hydrogen bonding networks. These spectral signatures will be crucial for identifying such a structure in larger segments of polyQ.

1.3 References

(1) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773-778.

(2) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. Ø.; Riekel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453-457.

(3) Petkova, A. T.; Yau, W.-M.; Tycko, R. *Biochemistry* **2006**, *45*, 498-512.

(4) Tycko, R. *Annu. Rev. Phys. Chem.* **2011**, *62*, 279-299.

(5) Perera, P.; Wyche, M.; Loethen, Y.; Ben-Amotz, D. *J. Am. Chem. Soc.* **2008**, *130*, 4576-4577.

(6) Gierszal, K. P.; Davis, J. G.; Hands, M. D.; Wilcox, D. S.; Slipchenko, L. V.; Ben-Amotz, D. *J. Phys. Chem. Lett.* **2011**, *2*, 2930-2933.

(7) Buchanan, L. E.; Carr, J. K.; Fluitt, A. M.; Hoganson, A. J.; Moran, S. D.; de Pablo, J. J.; Skinner, J. L.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5796-5801.

(8) Shim, S.-H.; Gupta, R.; Ling, Y. L.; Strasfeld, D. B.; Raleigh, D. P.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6614-6619.

(9) Rao, F.; Garrett-Roe, S.; Hamm, P. *J. Phys. Chem. B* **2010**, *114*, 15598-15604.

(10) Dutta, S.; Ren, Z.; Brinzer, T.; Garrett-Roe, S. *Phys. Chem. Chem. Phys.* **2015**, *17*, 26575-26579.

(11) Engel, G. S.; Calhoun, T. R.; Read, E. L.; Ahn, T. K.; Mancal, T.; Cheng, Y. C.; Blankenship, R. E.; Fleming, G. R. *Nature* **2007**, *446*, 782-786.

(12) Fassioli, F.; Olaya-Castro, A.; Scholes, G. D. *J. Phys. Chem. Lett.* **2012**, *3*, 3136-3142.

(13) Buchanan, E. G.; James, W. H., III; Choi, S. H.; Guo, L.; Gellman, S. H.; Muller, C. W.; Zwier, T. S. *J. Chem. Phys.* **2012**, *137*, 094301-094316.

(14) Ham, S.; Cha, S.; Choi, J. H.; Cho, M. *J. Chem. Phys.* **2003**, *119*, 1451-1461.

(15) Ham, S.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 6915-6922.

(16) Hayashi, T.; Zhuang, W.; Mukamel, S. *J. Phys. Chem. A* **2005**, *109*, 9747-9759.

(17) Jansen, T. L.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. *J. Chem. Phys.* **2006**, *125*, 044312.

(18) Torii, H.; Tasumi, M. *J. Chem. Phys.* **1992**, *96*, 3379-3387.

(19) Wang, L.; Middleton, C. T.; Zanni, M. T.; Skinner, J. L. *J. Phys. Chem. B* **2011**, *115*, 3713-3724.

(20) Fleisher, A. J.; Bird, R. G.; Zaleski, D. P.; Pate, B. H.; Pratt, D. W. *J. Phys. Chem. B* **2013**, *117*, 4231-4240.

(21) Fleisher, A. J.; Young, J. W.; Pratt, D. W. *Phys. Chem. Chem. Phys.* **2012**, *14*, 8990-8998.

(22) Daussy, C.; Marrel, T.; Amy-Klein, A.; Nguyen, C. T.; Bordé, C. J.; Chardonnet, C. *Phys. Rev. Lett.* **1999**, *83*, 1554-1557.

(23) Douberly, G. E.; Raston, P. L.; Tao, L.; Marshall, M. D. *J. Chem. Phys.* **2015**, *142*, 1-9.

(24) Abeysekera, C.; Joalland, B.; Ariyasingha, N.; Zack, L. N.; Sims, I. R.; Field, R. W.; Suits, A. G. *J. Phys. Chem. Lett.* **2015**, *6*, 1599-1604.

(25) Baraban, J. H.; Changala, P. B.; Mellau, G. C.; Stanton, J. F.; Merer, A. J.; Field, R. W. *Science* **2015**, *350*, 1338-1342.

(26) Changala, P. B.; Baraban, J. H.; Merer, A. J.; Field, R. W. *J. Chem. Phys.* **2015**, *143*, 084310.

(27) Barends, T. R. M.; Foucar, L.; Ardevol, A.; Nass, K.; Aquila, A.; Botha, S.; Doak, R. B.; Falahati, K.; Hartmann, E.; Hilpert, M.; Heinz, M.; Hoffmann, M. C.; Köfinger, J.; Koglin, J. E.; Kovacsova, G.; Liang, M.; Milathianaki, D.; Lemke, H. T.; Reinstein, J.; Roome, C. M.; Shoeman, R. L.; Williams, G. J.; Burghardt, I.; Hummer, G.; Boutet, S.; Schlichting, I. *Science* **2015**, *350*, 445-450.

(28) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205-241.

(29) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827-8839.

(30) Zwier, T. S. *J. Phys. Chem. A* **2006**, *110*, 4133-4150.

(31) Rizzo, T. R.; Park, Y. D.; Levy, D. H. *J. Chem. Phys.* **1986**, *85*, 6945-6951.

(32) Rizzo, T. R.; Park, Y. D.; Peteanu, L. A.; Levy, D. H. *J. Chem. Phys.* **1986**, *84*, 2534-2541.

(33) Gruenloh, C. J. Thesis, Purdue University, 1999.

(34) Buchanan, E. G. Thesis, Purdue University, 2014.

(35) Buchanan, E. G.; Sibert, E. L.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 2800-2811.

(36) Buchanan, E. G.; Walsh, P. S.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2013**, *138*, 204313-204311.

(37) Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Phys. Chem. Lett.* **2013**, *4*, 1644-1648.

(38) Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015**, *142*, 154303.

(39) Hossein-Nejad, H.; Olaya-Castro, A.; Scholes, G. D. *J. Chem. Phys.* **2012**, *136*, 024112.

(40) Pillsbury, N. R.; Kidwell, N. M.; Nebgen, B.; Slipchenko, L. V.; Douglass, K. O.; Cable, J. R.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2014**, *141*, 064316.

(41) Pillsbury, N. R.; Muller, C. W.; Meerts, W. L.; Plusquellic, D. F.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *113*, 5000-5012.

(42) Pillsbury, N. R.; Stearns, J. A.; Muller, C. W.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2008**, *129*, 114301.

(43) Fulton, R. L.; Gouterman, M. *J. Chem. Phys.* **1961**, *35*, 1059-&.

(44) Fulton, R. L.; Gouterman, M. *J. Chem. Phys.* **1964**, *41*, 2280-&.

(45) Ottiger, P.; Leutwyler, S. *Chimia* **2011**, *65*, 228-230.

(46) Ottiger, P.; Leutwyler, S. *J. Chem. Phys.* **2012**, *137*, 204303.

(47) Kopec, S.; Ottiger, P.; Leutwyler, S.; Köppel, H. *J. Chem. Phys.* **2012**, *137*, 184312.

(48) Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015**, *142*, 154304.

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

(50) Svendsen, A.; Lorenz, U. J.; Boyarkin, O. V.; Rizzo, T. R. *Rev. Sci. Instrum.* **2010**, *81*.

(51) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *J. Am. Chem. Soc.* **2007**, *129*, 13820-13821.

(52) Alauddin, M.; Gloaguen, E.; Brenner, V.; Tardivel, B.; Mons, M.; Zehnacker-Rentien, A.; Declerck, V.; Aitken, D. J. *Chem.--Euro. J.* **2015**, *21*, 16479-16493.

(53) Chin, W.; Compagnon, I.; Dognon, J. P.; Canuel, C.; Piuzzi, F.; Dimicoli, I.; von Helden, G.; Meijer, G.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 1388-1389.

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

(55) James, W. H., III; Buchanan, E. G.; Müller, C. W.; Dean, J. C.; Kosenkov, D.; Slipchenko, L. V.; Guo, L.; Reidenbach, A. G.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2011**, *115*, 13783-13798.

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

(57) Kusaka, R.; Zhang, D.; Walsh, P. S.; Gord, J. R.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2013**.

(58) Gellman, S. H. *Acc. Chem Res.* **1998**, *31*, 173-180.

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

(60) Walsh, P. S.; Kusaka, R.; Buchanan, E. G.; James, W. H.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 12350-12362.

(61) Guo, L.; Chi, Y.; Almeida, A. M.; Guzei, I. A.; Parker, B. K.; Gellman, S. H. *J. Am. Chem. Soc.* **2009**, *131*, 16018-16020.

(62) Guo, L.; Zhang, W.; Reidenbach, A. G.; Giuliano, M. W.; Guzei, I. A.; Spencer, L. C.; Gellman, S. H. *Angew. Chem., Int. Ed.* **2011**, *50*, 5843-5846.

(63) Mangiarini, L.; Sathasivam, K.; Seller, M.; Cozens, B.; Harper, A.; Hetherington, C.; Lawton, M.; Trottier, Y.; Lehrach, H.; Davies, S. W.; Bates, G. P. *Cell* **1996**, *87*, 493-506.

(64) Walsh, P. S.; Dean, J. C.; McBurney, C.; Kang, H.; Gellman, S. H.; Zwier, T. S. *Phys. Chem. Chem. Phys.* **2016**, *18*, 11306-11322.

(65) Kar, K.; Hoop, C. L.; Drombosky, K. W.; Baker, M. A.; Kodali, R.; Arduini, I.; van der Wel, P. C. A.; Horne, W. S.; Wetzel, R. *Journal of Molecular Biology* **2013**, *425*, 1183-1197.

(66) Kar, K.; Jayaraman, M.; Sahoo, B.; Kodali, R.; Wetzel, R. *Nature Structural & Molecular Biology* **2011**, *18*, 328-336.

(67) Thakur, A. K.; Wetzel, R. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 17014-17019.

CHAPTER 2. EXPERIMENTAL METHODS

2.1 Introduction

The work described in this dissertation uses a combination of both experimental and theoretical tools to study the hydrogen bonding motifs and inherent conformational preferences of a variety of molecules and molecular clusters. Many of these techniques and methodologies have been explained in great detail in the literature.[1-4] This chapter will serve to explain the general techniques and methodologies that are used to record and interpret single-conformation spectra. Details specific to a particular project will be presented in the chapter outlining that specific project.

2.2 Molecular Beams and Supersonic Expansions

The most desirable medium to make spectroscopic measurements has two criteria: the molecules of interest must exist in a well-defined set of quantum states, and have no interactions/collisions with the rest of the molecules in the system. When molecules occupy a large number of quantum states, the density of spectroscopic transitions will increase correspondingly. This density often times makes assigning and understanding the spectra virtually impossible. Similarly, strong interactions with surrounding molecules will additionally perturb and broaden the transitions of interest, making the spectral data difficult to interpret and assign.

Overcoming both of these challenges simultaneously is exceptionally difficult. Traditional spectroscopy performed in the solution-phase suffers from large contributions from intermolecular interactions, while a room temperature gas cell at thermal equilibrium will likely have a huge number of states populated.

Supersonic expansions and molecular beams overcome both of these challenges by cooling the temperature of the molecules and placing them in a collision free environment; they are at the heart of every experiment described in this dissertation.[5,6] A supersonic expansion is formed when high pressure gas (typically He, Ne, or Ar in our experiments) is expanded through an orifice into a low pressure (vacuum) environment. This process is shown in Figure 2.1.

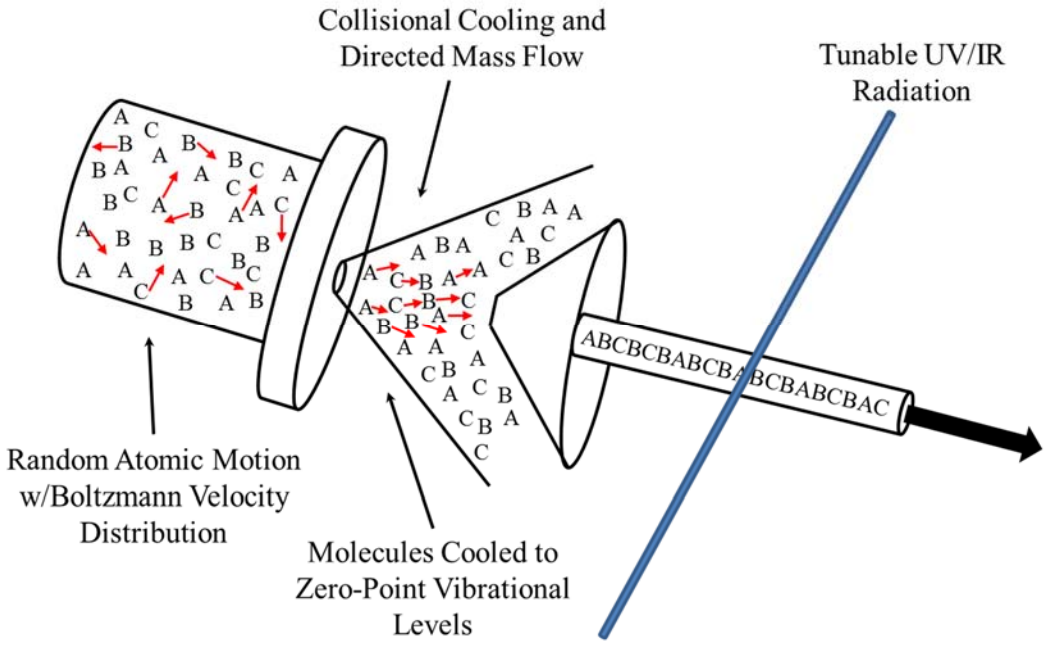


Figure 2.1. Schematic diagram of the supersonic expansion.

This expansion process takes random atomic motion, with a Boltzmann velocity distribution governed by the temperature of the pre-expansion nozzle, and converts it into directed mass flow down along the expansion axis. During this process the velocity distribution narrows considerably; the width of the velocity distribution is related to the translational temperature of the supersonic expansion. Translational temperatures of monoatomic gasses can reach temperatures below 1 K.[5]

When a small fraction of a molecule of interest is seeded into this expansion, it is possible to cool those molecules down via collisions with the cold backing gas atoms. The molecules vibrational and rotational degrees of freedom begin to equilibrate with the temperature of the expansion via collisions with the cold carrier gas. The rotation and vibration cooling rates are not the same, so the degree of cooling for each of these degrees of freedom is different (T_{rot} ~2-5 K, T_{vib} ~10-15 K).[5] In most cases the molecules are cooled down such that the zero-point vibrational levels of the lowest energy minima on the potential energy surface are the only vibrational quantum states with any population.[5,6]

The molecules would reach full equilibrium with the backing gas if it were not for one important fact: as the expansion expands, the density of the gas molecules drops and with it the collision rate. Although this limits the ultimate cooling capacity of the supersonic expansion, it does nicely provide a solution to the second problem facing most spectroscopic media: a collision free environment. Because there are no collisions downstream in the expansion, the cold molecules cannot condense and remain in the gas phase, even at temperatures where they would normally exist as a solid.[5,6]

The supersonic expansion fulfills both requirements for an ideal spectroscopic medium. The molecules are extremely cold, ensuring that only the lowest possible quantum states of the low-energy conformers are populated. This greatly simplifies the spectroscopy, making it possible to identify individual transitions and allowing for their assignment. The collision free environment strips away all intermolecular interactions that cause inhomogeneous broadening in solution-phase experiments, allowing for a detailed understanding of the molecule's unperturbed spectra. In practice this means that a UV spectrum which might be tens of nanometers wide in solution, has resolved vibronic transitions that are ~1-2 cm^{-1} (FWHM) wide in the supersonic expansion.

2.3 Sample Handing

2.3.1 Thermal Heating

The simplest method to introduce molecules into the gas phase is to heat them until sufficient vapor pressure is achieved so that it can be interrogated using various laser spectroscopy techniques. For many samples this is trivial. Solid or liquid sample is placed in glass wool and inserted into a glass holder. This glass holder is placed into a stainless steel sample cell (Swagelok, inline filter with filter removed). The glass holder is used to limit contact between the sample and the heated stainless steel surfaces.

The stainless steel sample cell is placed inline (behind) with a pulsed general valve (Parker-Hannifin Series 9, 500 or 800 μm orifice) which operates at 20 Hz. Typically a pure Neon or Helium backing gas is used at a pressure of ~2-4 bar. The

entirety of the valve assembly (valve face, general valve solenoid, and sample container) is then wrapped in a heating rope (OMEGALUX rope heater, Omega Engineering) which is resistively heated. The temperature is controlled with a Variac and monitored with a K-type thermocouple (Omega). By heating the entire valve assembly we limit the possibility that the molecules will condense before they exit the nozzle.

Using this approach we can routinely heat samples to ~250 °C. When heating to these high temperatures it is advisable to use a "high-temp" General valve solenoid. These "high temp" solenoids were made in-house by re-wiring a broken standard series 9 solenoid with insulation and wiring that is rated to such high temperatures. Generally these modified general valves achieve these high temperatures with a much lower failure rate compared to the unmodified general valves.

2.3.2 Laser Desorption

Many interesting peptides have negligible vapor pressures at the temperatures which can easily be achieved in our experiments (see Section 2.3.1). As such, these samples tend to thermally degrade before sufficient vapor pressure is attained to perform the experiment. Avoiding this degradation problem unlocks a wide variety of molecules for study. Pursuant to these goals, several different approaches have been taken to implement laser desorption sources.[7-10]

The source used in this work is shown in Figure 2.2, and has been described in great detail previously.[4,11] Sample is applied to the surface of a graphite rod, which in turn is mounted on a metal "boat." This boat is mounted on a metal rod and enters the vacuum chamber through a dual O-ring sealed load lock. This entire load lock assembly

is mounted on an O-ring sealed X-Y translation stage on the chamber, which allows for precise, real time positioning of the graphite rod below the general valve orifice (Parker General Valve Series 9, 800μm, 20 Hz) for maximum signal. This is the same valve assembly that was described above (Section 2.3.1) for thermal heating of samples.

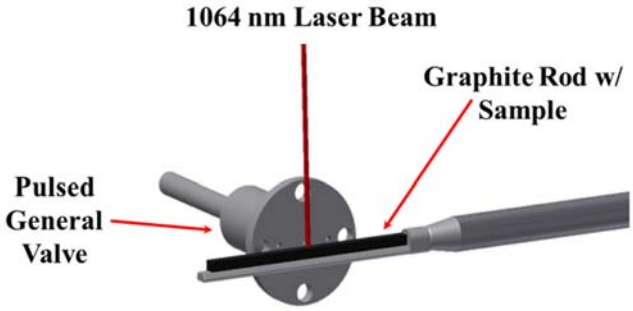


Figure 2.2. Drawing of the laser desorption source (Drawing courtesy of J.C. Dean).

The load lock and O-ring seal assembly allows the sample rod to be removed from vacuum without having to vent the chamber; an important capability to possess due to the frequent need to reload sample over the course of a day. Sample is desorbed from the graphite via a 1064 nm laser pulse (20 Hz, Continuum MiniLight, ~3-5 mJ/pulse), and entrained in the supersonic expansion. A pure Argon backing gas (typically ~4-6 bar) is used in these experiments to achieve maximum collisional cooling. The desorption laser and pulsed valve timing are controlled with a timing box (BNC 565) allowing for real time optimization of the delay that produces the coldest, most intense signal.

A home built linear actuator and speed controller push the rod externally at a rate of about 1 mm/min to move fresh sample into the path of the desorption laser. The speed of this linear actuator can be adjusted based on how quickly the sample is depleted from the surface of the graphite rod. Using this source, we have been able to introduce a ~660 Da peptide into a molecular beam and record its single-conformation spectra.[12]

2.3.3 Producing Water Clusters

A home built gas mixing manifold was built in order to seed high vapor pressure molecules into the backing gas (Figure 2.3). This manifold allows for control over the flow through two different sample cells, enabling the easy study of systems such as Benzene-$(H_2O)_n$ which involves two high vapor pressure molecules. Shut-off valves are included to enable one side, or both, to be closed off from the main backing gas flow, a feature that allows the manifold to be used full time. The sample cells are hollow, stainless steel containers that are filled with cotton wool. The liquid, volatile samples are then dropped on to the cotton wool. This reduces the possibility that liquid droplets will be formed in the gas lines. A similar set up was used previously by C.J. Gruenloh.[13]

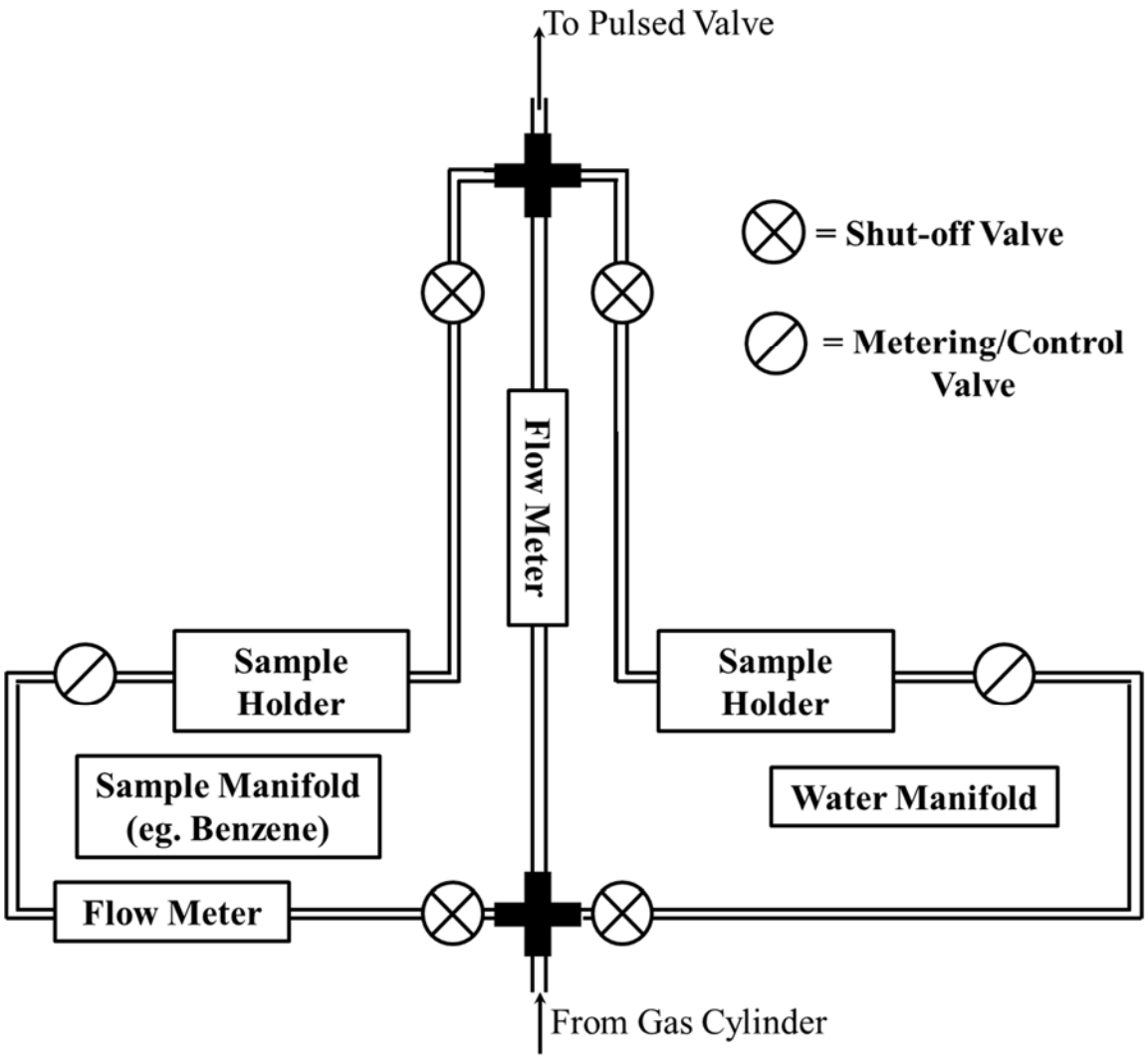


Figure 2.3. Drawing of the gas-mixing manifold that is used to introduce water vapor and/or highly volatile (ie. high vapor pressure) compounds into the expansion.

The manifold is constructed with 1/8" stainless steel tubing with Swagelok connectors and valves. The flow of the central line and the "sample" line are measured using Teledyne Hastings flow meters. The flow through the "water" line is measured as the drop in the main line flow when the water line is opened. For instance, if the total

flow is 10 sccm with the water line closed and 8 sccm when the water line is opened, the total flow through the water sample holder is 2 sccm.

The percentage of water added to the backing gas can be calculated using the following formula[13]:

$$\frac{Sample\ Flow\ (sccm)}{Total\ Flow\ (sccm)} \times \frac{Sample\ Vapor\ Pressure\ (torr)}{Backing\ Pressure\ (torr)} \times 100\ \% =\ \%\ Sample \qquad (2.1)$$

Typical percentages of water vapor in the backing gas range from roughly ~0.1 to ~0.5%.

The water/sample cell is kept at room temperature under normal operating conditions. If higher, or lower, water/sample concentrations are required the cell can be heated, or cooled, to increase, or decrease, the starting vapor pressure of the water/sample in the cell. It should be noted that if the water/sample cell is to be heated, the entirety of the gas manifold would also need to be heated to prevent condensation within the manifold. This method would be most beneficial in making large changes in the percentage of water/sample in the backing gas. For fine adjustments in the concentrations, it is most beneficial to use the needle/metering valves to change the flow through the water/sample cells. It should be noted that there is a delay between adjustment of the needle/metering valves and effect on the signal as the water/sample concentration will need to equilibrate in the gas lines.

2.4 Instrumentation and Hardware

2.4.1 Laser Sources in the Infrared (IR) and Ultraviolet (UV)

High-power, tunable laser radiation is crucial for conducting our experiments. To this end, several systems are available to generate the needed ultraviolet (UV) and infrared (IR) light.

Tunable UV radiation is generated by frequency doubling the output of a pulsed dye laser (Lambda Physik Scanmate, Radiant Dyes NarrowScan). The dye laser is pumped by a Q-switched Nd:YAG (Continuum Surelite II) operating in the either the second- (532 nm, 120-130 mJ per pulse) or third-harmonic (355 nm, 70-90 mJ per pulse). The final desired wavelength of light determines which laser dye is used, which in turn determines which harmonic is needed to pump the dye laser. The dye laser fundamental has a final resolution of 0.06 cm^{-1} at 580 nm, and is separated from the desired UV light with a Pellin Broca prism. UV laser energy per pulse was controlled by placing neutral density filters into the dye laser fundamental beam prior to frequency doubling in a BBO crystal. UV power is kept in the linear absorption regime.

Tunable infrared (IR) radiation is generated using a Laservision optical parametric converter (OPC) containing both an optical parametric oscillator and amplifier (OPO/OPA). The fundamental of a seeded Nd:YAG laser (1064 nm, ~640 mJ per pulse, 10 Hz, Contiuum Surelite EX) is used to pump both the OPO and OPA stages. The pump beam enters the OPC through a telescope (ensuring optimal beam waist) before passing through a 33/67 beam splitter: 33% of the pump energy is used to pump the OPO stage

while 67% is directed towards the OPA stage. The OPO pump beam is then frequency doubled to 532 nm (18794 cm^{-1}) in a KTP crystal.

The OPO stage is comprised of an output coupler, two angle tuned KTP crystals, and a grating-tuning mirror combination. The grating-tuning mirror set up acts as both a rear mirror for the OPO stage and spectral filter for the output light, increasing the resolution. The OPO crystals "split" the input pump photon into two photons whose sum is equal to the input photon energy (owing to the conservation of energy). Historically, the lower energy photon is termed the "idler" while the higher energy photon is termed the "signal". The frequency of these photons relative to one another is determined by the angle of the OPO crystals relative to one another; by angle tuning the crystals (with a simultaneous adjustment of the grating-tuning mirror) the photon energies can be scanned. The signal and idler beams possess opposite polarizations, allowing for the filtering of one beam with a stack of plates silicon polarizer.

The "idler" photon is then injected into the OPA stage where it is mixed with the residual (67%) 1064 nm pump beam in a series of four KTA crystals. In the OPA crystals difference frequency mixing occurs between the 1064 nm photon and the idler from the OPO stage. In this process the OPO idler becomes the signal of the OPA. The difference frequency mixing yields two OPA signal photons and one OPA idler photon. Again, the signal and idler photons will have opposite polarization which allows of the selection of either of the two beams. The full process yields an OPA idler photon that is completely tunable from ~2000 cm^{-1} to ~4100 cm^{-1}—which fully encompasses the hydride stretching region (~2800 to ~3800 cm^{-1})—with pulse energies of ~10-15 mJ.

The tuning range of this high-power table top infrared source can be further extended with the addition of a third mixing crystal. The $AgGaSe_2$ crystal is used to access the infrared region between ~1200 cm^{-1} and 1800 cm^{-1}; this region is particularly important to peptide systems because it covers the amide I (C=O stretch) and II (NH bend) regions. To generate this light, the signal and idler beams from the OPA are carefully overlapped and propagated ~6 feet before entering the $AgGaSe_2$ crystal, where they undergo mixing to generate the desired far-infrared light. The beams are propagated such a great distance to generate a more uniform beam intensity profile. The crystal itself is extremely delicate to "hot spots" in the beam profile, so great care is taken to ensure a uniform intensity profile. The two beams should barely cause a "back burn" on well-developed Kodak Linagraph paper. Typical pulse energies are ~300 μJ-1mJ per pulse.

2.4.2 Time-of-Flight (TOF) Mass Spectrometer

The molecular beam time-of-flight (TOF) chamber used in all of the experiments described in this dissertation is shown in Figure 2.4 (drawn by C.P. Rodrigo). This coupling of supersonic-jet cooling, TOF-MS, and laser spectroscopy allows for spectra to be recorded in not only a conformer-specific manner, but also mass-selectively. In this section we will follow the molecules through one experimental cycle.

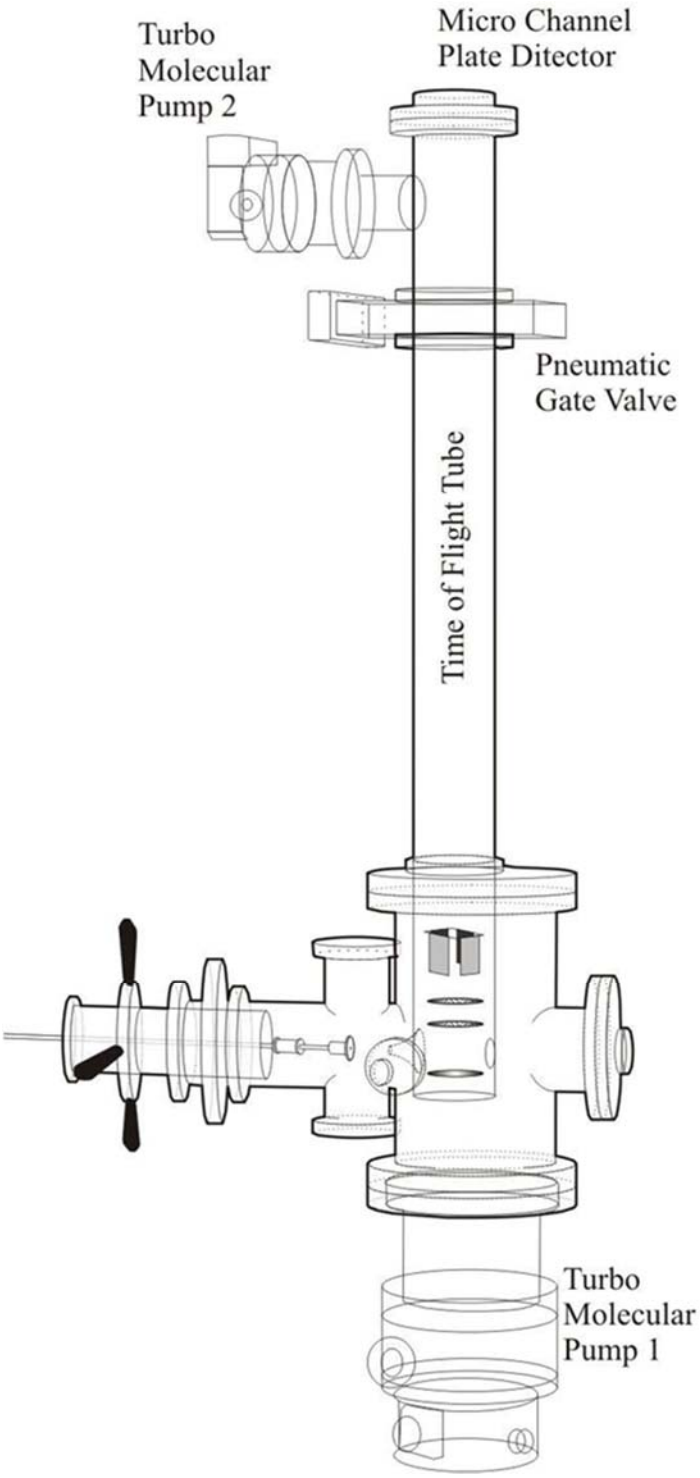


Figure 2.4. Drawing of the molecular beam, time-of-flight (TOF) mass spectrometer chamber (Drawing by C.P. Rodrigo).

A molecule of interest is entrained in an Argon, Neon, or Helium backing gas and expanded through a pin-hole orifice by a pulsed general valve (Parker-Hannifin Series 9, 500 or 800 μm) forming a supersonic expansion (see Section 2.2) with a 20 Hz repetition rate. The valve assembly is mounted in a "captains wheel" on a O-ring sealed X-Y translation stage on the chamber, allowing for real time positioning of the valve for optimal signal. The supersonic expansion is skimmed with a 2 or 3 mm skimmer (Beam Dynamics) downstream at the entrance of the TOF (figure 2.4). The ion extraction optics are surrounded by a metal shroud and the skimmer which generates two sections of the chamber which are differentially pumped: the source region, and the TOF. The source chamber typically reaches an operating pressure of ~10^{-5} mbar via an 880 L/s water cooled turbomolecular pump (TMU 1001, Pfeiffer) which is backed by a mechanical pump.

After the expansion is skimmed, the molecular beam travels into the extraction region of an orthogonal Wiley-McLaren type TOF (R.M. Jordan).[14,15] Once in the extraction region, the molecular beam interacts with various laser beams that are orthogonal to both the TOF axis, and the molecular beam axis. Ions generated by interactions with the laser radiation are immediately extracted and accelerated into a 1 m long flight tube by repeller and draw-off-grid held at constant voltages of +4270 V and +3950 V, respectively. After exiting the extraction region, but still near the base of the flight tube, there are two pairs of steering plates to which a potential can be applied. By altering these applied potentials, the ions' X-Y trajectories can be controlled to ensure optimal impact on the detector.

After traversing the 1 m long field free flight tube, the mass-selected ions will impact the detector. The detector in these experiments is a chevron-geometry dual microchannel plate detector (MCP, 25 mm, R.M. Jordan). The detector region is pumped by a second, air-cooled, 220 L/s turbomolecular pump (TMU 261, Pfeiffer) which is backed with a second mechanical pump. This turbo pump holds the detector and flight tube at an operating pressure of ~10^{-6}-10^{-7} mbar depending on the backing gas and backing pressure used.

The detector region can be sealed off from the lower portion of the flight tube and the source chamber with a pneumatic gate valve. This allows for the chamber to be vented to atmosphere while the MCP can remain at vacuum, extending the life of the MCP. The MCP plates themselves can, and should be, replaced when signal levels begin to fade. *Z*(carboxybenzyl)-Gly-OH is typically used as a calibration molecule for optimizing signal when starting a new project. With the old MCP plates, between 4300 and 4500 V of gain would need to be applied to achieve ~6 V of integrated signal on the main origin transition. With the new MCP plates ~6 V of integrated signal could be achieved with only 3100 to 3300 V of gain applied (~4 mJ/pulse of dye laser fundamental). The applied gain on the MCP has a range of values because signal levels depend on many different parameters and not just the detector gain. The rule of thumb for the applied gain is that for every 100 V applied signal increases by a factor of 2. With the addition of new MCP plates, improvement in signal between 2^{10} and 2^{14} (or a factor of 1000-4000) was achieved, indicating that the old MCP plates were far beyond their useful lifetime.

Raw signal from the MCP detector is passed through a DC-300 MHz pre-amplifier (Stanford Research Systems, SR445) twice providing 25X amplification. The amplified signal is then sent to a 300 MHz digital oscilloscope (Tektronix 3032B) and a gated integrator (Stanford Research Systems, SR250). When recording UV spectra, a frame-referenced adjustable mass-gate is set and data is acquired by the data acquisition computer via a GPIB connector from the oscilloscope to a PCI board. IR data is acquired using the gated integrator whose output is sent to a DAQ card (National Instruments) which is read via GPIB connector to a separate PCI board. The laser scanning and data acquisition are controlled via home-built LabVIEW code. All timings on the instrument are controlled by two timing boxes (BNC 565).

2.5 Single- and Double-Resonance Spectroscopic Methods

Multi-resonance laser spectroscopy methods lie at the heart of every project discussed in this dissertation. This section will provide an explanation on how laser pulses, in various schemes, can provide detailed spectroscopic information; schemes ranging from the basic single laser experiment to more complicated multi-resonance experiments used to record excited state IR spectra.

2.5.1 Single-Resonance Methods

Recording the vibronic excitation spectrum of the gas phase molecule is the first step in any experiment. This is accomplished using resonant two-photon ionization (R2PI).[16] R2PI has two variants: one-color resonant two-photon ionization (1C-R2PI) and

two-color resonant two-photon ionization (2C-R2PI). Figure 2.5 provides schematics for each method.

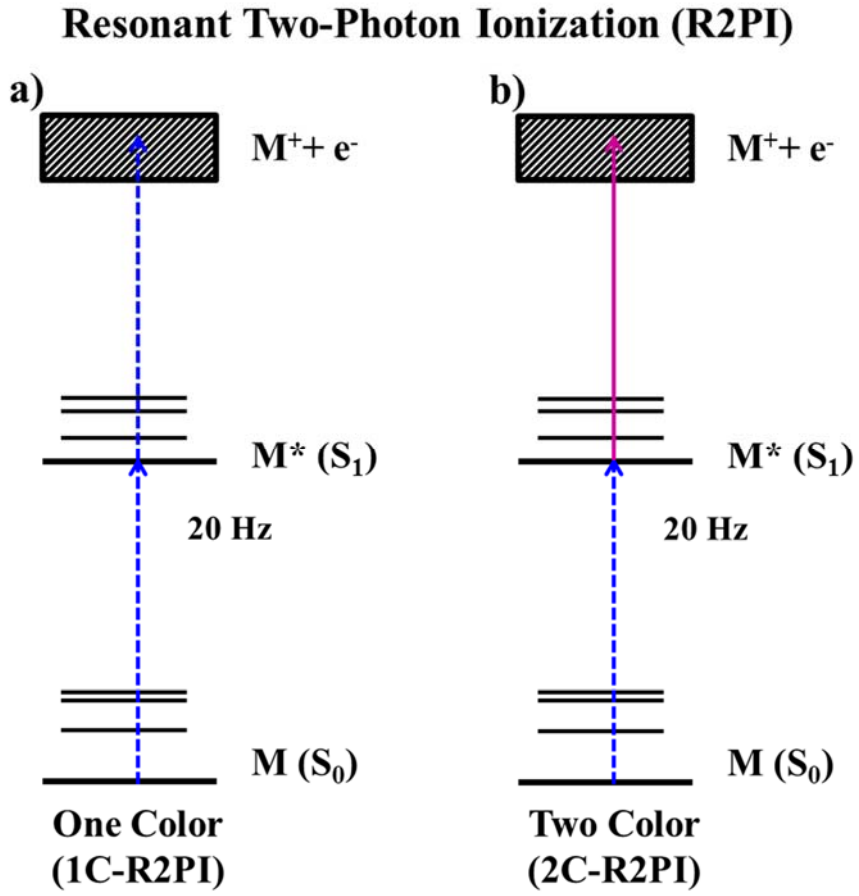


Figure 2.5. Schematics of one- (a) and two-color (b) R2PI.

R2PI is a process which requires absorption of two photons to ionize the species of interest. The first photon is absorbed by the ground state molecule, exciting an electron to an excited state. In most of our experiments this transition is a π-π^* transition from S_0 to S_1. After the molecule is excited a second photon is absorbed bringing the molecule to the ionization continuum where the electron is stripped from the molecule. This generates

the parent ion of interest which is detected. This process takes place in the extraction region of the TOF instrument described in Section 2.4.2 (Figure 2.4), allowing for the mass-selective detection of the ion of interest.

In 1C-R2PI both photons come in the same laser pulse. The process is not two-photon absorption, but rather a 1+1 process (Figure 2.5a). The laser wavelength is tuned through the electronic origin region of interest and a spectrum is recorded by plotting the ion signal as a function of laser wavelength. Special care is taken to keep the laser power low to prevent saturation effects and multiphoton processes (< ~0.5 mJ per pulse).

2C-R2PI involves a wrinkle on the standard 1C-R2PI experiment. In 2C-R2PI, two separate laser pulses are used to perform the R2PI process, $\lambda_1$and λ_2 (Figure 2.5b). These two laser pulses are spatially and temporally overlapped in the interaction region with the molecular beam. Typically, the lasers are brought in from opposite sides of the chamber facilitating easy spatial overlap of the two beams. Occasionally—like when performing RIDIR spectroscopy with 2C-R2PI—it is necessary to bring the UV lasers into the same side of the chamber. In such instances the UV beams can be combined in a beamsplitter. Temporal overlap of the laser pulses is controlled with a timing box (BNC 565). The first photon, λ_1, is scanned through the origin region of interest and is used to excite the molecules from the ground to the electronically excited state. This photon energy is typically kept very low, such that minimal 1C-R2PI signal is observed (< ~0.2 mJ per pulse). The second photon, λ_2, is set to the red of the origin transition such that it causes no ionization of the molecules. This laser is typically kept at the highest allowable energy (> ~0.5 mJ per pulse). In this set-up the total energy, 1+1', is great enough to cause ionization, but each laser on its own causes minimal, or ideally, no ionization.

2C-R2PI is used in a variety of situations. The first of these is when the ionization potential (IP) of a molecule is so high that two photons of the excitation wavelength has insufficient energy to cause ionization. Although this is the case for many molecules, it is not a problem in the molecules discussed in this dissertation.

Another common instance where 2C-R2PI is favorable is when the oscillator strength of a molecules electronic transition is very high. This is a problem because saturation of these transitions is easily accomplished; in order to achieve unsaturated conditions, the laser power must be decreased so far that signal levels drop to unusable levels.[17] This is an optimal instance for 2C-R2PI to be employed. The energy of the first photon is kept extremely low to avoid saturation, while the second photon energy remains high for optimal ionization, ensuring high, unsaturated signal levels. Similar arguments can be made where fragmentation upon photo-ionization is exceedingly efficient—a common problem in hydrogen bonded clusters.

2C-R2PI is also required when recording single-conformation infrared spectra in electronically excited states. The details regarding this situation will be discussed in Section 2.5.2.2.

2.5.2 Double-Resonance Methods

Double-resonance methods are required to record conformation-specific spectra in both the IR and the UV. In general, the first laser “tags” the conformer of interest while the second monitors the changes in signal induced by the first laser. The following sections will fully describe the implementation of these double-resonance techniques.

2.5.2.1 Single-Conformation UV Spectroscopy

Recording single-conformation UV spectra fulfills two important goals: identification of the number of conformers present in the molecular beam, and identification of clean, unique UV transitions which are required for recording single-conformation IR spectra. Pursuant to these goals, two methods are employed to record the conformation-specific UV spectra: UV-UV holeburning (UV-UV HB) and IR-UV holeburning (IR-UV HB) (Figure 2.6).

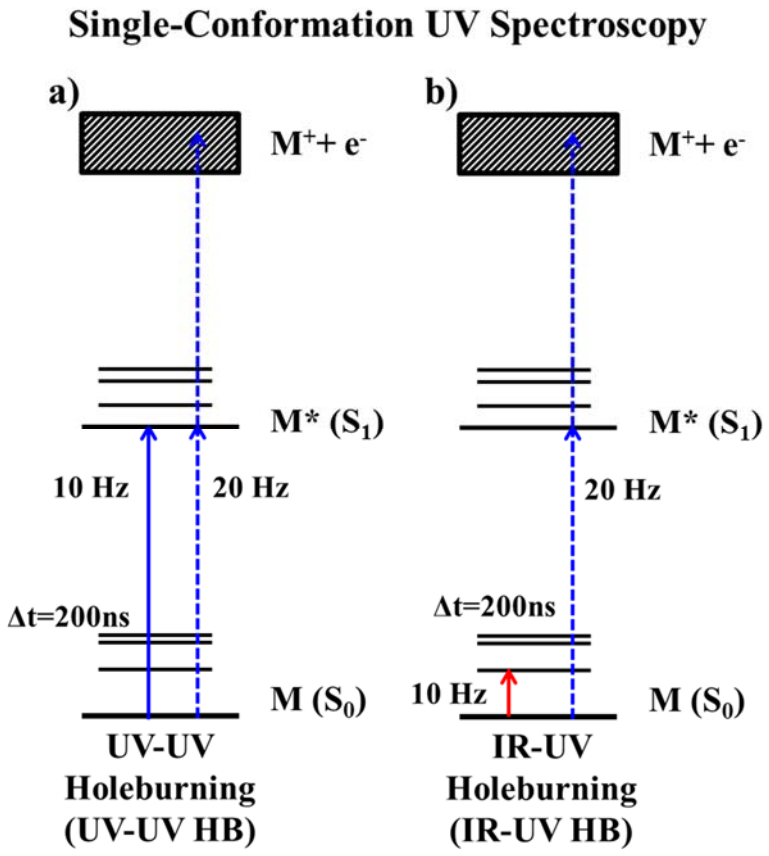


Figure 2.6. Schematics of UV-UV holeburning (a) and IR-UV holeburning (IR-UV HB).

In UV-UV HB, two UV lasers are spatially overlapped with one temporally preceding the other by ~200 ns (typical timing, actual timing can vary). The first laser (in time) operates at a repetition rate of 10 Hz and is termed the "burn" laser. The burn laser is fixed to an electronic transition that is believed to be unique to a single conformer. The second laser (in time) operates at 20 Hz and is termed the "probe" or "scan" laser (Figure 2.6a). In UV-UV HB two distinct ion packets will be generated in the TOF mass spectrum when both lasers are on resonance. The second of these ion packets, the one generated by the probe laser, will be monitored to generate the desired signal.

With the burn laser fixed to a single transition, the probe laser is scanned through the electronic origin region of interest. When the burn and probe lasers are resonant with a transition that shares a common ground state, the ion signal generated by the probe laser will be depleted. This depletion in signal is caused by the burn laser removing a significant portion of the population from the ground state—there are simply fewer molecules to be ionized. By mismatching the laser repetition rates, a real time subtraction between the "burn on" and "burn off" measurements can be recorded, using active baseline subtraction in a gated integrator (Stanford Research Systems, SR250). The holeburning spectrum is recorded by plotting the depleted ion signal as a function of probe wavelength. Typically, the burn laser power is left quite high so that a significant portion of the ground state population can be removed (~1 mJ per pulse).

A second method to record conformation-specific UV spectra is IR-UV holeburning (Figure 2.6b). The experimental procedure for this technique is the same as for UV-UV HB with one major difference: the burn laser is an IR laser. In IR-UV HB the burn laser serves the same purpose as in UV-UV HB except here a unique IR transition is

used to modulate the ground state population. The UV probe laser is than scanned through the electronic origin region of interest. When the IR and UV lasers are both resonant with transitions which share the same ground state, the signal generated by the UV laser is depleted. Similarly to UV-UV HB, this depletion signal can be monitored in real time and plotted as a function of probe laser wavelength to generate conformation specific UV spectra.

Special care should be taken when performing IR-UV HB that truly unique, single-conformer IR transitions are used as burn transitions. IR transitions are typically much broader than those in the UV, thus there is a higher likelihood of overlap between IR spectral features. When the burn laser transition is partially or completely overlapped with another conformation's transition, the resulting holeburning spectrum will not be single-conformation but rather will have contributions from both conformers. This results in spectral cross-contamination of one conformers' spectrum into another's, which can complicate the spectroscopic interpretation of the data. The likelihood of such overlap is significantly less in the UV as the transitions are much sharper (FWHM: ~1.5 cm^{-1}).

2.5.2.2 Single-Conformation IR Spectroscopy

Conformation-specific IR spectra are the most important spectroscopic data recorded for the projects detailed in this dissertation. Vibrational modes of a molecule are exquisitely sensitive to their environments. As such, single-conformation IR spectroscopy allows us to probe and assign the structure of the conformers found in the molecular beam. To obtain this data resonant ion-dip infrared (RIDIR) spectroscopy is used.[18-20]

Figure 2.7a provides the laser interactions that are needed to generate single-conformation IR spectra using RIDIR spectroscopy. To record a RIDIR spectrum two lasers are required: one UV and one IR. The UV and IR lasers are spatially overlapped through the extraction region of the TOF instrument described in Section 2.4.2 (Figure 2.4); the IR laser temporally precedes the UV by ~200 ns. The UV laser is operated at 20 Hz and is set to an electronic transition that belongs to a single conformation. This UV laser will generate a constant ion-signal which can be monitored. The IR laser, operating at 10 Hz, is scanned through an IR region of interest; the IR laser removes ground state population when resonant with a vibrational transition. When the IR and UV lasers are resonant with transitions that share a common ground state, the IR induced depletion of ground state population causes a corresponding depletion in the total ion signal generated by the UV laser. As with the holeburning techniques, the repetition rate mismatch between the lasers allows for the real time monitoring of the depletion signal via active baseline subtraction in the gated integrator (Stanford Research Systems, SR250). The depletion signal is then plotted as a function of IR laser wavelength to produce the conformation-specific IR spectrum.

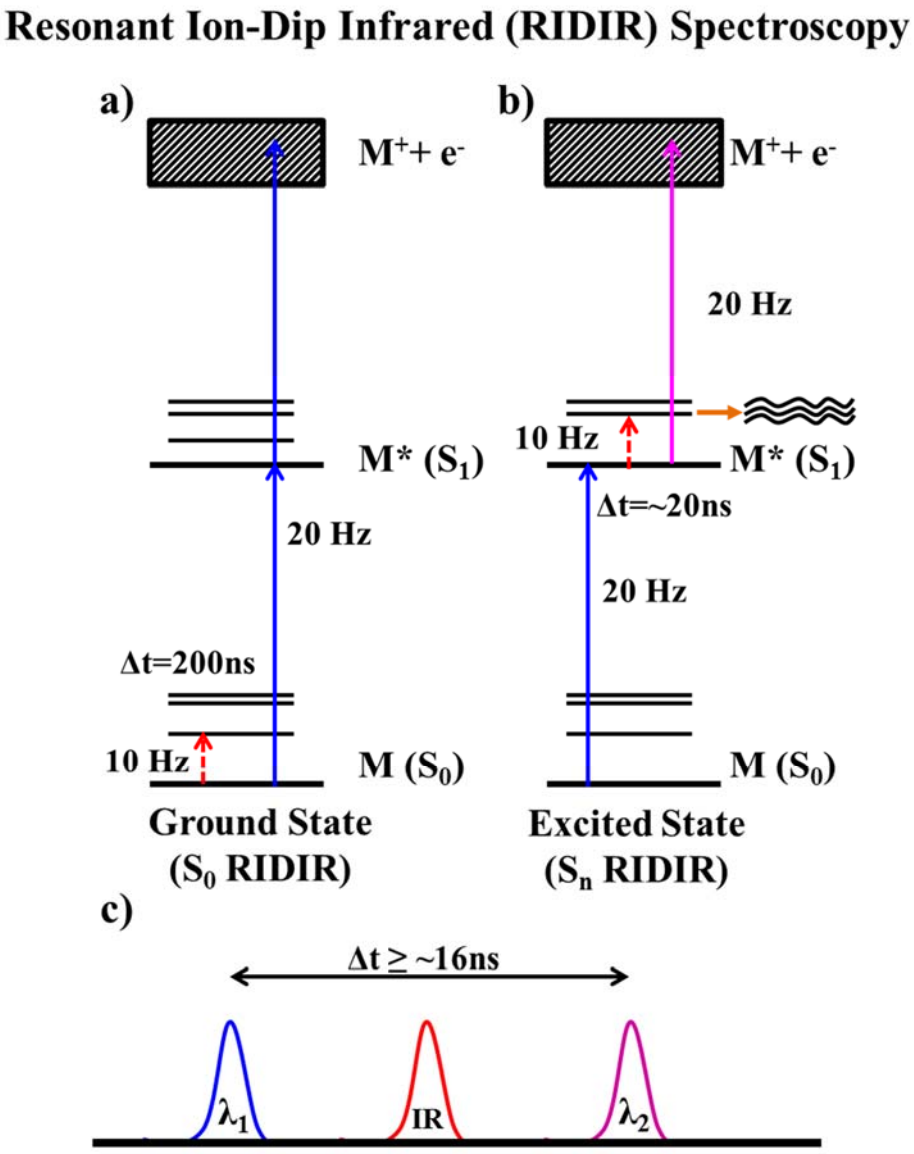


Figure 2.7. Schematics of ground- (a, S_0 RIDIR) and excited-state RIDIR (S_n RIDIR, ES-RIDIR) spectroscopy. Timing of laser pulses (c) used in the S_n RIDIR experiments.

The procedure described above is used to record IR spectra in the ground electronic state. It is possible to record similar RIDIR spectra in electronically excited states (S_n RIDIR, ES-RIDIR).[21-23] Although the underlying physical processes are the same in ground- and excited-state RIDIR spectroscopy, there are key experimental

differences between the two methods. The first step in recording ES-RIDIR (Figure 2.7b) is to achieve quality 2C-R2PI conditions as described in Section 2.5.1. The IR laser beam is spatially overlapped with the two UV beams used to generate the 2C-R2PI signal. Once 2C-R2PI signal is optimized, λ_1 and λ_2 are separated temporally as shown in figure 2.7c. This will generate two ion packets (signals) in the TOF mass spectrum which are separated by the temporal delay between λ_1 and λ_2: the first comes from weak 1C-R2PI from λ_1 and the second strong signal comes from "pure" 2C-R2PI caused by λ_2. In this case, λ_1 prepares a population of excited state molecules which are ionized some time later by λ_2. It should be noted that using this temporally separated method would be an excellent way to record "pure" 2C-R2PI with zero one-color contribution.

After λ_1 and λ_2 are temporally separated, the IR beam is temporally placed between them (Figure 2.7b and 2.7c). Once this pulse sequence is achieved, the IR laser is scanned while the ion packet generated by λ_2 is monitored for signal depletions in a similar manner to ground state RIDIR spectroscopy. This results in the recording of an ES-RIDIR spectrum.

One molecular property that is critical for this experiment is the excited state lifetime of the molecule of interest. In order for this experiment to work, the molecule's excited state lifetime must be long enough that λ_1 and λ_2 can be temporally separated such that the IR laser pulse has no temporal overlap with either of the two UV pulses. It is critical for λ_1 and λ_2 to be 100% separated; if this is not the case, it cannot be definitively stated whether ground- or excited-state depletions are being monitored. When using lasers with temporal profiles that are 6-8 ns wide, this means the lifetime must be greater than ~16 ns (Figure 2.7c).

2.6 Computational Methods

Theoretical calculations are vital for making the comparisons with the experimental spectra that lead to firm conformational assignments. Although a great deal of insight into the structures can be gleaned from simple spectroscopic intuition, having quantum mechanical calculations is required for a comprehensive conformational assignment.

Many of the molecules that will be discussed in this dissertation are large, possessing a large number of internal degrees of freedom and a complicated potential energy surface. As such, developing a comprehensive set of suitable starting structures by hand, using only chemical intuition is not feasible. To circumvent this issue, calculations are performed at a low level of theory to produce a starting set of local minimum structures. To generate these starting structures a conformational search algorithm, as implemented in MACROMODEL,[24] is used to thoroughly explore the potential energy surface. These calculations are performed using classical mechanics force fields, usually Amber*[25] for peptides, and the molecular mechanics force field (MMFFs) for water clusters. Structures within 50 kJ/mol of the calculated global minimum are retained by the calculation. This window can be expanded if needed.

These force field calculations generate a set of starting structures that are optimized using higher levels of *ab initio* and density functional theory (DFT) calculations. When starting a new project, the lowest 30-40 structures generated by the force field search are re-optimized using higher level quantum mechanical calculations. In many cases this range is expanded to the lowest ~100 structures to fully sample the potential energy surface and locate the global minimum structure at this higher level of

theory. All calculated structures in this dissertation were optimized using Truhlar's hybrid density functional[26] M05-2X and the Pople basis set 6-31+G(d) as implemented in Gaussian09.[27] Harmonic frequency calculations were performed on these optimized structures. The Gaussian keywords int(grid=ultrafine) and scf=tight are specified to ensure a tight convergence and the best possible vibrational frequencies. Because the calculated vibrational frequencies are harmonic, they are scaled in order to account for anharmonicity in the real vibrational potential. These scale factors are provided in the individual chapters.

2.7 References

(1) Clarkson, J. R. Thesis, Purdue University, 2006.

(2) James, W. H., III Thesis, Purdue University, 2009.

(3) Buchanan, E. G. Thesis, Purdue University, 2014.

(4) Dean, J. C. Thesis, Purdue University, 2014.

(5) Smalley, R. E.; Wharton, L.; Levy, D. H. *Acc. Chem Res.* **1977**, *10*, 139-145.

(6) Lubman, D. M.; Rettner, C. T.; Zare, R. N. *Journal of Physical Chemistry* **1982**, *86*, 1129-1135.

(7) Meijer, G.; deVries, M. S.; Hunziker, H. E.; Wendt, H. R. *Applied Physics B* **1990**, *51*, 395-403.

(8) Nir, E.; Plützer, C.; Kleinermanns, K.; de Vries, M. *Eur. Phys. J. D* **2002**, *20*, 317-329.

(9) Piuzzi, F.; Dimicoli, I.; Mons, M.; Tardivel, B.; Zhao, Q. *Chem. Phys. Lett.* **2000**, *320*, 282.

(10) Taherkhani, M.; Riese, M.; BenYezzar, M.; Müller-Dethlefs, K. *Rev. Sci. Instrum.* **2010**, *81*, 063101.

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

(12) Gord, J. R. Thesis, Purdue University, 2016.

(13) Gruenloh, C. J. Thesis, Purdue University, 1999.

(14) Wiley, W. C.; McLaren, I. H. *Rev. Sci. Instrum.* **1955**, *26*, 1150-1157.

(15) Lubman, D. M.; Jordan, R. M. *Rev. Sci. Instrum.* **1985**, *56*, 373.

(16) Lubman, D. M. *Mass Spectrom. Rev.* **1988**, *7*, 559-592.

(17) Dean, J. C.; Kusaka, R.; Walsh, P. S.; Allais, F.; Zwier, T. S. *J. Am. Chem. Soc.* **2014**, *136*, 14780-14795.

(18) Page, R. H.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988**, *88*, 4621-4636.

(19) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205-241.

(20) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827-8839.

(21) Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Phys. Chem. Lett.* **2013**.

(22) Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015**, *142*, 154303.

(23) Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015**, *142*, 154304.

(24) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440-467.

(25) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287-303.

(26) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289-300.

(27) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J.V.; Cioslowski, J.; Fox, D. J.; Revision C.01 ed.; Gaussian, Inc.: Wallingford CT, 2009.

CHAPTER 3. BINDING WATER TO A PEG-LINKED FLEXIBLE BICHROMOPHORE: IR SPECTRA OF DIPHENOXYETHANE-$(H_2O)_N$ CLUSTERS, N=2-4

3.1 Introduction

1,2-Diphenoxyethane (DPOE) is a prototypical flexible bichromophore in which the two phenyl rings responsible for the molecule's UV absorptions are joined by a flexible -O-CH_2-CH_2-O- linker characteristic of a polyethylene glycol (PEG) or polyethylene oxide (PEO) chain (Figure 3.1).[1-8]

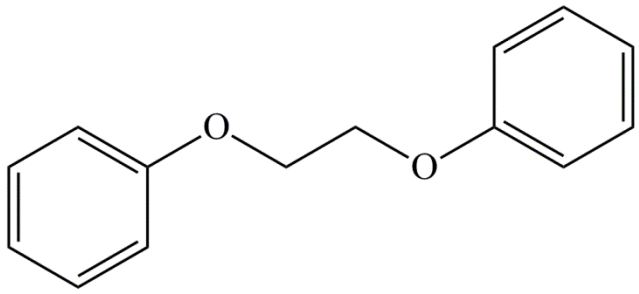


Figure 3.1. Chemical structure of 1,2-diphenoxyethane (DPOE).

The fact that this linkage is flexible imbues DPOE with a pair of close-lying multi-dimensional excited state surfaces that change the distance and orientation of the two aromatic chromophores, thereby shifting their relative energies and modulating the strength of their excitonic coupling. By bringing the molecule into the gas phase, and

cooling it in a supersonic expansion, the population is funneled into the zero-point levels of low-lying conformational minima where they can be studied via conformation-specific spectroscopy.

While much of the focus of past work[9-25] has been on the vibronic coupling between the close-lying S_1 and S_2 states, a necessary prerequisite for such understanding is establishing the inherent ground-state conformational preferences of the molecule, which serve as the starting geometry for excitation onto the excited state surfaces. To that end, the observed vibronic and infrared transitions of DPOE have been previously assigned to the *ttt* and *tgt* conformers shown in Figure 3.2a.[22,23] These conformers differ primarily in the dihedral angle τ about the central C-C bond, either *trans* (180^o) or *gauche* (60^o), with the former of C_{2h} symmetry and the latter C_2 symmetry. The two conformers have roughly equal populations in the supersonic expansion, as proven by infrared population transfer spectroscopy.

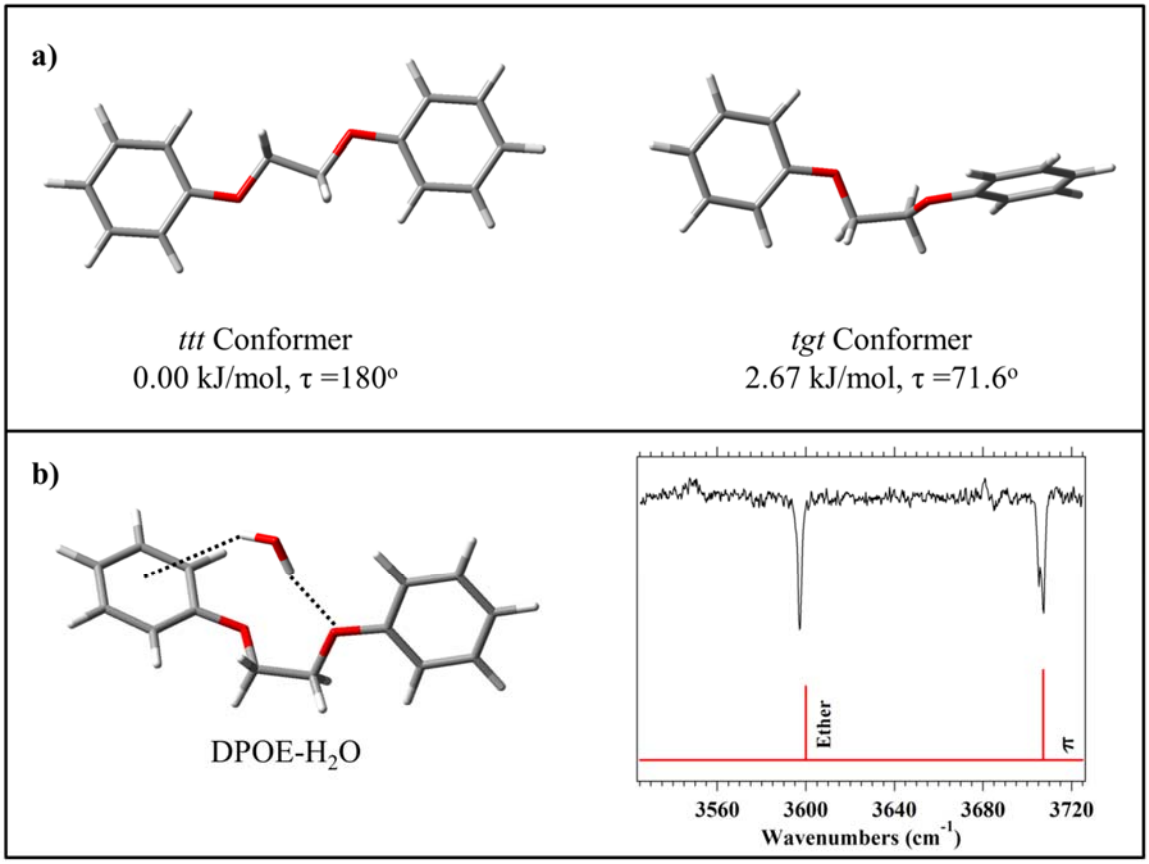


Figure 3.2. The two observed conformers of DPOE monomer, (a) *ttt* and (b) *tgt*, with their calculated central dihedral angles and relative energies (from Ref. [22]). (c) Single structure observed for DPOE-H_2O complex (from Ref. [24]).

Armed with this understanding about the isolated molecule, one can then address how the conformational preferences are modified in the presence of solvent, connecting more directly with the solution phase. As nature's solvent, water is an important solvent to consider in this regard. Indeed, DPOE is a fascinating solute for water because it offers several binding sites, most notably the two ether oxygens and the two π clouds of the two phenyl rings. As a first step along this path, we previously studied the infrared and ultraviolet spectroscopy of the 1:1 DPOE-H_2O complex containing a single H_2O molecule.[24] This single water molecule binds selectively to the *tgt* conformer of DPOE

(Figure 3.2b), donating both its OH groups to H-bonds to the ether oxygen of one phenoxy group and the π cloud of the other. This asymmetric solvent binding is reflected in the OH stretch infrared spectrum of the water molecule (Figure 3.2b, right), with OH stretch fundamentals that are now partially localized on the ether-bound OH (3597 cm^{-1}) and π-bound OH (3705/3707 cm^{-1}) groups, respectively. The asymmetric binding of this single water molecule also has important consequences for the excited electronic states, breaking the near-degeneracy of the S_1 and S_2 excited states, localizing the electronic excitation on one or the other aromatic ring, and providing an additional mechanism for vibronic mixing to occur between the two excited states.

In the present work, we extend our spectroscopic and structural investigations to larger DPOE-$(H_2O)_n$ cluster with n=2-4, focusing attention here on their ground state properties. In so doing, several fascinating questions arise. First, how are the conformational preferences of DPOE modified by the presence of these larger water clusters? Second, is the interaction with DPOE sufficiently strong that it breaks up the water cluster or re-shapes it in a significant way? Third, how does DPOE make use of the multiple binding sites for water? The influences these water clusters have on the excited states of DPOE is addressed in Chapter 4.

3.2 Experimental Methods

1,2-Diphenoxyethane (DPOE) (99%) was obtained from Sigma-Aldrich and used without further purification. The DPOE was placed into the valve assembly (described in Chapter 2) and heated to ~95 °C to obtain sufficient vapor pressure of DPOE. The DPOE vapor was entrained in a neon backing gas containing trace amounts of water controlled

by a needle valve/flow meter mixing system (~0.35-0.45% water, ~2-3 bar, ~0.1-0.12 (bar cm^3)/s) and expanded through the nozzle (500 μm), producing a supersonic expansion in which DPOE-$(H_2O)_n$ cluster formation occurs.

A detailed description of the instrumentation and experimental methods used in this study is provided in Chapter 2.[24,26-33] UV spectra were recorded in the appropriate mass channel using two-color resonant two-photon ionization (2C-R2PI).[31] Spectra were recorded over the entire S_0-S_1/S_2 origin regions, ~36200-36800 cm^{-1}. The wavelength used for the ionization step (λ_2, ~277-279 nm) was chosen to maximize two-color enhancement while minimizing production of broad background from fragmentation of higher mass ions. Conformation specific UV spectra were recorded using both UV-UV and IR-UV holeburning.[28-31]

Single-conformation IR spectra were recorded for each of the observed conformers through the use of resonant ion-dip infrared (RIDIR) spectroscopy.[27-29] Ground state RIDIR spectra were recorded in a configuration in which the IR light counter-propagates the molecular beam.[32,33] The counter-propagation geometry generates a column of excited clusters and depleted ground state clusters which can be interrogated using the UV laser. By using a delay between IR (10 Hz) and UV (20 Hz) of Δt=69μs, the molecules interrogated by the UV laser were passing through the 2 mm diameter skimmer when the IR pulse was fired, overlapping the IR light with all the cold molecules/clusters that are later interrogated via R2PI. The depletion in ion signal is recorded in the same fashion as the more traditional RIDIR spectroscopy arrangement. The setup of this measurement scheme is similar to that of the mass-resolved infrared population transfer experiments (MR-IRPT) previously published, but with a different

delay time that avoids collisional re-cooling.[32,33] All ground state single-conformation spectra were recorded in the mass channel that afforded the highest intensity signal.

3.3 Computational Methods

A series of starting geometries for the DPOE-$(H_2O)_n$ clusters were generated using a conformational search using the molecular mechanics force field (MMFFs) in MacroModel.[34] This search provided all of the possible stable geometries within a 50 kJ/mol window. Geometry optimizations and harmonic frequency calculations were performed on the ~30 lowest energy structures using the GAUSSIAN09 quantum chemistry program.[35] These calculations were performed using Truhlar's hybrid density functional[36] M05-2X with a 6-31+G(d) basis set. The harmonic vibrational frequencies are scaled by 0.952 in the OH stretch region and 0.928 in the CH stretch region, to account for anharmonicity.[22,24]

For the remainder of this chapter, the term 'binding energy' will refer to the energy required to remove the intact water cluster from the DPOE monomer (ie., E(DPOE-$(H_2O)_n$-{E(DPOE) + E[$(H_2O)_n$]}=E(Binding)), using optimized structures for the DPOE monomer and $(H_2O)_n$ clusters. In the [1:3] chain structure, the energy of the cyclized water trimer was used in calculated E(Binding) because no stable water trimer chain exists due to a barrierless isomerization to the water cycle. Binding energies are corrected for basis set superposition error (BSSE) and zero-point energy effects using previously published methods.[37,38]

3.4 Results and Analysis

3.4.1 R2PI Spectra

Figure 3.3 presents a set of 2C-R2PI spectra recorded in the $[1{:}n]^+$ mass channels (1 DPOE:n water) with n=0-5. The R2PI spectrum in the monomer mass channel (top trace) contains transitions due to the two observed DPOE monomer conformers (labeled *ttt* and *tgt,* Figure 3.2a) and the DPOE-H_2O complex ([1:1]), which fragments into the monomer mass channel with high efficiency following photoionization.[39,40] These previously reported spectra and their assignments are included here because they provide an essential foundation and point of comparison for the present work on the larger clusters with n=2-4.[22-24]

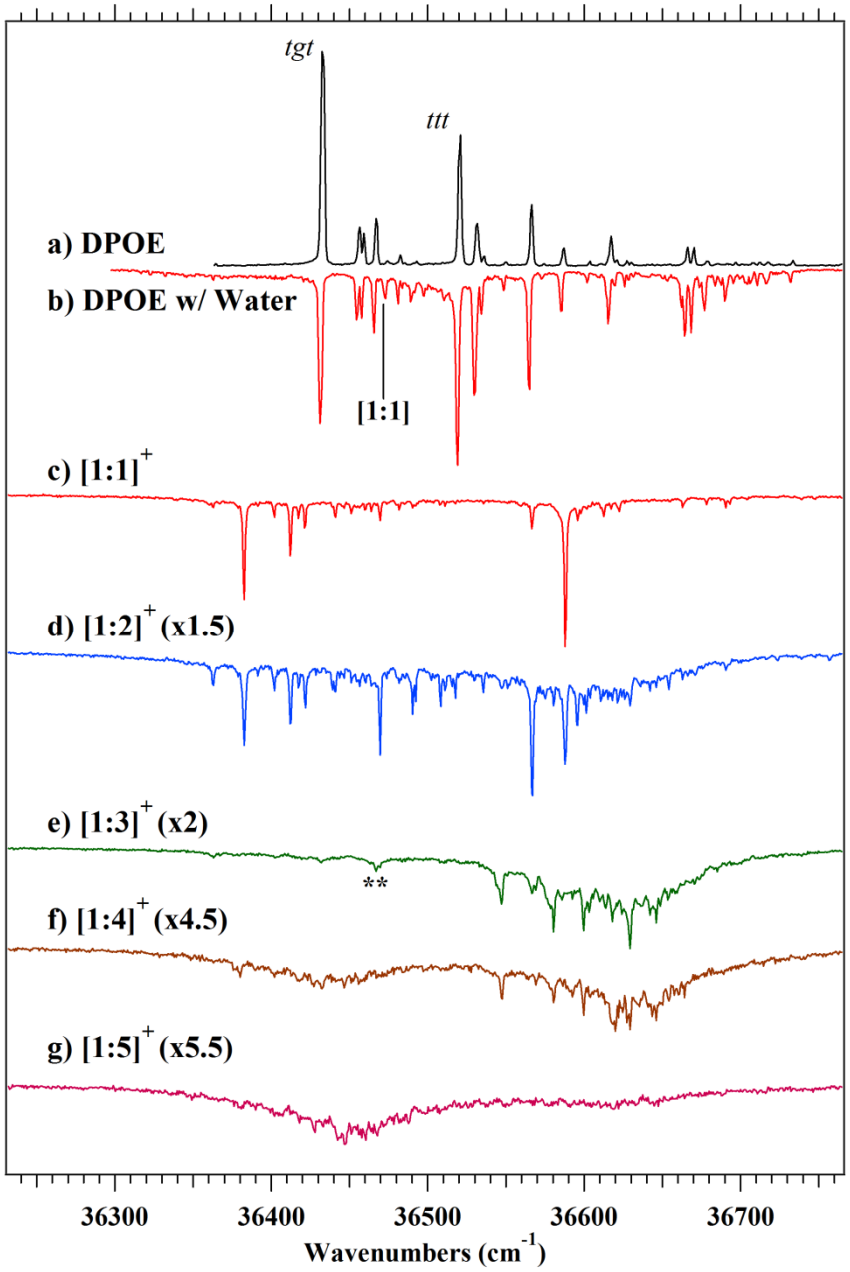


Figure 3.3. Overview 2C-R2PI spectra of DPOE-$(H_2O)_n$ clusters recorded in the $[1{:}n]^+$ mass channel. (a, b): DPOE monomer mass channel (a) with, and (b) without water present in the expansion. (c-g) n=1-5.

The R2PI spectra in the $[1:4]^+$ and $[1:5]^+$ are shown here in order to gain a perspective on the development of the R2PI spectrum with cluster size. However, clusters appearing in these mass channels (due to 1:5 and larger clusters) will not be characterized further in the main text, but expanded views of the 2C-R2PI spectra in these mass channels are provided in Figure 3.4 ($[1:4]^+$ mass channel) and Figure 3.5 ($[1:5]^+$ mass channel).

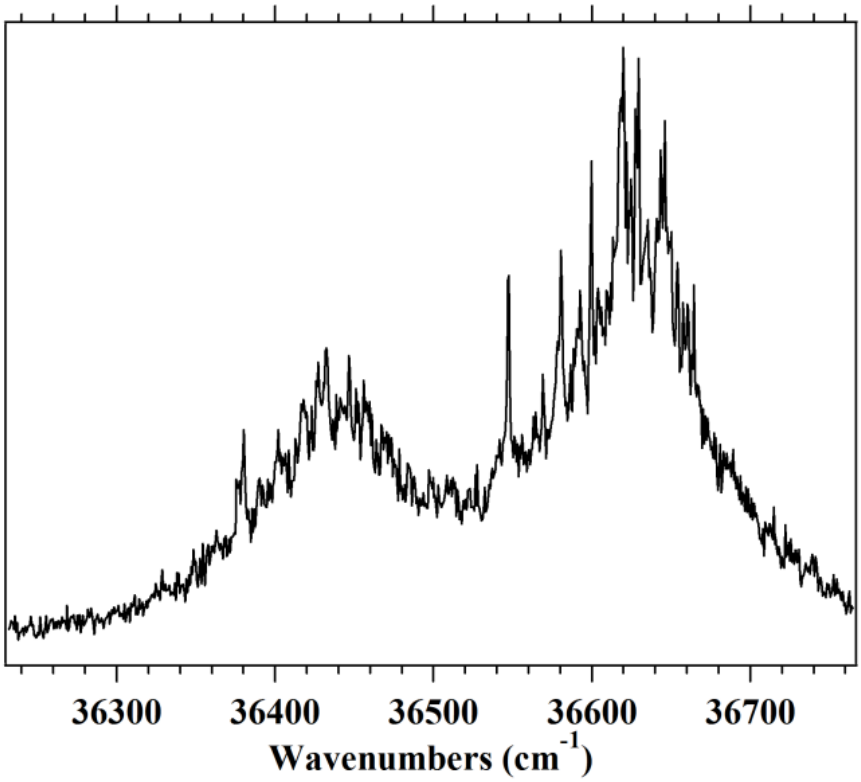


Figure 3.4. 2C-R2PI spectrum recorded in the $[1:4]^+$ mass channel (m/z=286).

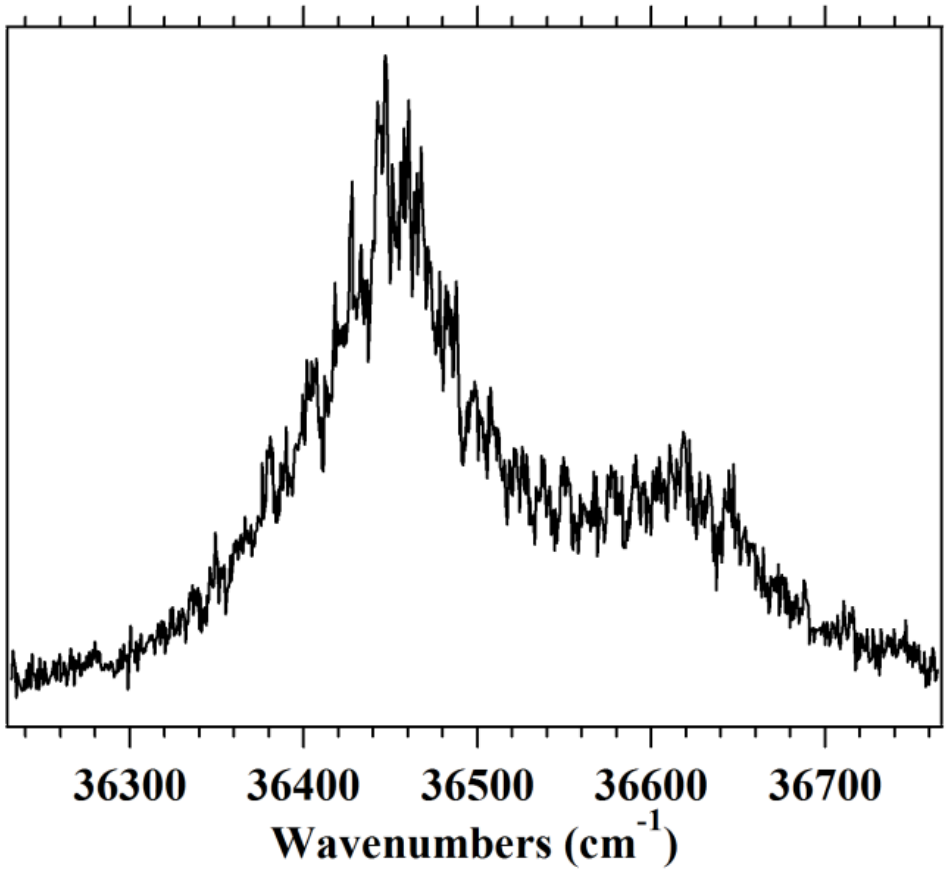


Figure 3.5. 2C-R2PI spectrum recorded in the $[1{:}5]^+$ mass channel (m/z=302).

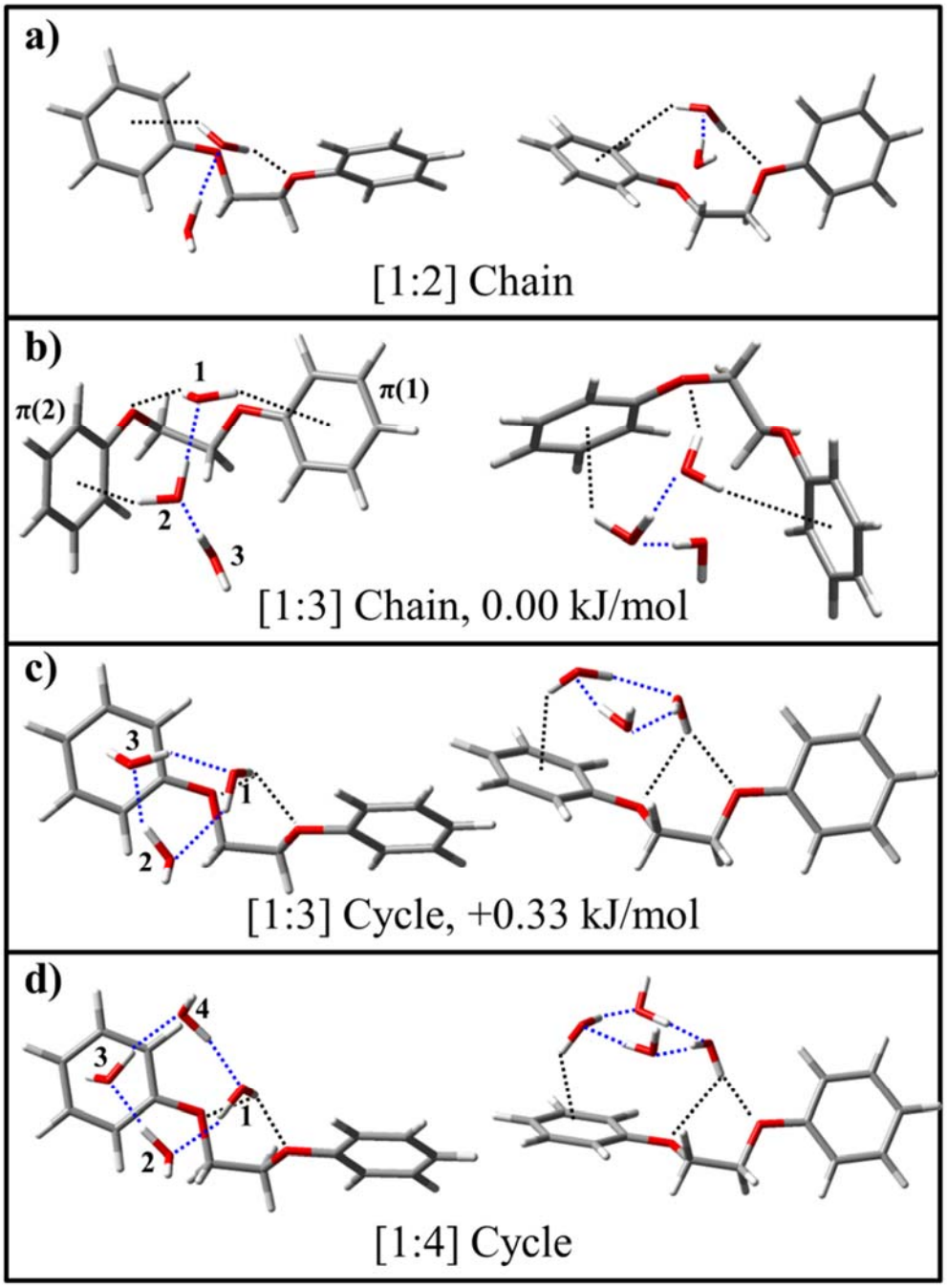


Figure 3.6. The assigned structures for the (a) [1:2], (b) [1:3] chain (conformer B), (c) [1:3] cycle (conformer A), and (d) the [1:4] cycle clusters. Structures are shown with their relative energies. Water molecules have been numbered for ease of discussion of the transition frequencies and hydrogen bond lengths. DPOE-water (black dotted lines) and water-water (blue dotted lines) hydrogen bonds have been labeled for clarity.

3.4.2 Calculated Ground State Structures of DPOE-$(H_2O)_n$, n=2-4

Figure 3.6(a,b,d) presents the calculated global minimum structures for DPOE-$(H_2O)_n$, n=2-4, respectively. Also included in the figure is the optimized structure of a second nearly isoenergetic structure for the 1:3 cluster (Figure 3.6c), calculated to be only 0.33 kJ/mol above the global minimum. According to the calculations, the lowest energy structure for the 1:2 cluster incorporates a water dimer in which the donor water forms an OH···O H-bond with the acceptor water that, in turn, binds to DPOE in the same way as in the 1:1 complex (Figure 3.2b). Like its 1:1 counterpart, the water dimer is bound in this structure to the *tgt* monomer; the lowest energy [1:2] cluster built off the *ttt* monomer is 6.95 kJ/mol higher in energy. The global minimum 1:3 cluster (Figure 3.6b) extends the water chain by adding a third water molecule onto the back-end of the 1:2 cluster. The calculations predict a significant re-orientation of the O-bound ring in forming a secondary π H-bond with interior water in the chain (water 2). Interestingly, a second 1:3 cluster structural type that incorporates a water trimer cycle (Figure 3.6c) is calculated to be only 0.33 kJ/mol higher in energy. This structure has its primary binding via a single OH that now binds more symmetrically to the two ether oxygens, while a secondary stabilization occurs via another of the water molecules in the trimer cycle forming a π H-bond with one of the aromatic rings. This same motif is retained in the global minimum 1:4 cluster structure (Figure 3.6d), which incorporates a tetramer cycle with secondary binding via a π H-bond the water across the cycle. Both the trimer and tetramer cycles are predicted to be bound preferentially to the *tgt* DPOE monomer structure which facilitates formation of this secondary π H-bond. The energy level diagram provided in Figure 3.7 nicely illustrates the competition between chain (blue lines) and cycle (red

lines) structures in the 1:2, 1:3, and 1:4 structures. The cross-over from a preference for forming chains to cycles at n=4 is clear from the figure.

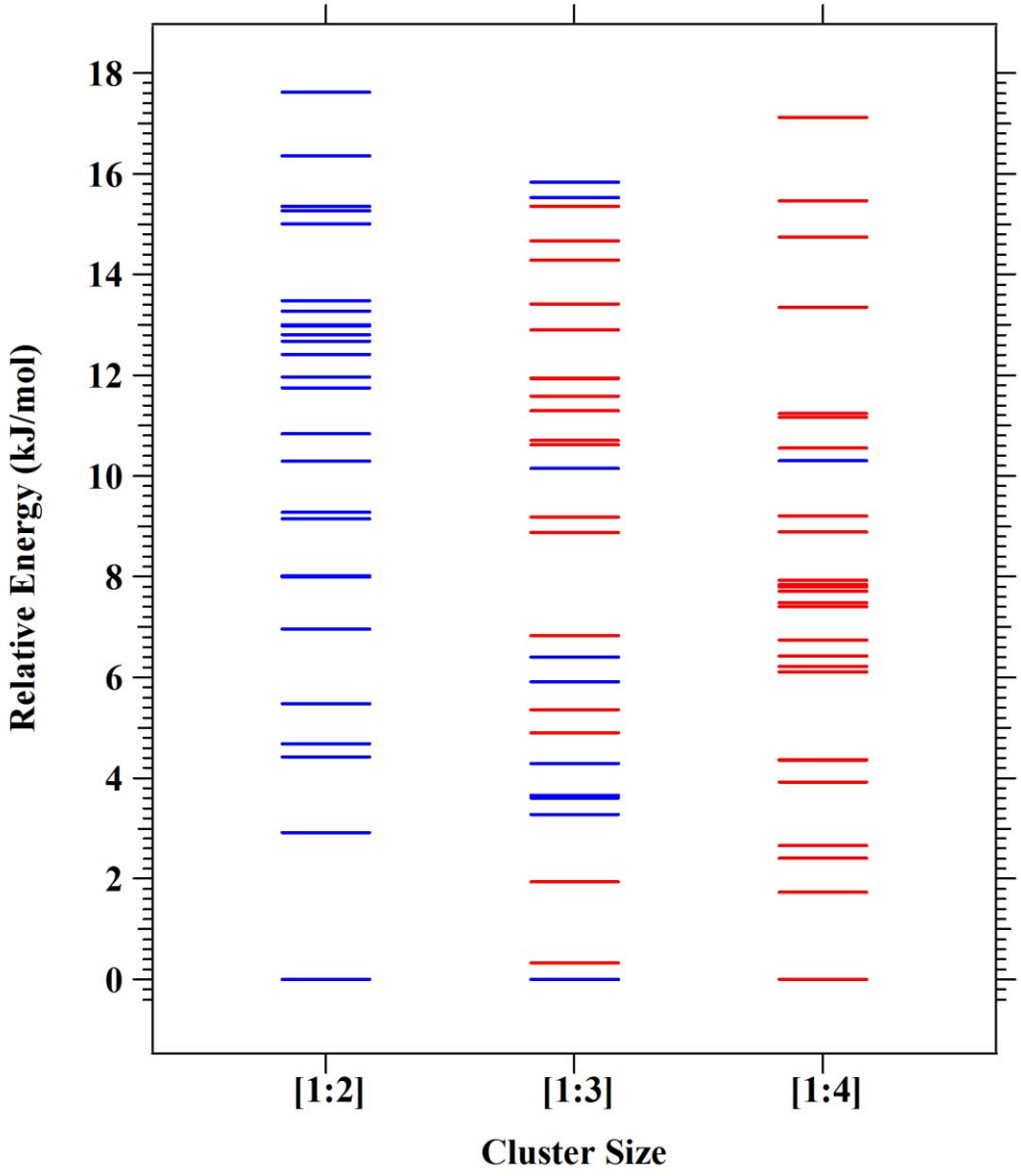


Figure 3.7. Energy level diagram comparing the relative energies of the DPOE-$(H_2O)_n$ (n=2-4) clusters. The blue lines correspond to structures in which the water molecule forms an H-bonded chain, while the red lines correspond to cyclic H-bonded water structures. Energies are zero-point corrected and calculated at the M052x/6-31+G(d) level of theory.

3.4.3 Holeburning and S_0 RIDIR Spectroscopy of DPOE-$(H_2O)_n$, n=2-4

3.4.3.1 DPOE-$(H_2O)_2$

Figure 3.8a presents a UV-UV holeburning scan in the $[1:1]^+$ mass channel, with UV holeburn laser fixed on what is proven later to be the S_1 origin transition (36383 cm^{-1}) of the 1:2 neutral cluster. Comparison of this scan with the 2C-R2PI spectrum in the $[1:1]^+$ mass channel shows that all observed transitions in this mass channel are due to a single isomeric structure for the [1:2] neutral cluster, appearing in the $[1:1]^+$ mass channel due to efficient photofragmentation following photoionization. In principle, the wavelength of the second photon used for ionization could be tuned further red to remove this fragmentation. However, the Franck-Condon factors from the S_1/S_2 states to the ion drop off with increasing wavelength, therefore reducing the photoionization efficiency, making this impractical. Such an efficient fragmentation following photoionization is a common occurrence in neutral water clusters bound to both simple aromatics and peptide species.[30,39-41] Thus for this and larger clusters, the most intense signals in a given mass channel ($[1:n]^+$) comes from the neutral [1:n+1] cluster. Figure 3.8b shows the analogous IR-UV holeburning scan, taken with the IR holeburning laser fixed at 3550 cm^{-1}. This IR absorption is unique to the 1:2 neutral cluster, producing an IR-UV holeburning scan essentially identical to the UV-UV holeburning scan above it. Further discussion of these spectra is reserved for Chapter 4 on the excited states of DPOE-$(H_2O)_n$.[42]

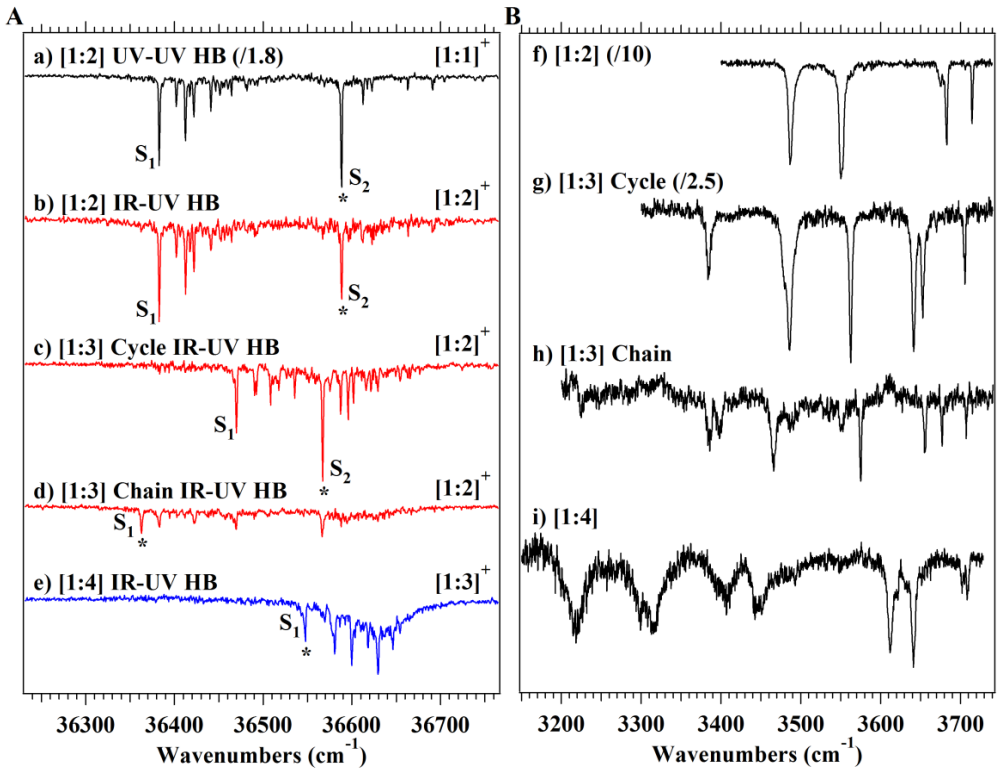


Figure 3.8. Panel A: Holeburning spectra recorded in the indicated mass channels. (a) UV-UV and (b) IR-UV holeburning scans for the 1:2 neutral cluster. (c-e): IR-UV holeburning spectra for the (c) [1:3] cycle, (d) [1:3] chain, and (e) [1:4] cluster. Transitions used to record the ground state RIDIR spectra are marked with asterisks. Panel B: Overview RIDIR spectra recorded for the (f) [1:2], (g) [1:3] cycle, (h) [1:3] chain, and (i) [1:4] clusters.

RIDIR spectra in the OH- and CH-stretching regions were recorded in order to make firm structural assignments. Panel B of Figure 3.8 presents an overview of the RIDIR spectra in the OH stretch region for DPOE-$(H_2O)_n$ clusters with n=2-4. This side-by-side comparison highlights the development in the OH stretch spectra with increasing water cluster size. The results for the 1:2 cluster in the OH-stretch region are shown in Figure 3.9a. The spectrum exhibits four well resolved transitions due to the four OH stretch fundamentals at 3715, 3683, 3550 and 3486 cm^{-1}.

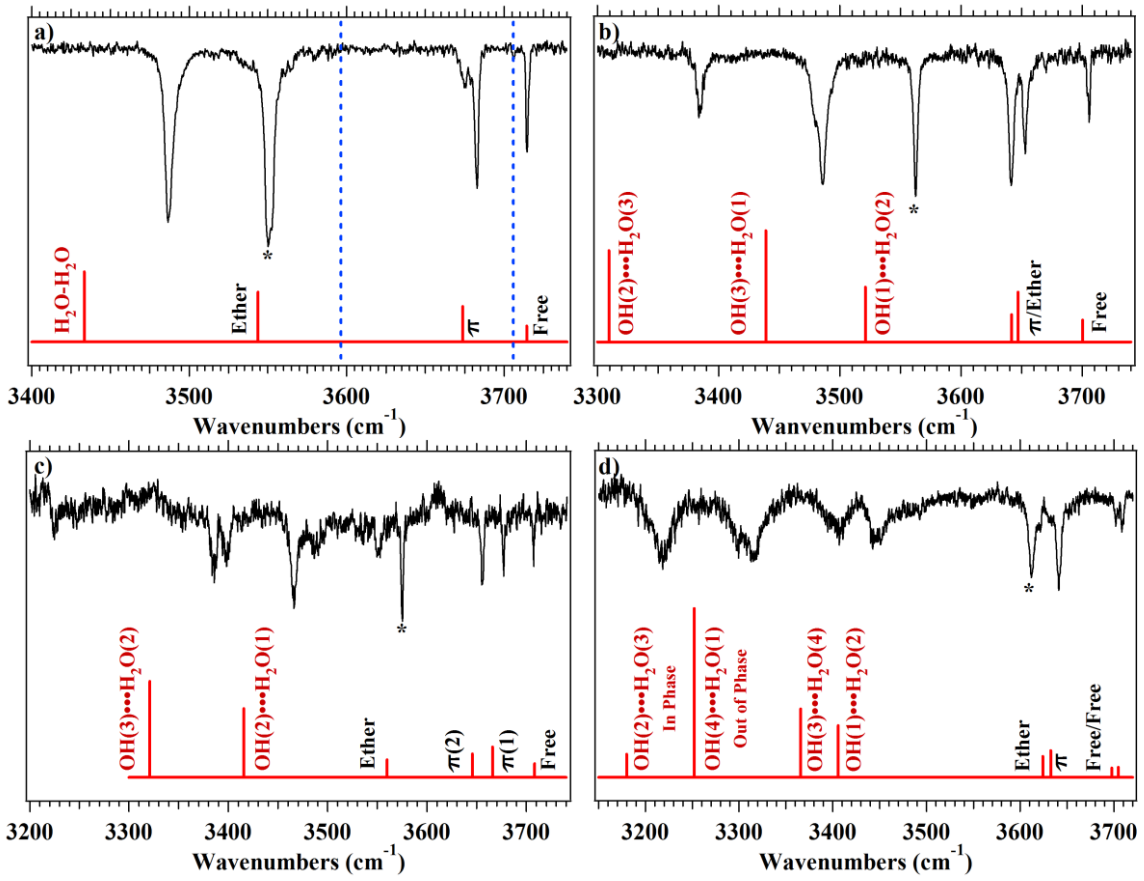


Figure 3.9. OH-stretch RIDIR spectra for (a) [1:2], (b) [1:3] cycle (conformer A), (c) [1:3] chain (conformer B), and (d) [1:4] cycle clusters. The stick spectra are those predicted by DFT M05-2X/6-31+G(d) calculations for the assigned structures. Dashed lines in panel (a) locate the wavenumber positions of the [1:1] cluster's π and ether bound OH-stretch fundamentals. Transitions marked by an asterisk were used to record IR-UV holeburning spectra.

Several conclusions regarding the structure of the DPOE-$(H_2O)_2$ cluster can be made even in the absence of calculations, by comparison with the DPOE-H_2O complex and other aromatic-$(H_2O)_n$ clusters.[24,39-41,43-48] The highest frequency transition at 3715 cm^{-1} is readily assignable to a free OH-stretch fundamental. The transitions at 3683 and 3550 cm^{-1} are similar to those observed in the DPOE-H_2O spectrum[24] (Figure 3.2b and dotted lines in Figure 3.9a), consistent with a similar primary binding site to DPOE in

which one water acts as a double donor bridge between an ether O-atom on one ring and the π cloud on the other ($OH\cdots\pi$ at 3683 cm^{-1} and $OH\cdots O$ at 3550 cm^{-1}). The lowest-frequency OH stretch transition at 3486 cm^{-1} is ascribable to a water-water H-bond, much as is present in benzene-$(H_2O)_2$ (3550 cm^{-1}).[39,40,48]

Table 3.1. Experimental and theoretical frequencies for the electronic and vibrational transitions (OH-stretching region) for the DPOE-$(H_2O)_n$ (n=1-4) clusters.

Structure		Vibrational Frequency (cm^{-1}) OH-Stretch								Electronic Frequency (cm^{-1})		
		Free	*Free*	*OH···O, Ether*	*OH···π*	*(1)→(2)*	*(2)→(3)*	*(3)→(4)*	*(3,4)→(1)*	$S_0 \rightarrow S_1$	$S_0 \rightarrow S_2$	$\Delta S_1/S_2$
DPOE-H_2O[a]	*Th.*[b]			3600	3707							
	Exp.			3597	3707/ 3705					36471	36660	189
DPOE-$(H_2O)_2$	*Th.*[b]	3714		3544	3674	3433						
	Exp.	3715		3550	3683	3486				36383	36588	205
A-DPOE-$(H_2O)_3$ Cycle	*Th.*[b]	3700		3642	3647	3521	3309		3439			
	Exp.	3706		3641	3653	3563	3384		3486	36470	36567	97
B-DPOE-$(H_2O)_3$ Chain	*Th.*[b]	3708		3560	3666, 3645	3416	3321					
	Exp.	3707		3575	3677, 3655	3466	3386/ 3398			36362		
DPOE-$(H_2O)_4$	*Th.*[b]	3704	3697	3624	3633	3406	3180	3366	3252			
	Exp.	3708	3702	3612	3641	3448	3218	3407	3312	36547		

[a]Vibrational frequencies taken from ref. [24].

[b]Vibrational frequencies calculated at the M05-2x/6-31+G(d) level of theory. Theoretical harmonic frequencies for the OH stretch vibrations were scaled by 0.952.

The DPOE-$(H_2O)_2$ structure is thus a simple extension of the DPOE-H_2O structure in which the second water simply hydrogen bonds to the first, retaining the *tgt* conformation for the DPOE. Indeed, calculations bear this out. The calculated vibrational frequencies and IR intensities, presented as a stick diagram below the experimental spectrum in Figure 3.9a, show excellent agreement with experiment. The assigned 1:2 structure is the calculated global minimum shown in Figure 3.6a. Table 3.1 provides a summary of the experimental and calculated vibrational frequencies, while the important bond distances derived from the calculations are collected in Table 3.2.

Table 3.2. Calculated bond lengths for the hydrogen bonds in DPOE-$(H_2O)_n$ and $(H_2O)_n$ (n=1-4) clusters in Angstroms (Å).

	Hydrogen Bond Distance (Å)						
Structure[a]	*OH···O, Primary Ether*	*OH···O, Secondary Ether*	*OH···π, Center of Ring*	*H_2O-H_2O Hydrogen Bonds*[b,c]			
				(1)→(2)	*(2)→(3)*	*(3)→(4)*	*(3,4)→(1)*
DPOE-H_2O	2.01	2.49	3.27				
DPOE-$(H_2O)_2$	1.95	2.52	3.18	1.86 *(2.81)*			
$(H_2O)_2$				1.91 *(2.87)*			
A-DPOE-$(H_2O)_3$ Cycle	2.23	2.41	2.43	2 *(2.87)*	1.82 *(2.75)*		1.91 *(2.79)*
$(H_2O)_3$ Cycle				1.88 *(2.76)*	1.87 *(2.76)*		1.9 *(2.77)*
B-DPOE-$(H_2O)_3$ Chain	2.12	2.35	π(1) 3.06 π(2) 2.48	1.84 *(2.82)*	1.8 *(2.78)*		
DPOE-$(H_2O)_4$	2.11	2.5	2.45	1.83 *(2.79)*	1.74 *(2.72)*	1.8 *(2.77)*	1.75 *(2.72)*
$(H_2O)_4$				1.77 *(2.74)*	1.77 *(2.74)*	1.77 *(2.74)*	1.77 *(2.74)*

[a]Stuctures optimized at the M05-2x/6-31+G(d) level of theory.
[b]Hydrogen bond numbering retained from Figure 3.6. Water homocycle orientation matches that of what is seen in DPOE-$(H_2O)_n$.
[c]OH···O bond lengths are given first while the O···O bond distance is the second, italicized number given in parenthesis.

The CH-stretch spectrum of the 1:2 cluster is shown in Figure 3.10c, where it is compared with that for the *tgt* DPOE monomer conformer (Figure 3.10a) and DPOE-H_2O complex (Figure 3.10b). Importantly, the excellent match with experiment confirms that it is the *tgt* conformer of DPOE that is used as the scaffold for binding the water dimer. Note that the experimental spectrum shows only minor perturbation from stretch-bend Fermi-resonance, distinctly less than is present in the spectra of the *tgt* DPOE monomer

and DPOE-H_2O complex shown above it in the figure.[22,24] The absence of R2PI transitions ascribable to the 1:2 cluster built off the *ttt* monomer conformer is in keeping with its absence in the DPOE-H_2O complex.[24]

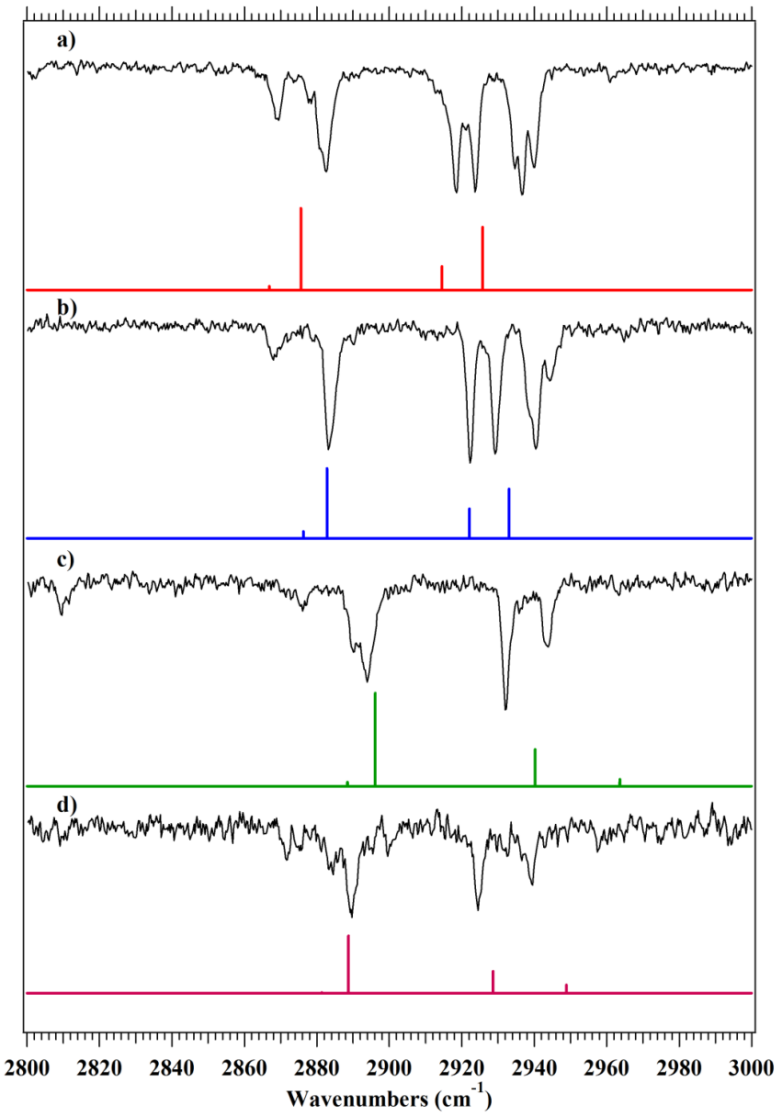


Figure 3.10. CH-Stretch RIDIR spectra for (a) DPOE (*tgt*), (b) the [1:1] complex, (c) the [1:2] cluster, and (d) the [1:3] cycle cluster. The stick spectra below the experimental traces are the predicted harmonic frequencies and IR intensities for the assigned cluster structure. The monomer and [1:1] cluster spectra have been previously published, and are shown for comparison (Refs. [22] and [24]).

3.4.3.2 DPOE-$(H_2O)_3$

The 2C-R2PI spectrum in the $[1:2]^+$ mass channel (Figure 3.3) also contains transitions due to the 1:3 clusters because of efficient fragmentation following photoionization. Using IR-UV hole-burning, two conformations of the 1:3 cluster have been identified (conformers A and B), with UV spectra presented in Figure 3.8c and d.

3.4.3.2.1 Major Conformer of DPOE-$(H_2O)_3$: Cyclic Water Trimer

Like the 1:2 cluster, the major 1:3 conformer also possesses two distinct regions of vibronic activity in its UV spectrum (Figure 3.8c), ascribable to the two electronic states of the DPOE bichromophore.

Figures 3.9b and 3.10d show the ground state RIDIR spectra of 1:3(A) in the OH- and CH-stretch regions, respectively. Based on the comparison with the DFT M05-2X calculations, conformer A is confidently assigned to the structure shown in Figure 3.6c, in which the water trimer bound to DPOE forms a H-bonded cycle, much as it does in the free water trimer.[49] This requires a re-structuring of the way in which the water molecule(s) interact with DPOE, most notably in that at most one OH group on each water molecule is available for binding to DPOE. Indeed, the calculated 1:3(cycle) uses two of these (previously) free OH groups to bind to the ether pocket and an aromatic ring π cloud.

The OH stretch spectrum reflects these structural changes in exquisite detail. As anticipated, there are six OH stretch fundamentals in the experimental spectrum, with frequencies summarized in Table 3.1. Much as in benzene-$(H_2O)_3$, the three water-water H-bonds in the trimer cycle produce OH stretch fundamentals shifted well down in

frequency, appearing in 1:3(A) at 3563, 3486, and 3384 cm^{-1}. Importantly, there is only a single "free" OH stretch transition, at 3706 cm^{-1}, and even this transition is shifted down slightly in frequency from a typical free OH-stretch (3710-3720 cm^{-1}), indicating a minor perturbation to the free OH(4) stretch. The remaining two OH stretch fundamentals appear as a closely-spaced pair at 3641 and 3653 cm^{-1}. According to the calculations, these arise from OH(1)···O(ether) and OH(3)···π interactions with DPOE(*tgt*). The frequencies of both are noticeably shifted from their values in the 1:1 and 1:2 clusters (Figure 3.9 and 3.8B). The ether-bound OH is shifted up in frequency by +44 and +91 cm^{-1} relative to the 1:1 and 1:2 clusters, reflecting a substantial weakening of this interaction. At the same time, the π-bound OH, now due to a second water molecule, is shifted down in frequency -54 and -30 cm^{-1} relative to the π-bound OH fundamentals of the double-donor bridge water that is primary binding site in the 1:1 and 1:2 clusters. Here the formation of a cycle is stabilized in part by enabling a stronger π H-bond with DPOE.

Finally, the close similarity of the alkyl CH stretch RIDIR spectrum of 1:3(A) (Figure 3.10d) with 1:2 (Figure 3.10c) confirms that 1:3(A) has the water cycle bound to the *tgt* DPOE conformer. Indeed, it is only possible for a single OH group of the cyclic trimer to bind into the pocket created by both the ether oxygens for the *tgt* conformer.

3.4.3.2.2 Minor Conformer of DPOE-$(H_2O)_3$: Water Trimer Chain

The IR-UV holeburning spectrum in Figure 3.8d reveals the presence of a second, minor conformation for DPOE-$(H_2O)_3$, 1:3(B). The UV transition at 36362 cm^{-1} is assigned to the S_0-S_1 origin of 1:3(B), which is used as detection wavelength for RIDIR scans. Figure 3.9c presents the ground state RIDIR spectrum of 1:3(B) in the OH stretch region, and compares this spectrum with the predicted spectrum for the global minimum 1:3 structure shown in Figure 3.6b. Note that this structure contains a water trimer chain bound to DPOE, with primary binding site to DPOE via a bridging water molecule, much as in the 1:1 and 1:2 clusters. The observed spectrum matches the predictions for this structure in significant detail. First, there is a set of four OH stretch transitions above 3550 cm^{-1}, including a free OH stretch transition (3707 cm^{-1}), two π-bound OH groups at 3677 and 3656 cm^{-1}, and an ether-bound OH (3575 cm^{-1}). The remaining two OH stretch fundamentals (3466 and 3386/3398 cm^{-1}) involve the two water-water H-bonds in the trimer chain. As the stick diagram indicates, the two OH stretch transitions of the primary binding H_2O molecule, have frequencies (3677 π/ 3575(ether)) in close proximity with their values in the 1:1 and 1:2 clusters (Table 3.1). More striking is the presence of a second π H-bond. The optimized structure shows this second π bond to be due to the middle H_2O in the chain, which forms with the opposite aromatic ring from π(1). In order to form π(2), the calculated structure (Figure 3.6b) has the θ_2= -85°, which re-orients the aromatic ring that acts as π(2) acceptor. This dihedral formally corresponds to a *tgg* structure that is not among the low energy structures for DPOE monomer, and is stabilized through facilitating formation of two π H-bonds.

Not surprisingly, there are at least three other 1:3 chain structures calculated to be within 4 kJ/mol of the global minimum (Figure 3.11). These structures have similar patterns to their calculated OH stretch spectra to the global minimum, and the present data cannot clearly distinguish between them. Indeed, it may be that more than one of these isomers contributes to the experimental spectrum, which sports weak "side-bands" near the main H-bonded OH stretch transitions, most notably producing a doublet in the lowest-frequency OH stretch. We postulate that more than one of the 1:3 chains may have S_0-S_1 origins at the UV wavelength used to record the RIDIR spectrum, and therefore contribute these extra transitions to the spectrum. In retrospect, the close frequency position of the S_1 origin of the 1:3 chain (-61 cm^{-1}) to that for the 1:2 chain (-40 cm^{-1}) is another strong structural indicator for the chain structure.

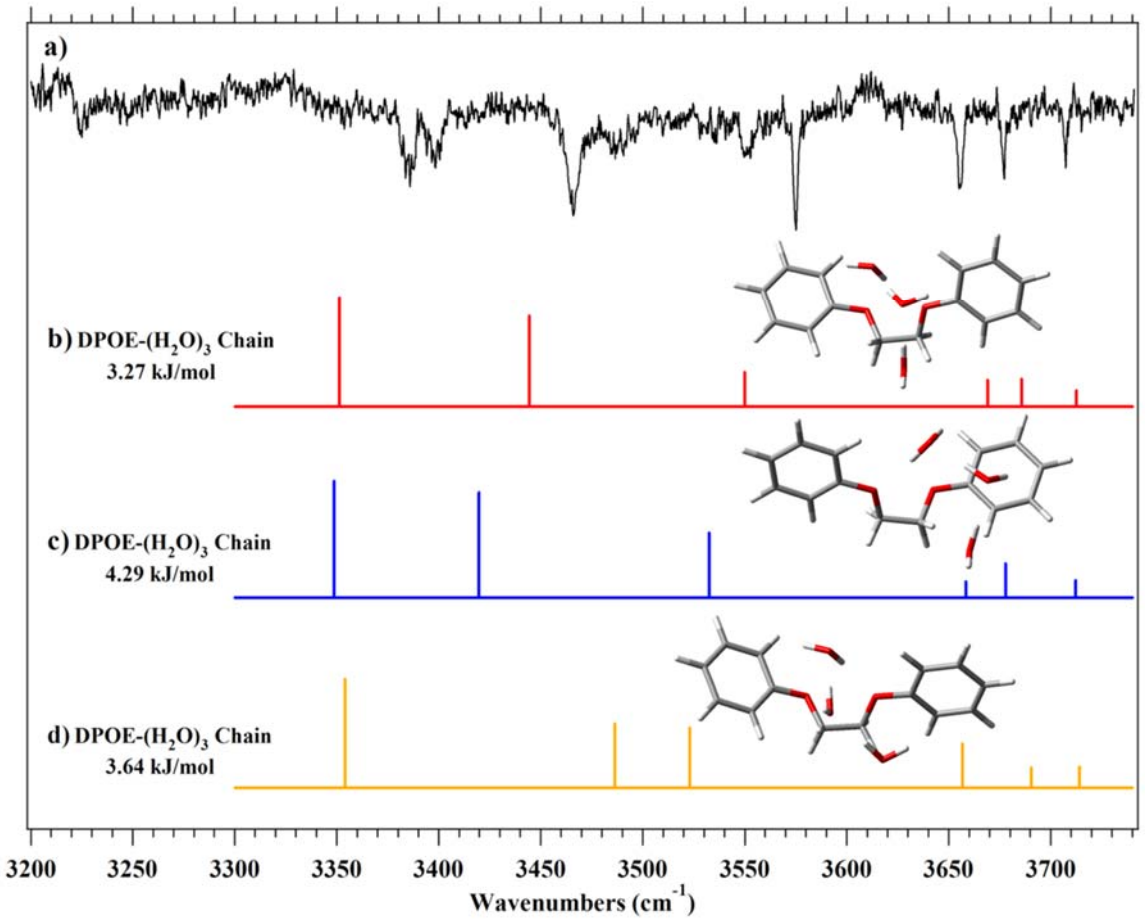


Figure 3.11. (a) Experimental OH-stretch RIDIR spectrum for the [1:3] chain cluster along with the structures and predicted spectra for three slightly higher water chain containing structures (b-d).

3.4.3.3 DPOE-$(H_2O)_4$

The 2C-R2PI spectrum in the $[1:3]^+$ mass channel is shown in Figure 3.3, with the IR-UV hole-burning spectrum taken with IR hole-burn laser at 3612 cm^{-1} shown in Figure 3.8e. As with previous sized clusters, the IR spectra will prove that the spectrum in the $[1:3]^+$ mass channel is due to the 1:4 neutral cluster, following loss of a single H_2O molecule during the photoionization process. The double-starred transition in Figure 3.3 does not burn out with the 1:4 cluster. Possible carries of this transition include a hot

band, a minor 1:4 cluster, or a higher order water cluster, which has not been explored further due to its weak intensity.

The RIDIR spectrum of the 1:4 cluster is shown in Figure 3.9d. The calculated spectrum for the global minimum structure of the 1:4 cluster is shown below it, providing an excellent match with experiment, and informing its interpretation. This structure incorporates a tetramer water cycle, which is bound to DPOE via a single OH bound to the ether oxygen(s), with further stabilization from a π H-bond with one of the aromatic rings. The spectrum consists of two free OH stretch fundamentals (3708 and 3702 cm^{-1}), a set of four broadened transitions at low frequency due to the four water-water H-bonds in the tetramer cycle (3448, 3407, 3312, and 3218 cm^{-1}), and a closely spaced pair of transitions due to the OH(3)···π and the OH(1)···O(ether) interactions with DPOE (3641 and 3612 cm^{-1}). These latter two transitions are very similar in appearance to the corresponding pair in the 1:3 cycle (Figure 3.9b, Figure 3.8h and i). Note that in the 1:4 cluster, the two OH groups used to form the H-bonds to the ether and π clouds are on water molecules on opposite sides of the tetramer (Figure 3.6d).

3.5 Discussion

1,2-diphenoxyethane has in its chemical structure (Figure 3.2a) at least three different attractive regions for binding to water molecule(s). The strongest binding is anticipated to involve the ether oxygens, which can act as acceptors in forming OH···O(ether) H-bonds. The adjacent phenyl rings can donate electron density to the ether oxygen, increasing its strength as an H-bond acceptor relative to H_2O itself. The π clouds of the two phenyl rings constitute second-tier binding sites for OH···π H-bonds,

with strengths similar to those formed in benzene-$(H_2O)_n$ clusters.[39,40,48] Finally, the aromatic and alkyl CH groups can form weak CH···O interactions with the O-atom of H_2O.

The DPOE-$(H_2O)_n$ clusters with n=1-4 form structures that make use of all three binding sites, which develop in fascinating ways with increasing $(H_2O)_n$ cluster size. The single water molecule in the DPOE-H_2O complex (Figure 3.2b) forms a bridge between the primary ether oxygen binding site on one ring and the π cloud on the other. Formation of such a bridge is only possible in the *tgt* DPOE conformer (Figure 3.2a), and in fact, there is no spectroscopic evidence for the presence of *ttt*-H_2O complexes. The calculated structure for *ttt*-H_2O (Figure 3.12) has the water bound to one of the ether oxygens, and is stabilized by a CH···O interaction. The calculations predict this structure to be 14.6 kJ/mol less strongly bound than the *tgt* structure. The absence of any DPOE(*ttt*)-H_2O complexes in the jet is likely to arise because the energy imparted to the complex by H-bond formation is sufficient to anneal the complex into its global minimum structure.

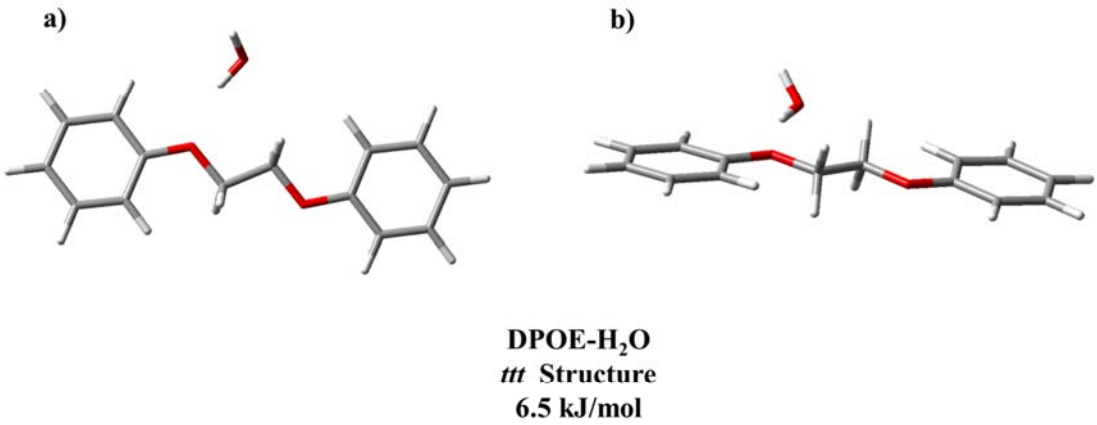


Figure 3.12. The lowest energy *ttt* DPOE-H_2O structure shown in two different orientations (a, b).

The calculated binding energy for the DPOE-H_2O complex is D_0 = 2530 cm^{-1} (30.3 kJ/mol), including a correction for basis-set superposition error and zero-point energy. Note that this value is more than twice the experimentally determined binding energy for the water dimer (1105 ± 4 cm^{-1}, 16.4 kJ/mol)[50], which is calculated to be 1380 cm^{-1} at our level of theory. The strong binding results from the single water molecule using its two OH groups to H-bond to the two best binding sites of DPOE, the ether oxygen (best) and π cloud (second best).[50]

It is not surprising, then, that the 1:2 cluster retains this same primary binding motif, with the second water forming the water dimer by donating a H-bond to the first (Figure 3.6a). The 1:2 cluster affords the first opportunity to explore DPOE's perturbation on the free water clusters. A comparison of the OH···O H-bonded OH stretch fundamental in DPOE-$(H_2O)_2$ is at 3486 cm^{-1}, shifted down in frequency by ~45 cm^{-1} relative to the free water dimer.[51,52] This shift reflects a cooperative strengthening of the water-water H-bond in the presence of the π H-bond, which the calculations capture ($\Delta\omega$= -126 cm^{-1}). The free water dimer has a calculated OH···O hydrogen bond length of 1.91 Å, which is shortened in the presence of DPOE to 1.86 Å (Table 3.2).

In order to quantify the strength of the total binding of the $(H_2O)_n$ clusters to the DPOE 'solute', the energy difference for removing the $(H_2O)_n$ cluster from DPOE to form the *tgt* monomer plus an optimized $(H_2O)_n$ cluster, was calculated for each sized cluster.

$$DPOE-(H_2O)_n \rightarrow DPOE(tgt)+(H_2O)_n \qquad \Delta E \qquad (3.1)$$

The results are summarized in Table 3.3. In the 1:2 cluster, this solute-$(H_2O)_2$ binding energy is 3350 cm^{-1} (40.1 kJ/mol).

Table 3.3. Calculated cluster dissociation energies.

Structure	Dissociation Energy[a,b]	
	(kJ/mol)	*(cm^{-1})*
DPOE-H_2O[c]	30.3	2530
DPOE-$(H_2O)_2$	40.1	3350
A-DPOE-$(H_2O)_3$ Cycle	36.9	3080
B-DPOE-$(H_2O)_3$ Chain	41.7	3490
DPOE-$(H_2O)_4$	36.7	3090

[a]All calculations performed using M05-2x/6-31+G(d).
[b]Dissociation energies are BSSE and zero-point corrected using existing methods in refs. [37] and [38]. Dissociation energy reported is the energy to remove the entire water homocycle from complexation with DPOE. Optimized water structures were used in calculating the cluster dissociation energy.
[c]Values and structures taken from ref. [24].

DPOE-$(H_2O)_3$ is an intriguing case because two structures have been observed and characterized, one with a water trimer cycle attached to DPOE, and the other a trimer chain. As already noted, these two structures are nearly isoenergetic, but derive their stability in quite different ways. In the 1:3 chain, there are just two water-water H-bonds, compensated for by a cooperatively strengthened DPOE-H_2O interaction in which water retains its double-donor bridge, as in the 1:1 and 1:2 clusters. In addition, the DPOE monomer reorients one of its rings to accommodate an optimal OH$\cdots\pi$ bond with the middle water in the chain.

In the 1:3 cycle, formation of the trimer cycle increases the number of water-water H-bonds from two to three, but these H-bonds are relatively weak due to the highly strained nature of the trimer cycle. In forming the cycle, all water molecules already have one OH group tied up in the cycle, and therefore only a single OH available for

binding to DPOE. This qualitatively changes the interaction with DPOE, which is now accomplished by using one OH to bind to the ether pocket, while another of the three available OH groups binds to one of the aromatic ring π clouds. Interestingly, the net binding energy of the trimer cycle to DPOE is only slightly less than in the 1:2 and 1:3 chain structures. This is a simple consequence of the presence of the two separate interactions that together combine to reach a similar binding energy.

Binding to DPOE distorts the trimer cycle, affecting the H-bonded OH stretch transitions dramatically relative to their values in the free water trimer. In the free water trimer, the three OH···O bond lengths are nearly equivalent (1.88, 1.87 and 1.90 Å). These bond lengths are changed substantially in the 1:3 cycle to 2.00, 1.82, and 1.91 Å, for the OH(1)→OH(2), OH(2)→OH(3), and OH(3)→OH(1) hydrogen bonds, respectively (Table 3.2, see Figure 3.6 for water numbering). The formation of the OH···O(ether) bond turns H_2O(1) into a double-donor water, which weakens the OH···O(water) H-bond it makes in the cycle. These structural distortions partially localize and split the three H-bonded OH stretch transitions (Figure 3.9b), turning on intensity in all three fundamentals.

In the 1:4 cycle, the net binding energy of the tetramer cycle to DPOE is calculated to be very similar to that in the 1:3 cycle (3080 cm^{-1} for 1:3 and 3090 cm^{-1} for 1:4, Table 3.3). Not surprisingly, the 1:4 cycle also utilizes both the ether pocket and π clouds in binding to DPOE. OH stretch transitions due to the π and ether-bound OH groups are now in close proximity, with a near-optimal π H-bond formed. Binding to DPOE creates the same asymmetric perturbation of the cyclic structure as in the 1:3 cycle, causing the nominally square water tetramer (1.77 Å H-bond lengths) to deform

into a more rectangular shape (Table 3.2). The distortions are smaller than in the 1:3 cycle due to the increased strength of the water-water H-bonds in the 1:4 cycle. The water molecule involved in the OH···O(ether) H-bond has its hydrogen bond to water lengthened to 1.83 Å. Secondary binding of water(3) to the π cloud causes minor further distortions. The OH stretch transitions due to the H-bonded OH groups within the cycle are again quite distorted from those in the free water tetramer, with significant localization on individual OH groups, due in part to the nearly perpendicular orientations of adjacent OH groups in the cycle.

3.6 Conclusions

1,2-diphenoxyethane offers at least four binding sites to water molecules, the two ether oxygens and two π clouds of this flexible PEG-linked aromatic bichromophore. The present study of DPOE-$(H_2O)_n$ clusters with n=2-4 has led to firm assignments for four structures, a single 1:2, two 1:3, and one 1:4 cluster, based on their OH and CH stretch infrared spectra. The calculated binding energy of $(H_2O)_n$ to DPOE is far in excess of a single water-water H-bond in the water dimer, indicating that solute-solvent and solvent-solvent H-bonds are in competition with one another.

All structures, including the 1:1 complex, involve exclusive binding of $(H_2O)_n$ to the *tgt* conformer of DPOE. The *gauche* central dihedral angle places the four primary binding sites in relatively close proximity where the water molecules can interact optimally with more than one of these sites.

The primary binding site used by the 1:2 and 1:3 chain structures is similar to that in the 1:1 complex[24], involving asymmetric binding of a single water molecule that forms two H-bonds, one with the ether oxygen on one phenoxy group and the second to the π cloud of the other ring.

The competition between water-water and DPOE-water binding strengths is seen clearly in the 1:3 cluster, where two structures of similar stability are observed that incorporate cyclic and chain forms of the water trimer. Both structures contain a total of five H-bonds. The 1:3 chain structure has just two water-water H-bonds, but the terminal water molecule in the chain is thereby able to form two H-bonds to DPOE (one water-ether, and one OH···π), much as it does in the 1:1 and 1:2 clusters. The fifth H-bond is formed by the interior water in the trimer chain with the other aromatic ring (a second OH···π). The 1:3 cycle structure forms three water-water H-bonds and two of its otherwise free OH groups to bind to the ether oxygen(s) and π cloud of one ring. While this structure incorporates the optimal $(H_2O)_3$ structure, it also distorts it significantly, partially localizing the three OH stretch fundamentals in the cycle, and turning on intensity in all of them.

Finally, the 1:4 cluster is exclusively found with the water tetramer in a cycle, giving up strength in its interaction with DPOE in favor of forming four strong OH···O H-bonds in the cycle. Distortions in this cycle are apparent from the OH stretch spectrum, but to a lesser degree than in the 1:3 cluster.

This study serves as an important foundation for understanding the UV and IR spectroscopy of the close-lying S_1 and S_2 excited states of DPOE, which is taken up in the adjoining paper.

3.7 References

(1) Jha, A. K.; Freed, K. F. *J. Chem. Phys.* **2008**, *128*, 034501.

(2) Inomata, K.; Abe, A. *J. Phys. Chem.* **1992**, *96*, 7934-7937.

(3) Abe, A.; Inomata, K. *Journal of Molecular Structure* **1991**, *245*, 399-402.

(4) Borodin, O.; Smith, G. D. *J. Phys. Chem. B* **2003**, *107*, 6801-6812.

(5) Goutev, N.; Ohno, K.; Matsuura, H. *J. Phys. Chem. A* **2000**, *104*, 9226-9232.

(6) Yoshida, H.; Tanaka, T.; Matsuura, H. *Chemistry Letters* **1996**, *25*, 637-638.

(7) Yoshida, H.; Kaneko, I.; Matsuura, H.; Ogawa, Y.; Tasumi, M. *Chem. Phys. Lett.* **1992**, *196*, 601-606.

(8) Anderson, P. M.; Wilson, M. R. *Mol. Phys.* **2005**, *103*, 89-97.

(9) Chattoraj, M.; Paulson, B.; Shi, Y.; Closs, G. L.; Levy, D. H. *Journal of Physical Chemistry* **1993**, *97*, 13046-13051.

(10) Zehnacker, A.; Lahmani, F.; Desvergne, J. P.; Bouas-Laurent, H. *Chem. Phys. Lett.* **1998**, *293*, 357-365.

(11) Zehnacker, A.; Lahmani, F.; van Walree, C. A.; Jenneskens, L. W. *J. Phys. Chem. A* **2000**, *104*, 1377-1387.

(12) Lee, J. K.; Judge, R. H.; Boo, B. H.; Lim, E. C. *J. Chem. Phys.* **2002**, *116*, 8809-8816.

(13) East, A. L. L.; Cid-Aguero, P.; Liu, H. S.; Judge, R. H.; Lim, E. C. *J. Phys. Chem. A* **2000**, *104*, 1456-1460.

(14) Pillsbury, N. R.; Stearns, J. A.; Muller, C. W.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2008**, *129*, 114301.

(15) Pillsbury, N. R.; Muller, C. W.; Meerts, W. L.; Plusquellic, D. F.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *113*, 5000-5012.

(16) Pillsbury, N. R.; Kidwell, N. M.; Nebgen, B.; Slipchenko, L. V.; Douglass, K. O.; Cable, J. R.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2014**, *141*, 064316.

(17) Stearns, J. A.; Pillsbury, N. R.; Douglass, K. O.; Muller, C. W.; Zwier, T. S.; Plusquellic, D. F. *J. Chem. Phys.* **2008**, *129*, 224305-224318.

(18) Rodrigo, C. P.; Mueller, C. W.; Pillsbury, N. R.; James, W. H., III; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2011**, *134*, 164312.

(19) Ottiger, P.; Leutwyler, S. *Chimia* **2011**, *65*, 228-230.

(20) Ottiger, P.; Leutwyler, S. *J. Chem. Phys.* **2012**, *137*, 204303.

(21) Kopec, S.; Ottiger, P.; Leutwyler, S.; Köppel, H. *J. Chem. Phys.* **2012**, *137*, 184312.

(22) Buchanan, E. G.; Sibert, E. L.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 2800-2811.

(23) Buchanan, E. G.; Walsh, P. S.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2013**, *138*, 204313-204311.

(24) Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Phys. Chem. Lett.* **2013**, *4*, 1644-1648.

(25) Nebgen, B.; Emmert, F. L.; Slipchenko, L. V. *J. Chem. Phys.* **2012**, *137*, 084112.

(26) Lubman, D. M. *Mass Spectrom. Rev.* **1988**, *7*, 559-592.

(27) Page, R. H.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988**, *88*, 4621-4636.

(28) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205-241.

(29) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827-8839.

(30) Walsh, P. S.; Kusaka, R.; Buchanan, E. G.; James, W. H.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 12350-12362.

(31) Dean, J. C.; Buchanan, E. G.; James, W. H., III; Gutberlet, A.; Biswas, B.; Ramachandran, P. V.; Zwier, T. S. *J. Phys. Chem. A* **2011**, *115*, 8464-8478.

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

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

(34) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440-467.

(35) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J.V.; Cioslowski, J.; Fox, D. J.; Revision C.01 ed.; Gaussian, Inc.: Wallingford CT, 2009.

(36) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289-300.

(37) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553-566.

(38) van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. *Chem. Rev.* **1994**, *94*, 1873-1885.

(39) Pribble, R. N.; Zwier, T. S. *Science* **1994**, *265*, 75-79.

(40) Gotch, A. J.; Zwier, T. S. *J. Chem. Phys.* **1992**, *96*, 3388-3401.

(41) Buchanan, E. G.; Zwier, T. S. *J. Phys. Chem. A* **2014**, *118*, 8583-8596.

(42) Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015**, *142*, 154304.

(43) Shubert, V. A.; Müller, C. W.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *113*, 8067-8079.

(44) Kusaka, R.; Inokuchi, Y.; Ebata, T. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4452-4459.

(45) Kusaka, R.; Inokuchi, Y.; Ebata, T. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6238-6244.

(46) Kusaka, R.; Inokuchi, Y.; Ebata, T. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9132-9140.

(47) Gierszal, K. P.; Davis, J. G.; Hands, M. D.; Wilcox, D. S.; Slipchenko, L. V.; Ben-Amotz, D. *J. Phys. Chem. Lett.* **2011**, *2*, 2930-2933.

(48) Pribble, R. N.; Zwier, T. S. *Faraday Discuss.* **1994**, *97*, 229-241.

(49) Pugliano, N.; Saykally, R. J. *Science* **1992**, *257*, 1937-1940.

(50) Rocher-Casterline, B. E.; Ch'ng, L. C.; Mollner, A. K.; Reisler, H. *J. Chem. Phys.* **2011**, *134*, 211101.

(51) Coker, D. F.; Miller, R. E.; Watts, R. O. *J. Chem. Phys.* **1985**, *82*, 3554-3562.

(52) Huang, Z. S.; Miller, R. E. *J. Chem. Phys.* **1989**, *91*, 6613-6631.

CHAPTER 4. SOLVENT-MEDIATED INTERNAL CONVERSION IN DIPHENOXYETHANE-$(H_2O)_N$ CLUSTERS, N=2-4

4.1 Introduction

Kasha's rule stands as one of the cornerstones of solution phase photochemistry[1-4], summarizing the common observation that, for many molecules, independent of the excitation wavelength used, emission occurs primarily from the first excited electronic state. The rule arises because, in many circumstances, the time scale for internal conversion from high lying electronic states $\{S_n\}$ to the lowest excited state of the same multiplicity, S_1, is much shorter than the timescale for fluorescence itself. In solution, vibrational relaxation within a given electronic state is also very fast (ps timescale), leading to a pile-up of thermalized population in the lowest excited state, from which emission occurs. More than 40 years ago, Bixon and Jortner provided a theoretical description of such radiationless transitions.[5,6] Their expression for the rate of internal conversion showed clearly that the small energy difference between the manifold of excited singlet electronic states provides a mechanism for fast internal conversion down to S_1, where fluorescence can compete with the slower internal conversion to S_0 that results from the large energy gap between S_1 and S_0.

In the simplest case in which the molecule contains a single electronic chromophore, internal conversion occurs between states of the same multiplicity localized on this chromophore, where the geometry changes and electronic energy gaps

are those of that single chromophore. When more than one chromophore is present, either as separate parts of the same molecule, or as separate molecules within the sample, electronic energy transfer between chromophores is possible, leading to fascinating and important new possibilities. Electronic energy transfer between two chromophores, often referred to as Förster resonant energy transfer (FRET), is widely used by the biophysics community to probe the time-dependent changes in distance between donor and acceptor chromophores, with the rate of electronic energy transfer sensitively dependent on the distance between them, their electronic energy gap, and relative orientation.[7-9] In the context of light harvesting and photosynthesis, resonant energy transfer occurs within an array of chromophores in close spatial proximity, funneling electronic excitation from its excitation point to a distant reaction center, where that energy can be used to generate charge separation. Recent advances in 2D electronic spectroscopy have provided a means of probing these energy transfer mechanisms on the femtosecond timescales, leading to qualitative new insights to the pathways for electronic excitation and the possible role of coherences in electronic energy transport.[10-15] Several recent studies have sought to determine whether these coherences play an integral role in the energy transfer process, and to what extent particular vibrational modes will facilitate the energy transfer.[16,17]

Most biologically relevant internal conversion or electronic energy transfer processes occur in aqueous solution at room temperature, where competing intramolecular and inter-chromophore photophysical pathways are modified or possibly even re-directed by interactions with aqueous solvent. This has led several groups to study the laser spectroscopy of coupled chromophores in the form of both molecular

dimers and as covalently linked bichromophores, prepared under jet-cooled, isolated conditions.[18-27] These experiments enable the in-depth study of the close lying excited states and the underlying vibronic coupling on a state-specific basis devoid of solvent effects. The results serve as a testing ground for emerging theories of excitonic and vibronic coupling.[28-34]

Following this bottom up approach, our group has recently completed a study of this type on a prototypical weakly coupled bichromophore, 1,2-diphenoxyethane (DPOE), whose structure is shown in Figure 4.1.[35-37]

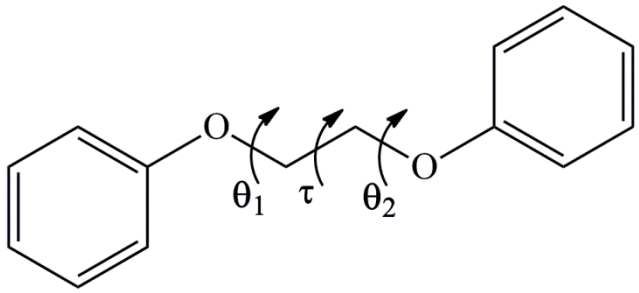


Figure 4.1. Chemical structure of 1,2-diphenoxyethane (DPOE) with important dihedral angles labeled.

There are two conformations (*ttt* and *tgt*) of DPOE present under jet cooled conditions in the gas phase, the latter of which is shown in Figure 4.2a.[35] Using a combination of vibronic spectroscopy, mass-resolved stimulated emission pumping (MR-SEP), and rotationally-resolved high-resolution ultraviolet spectroscopy, it was determined that the two excited states of both conformations are very weakly coupled, leading to an excitonic splitting of only ~1 cm^{-1} (Figure 4.2a).[36] Both the vibronic

spectroscopy and the direction of the transition dipole moments of the S_0-S_1 and S_0-S_2 origin transitions prove that the two states are electronically delocalized over the two aromatic rings, producing excited states that are simple sums and differences of the two locally excited states. The 1.0 cm^{-1} splitting is best interpreted in an adiabatic picture as a tunneling splitting on the lower adiabatic potential energy surface, as shown schematically at the bottom of Figure 4.2a, with the magnitude of the splitting determined by the nuclear reconfiguration and interchromophore electronic coupling that accompanies tunneling between |A*B> and |AB*>.

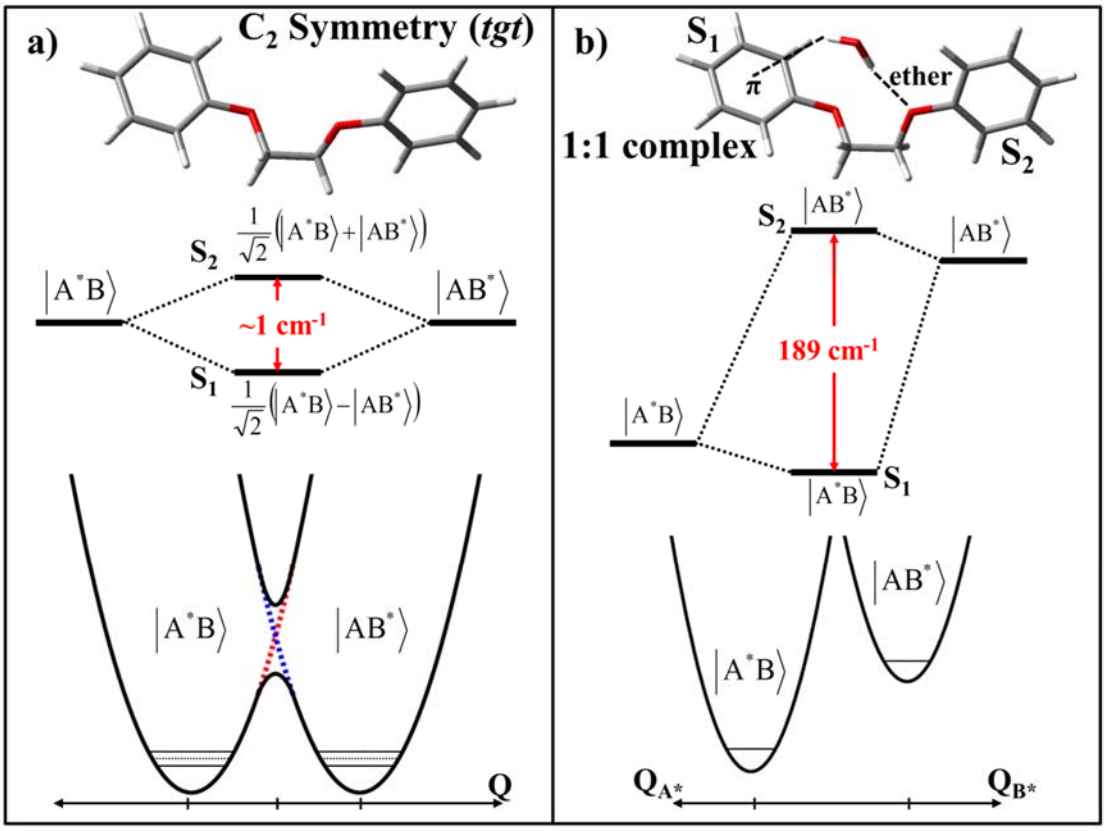


Figure 4.2. Structure, energy level diagram, and 1D potential energy curves for the S_1 and S_2 states of (a) the *tgt* conformer of the DPOE monomer, and (b) the DPOE-H_2O complex. Energy level diagrams and potential energy curves are not to scale.

The small S_1/S_2 splitting present in DPOE monomer makes it an interesting testing ground for understanding the effects of individual solvent molecules on the two close lying electronic excited states and the coupling between them. For such studies, the supersonic expansion is used not only to cool the DPOE monomer, but to prepare DPOE-$(H_2O)_n$ molecular clusters that can be spectroscopically interrogated on a size- and conformation-specific basis. The purpose of this chapter is to probe the effect of stepwise solvation on the excited states of this weakly coupled bichromophore by water, focusing primarily on the DPOE-$(H_2O)_n$ (n=2-4) clusters.

This work builds directly on the previously published results on the DPOE-H_2O complex (n=1).[37] The binding of a single water molecule to the DPOE monomer induces a site splitting of 189 cm^{-1} between the two excited states that is produced by the asymmetric binding of the water molecule to the two electronic chromophores (Figure 4.2b). In DPOE-H_2O, the water binds to the *tgt* monomer via a double-donor interaction, serving as a bridge between the two phenoxy chromophores, with one OH group forming an OH···π hydrogen bond with one aromatic ring, and the other an OH···O(ether) hydrogen bond to the phenoxy oxygen of the other chromophore of DPOE, as shown in Figure 4.2b.[37] This asymmetric binding localizes the electronic excitation in one or the other ring in the S_1 (|A*B>) and S_2 (|AB*>) states. It was possible to identify which ring was responsible for each excited state by studying the excited state resonant ion-dip infrared (ES-RIDIR) spectra. Furthermore, the upper level accessed by excitation of the S_2 origin bears striking evidence for its mixed electronic character both in its emission and in its infrared spectrum, due to coupling of the S_2 zero-point level with vibronic levels associated with the S_1 state.[37]

Here, we report and analyze similar spectra recorded for the DPOE-$(H_2O)_n$ clusters with n=2-4, in order to understand the electronic state changes associated with increased solvation. In Chapter 3, we focused attention on the spectroscopic characterization of the ground state structures of the clusters, using ground state resonant ion-dip infrared (RIDIR) spectroscopy in the OH- and CH-stretch regions to make structural assignments based on comparison with the predictions of calculations.[38]

The assigned structures are presented in Figure 4.3. Distances (in Å) of importance in interpreting the excited state properties are included in the figure, taken from the ground state optimized structures. The IR spectra provide evidence that all of the observed clusters utilize the same *tgt* conformation as the DPOE-H_2O complex, thereby forming a pocket in which the water molecules can form H-bonds with the various H-bonding sites on DPOE (ether oxygens and aromatic π cloud). In all but one of the observed structures, the water clusters adopt structures similar to those of the free water clusters, forming a water dimer (Figure 4.3a), and cyclic water trimer (Figure 4.3c) and tetramer (Figure 4.3d) bound to DPOE. In addition, a 1:3 chain structure (Figure 4.3b) is formed that builds off the 1:1 and 1:2 structures by adding a third molecule onto the back end of the growing chain.

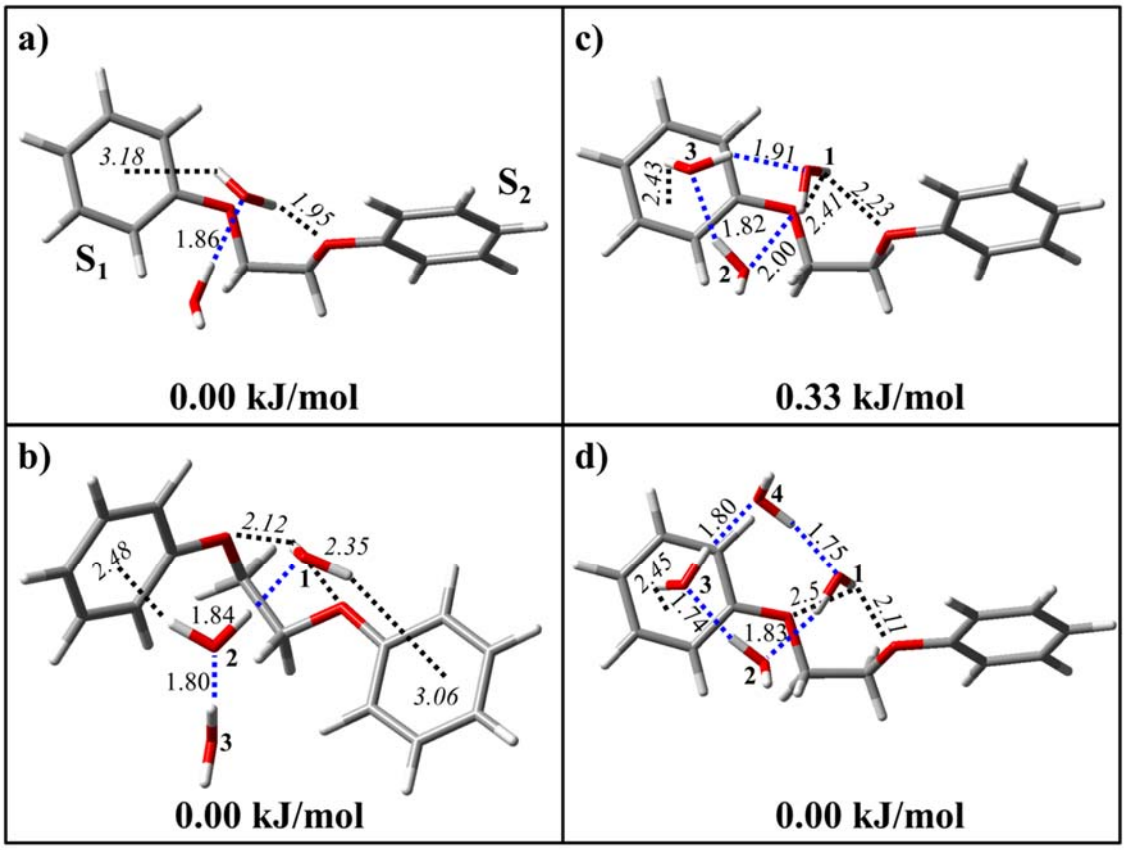


Figure 4.3. Assigned conformations of the (a) [1:2], (b) [1:3] chain, (c) [1:3] cycle, and (d) [1:4] clusters. Hydrogen bond distances for each of the clusters are provided for both the water-water (blue dotted lines, normal numbers) and DPOE-water (black dotted lines, italic numbers). Relative energies of the clusters are also provided.

In what follows, we focus attention on the consequences of each solvent-bound 'architecture' for the two close-lying excited states of DPOE. We shall see that the energy splitting between the S_1 and S_2 states is modulated by the changing way in which the water clusters interact with the DPOE molecule as a function of cluster size. In the 1:2 and 1:3 cyclic structures, the nature of these excited states is once again probed via ES-RIDIR spectroscopy, showing evidence for a molecular-scale version of Kasha's rule. The role played by the water clusters in localizing electronic excitation, shifting the

electronic states relative to one another, and inducing coupling between them will be considered.

4.2 Experimental Methods

A full description of the experimental apparatus and the methods for generating water clusters has been described previously in Chapters 2 and 3.[37-43] Chapter 3 provides a detailed explanation of the ground state conformational assignments using RIDIR spectroscopy. RIDIR spectra of the excited states of the DPOE-$(H_2O)_n$ clusters was accomplished using the method of excited state RIDIR (ES RIDIR) spectroscopy which is described extensively in Chapter 2 and the in literature.[37,44] Briefly, λ_1 (<~0.3 mJ/pulse) is fixed to a vibronic transition that is specific to a particular cluster size and isomeric form while λ_2 (~277-279 nm, ~1 mJ/pulse) was used to ionize the molecules in a 2C-R2PI scheme. λ_1 and λ_2 were separated by ~23 ns to facilitate the recording of ES RIDIR spectra which were recorded in the OH-stretching region (~3100-3750 cm^{-1}).

When IR absorption occurs from the excited state level accessed by λ_1, it leads to depletion in the ion signal created by λ_2, due to fast non-radiative processes at these energies (~3500 cm^{-1}) above the S_1 origin. Cluster fragmentation is one of the possible excited state processes, but is not required for the method to work. Since the ion signal is size- and isomer-selective, the IR depletions recorded are also specific to this same species.

4.3 Computational Methods

The computational process used to predict the ground state DPOE-$(H_2O)_n$ cluster structures is available in Chapter 3.[38] Only calculations important to the excited state analysis are described here. Density functional theory (DFT) calculations were carried out using GAUSSIAN09[45] and Truhlar's hybrid density functional[46] M05-2X with a 6-31+G(d) basis set. This functional incorporates dispersion, and has demonstrated accuracy in the prediction of vibrational frequencies and IR intensities.[47,48] Vertical excitation energies were calculated using time-dependent density functional theory (TD-DFT) and scaled such that the calculate $S_0 \rightarrow S_1$ origin transition in DPOE monomer matched experiment (0.835).

4.4 Results and Analysis

4.4.1 R2PI and Holeburning Spectra

The R2PI spectra and overview hole-burning spectra for DPOE-$(H_2O)_n$ clusters were presented first in Chapter 3.[38] Figures 4.4–4.6 show these spectra in an alternate format that displays more clearly the transitions ascribed to the S_0-S_1 and S_0-S_2 states which are the principal focus of the present work. R2PI spectra were recorded in the $[1:1]^+$, $[1:2]^+$, and $[1:3]^+$ mass channels (m/z=232, 250, and 268, respectively). Table 4.1 summarizes the UV transition frequencies, experimentally determined excited state splitting, and the predicted vertical excited state splitting from TD-DFT calculations.

These results are summarized pictorially in Figures 4.7–4.10 with the calculated molecular orbitals associated with the transitions also provided.

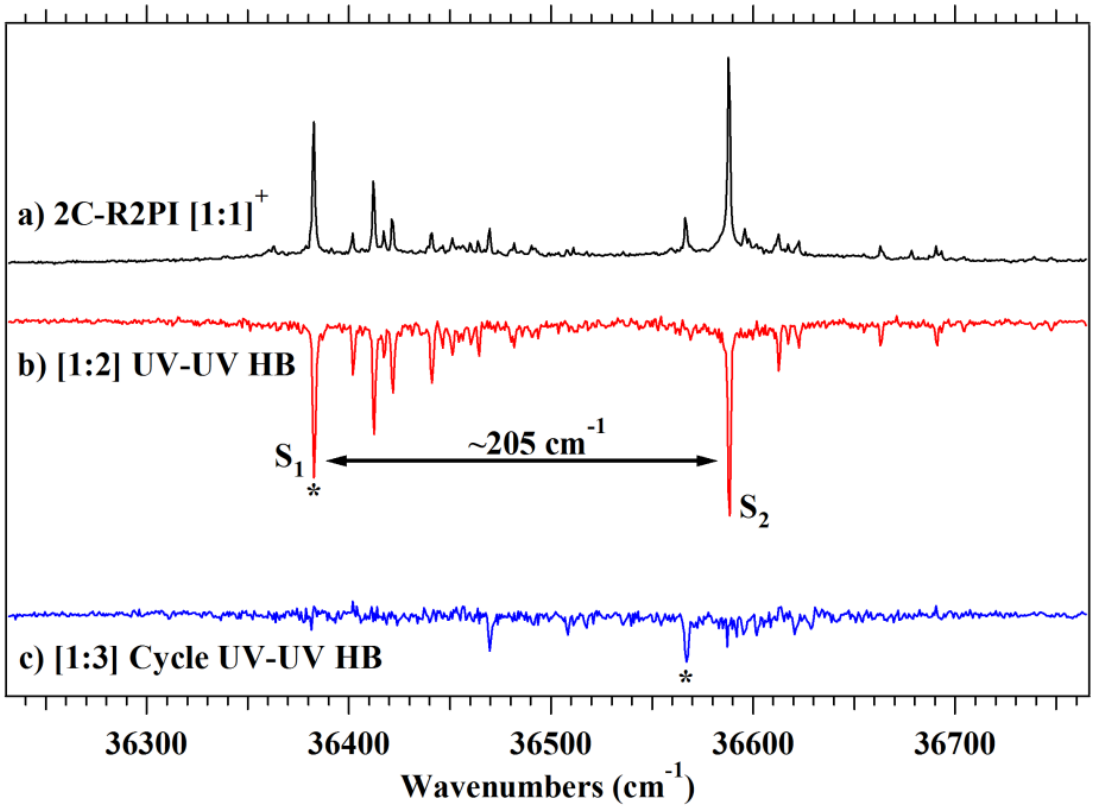


Figure 4.4. (a) 2C-R2PI spectrum recorded in the $[1:1]^+$ mass channel (m/z=232). (b, c): UV-UV holeburning spectra revealing the single [1:2] cluster (b) and the major [1:3] cycle cluster (conformer A, c) which appear in the $[1:1]^+$ mass channel due to efficient photofragmentation. Starred transitions were used to record the UV-UV holeburning spectra. Nominal S_1 and S_2 origin transitions are labeled along with the excited state splitting.

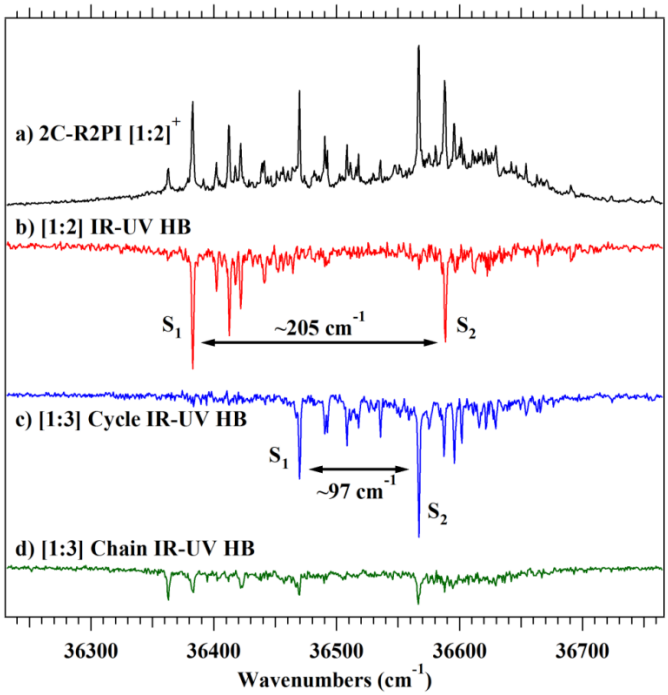


Figure 4.5. (a) 2C-R2PI spectrum recorded in the $[1:2]^+$ mass channel (m/z=250). (b-d): IR-UV holeburning spectra for the [1:2] cluster (b), major [1:3] cycle cluster (conformer A, c), and minor [1:3] chain cluster (conformer B, d). Nominal S_1 and S_2 origin transitions are indicated along with the excited state splitting.

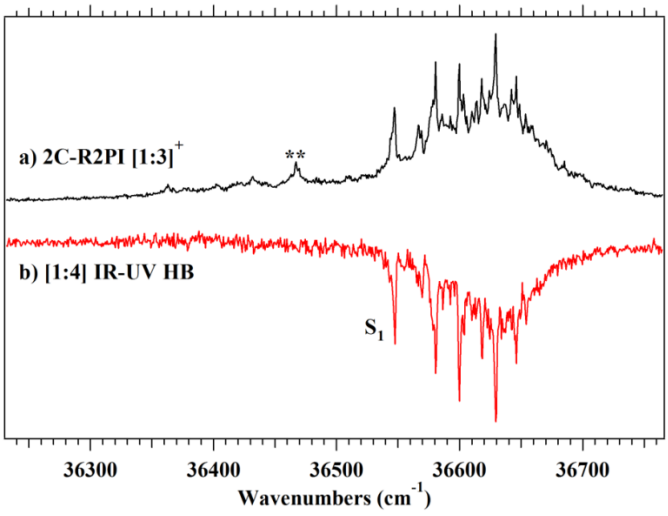


Figure 4.6. (a) 2C-R2PI spectrum recorded in the $[1:3]^+$ mass channel (m/z=268). (b) IR-UV holeburning spectrum of the [1:4] cluster. The double starred transition (**) does not burn with the main conformer. Due to low signal intensity, the carrier of this transition was not pursued further.

The hole-burning spectrum for the 1:2 cluster (UV-UV in Figure 4.4b and IR-UV in Figure 4.5b) has two dominant transitions which appear at 36383 and 36589 cm^{-1}, respectively, with unique low-frequency vibronic activity built off of each of them (Table 4.1). The hole-burning spectra prove that these transitions are not due to two different conformations, but that both belong to the same 1:2 ground state structure. The similar splitting (205 cm^{-1}) to that determined for the 1:1 complex (189 cm^{-1}), points to these two transitions as the S_0-S_1 and S_0-S_2 origins, respectively. The S_1 origin is red-shifted by -40 cm^{-1} from the *tgt* monomer origins.[36,37] The S_1 origin shows low-frequency vibronic activity in several low-frequency modes involving large amplitude motions of the phenoxy rings, with close analogs in the *tgt* monomer.[36,37] Note that the Franck-Condon activity built off the S_1 origin is considerably more intense than that built off the S_2 origin (Figure 4.4b), suggesting a stronger interaction of the water molecules in the S_1 state than the S_2 state with the aromatic ring π cloud involved in the $\pi\pi^*$ transitions.

Table 4.1. Summary of experimental and theoretical electronic transition frequencies and splittings, and important structural parameters for the DPOE-$(H_2O)_n$ (n=0-4).

Structure	Origin (cm^{-1})		Exp. Splitting (cm^{-1})	Calc. Splitting (cm^{-1})[a]	Shift from *tgt* Monomer	Dihedral Angle (°)	Hydrogen Bond Length (Å)	
	S_1	S_2	S_1-S_2	S_1-S_2	S_0-S_1 Origin (cm^{-1})	(θ_1, τ, θ_2)	(π_1/π_2)	($Ether_1/Ether_2$)
DPOE (tgt)[b]	36422.91	36423.93	1.02	9	0	(-177.9, +71.6,-177.9)		
DPOE-H_2O[b]	36471	36660	189	300.3	+48	(-170.8, +66.4,-177.5)	3.27	2.01
DPOE-$(H_2O)_2$	36383	36588	205	276.5	-40	(-170.9, +66.9,-179.4)	3.18	1.95
DPOE-$(H_2O)_3$ A-Cycle	36470	36567	97	194.7	+47	(-175.5, +65.7,-179.8)	2.43	2.23/2.41
DPOE-$(H_2O)_3$ B-Chain	36362			207	-61	(-88.3, -69.1,+169.9)	3.06/2.48	2.12/2.35
DPOE-$(H_2O)_4$	36547			166.5	+124	(+176.1, +64.7,+177.9)	2.45	2.11/2.5

[a] Vertical excited state splittings at the optimized ground state geometries, at the TDDFT M052x/6-31+G(d) level of theory. Calculated excitation frequencies were scaled by 0.835.

[b]From refs. [35-37].

The S_0-S_1 origin transition for the 1:3 conformer A (Figures 4.4c and 4.5c) assigned to the 1:3 cycle, appears at 36470 cm^{-1}, blue-shifted by +47 cm^{-1} from the *tgt* monomer origin, and 87 cm^{-1} above the corresponding S_1 origin of the 1:2 cluster.[36,37] The low-frequency vibronic activity built off the S_1 origin in 1:3(A) is again quite similar to that in the DPOE *tgt* monomer and smaller water-containing clusters. The second strong transition appears just 97 cm^{-1} above the S_0-S_1 origin (36576 cm^{-1}). This transition is not part of a vibronic progression, but has unique low-frequency vibronic activity that marks it as the S_0-S_2 origin of 1:3(A). The S_1/S_2 splitting of 1:3(A) is thus less than half the splitting in the 1:2 cluster. In addition, this second origin shows markedly more vibronic activity than the S_0-S_2 origin in the 1:2 cluster. Both the small splitting and change in S_2 vibronic activity suggest a change in the environment around the S_2 chromophore.

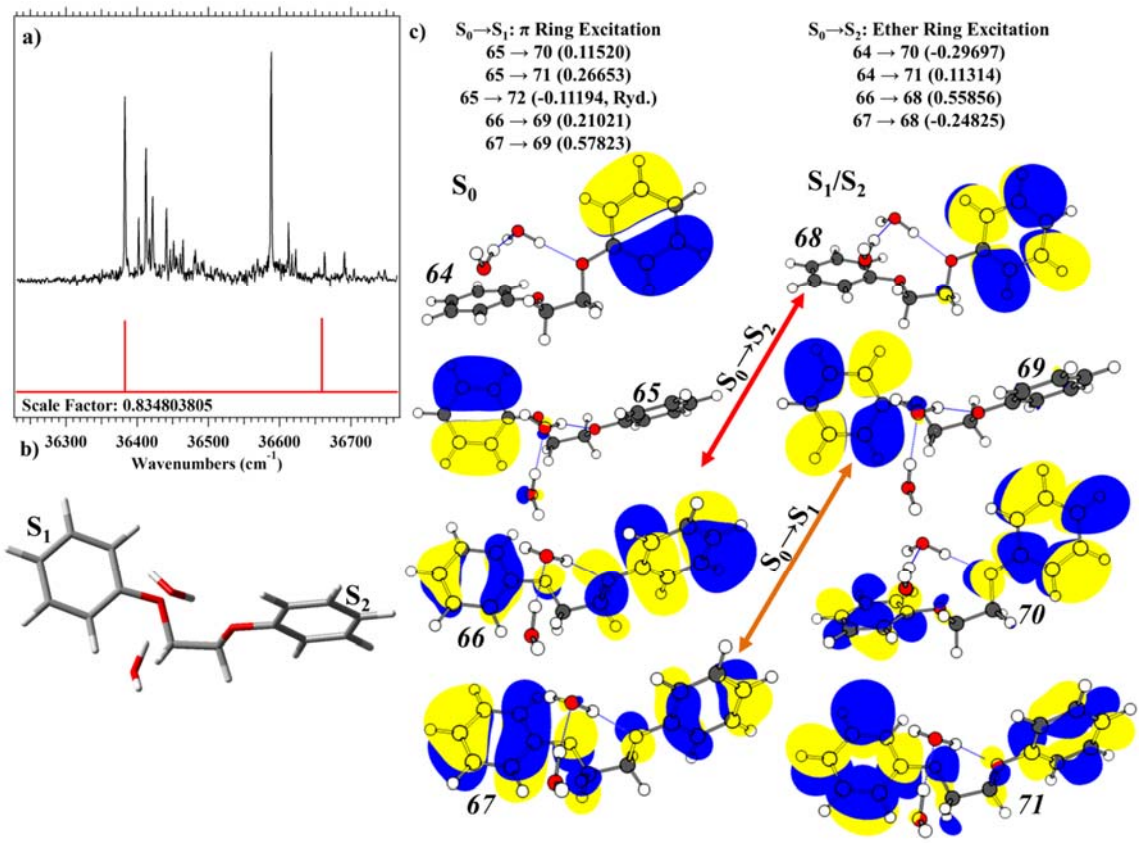


Figure 4.7. The calculated UV spectrum for the [1:2] cluster (a), along with the structure of the [1:2] cluster (b). The calculated weights and molecular orbitals for excitation to S_1 and S_2.

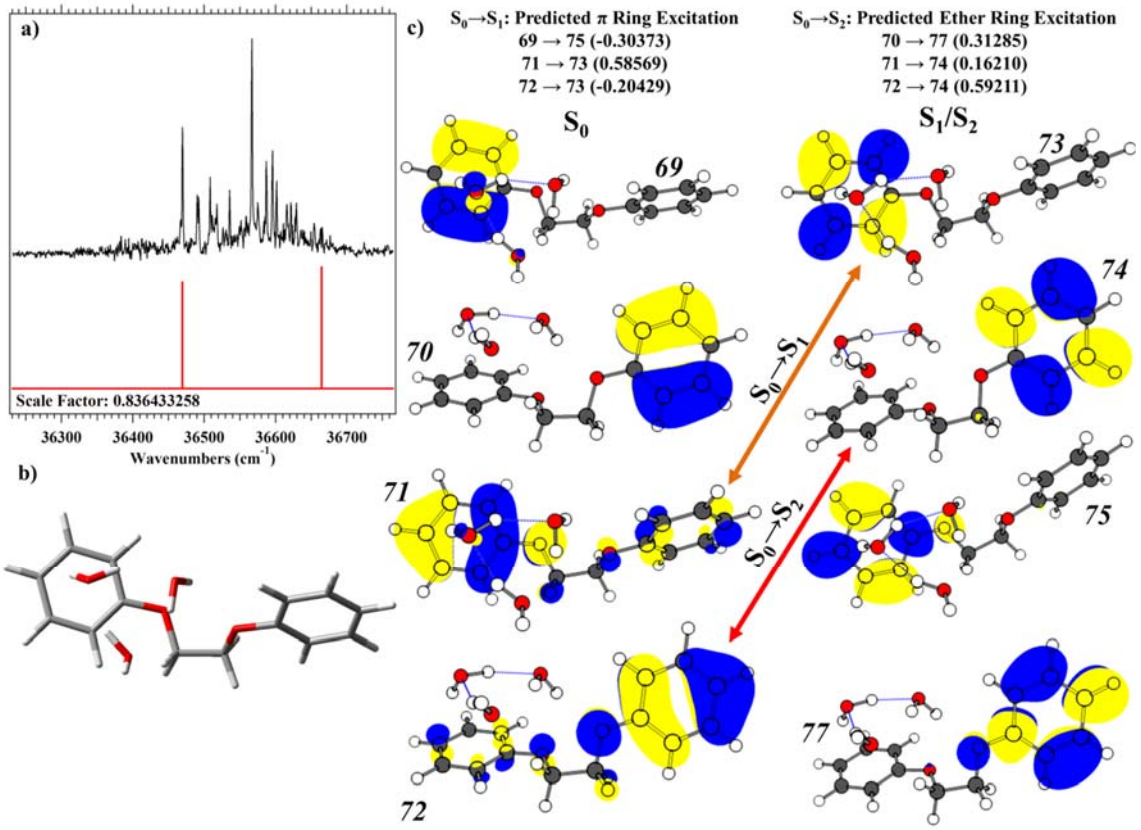


Figure 4.8. The calculated UV spectrum for the [1:3] cycle cluster (a), along with the structure of the [1:3] cycle cluster (b). The calculated weights and molecular orbitals for excitation to S_1 and S_2.

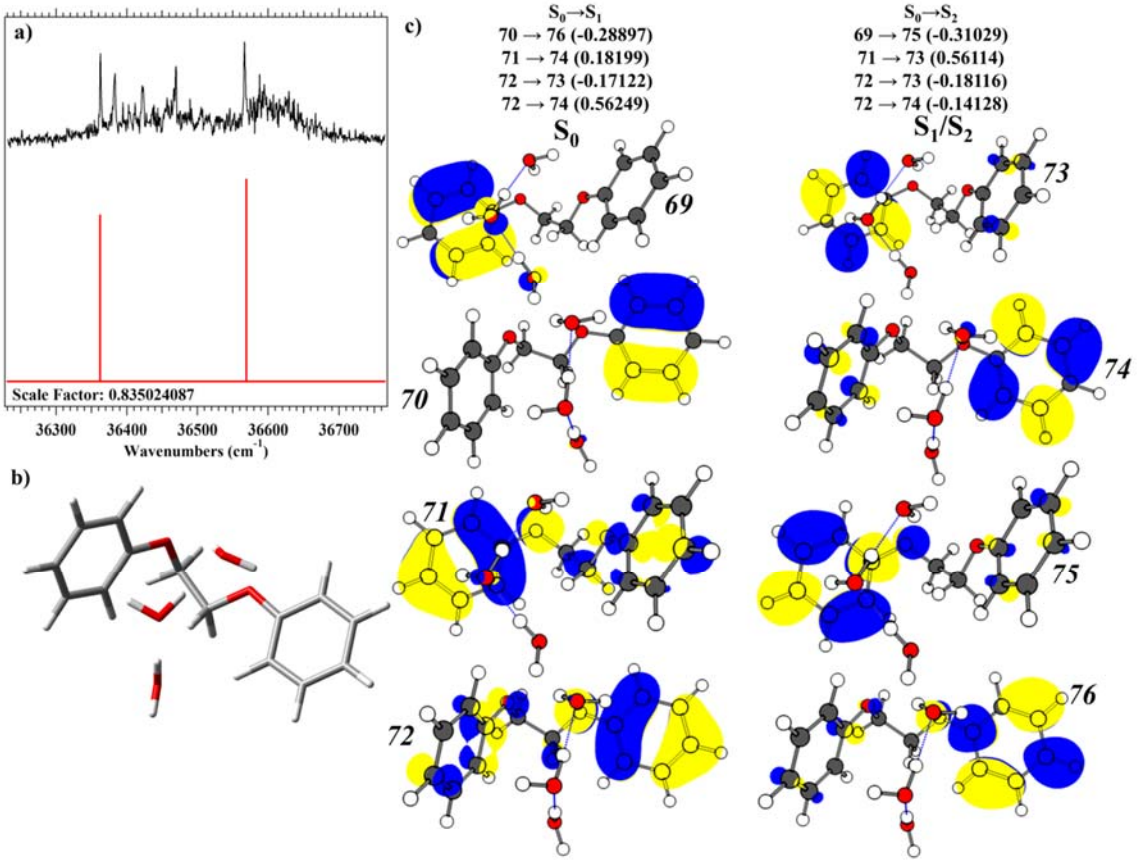


Figure 4.9. The calculated UV spectrum for the [1:3] chain cluster (a), along with the structure of the [1:3] chain cluster (b). The calculated weights and molecular orbitals for excitation to S_1 and S_2.

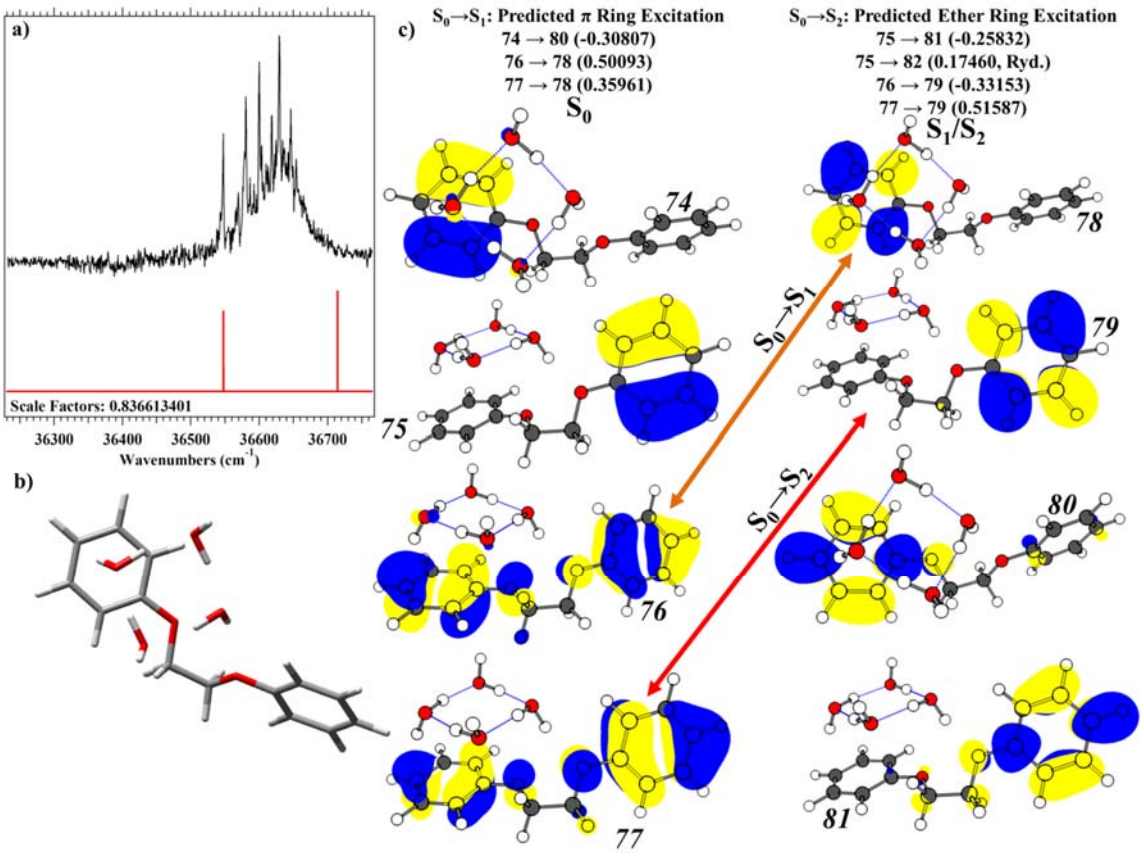


Figure 4.10. The calculated UV spectrum for the [1:4] cluster (a), along with the structure of the [1:4] cluster (b). The calculated weights and molecular orbitals for excitation to S_1 and S_2.

The UV spectrum assigned to 1:3 conformer B (Figure 4.5d) appears -61 cm^{-1} from the DPOE *tgt* monomer origin, close to that for the 1:2 structure, and more than 100 cm^{-1} red of the S_1 origin of the 1:3 cycle. Due to low signal intensity and the presence of large transitions due to 1:3(A) in the same region, it was difficult to assign other vibronic transitions due to 1:3(B), nor to locate the S_2 origin with certainty.

The UV spectrum of the 1:4 cluster (Figure 4.6b) is markedly different than those observed for the 1:1, 1:2, and 1:3 clusters, most notably in not dividing in to well-defined regions due to the S_1 and S_2 origins, but instead coalescing into a single dense clump of vibronic transitions. The S_0-S_1 origin is tentatively assigned to the transition at 36547 cm^{-1}, marked by an asterisk in the hole-burn spectrum. Its frequency is blue-shifted from the *tgt* monomer origin by +124 cm^{-1}, marking it as the furthest blue of any member of the 1:n series. Progressions in low frequency modes of 19 and 33 cm^{-1} are observed. It is likely that the S_2 origin of the 1:4 cluster is also amidst this same set of vibronic transitions, most likely in the largest intensity transition at +82 cm^{-1}, making the S_1/S_2 splitting smallest of the clusters studied.

4.4.2 Excited State IR Spectra

One of the principal motivations of the present work is to provide a direct link between a particular solvent structure bound to a symmetric bichromophore and its effect on the nature, energetics, and coupling between the two closely-spaced electronically excited states (S_1/S_2). Excited state RIDIR spectroscopy provides a unique and powerful probe of these effects through the changes induced by electronic excitation on the OH stretch spectra of the water molecules in the DPOE-$(H_2O)_n$ cluster.

4.4.2.1 DPOE-$(H_2O)_2$

The ES-RIDIR spectra following excitation of the S_1 and S_2 origins of the 1:2 cluster are shown as blue and red traces, respectively, in Figure 4.11b. For direct comparison, the ground state RIDIR spectrum is included below the others (black trace).

Also included for comparison are the corresponding spectra of the 1:1 complex (Figure 4.11a), which have been reported previously.[37]

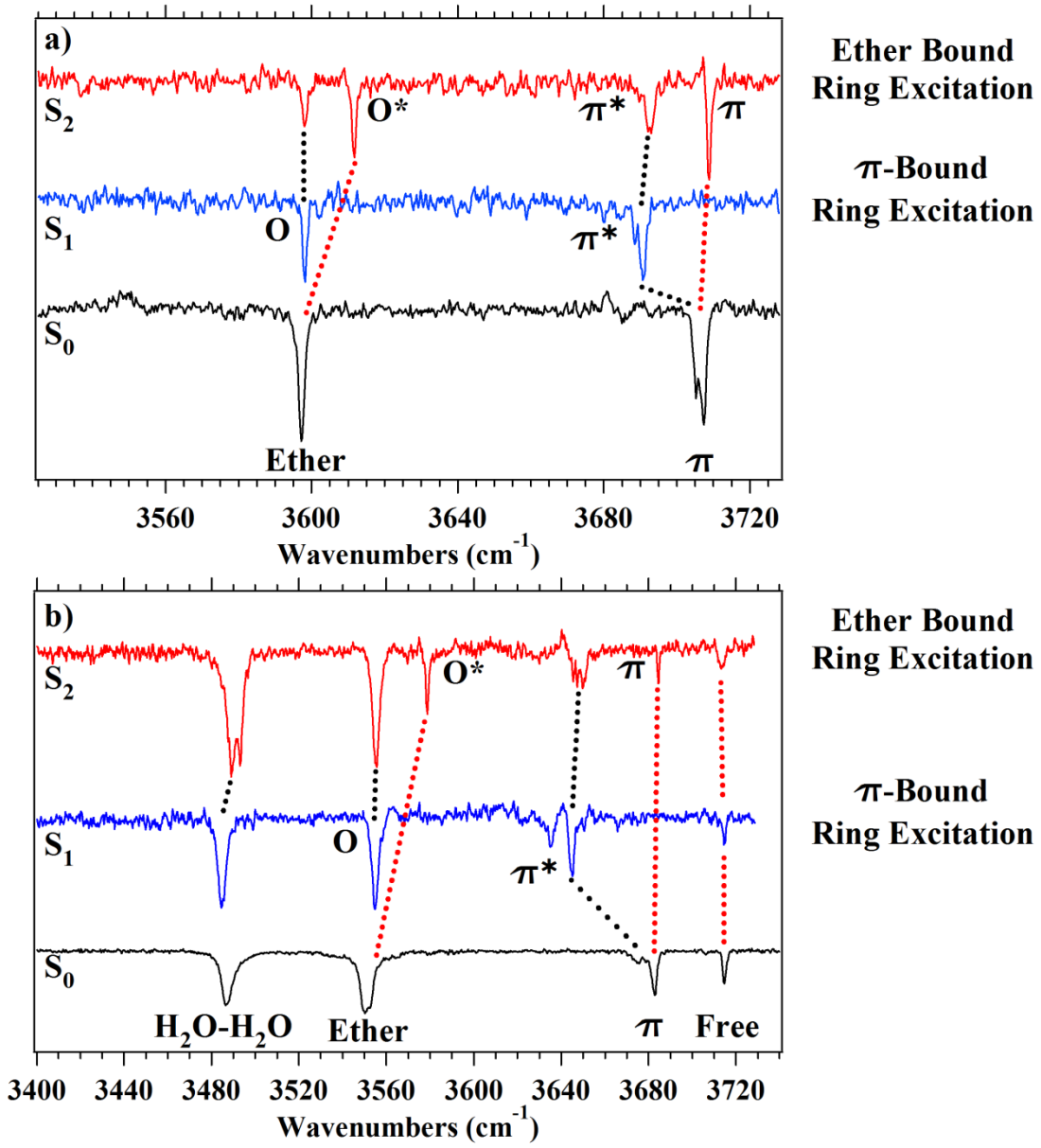


Figure 4.11. Comparison of the OH-stretch RIDIR spectra recorded in the S_0 (black trace), S_1 (blue trace), and S_2 (red trace) electronic states for the (a) [1:1] complex, and (b) [1:2] cluster. The spectra for the [1:1] complex are taken from ref. [37].

Comparison of the S_0 and S_1-state RIDIR spectra of the 1:2 cluster show shifts of less than five cm^{-1} in three of the four OH stretch transitions: the free OH stretch (3715 S_0 /3714 S_1), OH···O (ether) OH stretch (3550/3555 cm^{-1}), and water-water H-bonded OH stretch (3486/3485 cm^{-1}). By contrast, the OH···π hydrogen bonded transition shifts to lower frequency by -38 cm^{-1}, from 3683 cm^{-1} to 3645 cm^{-1}. The selective shift of the π-bound OH stretch (labeled π*) marks the S_1 state as being localized primarily on the π-bound ring, as indicated in Figure 4.3a. Table 4.2 summarizes these frequencies.

Table 4.2. Vibrational frequencies from the experimental RIDIR spectra in S_0, S_1, and S_2 electronic states for DPOE-$(H_2O)_n$ (n=1-3). S_2 ES-RIDIR spectra show mixed character due to Kasha's and Fermi's Golden rule.

Structure/State		Vibrational Frequency (cm^{-1})					
		Free	OH···O, Ether	OH···π	(1)→(2)[c]	(2)→(3)[c]	(3)→(1)[c]
DPOE-H_2O[a,b]	Water Label		OH(1)-Sym.	OH(1)-Asym.			
	S_0		3597	3707/ 3705			
	S_1		3598	3691			
	$S_2(S_1)$		3612 (3598)	3709 (3693)			
DPOE-$(H_2O)_2$[b]	Water Label	OH(2)	OH(1)-Sym.	OH(1)-Asym.	OH(2)		
	S_0	3715	3550	3683	3486		
	S_1	3714	3555	3645/ 3635	3485		
	$S_2(S_1)$	3713 (3714)	3579 (3555)	3684 (3650)	3493 (3489)		
A-DPOE-$(H_2O)_3$[b]	Water Label	OH(2)	OH(1)-Asym.	OH(3)-Asym.	OH(1)-Sym.	OH(2)-Sym.	OH(3)-Sym.
	S_0	3706	3641	3653	3563	3384	3486
	S_1	3704	3646	3652	3560	3380	3486
	$S_2(S_1)$	3704	3653	3661	3563	3382	3486

[a]Values taken from Ref. [37] .
[b]S_1 transitions appearing in S_2 are shown in parenthesis.
[c]Water numbering scheme from Figure 4.3.

The ES-RIDIR spectrum of the S_2 origin (Figure 4.11b, red) is striking in that it contains at least six resolved transitions, rather than the four anticipated for the 1:2 cluster, with its four OH groups. In order to understand this spectrum, it is best first to briefly reconsider the corresponding S_2-state RIDIR spectrum of the 1:1 complex (Figure 4.11a).[37] In the 1:1 complex, there are clearly four OH stretch fundamentals rather than the expected two. The frequencies of two of these transitions, labeled π* and O, are near those found in the S_1-state spectrum below it. The labeling scheme emphasizes that in the S_1 state it is the π-bound OH stretch that shifts relative to the ground state. The remaining two observed transitions are shifted in frequency, with the "extra" transition in the ether-bound region (labeled O*) shifted up in frequency, and the additional transition in the π-bound region (labeled π) very near the ground state frequency of the π-bound OH group. Based on this and other evidence, it is clear that the (nominal) S_2 origin has mixed electronic character, with excited state wave function that is an admixture of $S_2|0^0>$ and $S_1|\{v\}>$ character. Thus, in the 1:1 complex, the OH stretch spectrum of the water molecule attached to DPOE reports in exquisite fashion on the mixed electronic character of the S_2 origin, which is imbedded in $S_1(\{v\})$ levels with which it is vibronically coupled and mixed.

Against this backdrop, the interpretation of the "extra" transitions in the ES-RIDIR spectrum of the S_2 origin of the 1:2 cluster is clear. Recall first that the 1:2 cluster has the same primary binding site as 1:1, with the first water molecule bound as a double-donor bridge, with one OH bound to the ether oxygen of one ring, and the other the π cloud of the opposite ring. Four of the observed transitions appear at frequencies close to the S_1 origin spectrum below it. These transitions, which are connected to their S_1

counterparts by black dotted lines, are due to the $S_1|\{v\}>$ character of the excited state wave function. Two others, labeled O* and π, arise from the $S_2|0^0>$ portion of the wave function. The ether-bound OH (O*) is shifted up in frequency by +29 cm^{-1}, marking the S_2 state as the O-bound ring, while the π-bound OH stretch transition is close to its ground-state counterpart, consistent with localized electronic excitation of the O-bound ring.

Strictly speaking, there should be a total of eight transitions in the spectrum, four for the $S_2|0^0>$ character of the wave function and four for the $S_1|\{v\}>$ character. Indeed, careful inspection does show that even the water-water H-bonded OH stretch is a partially resolved doublet, indicating a small shift in the strength of the water-water H-bond depending on which aromatic ring is electronically excited. The free OH stretch transition does not split, but is broadened, suggesting that it too is composed of two transitions.

4.4.2.2 DPOE-$(H_2O)_3$ Cycle

Figure 4.12 shows the corresponding spectra for the main 1:3 conformer, the 1:3(cycle), whose structure is shown in Figure 4.3c. The ES-RIDIR spectrum with S_1 origin as intermediate state (blue trace) shows much less dramatic shifts in the frequencies of the six OH stretch fundamentals than in the 1:1 and 1:2 clusters. Not unexpectedly, the free OH stretch and three OH···O fundamentals of the trimer cycle are essentially unshifted from their ground state frequencies. A close-up view of the pair of transitions ascribed to the π-bound and ether-bound OH groups is shown in Figure 4.12b. Note that these two OH groups are now on different water molecules (Figure 4.3c). Their

OH stretch fundamentals are shifted in frequency from their ground state values, but only by small amounts (3641/3653 cm^{-1} in S_0 to 3646/3652 cm^{-1} in S_1). According to the ground state calculations, the π and ether-bound OH stretch fundamentals are partially delocalized between these two OH groups due to their near-degeneracy, producing in-phase and out-of-phase pairs. On this basis, it is not possible to determine clearly which ring is excited in S_1 and which in S_2; although calculations predict that electronic excitation in the S_1 state is localized on the π-bound ring.

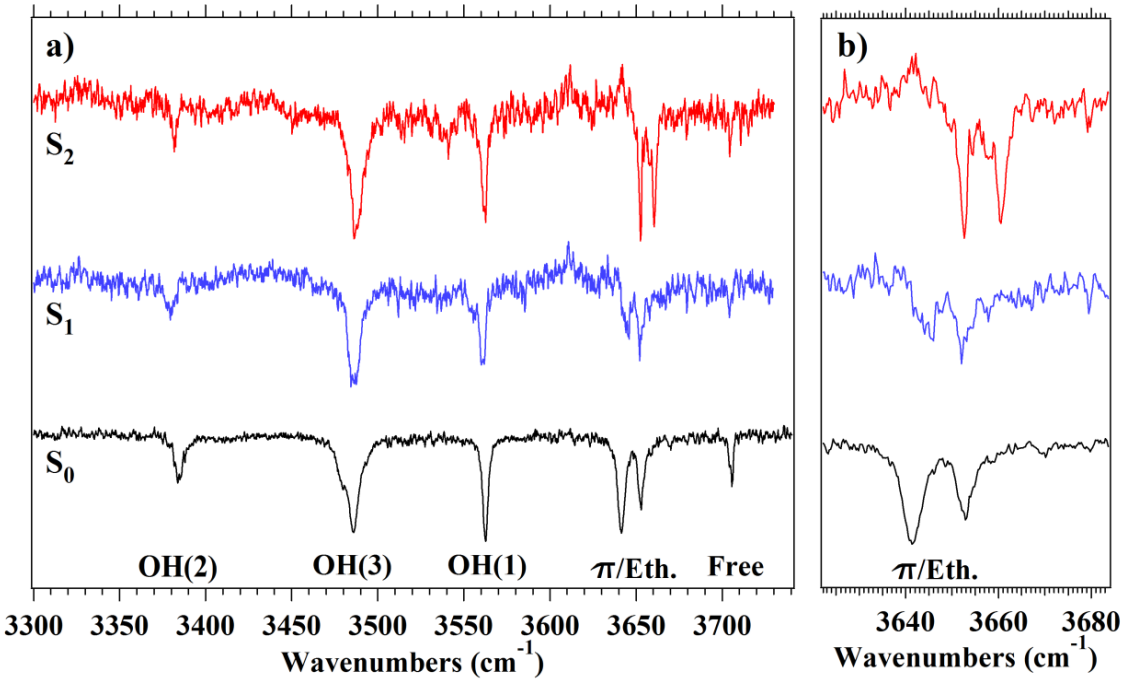


Figure 4.12. Comparison of the OH-stretch RIDIR spectra (a) recorded in the S_0 (black trace), S_1 (blue trace), and S_2 (red trace) electronic states for the [1:3] cycle (conformer A). (b) Expanded view of the π/ether bound transitions in the [1:3] cycle.

The ES-RIDIR spectrum of the S_2 origin now shows just six OH stretch transitions, with a noticeable absence of extra transitions that were the characteristic signature of a mixed S_2 0^0/ $S_1(v)$ excited state level. Importantly, even though the shifts are small, the π-bound/ether-bound doublet (3661 and 3653 cm^{-1}) has frequencies in the S_2 state that are clearly distinct from the S_1 state, as shown in close-up view in Figure 4.12b. Both the lack of extra transitions and the unique positions of the π/ether doublet lead to the unmistakable conclusion that the S_2 origin of the 1:3 cycle does not undergo significant mixing with $S_1(v)$ levels. This represents a striking change from the 1:1 and 1:2 clusters, where S_1/S_2 mixing at the (nominal) S_2 origin is clear, and in the 1:2 cluster dominates the spectrum. The reasons for this change are considered further in the Discussion section (Section 4.5).

Corresponding spectra of the 1:3 chain and 1:4 cycle were not attempted due to their weak intensity, which hindered detection with adequate signal-to-noise ratio in depletion.

4.5 Discussion

4.5.1 Perturbations Imposed by Water Clusters on DPOE Excited States

One of the principal goals of the present study is to understand how binding water clusters, of increasing size, influences the two close-lying excited states of DPOE. The size- and structure-specific UV spectra have identified S_1 and S_2 origin transitions (Figures 4.4 and 4.5), while ES-RIDIR spectra (Figure 4.11 and 4.12) follow the

development in the OH stretch IR spectra of the water molecules in each cluster with electronic state.[36,37]

4.5.1.1 Electronic Frequency Shifts and S_1/S_2 Splittings

Table 4.1 summarizes the absolute and relative wavenumber positions of the S_1 and S_2 origins of DPOE-$(H_2O)_n$, n=1-4, and summarizes key structural parameters of the clusters that influence the positions of these origins. The *tgt* conformer of DPOE is retained in all but one of the observed DPOE-$(H_2O)_n$ clusters, and even in the one exception (1:3 chain), the structure is a slightly modified version in which one of the rings re-orients dihedral θ_1 (Table 4.1) in order to accommodate a secondary stabilization via a π H-bond. Since the S_1 and S_2 states of the *tgt* DPOE monomer are split by only 1.0 cm^{-1}, we use the position of these monomer origins as the zero of our relative frequency scale (36423 cm^{-1} = "0").[36,37]

Consider the electronic frequency shifts of the S_0-S_1 origins induced by the water clusters. There are several contributions to these shifts. First, due to changes in the vibrational frequencies, the zero-point energy of the cluster can change with electronic state, producing a shift in the electronic origin transition. Second, binding a water cluster to DPOE can distort the DPOE monomer conformation, thereby changing the site energies of the S_1 and S_2 states due to this conformational distortion. Third, in each sized water cluster, there is a single water molecule that is the primary means of binding to DPOE through the hydrogen bonds it forms. This primary site binding will influence the relative energies of each aromatic ring in ground and excited states (and therefore shift

and split the S_1 and S_2 state origins). Finally, secondary binding sites will produce secondary interactions that can either reinforce or counteract the primary binding shifts.

These factors provide a framework for understanding the observed electronic frequency shifts. The clusters in this size regime fall into one of two H-bonding architectures involving H-bonded chains and cycles. There are a series of water chains with n=1-3 that maintain the asymmetric double-donor binding of the primary water molecule that serves as a bridge, donating one OH group to the ether oxygen of one ring, and the other OH to the π cloud on the opposite ring. We put in this series DPOE-H_2O (Figure 4.2b), DPOE-$(H_2O)_2$ (Figure 4.3a), and DPOE-$(H_2O)_3$ chain (conformer B, Figure 4.4b). These have S_1 origins that have frequency shifts of +48, -40, and -61 cm^{-1} from the DPOE *tgt* monomer. The degree of asymmetry induced by this primary binding site is sufficient to completely localize the electronic excitation, with the S_1 state arising from the π-bound ring and S_2 from the O-bound ring.

The fact that the frequency shift in DPOE-H_2O is to the blue (+48 cm^{-1}) is somewhat surprising, given that the OH···π fundamental in the infrared decreases in frequency in the S_1 state spectrum (Figure 4.11a). In single chromophore X···Y complexes, the electronic frequency shift is a measure of the difference in binding energy of the complex between ground and excited states.[49] A decrease in frequency of the OH stretch fundamental would suggest a strengthening, leading to a red-shift. However, Table 4.1 shows that binding of the first H_2O molecule induces an asymmetry in the backbone conformation by modest re-orientations of the two rings (θ_1= -170.8° and θ_2= -177.5°). With the two aromatic rings in different local environments, their site frequencies are different, and contribute to the observed electronic frequency shift. This

change is likely responsible for the small blue shift in the S_0-S_1 origin relative to that in DPOE monomer.

In the 1:2 and 1:3 chains, the S_0-S_1 origins appear at -40 and -61 cm^{-1}, shifted well to the red of that in DPOE-H_2O (+48 cm^{-1}). Since the asymmetric primary H_2O binding is retained, and the DPOE dihedral angles are both quite similar to that in DPOE-H_2O (Table 4.1), we postulate that it is the interactions of the secondary water molecule(s) that produces this red-shift. Indeed, the second water molecule positions its O-atom so as to sit in a pocket created by the aromatic and alkyl CH groups on the π-bound ring, consistent with a red-shift in the electronic origin upon excitation of this ring.[50-56]

Commensurate with the asymmetric primary binding in DPOE-(H_2O) and DPOE-$(H_2O)_2$ is a large splitting between S_1 and S_2 states of 189 cm^{-1} in the former, and 205 cm^{-1} in the latter. The principal interaction with the second ring responsible for the S_2 state is via the OH···O H-bond with the phenoxy oxygen, producing a large blue shift relative to S_1. This would suggest a weakening of the OH···O H-bond with electronic excitation, which is indeed observed via a shift to higher frequency in the OH stretch fundamental of the OH···O group in the S_2 ES-RIDIR spectrum of the 1:2 cluster (Figure 4.11b). While we were not able to locate the S_2 state in the 1:3 chain, the calculated vertical S_1/S_2 splitting is similar to those in DPOE-H_2O and DPOE-$(H_2O)_2$ (Table 4.1 and Figure 4.9).

Dramatic changes in the UV spectroscopy accompany the switch from water chains to cycles at n=3. The 1:3 cycle has its S_0-S_1 origin shifted 108 cm^{-1} to the blue of the 1:3 chain (from -61 cm^{-1} to +47 cm^{-1}), while in the 1:4 cycle is shifted +124 cm^{-1} to the blue of the DPOE monomer (Table 4.1). This reflects the switch from chain to cycle,

with primary binding to DPOE through a single OH group of one water in the cycle (Figure 4.3c and d). This water molecule now binds more symmetrically to the two ether oxygens, with OH···O distances of 2.23/2.41 Å in the 1:3 cycle, and 2.11/2.50 Å in the 1:4 cycle. Secondary binding is now principally via a π H-bond of another water molecule in the cycle, with the size of the cycle a good match with the distance between these two sites, especially in the 1:4 cycle. The combined effects of a stronger π H-bond and a more nearly symmetric interaction with the ether on the same ring shift the S_1 origin to the blue. Vertical TDDFT calculations (Figures 4.7–4.10) confirm that the S_1 state is localized on the aromatic ring with the π-bound water, while S_2 is localized on the opposite ring, which has the shorter OH···O H-bond.

The switch from chain to cycle also dramatically reduces the S_1/S_2 splitting. In the 1:3 cycle, this splitting is only 97 cm^{-1}, about half that in the 1:1 and 1:2 clusters. In the 1:4 cycle, we cannot clearly identify the S_2 origin, but it must be somewhere in the set of transitions that extend only ~100 cm^{-1} above the S_1 origin. The S_2 state is localized on the opposite ring, with principal interaction via the ether-bound water, which is still strong, but now somewhat more symmetrically distributed between the two ether oxygens. The reduction in S_1/S_2 splitting is the most obvious and important consequence of the more symmetric binding of the primary water molecule in forming the cycle. At the same time, the secondary π H-bond shifts the S_1 state further to the blue, closing the gap even more. As Table 4.1 shows, calculated vertical excited state splittings from TD-DFT calculations show the same trends with size and cluster structure as observed experimentally (Figures 4.7–4.10).

4.5.1.2 Development in S_1/S_2 Mixing with Water Cluster Size

The excited-state RIDIR spectra of Figures 4.11 and 4.12 afford a unique opportunity to follow the ways in which the OH stretch IR spectrum reports on the electronic character of the excited state vibronic level accessed by the resonant UV laser. When the S_1 zero-point level is used as intermediate state, the OH stretch transitions of the attached water molecule(s) reveal the type and strength of interaction with each aromatic moiety, principally through the frequency shift that electronic excitation induces in the π-bound OH fundamental, and the lack of a corresponding shift in the ether-bound OH fundamental. This is the principal experimental evidence that the S_1 state is localized on the π-bound aromatic ring.

When the S_2 origin serves as intermediate state, the excited state OH stretch infrared spectra show some transitions that are clearly identified with the $S_2|0^0>$ state that provides the oscillator strength from the ground state zero-point level in UV excitation. In addition, however, the spectrum shows a set of OH stretch fundamentals that are S_1-like. This doubling of the spectrum is immediately evident in the spectra of the S_2 origins of the 1:1 and 1:2 clusters, but is absent in the 1:3 cycle.

Thus, in the 1:1 and 1:2 clusters, the intermediate state is of mixed character, but this mixing is not occurring measurably in the corresponding level in the 1:3 cycle. To be more specific, in the 1:2 cluster, the transition +205 cm^{-1} above the S_1 origin, which is assigned in the hole-burning spectra of Figures 4.4b and 4.5b to the S_0-S_2 origin, in fact has an excited state wave function that is an admixture of components arising from vibronic coupling between the two manifolds:

$$\Psi(+205) = c_0\psi\left(S_2\left|0^0\right\rangle\right) + \sum_{i=1} c_i \cdot \psi\left(S_1\left|v_i\right\rangle\right) \tag{4.1}$$

The water molecules report on this excited state mixing through the unique OH stretch absorptions associated with each excited state. In this sense, the water molecules project the S_2 and S_1 components of the wave function onto corresponding levels with OH(v=1) far above the electronic origins, producing twice the number of OH stretch transitions. Some of these OH stretch frequencies are strongly affected by electronic excitation, and are thereby shifted from one another. Others are affected much less, and in some cases are not shifted measurably in frequency, and thus the S_1 and S_2 components of these transitions overlap in the spectrum.

Furthermore, the relative intensities of the OH stretch transitions assigned to $S_2|0^0>$ and $S_1(v)$ character of the wave function reflect the sizes of the mixing coefficients. In this sense, the excited state IR spectrum provides a window into the internal conversion from the S_2 zero-point level into the $S_1(v)$ levels induced by a known solvent structure carried out under collision-free conditions. Table 4.3 reports the ratio of integrated intensities of the transitions ascribed to S_2 and S_1 character in the OH stretch spectra of the 1:1 and 1:2 clusters. We associate this ratio with the ratio of the sum of the modulus squares of the coefficients in equation (4.1):

$$\frac{I(S_1(\{v\}))}{I(S_2\left|0^0\right\rangle)} = \frac{\sum_{i=1}\left|c_i\right|^2}{\left|c_0\right|^2} \tag{4.2}$$

While this intensity ratio is 0.94 in the 1:1 complex (Table 4.3), it increases to ~4.5 in the 1:2 cluster, but then drops to near zero in the 1:3 cycle. Based on the noise level in the scans, we estimate an upper bound for the integrated S_1/S_2 ratio at ~0.1.

Table 4.3. Ratio of integrated intensities and relative populations derived from ES-RIDIR spectra, calculated density of states, and number of low frequency vibrations.

Structure	Ratio of Integrated Intensity[a] $I(S_1)/I(S_2)$	Percent Character S_1:S_2	Calc. Density of States[b] (per cm^{-1})	Number of Normal Modes[c] (100/200 cm^{-1})
DPOE (tgt)[d]				4/7
DPOE-H_2O[d]	0.94	43:57	6	6/10
DPOE-$(H_2O)_2$[e]	4.5	82:18	20	7/12
A-DPOE-$(H_2O)_3$ Cycle	<0.1	<10 : >90	0.9	8/14

[a]Intensity ratios extracted from the excited state RIDIR spectrum.

[b]Density of S_1 vibrational states at the S_2 origin, calculated using the ground state frequencies at the M052x/6-31+G(d) level of theory.

[c]Number of normal modes with vibrational frequencies less than 100/200 cm^{-1}, calculated at the M052x/6-31+G(d) level of theory.

[d]Refs. [35-37].

[e]1σ errors for integrated ratio of ±0.6 and ±2% for the relative populations.

To understand why, one must consider the vibronic coupling mechanism at play in these clusters under collision-free conditions. Previous state-to-state studies of internal mixing in isolated bichromophores and aromatic dimers[19-25,30,31] have stimulated the development and refinement of theoretical models[32-34] that identify the Franck-Condon active vibrations as the principal modes engaged in vibronic coupling. In its simplest form, the $S_2|0^0>$ level will be mixed with $S_1|X^1>$ levels through Herzberg-Teller coupling, with its $\Delta v=\pm 1$ selection rules. Many of these Franck-Condon active vibrations are ring

modes directly associated with the two chromophores. In addition, S_1 vibronic levels that are near in energy to the S_2 0^0 level will engage in mixing through higher-order terms in the vibronic Hamiltonian.[23,27,34] While these higher-order coupling matrix elements are significantly weaker, their near-degeneracy makes them competitive with first-order Herzberg-Teller coupling that reaches across larger energy differences.

In the DPOE-$(H_2O)_n$ clusters of interest here, the asymmetry produced by the water clusters splits the S_1 and S_2 states and localizes the electronic excitation on one or the other ring. The Franck-Condon activity in the ring modes is unlikely to change significantly by complexation with water. Indeed, the previous study of DPOE-H_2O[37] included dispersed fluorescence scans from the S_1 and S_2 origins that confirmed similar Franck-Condon activities in the ring modes. On that basis, one wouldn't anticipate much change in the degree of S_1/S_2 mixing with developing water cluster size.

On the other hand, the nature of the near-degenerate S_1 levels and the density of S_1 vibrational states at the energy of the S_2 0^0 level depends sensitively on the size, shape, and mode of binding of the water cluster. With the addition of each successive water molecule, six more intermolecular vibrational modes are formed, many of which have vibrational frequencies low enough to add to the density of $S_1(v)$ states at the energy of the S_2 0^0. While the S_1/S_2 splitting is quite similar in the 1:1 and 1:2 clusters (189 and 205 cm^{-1}, respectively), the density of states increases from 6 to 20 states/cm^{-1} (Table 4.3) for those same clusters. These estimates involve direct state counting using ground state vibrational frequencies for the clusters, a reasonable first approximation given the similar geometry of the S_0 and S_1 states based on the modest Franck-Condon activity observed. The increased density of states is likely the principal reason for the increased $S_1(v)$

character in the wave function of the 1:2 cluster relative to 1:1. Similar reasoning provides a qualitative explanation for the lack of mixing in the 1:3 cycle. In this case, the smaller splitting between S_1 and S_2 origins (98 cm^{-1}) drops the density of S_1 vibrational states at the S_2 origin to <1 state/cm^{-1}, reducing the efficiency of internal conversion with near-degenerate levels of $S_1(v)$.

The presence of the water molecule(s) has significant effects on vibrational energy levels present in the first 200 cm^{-1} above the electronic origin. As Table 4.3 shows, using the S_0 state as a guide, the calculations predict that the DPOE(*tgt*) monomer has four normal modes with frequencies of 100 cm^{-1} or lower and seven lower than 200 cm^{-1}, while the 1:1, 1:2 chain, and 1:3 cycle have 6/10, 7/12, and 8/14 vibrations in the 100/200 cm^{-1} range, respectively. In the water-containing clusters, most of the vibrational modes have both intramolecular and intermolecular components to them, since the low frequencies torsions of DPOE necessarily modulate the positions of the water molecules. Several of the low frequency vibrations of DPOE are actually stiffened by the presence of the water molecules positioned in the space between the two phenyl rings.

We postulate, then, that the water molecules provide a mechanism by which the low-frequency DPOE vibrations are mixed throughout the full density of vibrational states, bringing key modes with strong coupling strengths and low-order resonances into close proximity with the S_2 $|0^0>$ level. In this sense, the water molecules play an active role in promoting internal conversion between the S_2 and S_1 states. Since the two states are electronically localized on one or the other ring, this internal conversion amounts to electronic energy transfer between the two rings.

4.6 Conclusion

The present study on DPOE-$(H_2O)_n$ clusters with n=1-4 have provided a molecular-scale, state-to-state picture of Kasha's rule under the unique conditions uncovered by laser excitation of the individual-sized clusters. Firm structural assignments for the clusters as a function of the number of water molecules provided well-characterized starting structures for excitation to either member of the near-degenerate pair of electronic states.[38] The H_2O molecules play an active role in every aspect of the excited state phenomena of these bichromophores. Their asymmetric binding to the two UV chromophores splits the near-degeneracy of the two weakly coupled $\pi\pi^*$ excited states by an amount that is unique to the type of solvent structure and its mode of binding to DPOE. This localizes the electronic excitation on one or the other aromatic ring. The water molecules then also provide a mechanism for increasing the density of vibrational states built off the S_1 origin at the energy of the S_2 origin, and distribute the DPOE torsional motion into a dense manifold of mixed torsional/intermolecular vibrational levels that facilitate internal conversion.

Finally, through the unique lens of excited state infrared spectroscopy, the water molecules report on the nature of the excited states and the degree of $S_2|0^0>$ and $S_1|v>$ character of the wave function at the nominal S_2 origin. This degree of mixing tracks the total density of vibrational states at the energy of the S_2 origin. Of the three clusters probed in this way, the 1:2 cluster has the greatest density of states, due to its comparatively large S_1/S_2 splitting (205 cm^{-1}) and its large density of S_1 vibrational states at the energy of the S_2 origin. The 1:3 cycle, with its much smaller S_1/S_2 splitting

(98 cm^{-1}), undergoes a precipitous drop in the vibrational state density, effectively shutting off mixing between the S_2 0^0 and nearby $S_1(v)$ levels.

The mixing of intermolecular modes with low-frequency torsional vibrations of DPOE itself, leads to what is effectively internal conversion of S_2 to $S_1(v)$, but does so without irreversible energy transfer to the solvent, as would occur in bulk solution. However, it is hoped that the molecular, state-to-state view afforded by this work will be of value to those seeking to understand, and perhaps direct and control, internal conversion and electronic energy transfer in other environments. Furthermore, the qualitative explanations offered in this work for the effect of the solvent molecules on the vibronic states of this model bichromophore set the stage for quantitative theoretical modeling using methods already in place for the isolated molecules.[33]

4.7 References

(1) Kasha, M. *Discussions of the Faraday Society* **1950**, *9*, 14-19.

(2) Kasha, M.; McGlynn, S. P. *Annu. Rev. Phys. Chem.* **1956**, *7*, 403-424.

(3) Klessinger, M.; Michl, J. *Excited States and Photochemistry of Organic Molecules*; VCH Publishers, Inc.: New York, 1995.

(4) Struve, W. S. *Fundamentals of Molecular Spectroscopy*; John Wiley and Sons, Inc.: New York, 1989.

(5) Bixon, M.; Jortner, J. *J. Chem. Phys.* **1968**, *48*, 715-726.

(6) Bixon, M.; Jortner, J. *J. Chem. Phys.* **1969**, *50*, 4061-4070.

(7) Forster, T. *Naturwissenschaften* **1946**, *33*, 166-175.

(8) Forster, T. *Discussions of the Faraday Society* **1959**, *27*, 7-17.

(9) Scholes, G. D. *Annu. Rev. Phys. Chem.* **2003**, *54*, 57-87.

(10) Hybl, J. D.; Albrecht Ferro, A.; Jonas, D. M. *J. Chem. Phys.* **2001**, *115*, 6606-6622.

(11) Jonas, D. M. *Annu. Rev. Phys. Chem.* **2003**, *54*, 425.

(12) Panitchayangkoon, G.; Voronine, D. V.; Abramavicius, D.; Caram, J. R.; Lewis, N. H. C.; Mukamel, S.; Engel, G. S. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 20908-20912.

(13) Pelzer, K. M.; Griffin, G. B.; Gray, S. K.; Engel, G. S. *J. Chem. Phys.* **2012**, *136*, 164508-164508-164506.

(14) Engel, G. S.; Calhoun, T. R.; Read, E. L.; Ahn, T. K.; Mancal, T.; Cheng, Y. C.; Blankenship, R. E.; Fleming, G. R. *Nature* **2007**, *446*, 782-786.

(15) Lee, H.; Cheng, Y. C.; Fleming, G. R. *Science* **2007**, *316*, 1462-1465.

(16) Fassioli, F.; Olaya-Castro, A.; Scholes, G. D. *J. Phys. Chem. Lett.* **2012**, *3*, 3136-3142.

(17) Hossein-Nejad, H.; Olaya-Castro, A.; Scholes, G. D. *J. Chem. Phys.* **2012**, *136*, 024112.

(18) Chattoraj, M.; Paulson, B.; Shi, Y.; Closs, G. L.; Levy, D. H. *Journal of Physical Chemistry* **1993**, *97*, 13046-13051.

(19) Zehnacker, A.; Lahmani, F.; Desvergne, J. P.; Bouas-Laurent, H. *Chem. Phys. Lett.* **1998**, *293*, 357-365.

(20) Zehnacker, A.; Lahmani, F.; van Walree, C. A.; Jenneskens, L. W. *J. Phys. Chem. A* **2000**, *104*, 1377-1387.

(21) Lee, J. K.; Judge, R. H.; Boo, B. H.; Lim, E. C. *J. Chem. Phys.* **2002**, *116*, 8809-8816.

(22) East, A. L. L.; Cid-Aguero, P.; Liu, H. S.; Judge, R. H.; Lim, E. C. *J. Phys. Chem. A* **2000**, *104*, 1456-1460.

(23) Pillsbury, N. R.; Stearns, J. A.; Muller, C. W.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2008**, *129*, 114301.

(24) Pillsbury, N. R.; Muller, C. W.; Meerts, W. L.; Plusquellic, D. F.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *113*, 5000-5012.

(25) Rodrigo, C. P.; Müller, C. W.; Pillsbury, N. R.; James, W. H.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2011**, *134*, 164312.

(26) Chou, S. G.; Rodrigo, C. P.; Müller, C. W.; Douglass, K. O.; Zwier, T. S.; Plusquellic, D. F. *J. Phys. Chem. A* **2011**, *115*, 9643-9652.

(27) Stearns, J. A.; Pillsbury, N. R.; Douglass, K. O.; Muller, C. W.; Zwier, T. S.; Plusquellic, D. F. *J. Chem. Phys.* **2008**, *129*, 224305-224318.

(28) Fulton, R. L.; Gouterman, M. *J. Chem. Phys.* **1961**, *35*, 1059-&.

(29) Fulton, R. L.; Gouterman, M. *J. Chem. Phys.* **1964**, *41*, 2280-&.

(30) Ottiger, P.; Leutwyler, S. *Chimia* **2011**, *65*, 228-230.

(31) Ottiger, P.; Leutwyler, S. *J. Chem. Phys.* **2012**, *137*, 204303.

(32) Kopec, S.; Ottiger, P.; Leutwyler, S.; Köppel, H. *J. Chem. Phys.* **2012**, *137*, 184312.

(33) Nebgen, B.; Emmert, F. L.; Slipchenko, L. V. *J. Chem. Phys.* **2012**, *137*, 084112.

(34) Pillsbury, N. R.; Kidwell, N. M.; Nebgen, B.; Slipchenko, L. V.; Douglass, K. O.; Cable, J. R.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2014**, *141*, 064316.

(35) Buchanan, E. G.; Sibert, E. L.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 2800-2811.

(36) Buchanan, E. G.; Walsh, P. S.; Plusquellic, D. F.; Zwier, T. S. *J. Chem. Phys.* **2013**, *138*, 204313-204311.

(37) Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Phys. Chem. Lett.* **2013**, *4*, 1644-1648.

(38) Walsh, P. S.; Buchanan, E. G.; Gord, J. R.; Zwier, T. S. *J. Chem. Phys.* **2015**, *142*, 154303.

(39) Lubman, D. M. *Mass Spectrom. Rev.* **1988**, *7*, 559-592.

(40) Page, R. H.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988**, *88*, 4621-4636.

(41) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205-241.

(42) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827-8839.

(43) Walsh, P. S.; Kusaka, R.; Buchanan, E. G.; James, W. H.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 12350-12362.

(44) Stearns, J. A.; Das, A.; Zwier, T. S. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2605-2610.

(45) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J.V.; Cioslowski, J.; Fox, D. J.; Revision C.01 ed.; Gaussian, Inc.: Wallingford CT, 2009.

(46) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289-300.

(47) Doemer, M.; Guglielmi, M.; Athri, P.; Nagornova, N. S.; Rizzo, T. R.; Boyarkin, O. V.; Tavernelli, I.; Rothlisberger, U. *Int. J. Quantum Chem.* **2013**, *113*, 808-814.

(48) Buchanan, E. G.; James, W. H., III; Choi, S. H.; Guo, L.; Gellman, S. H.; Muller, C. W.; Zwier, T. S. *J. Chem. Phys.* **2012**, *137*, 094301-094316.

(49) Bürgi, T.; Droz, T.; Leutwyler, S. *Chem. Phys. Lett.* **1995**, *246*, 291-299.

(50) Buchanan, E. G.; Zwier, T. S. *J. Phys. Chem. A* **2014**, *118*, 8583-8596.

(51) Venkatesan, V.; Fujii, A.; Mikami, N. *Chem. Phys. Lett.* **2005**, *409*, 57-62.

(52) Amicangelo, J. C.; Gung, B. W.; Irwin, D. G.; Romano, N. C. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2695-2705.

(53) Kusaka, R.; Inokuchi, Y.; Ebata, T. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4452-4459.

(54) Kusaka, R.; Inokuchi, Y.; Ebata, T. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6238-6244.

(55) Kusaka, R.; Inokuchi, Y.; Ebata, T. *Phys. Chem. Chem. Phys.* **2009**, *11*, 9132-9140.

(56) Shubert, V. A.; Müller, C. W.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *113*, 8067-8079.

CHAPTER 5. CYCLIC CONSTRAINTS ON CONFORMATIONAL FLEXIBILITY IN γ-PEPTIDES: CONFORMATION SPECIFIC IR AND UV SPECTROSCOPY

5.1 Introduction

Proteins are composed of α-amino acid residues, in which one carbon atom (Cα) separates adjacent amide groups. There has been increasing interest in extending the set of protein-like oligoamides to include other building blocks, including β-amino acid residues (two carbons between amide groups) and γ-amino acid residues (three carbons between amide groups). These protein-inspired oligomers can also contain heterogeneous backbones, in which α, β and/or γ residues are combined. These synthetic foldamers are designed to adopt secondary structures comparable to those found in biopolymers (e.g., helices, sheets, reverse turns), and to display properties beyond those available to conventional peptides and proteins, including resistance to enzymatic degradation and greater conformational stability.[1-6] Among the structural elements that have benefited most from this development are helices, which can be formed with a variety of helical pitches, radii, and side-chain architecture.

As Figure 5.1d illustrates, the conformation of a γ residue is determined by four dihedral angles. Two are analogous to the Ramachandran angles used to determine α-amino acid residue conformation (ϕ,ψ), and the other two arise from the extended γ residue backbone (θ,ζ). This increase in backbone complexity in a γ residue relative to an α residue might be anticipated to result in a wider range of competing low-energy

structures, which could lead to a greater diversity in folding preferences for γ-peptides relative to α-peptides. The presence of three positions for substitution within each γ residue also offers a richer palette of side chain-based strategies to control γ residue conformational propensity relative to the options available for α residues. Several previous studies focused attention on γ-peptides with side chains similar to those in α-peptides.[7-13] Alternatively, ring-based constraints have been built into γ residues that are intended to promote secondary structure specificity.[14] One promising substitution pattern combines a cyclohexane ring that incorporates the Cβ-Cγ backbone bond with ethyl substitution at Cα. These substituents create a total of three chiral centers, and an expeditious synthetic route is available for γ_{ACHC} in which all three centers have *S* configuration or all three have *R* configuration. This substitution pattern leads to a propensity for a g+, g+ (θ,ζ) torsion angle sequence, which imparts overall curvature across the γ residue.

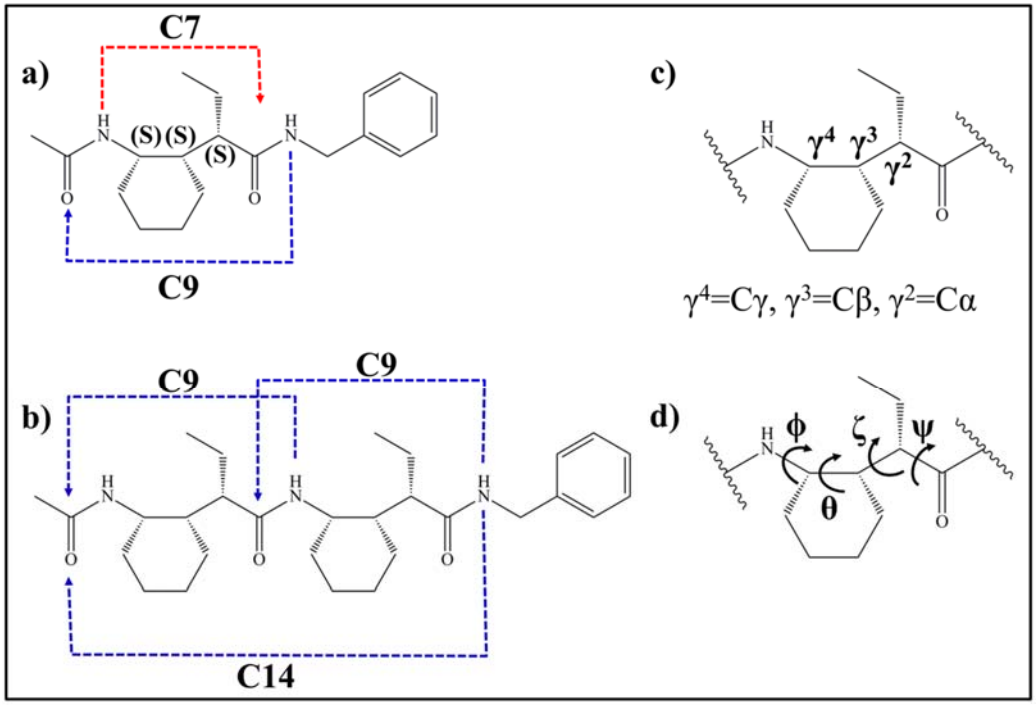


Figure 5.1. Structures of (a) Ac-γ_{ACHC}-NHBn and (b) Ac-$(\gamma_{ACHC})_2$-NHBn with possible intramolecular H-bonds and ring-sizes included. Hydrogen bonds shown in blue go from C- to N-terminus while those in red are formed from N- to C- terminus. (c) Carbon labeling scheme for the carbons in the backbone and (d) the naming convention for the dihedral angles. The hydrogen bonded rings are labeled as Cn, where n indicated the number of atoms contained in the ring.

Several studies have been carried out in condensed phases to determine the conformations adopted by short γ-peptides, and to discover how conformational behavior evolves as the length of the γ-peptide chain grows. One of the most well-characterized γ-peptide secondary structures, the 14-helix, contains 14-atom, *i, i+3* C=O···H-N H-bonds.[7,8,14-17] The recent work of Guo et al. has provided X-ray crystal structures and NMR spectral signatures (NOEs) that arise from the γ-peptide 14-helix, which has 2.6 residues per turn (2.6_{14}).[14] The 14-helix has been detected in nonpolar solution for

tetramers and longer oligomers.[7,8] By comparison, short gabapentin-containing γ-peptides, disubstituted at Cβ, prefer C9 "nearest-neighbor" H-bonds in condensed phases.[14,18,19]

Investigations of the conformational preferences of short peptidic foldamers in the gas phase offers a perspective complementary to that available from condensed-phase studies, which can deepen insight regarding the delicate balance of forces that operate within these novel backbones. By removing solvent, the gas phase studies reveal inherent conformational preferences. Such studies also offer a direct connection with high-level *ab initio* computational studies, and can contribute to the development of molecular mechanics force fields that correctly simulate non-natural foldamer sub-units.

As a part of our ongoing studies of the inherent conformational preferences of synthetic foldamers,[20-23] we have recently carried out a systematic investigation of the preferred conformations of a series of gas-phase γ-peptides cooled in a supersonic expansion.[24-27] Our initial studies of the prototypical capped diamide Ac-γ^2-hPhe-NHMe (Figure 5.2a) showed that the majority of the population formed C9 H-bonds.[24] However, about 20% of the population existed in an amide-stacked arrangement, with no intramolecular H-bonds (Figure 5.2b).[24,27] The inter-amide stacking led to favorable electrostatic and dispersive forces between the amide groups. The three-carbon spacer in the backbone allowed for the adoption of an optimal stacked geometry.[24] This surprising result led us to carry out follow-up studies with three γ-peptide derivatives that were designed to test the robustness of amide-stacking.[25] Indeed, when the C9 conformation was blocked in Ac-γ^2-hPhe-$N(Me)_2$, all the population was funneled into amide-stacked structures. Introduction of a bulky isopropyl group at the C-terminus disrupted amide

stacking, resulting in the observation of only C9 H-bonded rings. When gabapentin substitution (at Cβ) was combined with –N(Me)$_2$ capping, only C7 H-bonded conformations were observed.[25]

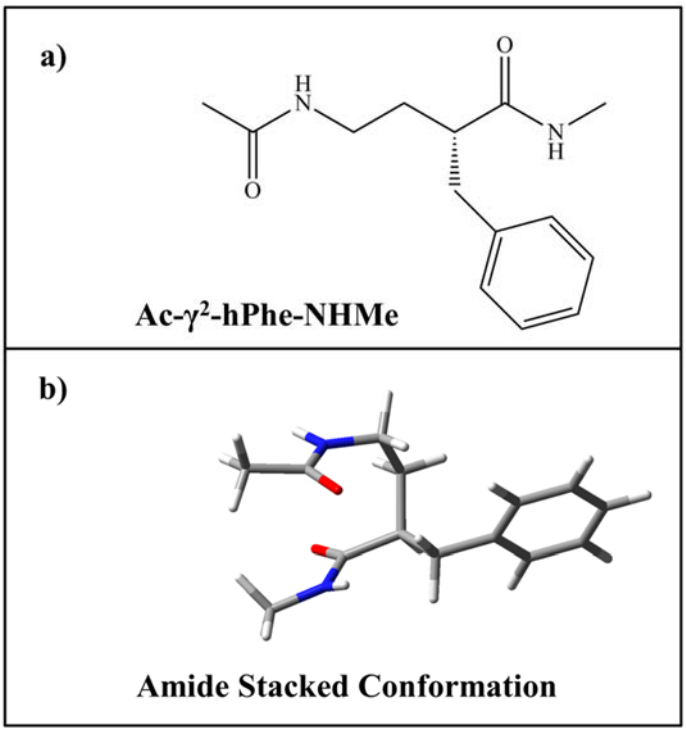


Figure 5.2. (a) The chemical structure of Ac-γ^2-hPhe-NHMe and (b) a typical amide stacked conformation of Ac-γ^2-hPhe-NHMe [24].

Analogous studies of capped γ-peptide triamides Ac-γ^2-hPhe-γ^2-hAla-NHMe and Ac-γ^2-hAla-γ^2-hPhe-NHMe offered a first look at the minimum length necessary for formation of a C14 H-bonded ring.[26] The results of that study revealed the presence of a variety of structures, including four C9/C9 sequential double ring conformers, a single C9/C14 bifurcated double-ring, and two C7/C7/C14 H-bonded cycles. While C14 rings were represented among the observed structures, they were not among the dominant

forms and always appeared in structures stabilized by additional H-bonds. Thus, γ residue substitution with a single methyl or benzyl group at Cα was seen to enable folding into diverse compact conformations featuring backbone dihedral angles that differ substantially from those of the 14-helix.[26]

In the present manuscript, we extended our single-conformation spectroscopy of γ-peptides in two important new directions. First, we study two prototypical constrained γ-peptides that incorporate the γ_{ACHC} amino acid that has been shown to support 14-helical folding in solution and in the crystalline state (Figure 5.1).[14] Here we seek to probe whether these constrained γ amino acid residues are sufficiently pre-organized to enforce a C14 H-bond in the smallest oligomer (the capped triamide, with two γ residues) that can form a hydrogen-bonded ring of this size. Second, we compare the conformations observed for the constrained, capped γ-peptide diamide in isolation and in the presence of a single bound H_2O molecule. With this comparison we seek to determine whether the H_2O molecule, with its strong intermolecular H-bonding potential, is a sufficient perturbation to re-order the conformational preferences of the γ-peptide.

5.2 Methods

5.2.1 Experimental Methods

Enantiomerically pure cyclically-constrained γ-amino acids were synthesized by established methods.[28] Peptide coupling was performed as described previously[14], with HOAt included as a coupling additive. Successive hydrogenolyses of the benzyl ester and nitro group, acetylation, and coupling to benzylamine via standard methods afforded $\gamma\gamma_{ACHC}$ in moderate yield.

A detailed description of the experimental apparatus has been provided in Chapter 2.[29] Briefly, sample is wrapped in glass wool, placed in a quartz insert and loaded into a stainless steel sample holder located behind a general valve (Parker, Series 9, 500μm) operating at 20 Hz. The entire valve assembly was heated to ~190 and ~235 ˚C for γ_{ACHC} and $\gamma\gamma_{ACHC}$, respectively. The gaseous analytes were entrained in a pure neon backing gas (~2-3 bar, ~0.05-0.1 bar·cm^3/s) and expanded through the general valve orifice to form a supersonic expansion. The γ_{ACHC}-H_2O complex was formed with trace H_2O present in the sample or gas handling lines.

Detailed descriptions and experimental setups for the single- and double-resonance methods has been thoroughly explained in Chapter 2, and as such only experiment specific information is provided here.[30-32] The UV spectrum in the S_0-S_1 origin region was recorded using one-color resonant two-photon ionization (R2PI) where tunable UV light was generated using the doubled output of a Nd:YAG pumped tunable dye laser (Radiant Dyes NarrowScan, Coumarin 540A). Conformation specific IR spectra were recorded using resonant ion-dip infrared spectroscopy (RIDIRS). RIDIR spectra

were recorded in the hydride stretch (~3200-3750 cm^{-1}, ~3-10mJ, 0.1 cm^{-1}) and amide I/II and water bend (~1400-1800 cm^{-1}, ~300-550 μJ, 1 cm^{-1}) regions.[30-32] Conformation-specific UV spectra are recorded using infrared-ultraviolet holeburning (IR-UV HB).

5.2.2 Computational Methods

A set of starting structures for each molecule was generated by performing a conformational search with the MACROMODEL suite of programs.[33] The Amber* force field[34] was used to find stable minima within a 50 kJ/mol window of the global minimum. A sub-set of these calculated structures, typically the 20-30 lowest energy structures, were then optimized using higher levels of theory. Geometry optimizations and harmonic frequency calculations were performed using density functional theory (DFT). To best account for any dispersive interactions, Truhlar's hybrid density functional M05-2X[35] was used with the 6-31+G(d) basis set in the Gaussian09 program.[36] Based on accepted values[37], NH stretch, OH stretch, and amide I/II vibrational modes are scaled by 0.94, 0.95, and 0.96, respectively, to account for anharmonicity in the calculated harmonic vibrational modes. Relaxed potential energy scans of the internal rotation of the ethyl group were performed using opt=modredundant in the Gaussian09 program package.[36]

5.2.3 Nomenclature

Dihedral angles along the γ-peptide backbone will be labeled using the standard nomenclature shown in Figure 5.1. For each γ-peptide sub-unit, four dihedral angles (ϕ, θ, ζ and ψ) are needed, the two Ramachandran angles (ϕ,ψ) and two interior dihedrals (θ,ζ) associated with the three-carbon spacer. An additional dihedral angle describing the ethyl side chain (ξ) is defined by Cβ-Cα-$C_{ethyl,\alpha}$-$C_{ethyl,\beta}$. Finally, the phenyl ring position will be characterized by the C-N-C-C_{phenyl} dihedral angle χ. We will use the gauche+ (g+), gauche- (g-), and anti (a) designations to describe dihedrals with values $\chi \sim +60^{o}$, -60^{o}, and $\pm 180^{o}$, respectively. The size of an H-bonded ring closed by an NH···O=C H-bond is denoted by Cn, where 'n' denotes the number of atoms in the ring. A C9 structure in which the ethyl side-chain is 'anti' and the phenyl side chain gauche- would then be labeled as C9(a, g-).

5.3 Results and Analysis

5.3.1 Ac-γ_{ACHC}-NHBn and (Ac-γ_{ACHC}-NHBn)-H_2O (γ_{ACHC} and γ_{ACHC}-H_2O)

5.3.1.1 UV Spectroscopy

R2PI and IR-UV HB scans monitoring the monomer mass channel are present in Figure 5.3, spanning a wavelength region that encompasses the S_0-S_1 origin of the C-terminal benzyl cap (37400-37720cm^{-1}). Conformation-specific UV spectra were recorded through IR-UV hole-burning spectroscopy, using IR hole-burn transitions

identified by RIDIR spectroscopy (Sec. 5.3.1.2) that are unique to each species. All major transitions in the R2PI spectrum of Figure 5.3a arise from two species, with S_0-S_1 origin transitions at 37557 cm^{-1} (Figure 5.3b) and 37571 cm^{-1} (Figure 5.3c). As we shall see shortly, the spectrum in Figure 5.3b is ascribed to a single conformer of γ_{ACHC}, the only one observed in the expansion. The spectrum in Figure 5.3c will be assigned based on its IR spectrum to a γ_{ACHC}-H_2O complex, appearing in the monomer mass channel due to efficient fragmentation following photoionization.[38,39] Both spectra show Franck-Condon intensity in several low-frequency vibrations, which are more prominent in the γ_{ACHC} monomer spectrum. The S_0-S_1 origin for the γ_{ACHC}-H_2O complex is blue shifted from the monomer origin by only 14 cm^{-1}, suggesting a similar local environment for the phenyl ring in the monomer and water-containing complex. The small transition ~43 cm^{-1} to the red of the γ_{ACHC} origin band did not burn out in either of the hole-burn scans in Figure 5.3b,c. It could be a minor conformer, a hot band, or a larger γ_{ACHC}-$(H_2O)_n$ cluster. It was too weak to spectroscopically characterize, and therefore was not pursued further.

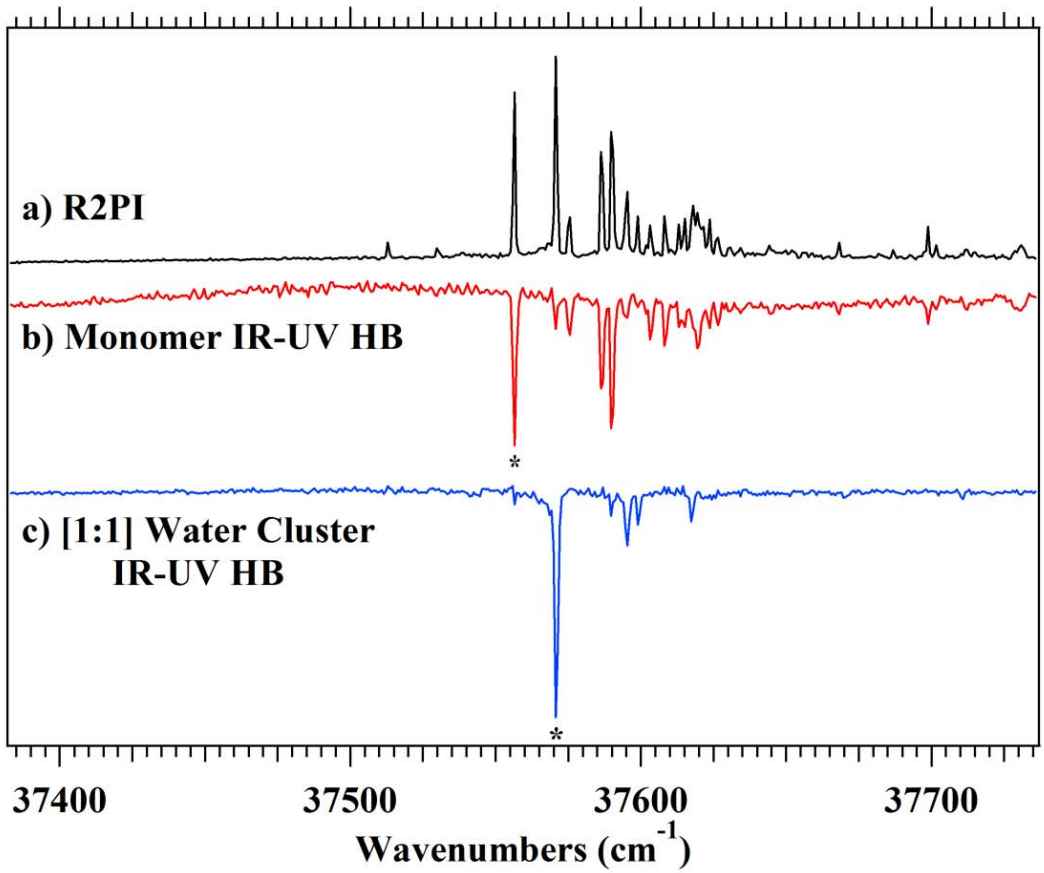


Figure 5.3. (a) R2PI spectrum taken in the monomer mass channel. IR-UV holeburning spectra for the (b) Ac-γ_{ACHC}-NHBn and (c) [1:1] water cluster recorded with IR set to 3476 and 3524 cm^{-1}, respectively. Peaks marked with an asterisk indicate vibronic transitions that were used to record RIDIR spectra.

5.3.1.2 RIDIR Spectroscopy

Figure 5.4 shows RIDIR spectra in the hydride stretch (Figures 5.4a and c) and amide I/II (Figures 5.4b and d) region for the γ_{ACHC} and γ_{ACHC}-H_2O conformers observed in the molecular beam. Several important pieces of information can be gleaned immediately from the spectra. The NH stretch RIDIR spectrum for γ_{ACHC} (Fig. 5.4a) has two NH stretch transitions due to the two amide NH groups appearing at 3476 cm^{-1} and 3301 cm^{-1}. The transition at 3476 cm^{-1} is readily assigned as a free amide NH stretch, while the intense, broadened, and shifted band at 3301 cm^{-1} indicates the presence of a strong intramolecular H-bond. The over-all appearance of the spectrum is strongly reminiscent of that in Ac-γ^2-hPhe-NHMe[24], pointing to the formation of a C9 H-bonded ring. The amide I (amide C=O stretch) transitions at 1704 and 1690 cm^{-1} show little shift due to H-bonding, while the amide II (amide N-H bend) fundamental at 1558 cm^{-1} is shifted up in frequency by 35 cm^{-1} from the free NH bend at 1523 cm^{-1}.[37]

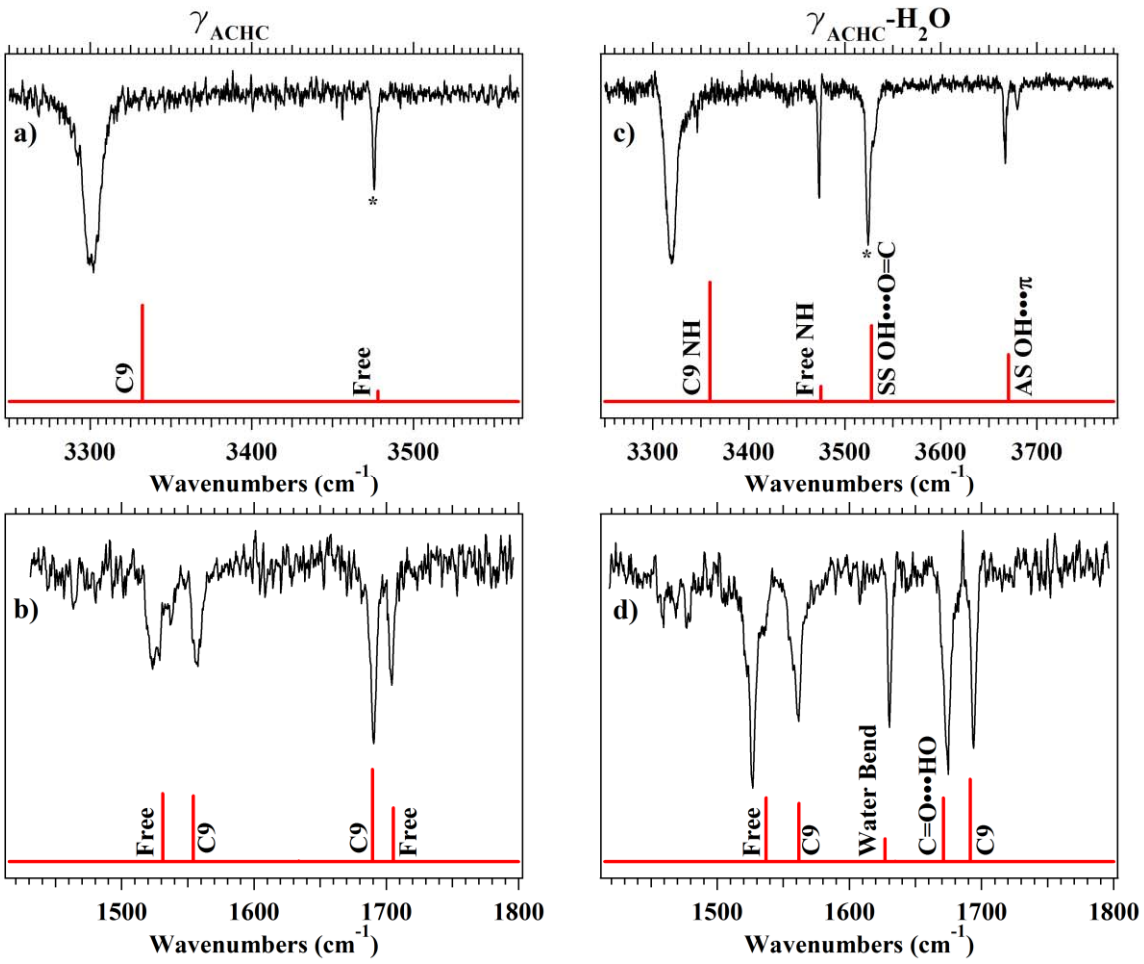


Figure 5.4. RIDIR spectra for (a, b) Ac-γ_{ACHC}-NHBn and (c, d) the [1:1] water complex in the hydride stretch (upper) and amide I/II regions (lower). Calculated scaled harmonic vibrational frequencies (see text) and IR intensities of the assigned conformations at the DFT M05-2X/6-31+G(d) level of theory are shown as stick diagrams in red. Transitions marked by an asterisk were used in recording IR-UV holeburning spectra (see Figure 5.3).

Table 5.1. Structural parameters and vibrational frequencies for the assigned structures γ_{ACHC}, $\gamma\gamma_{ACHC}$, and γ_{ACHC}-H_2O at the DFT M05-2X/6-31+G(d) level of theory.

	γ_{ACHC}		$\gamma\gamma_{ACHC}$		γ_{ACHC}-H_2O	
Backbone Dihedral	(-106 , 66 , 75 , -103)		(-158 , 61 , 67 , -115)		(-104 , 66 , 72 , -113)	
Angles (˚)	(g- , g+ , g+ , g-)		(a , g+ , g+ , g-)		(g- , g+ , g+ , g-)	
(ϕ, θ, ζ, ψ)			(-161 , 59 , 56 , -132)			
			(a , g+ , g+ , a)			
Sidechain Dihedral						
Angles (˚)	(-177 , -78)		(-175 , -173 , -116)		(-176 , -73)	
(Et_{Nterm},Et_{Cterm},Phenyl)	(a , g-)		(a , a , g-)		(a , g-)	
rC=O-NH (Å)	1.93		1.96		1.95	
C=O←H (˚)	126		138		128	
O→HN (˚)	165		172		154	
rC=O-HOH (Å)					1.95	
	Exp.	Theory	Exp.	Theory	Exp.	Theory
NH Stretch (cm^{-1})	3476	3478	3469.1	3469.5	3473	3475
	3301	3332	3467.8	3467.9	3319	3359
			3336	3356		
OH Stretch (cm^{-1})					3667	3670
					3524	3528
Amide I (cm^{-1})	1704	1705	1703	1704	1693	1691
	1690	1690	1697	1698	1675	1671
			1689	1687		
Amide II (cm^{-1})	1558	1554	1561	1562	1561	1562
	1523	1531	1516	1528	1526	1537
			1501	1509		
Water Bend (cm^{-1})					1630	1627

The calculated global minimum structure for γ_{ACHC} is the C9 single-ring structure shown in Figure 5.5a, labeled as C9(a, g-). Table 5.1 shows the structural parameters for the assigned C9 conformation along with the predicted and experimental vibrational frequencies for γ_{ACHC} C9(a, g-). The calculations predict an NH···O=C H-bond distance of 1.93Å, indicating the presence of a strong H-bond. The (θ,ζ) dihedral angles along the γ-peptide backbone for this global minimum structure are (g+, g+), constituting a turn in the backbone. The (ϕ,ψ) angles are both near -106°, which orients the two amide groups with their exterior NH and C=O groups pointing toward one another to form a C9 ring.

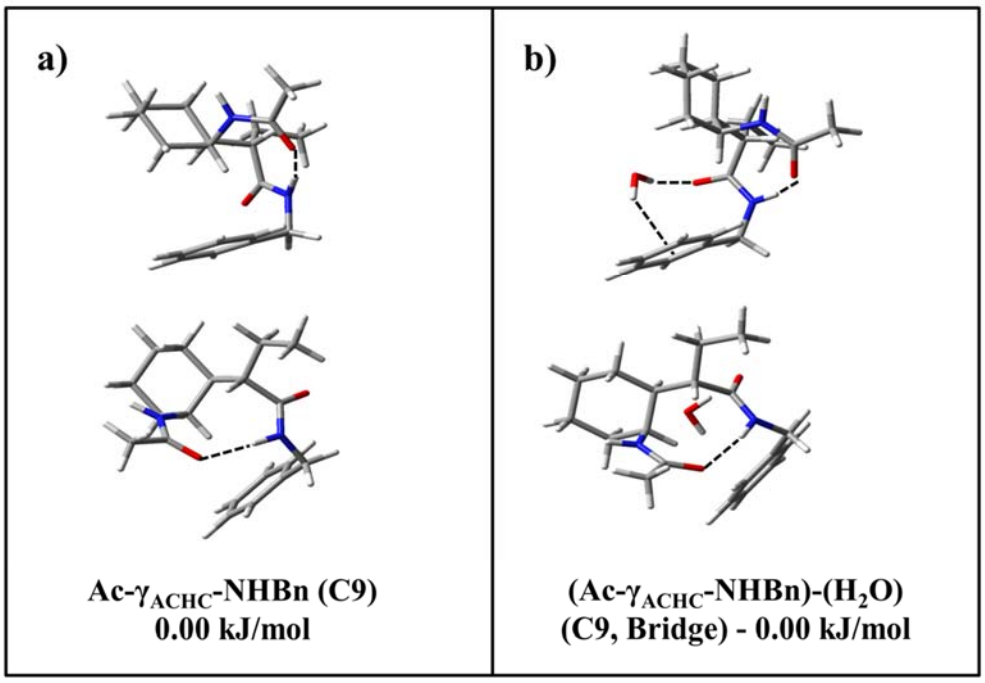


Figure 5.5. Assigned structures for the experimentally observed C9 conformations of (a) Ac-γ_{ACHC}-NHBn and (b) the [1:1] water complex. Both structures are the global minima predicted by calculations.

The calculated scaled vibrational frequencies and IR intensities are also shown as stick diagrams in Figure 5.4a and b. The calculated IR spectrum for this conformation reproduces the experimental spectrum well both in absolute frequency and relative intensity. The two NH stretch modes are highly localized on the free and H-bonded NH groups, as anticipated. Despite the small splitting between the two amide I fundamentals, the band at 1690 cm^{-1} is also predicted by the calculation to be largely localized on the C9 H-bonded C=O, while the 1704 cm^{-1} transition is localized on the free C=O group. The amide II vibrations are also quite localized, and show the opposite response to H-bond formation, in which the H-bonded NH bend is shifted up in frequency from the free amide NH bend, as is anticipated for an NH bend in an NH···O=C H-bond.

Calculations were also performed on the low energy structures of several other conformational families. Several of these additional low energy structures are compared with the experimental IR spectrum in Figure 5.6. This figure (Figure 5.6) also presents the structures of these calculated low-energy conformations. Two single-ring structures, a C7 conformation (13.19 kJ/mol), and a more extended C9(a, a) conformation (6.50 kJ/mol), reproduce the general patterns present in the experimental spectra, but show H-bonded NH stretch transitions that are poorer matches with experiment than the C9 global minimum. These structures are also significantly higher in energy than C9 global minimum. Conformations with no hydrogen bonds (8.26 kJ/mol) or possessing an amide stacked geometry (22.11 kJ/mol), are clearly inconsistent with experiment.

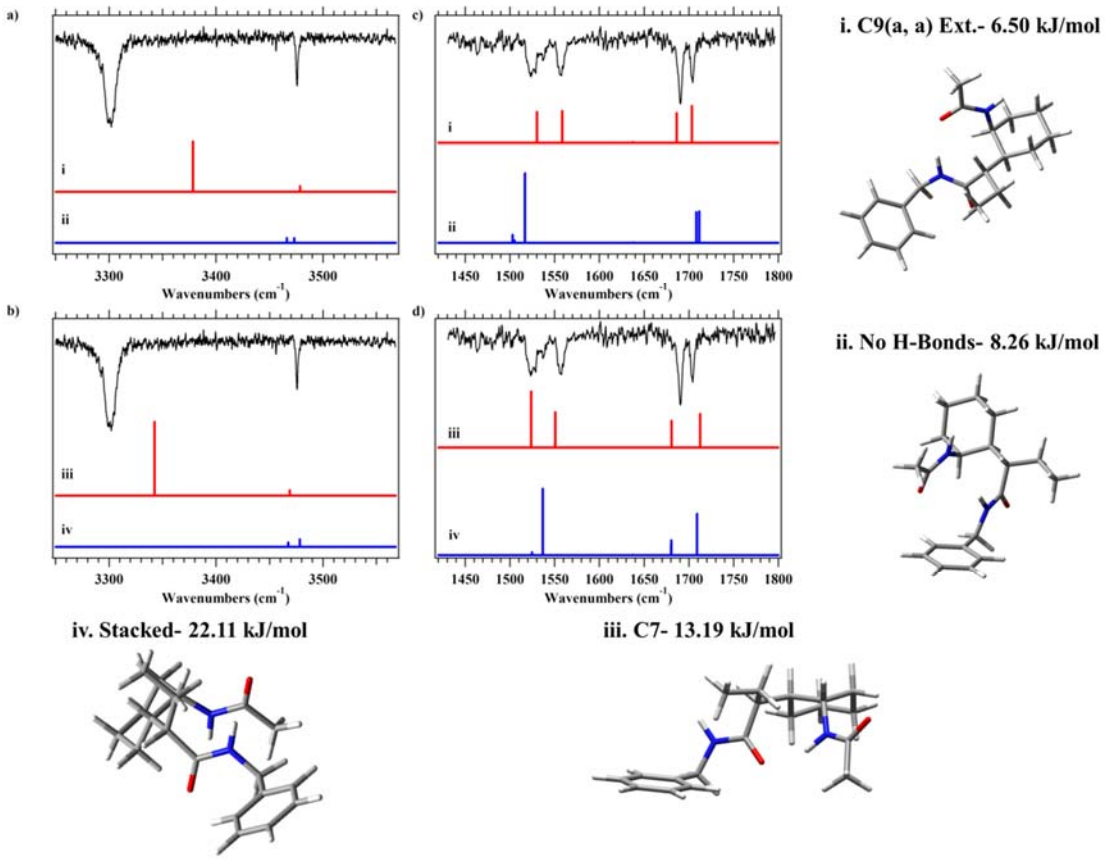


Figure 5.6. The experimental RIDIR and calculated stick spectra in the NH stretch (a,b) and amide I/II regions (c,d) for γ_{ACHC}. The structures and corresponding spectra are labeled with roman numerals around the outside of the figure (i-iv).

Figure 5.4c shows the γ_{ACHC}-H_2O RIDIR spectrum in the NH and OH stretch regions. Comparison of this spectrum with that of the bare monomer (Figure 5.4a) indicates that the γ_{ACHC} NH stretch fundamentals only undergo minor shifts in frequency upon complexation with H_2O. The hydrogen bonded NH stretch shifts from 3301 to 3319 cm^{-1}, while the free NH stretch shifts down in frequency by only 3 cm^{-1}, from 3476 to 3473 cm^{-1}. These minor shifts indicate that the γ_{ACHC} backbone is only minimally perturbed by the addition of the water molecule, maintaining the same C9 conformation, which may be slightly weakened.

The OH stretch fundamentals of the water molecule, which appear at 3667 and 3524 cm^{-1} in the γ_{ACHC}-H_2O complex, also tell us much about the way in which H_2O binds to the γ_{ACHC} monomer. The transition at 3667 cm^{-1} is well below that of a free OH (3710-3725 cm^{-1}), but right near where OH···π H-bonds are generally found.[38-40] The transition at 3524 cm^{-1} is in a frequency region typical of an OH···O=C H-bond.[41,42]

The amide I/II RIDIR spectrum for the γ_{ACHC}-H_2O complex is shown in Figure 5.4d. At first glance, the amide I/II regions show a strong resemblance to their bare-molecule counterparts. In the amide II region, this resemblance is borne out by quantitative comparison, with frequencies (1526/1561 cm^{-1}) within a few cm^{-1} of those in the γ_{ACHC} monomer. In the amide I region, however, the two C=O stretch fundamentals in γ_{ACHC}-H_2O appear at 1675 and 1693 cm^{-1}. This places the upper frequency C=O stretch very near the position of the lower-frequency transition in the monomer, while the band at 1675 cm^{-1} is well below either monomer transition frequency, indicating that both C=O groups are now in strong H-bonds. Finally, the water bending fundamental appears at 1630 cm^{-1}, positioning itself in the gap between the amide I and amide II regions, and at a frequency 35 cm^{-1} above that in the H_2O molecule (1595 cm^{-1}).[43]

The calculated global minimum structure for the γ_{ACHC}-H_2O complex (Figure 5.5b) has the solute molecule γ_{ACHC} retaining its intramolecular C9 H-bond in the same C9(a, g-) structure found in the monomer, with water forming a H-bonded bridge between the C-terminal carbonyl and the π-cloud. The calculated IR spectrum for this structure is shown in Figures 5.4c and d. The calculated frequencies and intensity patterns match experiment in significant detail, tracking the small shift produced by the (relatively) weak π H-bonded OH and the position and intensity of the C=O bound OH in

the bridging H_2O molecule, and retaining the signatures of the C9 intramolecular H-bond in the NH stretch and NH bend regions. Even the slight weakening of the C9 H-bond reflected in its shift to higher frequency is reproduced by the lengthening of this bond from 1.93 to 1.96 Å. The water-carbonyl hydrogen bond also is similar in length (1.95 Å), while the water-π hydrogen bond is substantially longer at ~2.7 Å.

5.3.2 Ac-γ_{ACHC}-γ_{ACHC}-NHBn ($\gamma\gamma_{ACHC}$)

5.3.2.1 UV Spectroscopy

The R2PI and IR-UV holeburning spectra for $\gamma\gamma_{ACHC}$ in the S_0-S_1 origin region are shown in Figure 5.7. The R2PI spectrum in Figure 5.7a shows a single dominant transition at 37487 cm^{-1}, with vibronic progressions built off of it. The IR-UV HB in Figure 5.7b confirms that essentially all the R2PI transition intensity is due to a single conformer in the molecular beam. Several transitions just to the red of the $\gamma\gamma_{ACHC}$ origin do not burn out with the major conformer. Due to low signal intensities no further conformation-specific spectroscopy was carried out on these transitions.

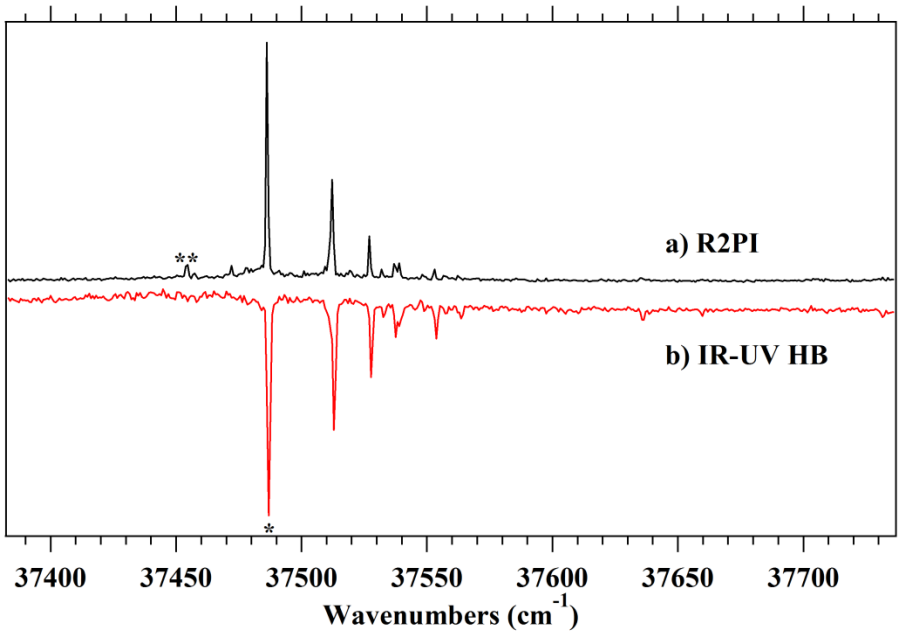


Figure 5.7. (a) R2PI spectrum of Ac-(γ_{ACHC})$_2$-NHBn taken in the monomer mass channel. (b) IR-UV holeburning spectrum recorded with IR holeburn laser fixed at 3336 cm^{-1}. Peak marked with a (*) was used to record the RIDIR spectrum, while the peak that did not burn out (marked with **) is a hot band or minor conformer not pursued further due to its weak intensity.

The S_0-S_1 origin for the sole $\gamma\gamma_{ACHC}$ conformer is red-shifted ~70 cm^{-1} from that of γ_{ACHC}. The observed Franck-Condon activity can be assigned to two low-frequency modes of frequency 26 and 41 cm^{-1} that involve rocking the $\gamma\gamma_{ACHC}$ peptide against the benzyl cap.

5.3.2.2 RIDIR Spectroscopy

Figure 5.8a and b present RIDIR spectra of the single conformer of $\gamma\gamma_{ACHC}$ in the NH stretch and amide I/II regions, respectively. The inset in Figure 5.8a shows that the free amide NH stretch region contains two transitions separated from one another by only

1.4 cm^{-1} (3467.8 and 3469.1 cm^{-1}). The band at 3336 cm^{-1} is due to a single, strong amide-amide H-bond. Based on the NH stretch spectrum, one can eliminate many of the possible structures taken up by unconstrained capped triamides studied in previous work, including C9/C9 sequential double rings, C14/C9 bifurcated double rings, and the triply hydrogen bonded C7/C7/C14 cycle.[26]

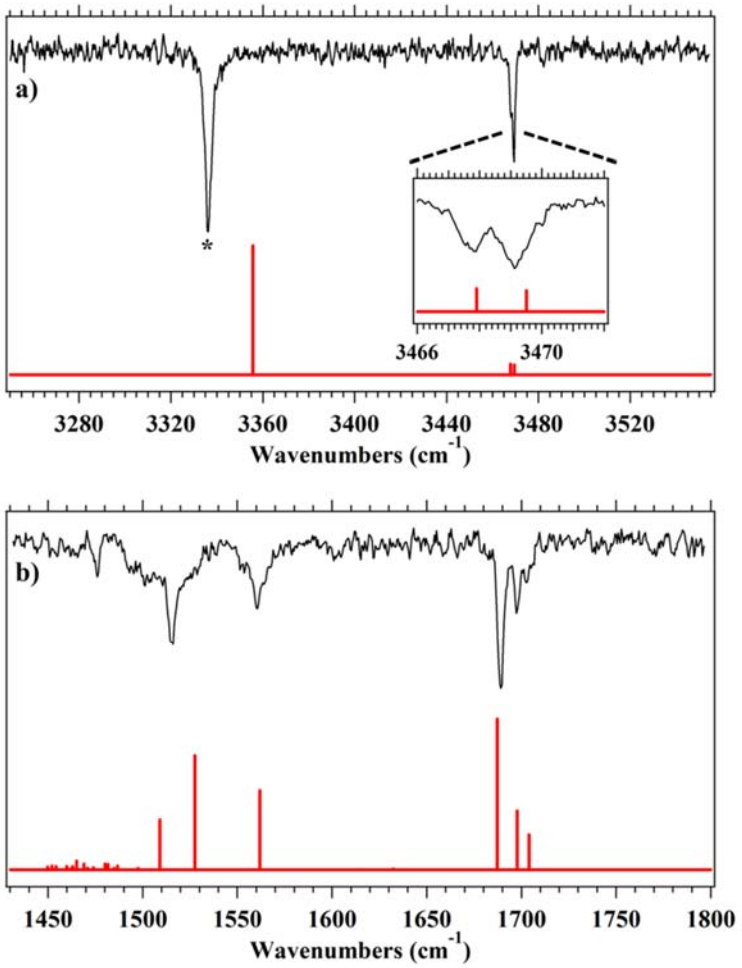


Figure 5.8. RIDIR spectrum for the observed conformer of Ac-(γ_{ACHC})$_2$-NHBn in the (a) NH stretch region, and (b) amide I/II regions. An expanded view of the free NH stretch region is shown as an inset. Calculated scaled harmonic vibrational frequencies (see text) and IR intensities of the assigned C14 conformation at the DFT M05-2X/6-31+G(d) level of theory are shown as stick diagrams in red. The transition marked by an asterisk was used in recording IR-UV holeburning spectra (see Figure 5.7).

Figure 5.8b shows the RIDIR spectrum in the amide I/II region. Due to spectral congestion in this region, the spectrum was recorded at higher resolution (~0.15 cm^{-1}) using a seeded Nd:YAG pump laser. A closely spaced triplet at 1703, 1697, and 1689 cm^{-1} is observed, with an asymmetric intensity pattern that suggests the presence of significant coupling between the C=O groups. Two amide II fundamentals at 1516, and 1561 cm^{-1} are readily resolved, while the third (~1501 cm^{-1}) is part of a congested and broadened shoulder on the low frequency side of the 1516 cm^{-1} band due to contributions from the CH bending modes which are in the same region.

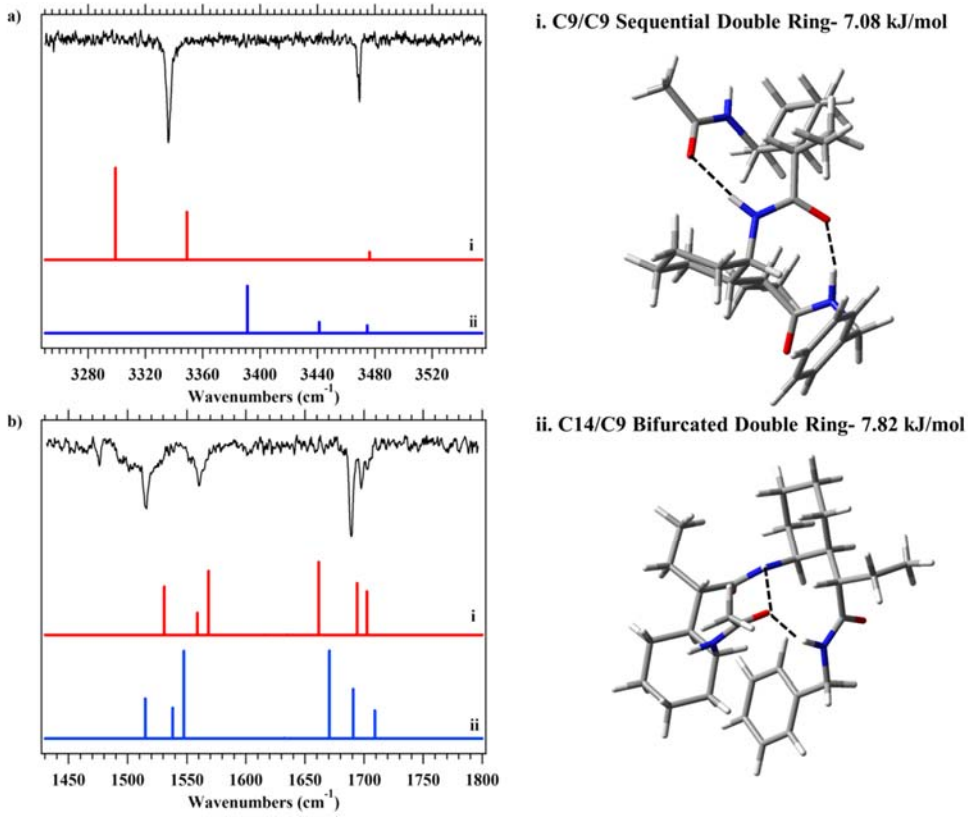


Figure 5.9. The experimental RIDIR spectra in the NH stretch (a) and amide I/II region (b) along with predicted spectra for two low-energy conformational families of $\gamma\gamma_{ACHC}$. The C9/C9 sequential double ring (i) and C14/C9 bifurcated double ring (ii) families are shown at right of the RIDRIS spectra.

The calculated global minimum conformation for $\gamma\gamma_{ACHC}$ is a C14 structure with the C-terminal NH forming a 14-membered H-bonded ring to the N-terminal carbonyl. The stick spectra shown below the experimental spectra in Figure 5.8 are those predicted by calculation for this global minimum structure. The match with experiment is excellent. Representative structures from two other low-energy families are shown and compared to experiment in Figure 5.9. The lowest energy C9/C9 sequential double rings (7.08 kJ/mol) and C14/C9 bifurcated double rings (7.82 kJ/mol) are significantly poorer fits to experiment, and are also higher in energy, enabling a firm assignment of the observed structure as a C14 single ring.

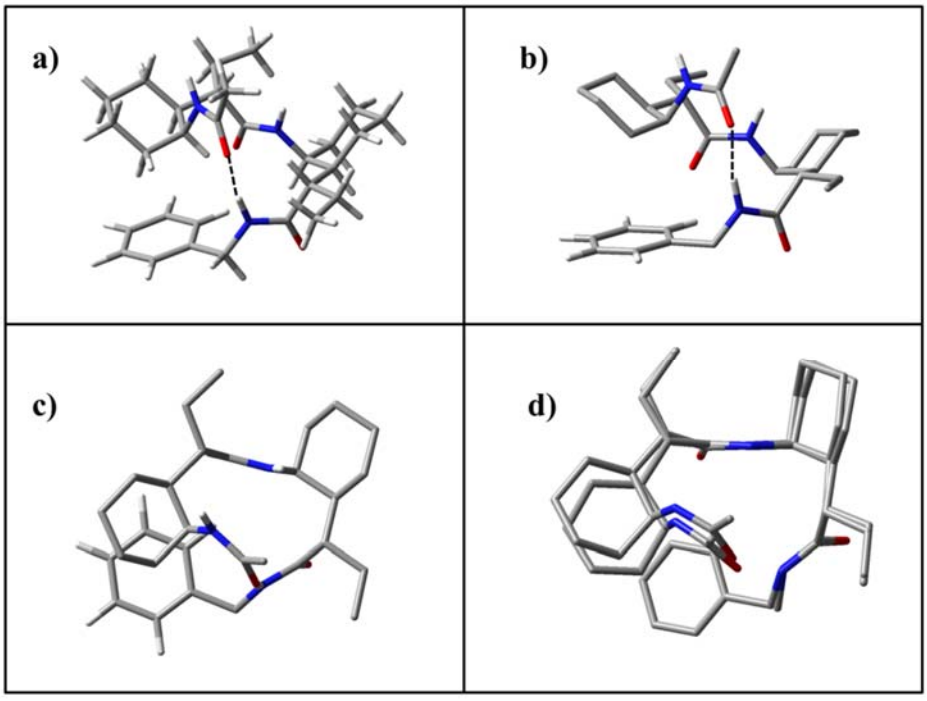


Figure 5.10. Structure of the assigned C14 single-ring structure for $\gamma\gamma_{ACHC}$ shown (a) with hydrogens present, and (b, c) two different views shown without hydrogens. (d) Overlay of the assigned structure with a segment of the crystal structure of compound 6 taken from ref. [14]. The Crystal structure was obtained online from the WebCSD.[44]

The assigned C14 single-ring structure is displayed in Figure 5.10 both with (Fig. 5.10a) and without hydrogens shown explicitly (Fig. 5.10b and c). The backbone engages in one full turn of a 14-helix that is held together principally by the 'end-on' C14 H-bond, with a calculated H-bond distance of 1.96 Å. The three amide groups are all oriented in the same direction, leading to a large dipole moment of 8.33 D. The cyclohexane rings and ethyl groups lock in (g+, g+) backbone geometries about both γ-peptide sub-units (Table 5.1). The two ethyl groups are both in the *anti*-configuration, while the phenyl ring takes up the g- orientation, much as it did in γ_{ACHC} and γ_{ACHC}-H_2O.

As anticipated, the intensity pattern in the amide I region is due to strong coupling between the C=O groups. The amide I transitions at 1689 and 1703cm^{-1} are caused by in- and out-of-phase combinations of the amide carbonyls involved in the C14 hydrogen bond (N- and C-terminal carbonyls) respectively, with the in-phase combination gaining intensity from the additive effect of the in-phase motion of the carbonyl stretches. The interior carbonyl group is minimally coupled to the other two groups, appearing at 1697cm^{-1}.

The amide II vibrations show a similar behavior, with the two free NH groups coupling to one another to form in- and out-of-phase combinations with frequencies of 1516 and 1501cm^{-1}. The H-bonded NH group (C-terminal) is shifted far up in frequency to 1561 cm^{-1}, and is largely uncoupled from the other two.

5.4 Discussion

5.4.1 Comparing Constrained and Unconstrained γ-Peptides in the Gas-Phase

In the introduction, we summarized the conformational preferences of a series of capped γ-peptides that had been studied previously by our group using single-conformation spectroscopy. Here we compare these preferences with the results obtained in the present work on the cyclically constrained analogs Ac-γ_{ACHC}-NHBn (γ_{ACHC}) and Ac-γ_{ACHC}-γ_{ACHC}-NHBn ($\gamma\gamma_{ACHC}$) in order to determine the effects of the cyclohexyl/ethyl group constraints on their folding.

5.4.1.1 Nearest-Neighbor Interactions in Ac-γ_{ACHC}-NHBn

We have determined that γ_{ACHC} has a strong energetic preference for C9 structures, funneling all its population into a single C9 ring structure by virtue of the built-in preference for a (g+, g+) structure due to the combined presence of ACHC and ethyl groups. Table 5.2 summarizes the preferences of the γ-peptides studied previously. The prototypical γ-dipeptide Ac-γ^2-hPhe-NHMe (γ) forms C9 and amide-stacked structures.[24] Ac-γ^2-hPhe-NMe_2 (γNMe_2) blocks C9 formation, leading to the formation exclusively of amide stacked structures. Ac-γ^2-hPhe-NH(iPr) (γiPR) breaks up the amide stack through the bulky nature of the isopropyl group, producing only C9 structures. Finally, incorporation of a gabapentin moiety, with its *spiro*-cyclohexyl group at Cβ (γ_{GABA}), when combined with-$N(Me)_2$ capping, leads exclusively to C7 structures.[25]

Table 5.2. Structural Parameters for the γ-peptide conformations studied in previous works.

Conformation[1]	Dihedral Angles (˚) (ϕ, θ, ζ, ψ)
Ac-γ^2-hPhe-NHMe[a,b]	
*Conformer A-C9(a)**	(99 , -70 , -73 , 105)
Conformer B-C9(g-)	(98 , -67 , -78 , 99)
*Conformer C-S(a)***	(-102 , 55 , -76 , 136)
Ac-γ^2-hPhe-NH(iPr)[a]	
*Conformer A-C9**	(99 , -71 , -73 , 107)
Gabapentin Derivative 1[a]	
Conformer A-C7[†]	(94 , 47 , 49 , 98)
Ac-γ^2-hPhe-γ^2-hAla-NHMe[b]	
Conformer A- C9/C9(a)	(98 , -70 , -72 , 100)
	(96 , -68 , -75 , 95)
Conformer B- C9/C9(g-)	(97 , -68 , -76 , 96)
	(96 , -67 , -75 , 94)
Conformer C- C7/C7/C14(g-)[‡c]	(-158 , 75 , -61 , 139)
	(81 , 73 , -52 , -90)
Ac-γ^2-hAla-γ^2-hPhe-NHMe[b]	
Conformer A- C9/C9(a)	(98 , -69 , -72 , 100)
	(99 , -66 , -75 , 95)
Conformer B- C9/C14(a)	(99 , -71 , -74 , 102)
	(-159 , -61 , -55 , 113)
Conformer B'- C9/C14(a)'	(-69 , 162 , -58 , -31)
	(-78 , -63 , -74 , 108)
Conformer C- C7/C7/C14(g-)[‡c]	(85 , 60 , -66 , -78)
	(-148 , 55 , -80 , 156)
C14[d]	(-154.5 , 60.2 , 59.5 , -126.8)
γAA[e]	(140.4 , -67.1 , -54.6 , 133.6)
H_9^I-Helix[f]	(98 , -70 , -75 , 98)
H_{14}^I-Helix[g]	(136.2 , -62 , -66 , 141)

[1]Naming schemes have been retained from the original references.
[a] Reference [25]
[b] Reference [26]
[c] Calculations performed using M05-2x/6-311++G(d,p)
[d] Reference [14]
[e] Reference [14], originally published in [15]
[f] Reference [17], Average of all dihedrals
[g] Reference [17], Average of last five residues
***** C9 H-bond lengths are 2.00 and 2.03Å respectively
[**] Distance between the planes of the amide stacked groups: 2.9 and 3.05Åfor the interior and exterior N-C spacing respectively.
[†]C7 hydrogen bond length is 2.04Å
[‡]C7/C7/C14 hydrogen bond lengths for the two structures are (2.166, 2.166, 2.057) and (2.152, 2.232, 2.027) respectively.

Comparison of the dihedral angles of the C9 structures (γ_{ACHC} in Table 5.1, γ and γiPr in Table 5.2) shows a general similarity between their backbone geometries, and hence to their C9 H-bonded rings. However, the presence or absence of constraints imposed among the γ-peptide backbone does lead to small changes in the approach geometry and H-bond distance within the set. For instance, the C9 NH···O distance is 1.93 Å in γ_{ACHC}, 2.00 Å in γ, and 2.03Å in γiPr. The constrained C9 ring, with its particularly short H-bond, also has the lowest frequency NH stretch fundamental (3301 cm^{-1}), ~50 cm^{-1} lower than the others.

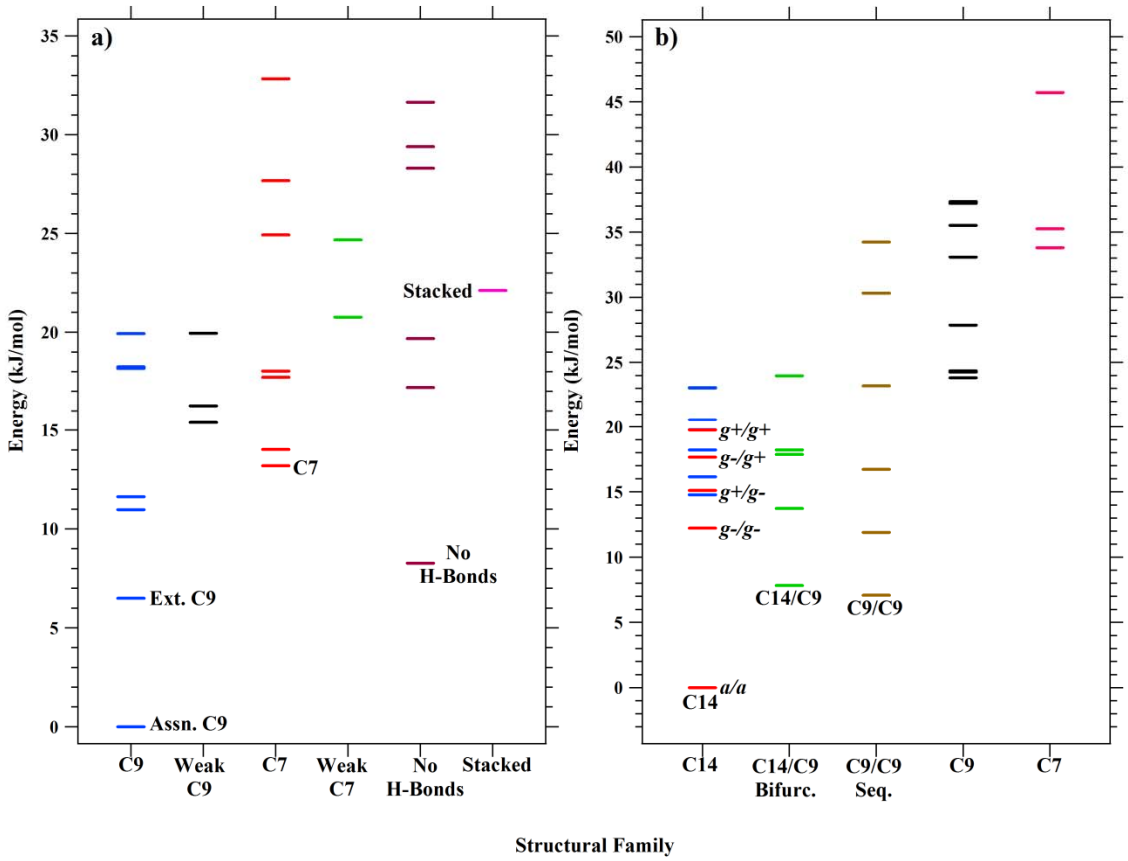


Figure 5.11. Energy level diagrams for the low-energy conformations of (a) Ac-γ_{ACHC}-NHBn and (b) Ac-$(\gamma_{ACHC})_2$-NHBn, grouped by structural family.

This fact is borne out by the energy level diagram for γ_{ACHC} shown in Figure 5.11a, which groups the conformational minima into structural families. The experimentally observed C9 structure (the global minimum), and its counterpart with a different phenyl ring orientation (6.50 kJ/mol), are two of only three structures in the first 10 kJ/mol. Interestingly, a structure with no intramolecular H-bonds is predicted by calculation to be third most stable (8.26 kJ/mol). This structure retains the central (g+, g+) preference in (θ, ζ), but orients the N-terminal amide group away from the optimal C9 hydrogen bond. By contrast, the C7 and amide stacked structures take up (g+, g-) geometries about (θ, ζ), and are pushed up in energy to 13.19 and 22.11 kJ/mol. The high energy of the amide stacked structure is notable, since in the unconstrained backbone of Ac-γ^2-hPhe-NHMe, the amide stacked configuration was able to compete with the C9 structure for stability (only 1.31 kJ/mol higher).[24]

Thus, the combined constraints present in the cyclohexyl ring/ethyl group combination lock in the C9 structure as the preferred nearest-neighbor interaction, in part by stabilizing the C9 ring, and in part by raising the energies of the alternative nearest-neighbor structures (C7 or amide stacked).

In future work, it would be interesting to consider a γ-peptide derivative in which the ACHC ring was retained, but the ethyl group removed, so that the ζ dihedral angle was left unconstrained. Calculations predict that the amide planes are more parallel for γ_{ACHC} (Table 5.3) than they are for the observed stacked conformation in γ (Table 5.2).[24,25] Preliminary results calculated at DFT M05-2X/6-31+G(d) show that when the ethyl groups are removed, the stacked geometry falls to 14.65 kJ/mol above the C9 with the ethyl group similarly removed.

Table 5.3. Structural parameters for other representative structures of γ_{ACHC} and $\gamma\gamma_{ACHC}$ optimized using M05-2X/6-31+G(d) level of theory.

	γ_{ACHC} (C9 Ext.)	γ_{ACHC} (No H-Bonds)	γ_{ACHC} (Stacked)
Backbone Dihedral Angles (˚) (ϕ, θ, ζ, ψ)	(-104 , 65 , 77 , -106) (g- , g+ , g+ , g-)	(-146 , 57 , 49 , -113) (a , g+ , g+ , g-)	(-115 , 63 , -72 , 117) (g- , g+ , g- , g+)
Sidechain Dihedral Angles (˚) (Et_{Nterm},Et_{Cterm},Phenyl)	(-178 , 99) (a , g+)	(-179 , 176) (a , a)	(56 , -96) (g+ , g-)
rC=O-NH (Å, [Cn_1,Cn_2])	1.97		
C=O←H (˚, [Cn1,Cn2]))	130		
O→HN (˚, [Cn1,Cn2]))	156		
Interior N-C (Å)			2.99
Exterior N-C (Å)			2.99

	γ_{ACHC} (C7)	$\gamma\gamma_{ACHC}$ (C9/C9)	$\gamma\gamma_{ACHC}$ (C14/C9)
Backbone Dihedral Angles (˚) (ϕ, θ, ζ, ψ)	(-161 , 60 , -92 , -121) (a , g+ , g- , a)	(-102 , 65 , 80 , -101) (g- , g+ , g+ , g-) (-103 , 64 , 78 , -101) (g- , g+ , g+ , g-)	(-109 , 63 , 78 , -123) (g- , g+ , g+ , a) (-166 , 56 , 53 , -130) (a , g+ , g+ , a)
Sidechain Dihedral Angles (˚) (Et_{Nterm},Et_{Cterm},Phenyl)	(172 , 92) (a , g+)	(-178 , -177 , -79) (a , a , g-)	(-175 , -50 , -99) (a , g- , g-)
rC=O-NH (Å, [Cn_1,Cn_2])	1.9	(1.97 , 1.89)	(2.11 , 2.41)
C=O←H (˚, [Cn1,Cn2]))	109	(131 , 125)	(129 , 120)
O→HN (˚, [Cn1,Cn2]))	155	(162 , 169)	(160 , 135)

In order to map out the conformational preferences of the ethyl group more explicitly, we carried out a relaxed potential energy scan for ethyl group hindered rotation for the C9 structure. The result is shown in Figure 5.12. As might be anticipated for an alkyl chain, the potential energy curve for the ethyl group shows three minima, with the *anti*-configuration substantially lower in energy than the two local minima associated with g+ and g- configurations (~12 kJ/mol higher in energy). The barrier between g+ and g- is quite large due to significant steric hindrance between the ethyl group and the ACHC ring. The relaxed potential energy scan also demonstrates that the barrier between *anti* and g+ is smaller than between *anti* and g-. This is likely caused by a steric interaction between the rotating ethyl group and the C-terminal carbonyl, which increases the barrier to internal rotation.

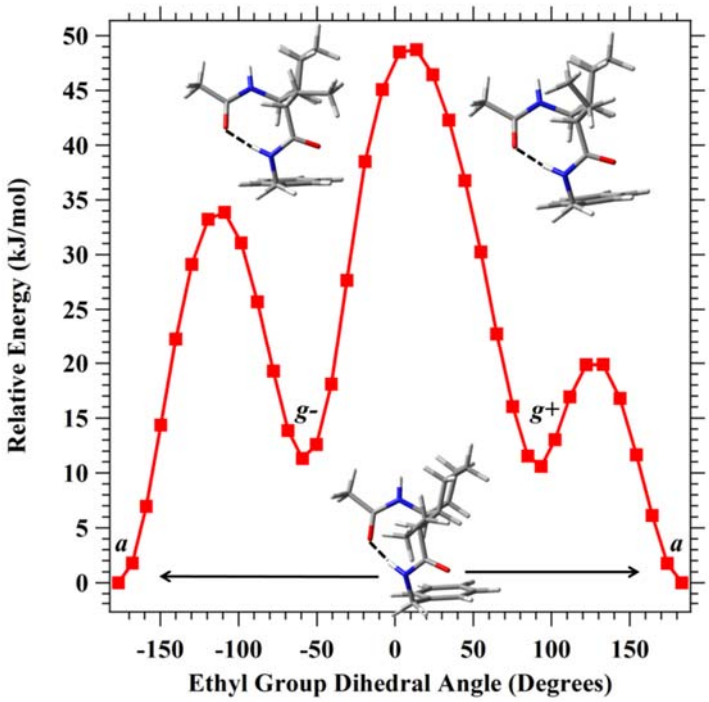


Figure 5.12. Relaxed potential energy scan about the Cα-Cβ bond of the ethyl side chain of Ac-γ_{ACHC}-NHBn. Three conformational minima *(anti, gauche+,* and *gauche-*) have the structures shown. The energies are not zero-point corrected.

5.4.1.2 Next Nearest-Neighbor Preferences in Ac-γ_{ACHC}-γ_{ACHC}-NHBn

The cyclically constrained capped triamide Ac-γ_{ACHC}-γ_{ACHC}-NHBn also shows strong preferences for a single conformation, with its entire conformational population residing in the gas phase in a C14 single-ring structure that is stabilized by a next-nearest-neighbor H-bond. Figure 5.11b shows an energy-level diagram for the low-energy conformations of $\gamma\gamma_{ACHC}$, separated into structural families. The experimentally observed C14 single-ring structure for $\gamma\gamma_{ACHC}$ is more than 5 kJ/mol below all others. It contains (a, a, g-) orientations for the ethyl group and phenyl ring orientations of the two γ-peptide sub-units, respectively. Other C14 single-rings that differ in the ethyl or phenyl ring positions are more than 10 kJ/mol higher in energy. The lowest energy C14/C9

bifurcated double ring and C9/C9 sequential double ring conformers are ~7 kJ/mol above the C14 single-ring global minimum. Both of these conformations adhere to the (g+, g+) backbone requirement, but show distinct differences in the (ϕ,ψ) angles that dictate the amide group orientations as these groups re-direct themselves to form additional nearest-neighbor H-bonds (Figure 5.9). The rigidity associated with the ACHC/ethyl constraints restricts the ability of the γ-peptide scaffold to form nearest neighbor H-bonds without introduction of a steric penalty, but facilitates maximal alignment of the C14 H-bond that closes one loop of the 14-helix.

An alternative viewpoint on the effects of the constrained backbone in $\gamma\gamma_{ACHC}$ is gained by considering the conformational preferences of two unconstrained triamides studied previously by our group: Ac-γ^2-hPhe-γ^2-hAla-NHMe (γFA) and Ac-γ^2-hAla-γ^2-hPhe-NHMe (γAF). Unlike $\gamma\gamma_{ACHC}$, γFA and γAF distribute their population between a total of seven conformations from three structural families: four C9/C9 sequential double rings, two C7/C7/C14 H-bonded cycles, and a single C9/C14 bifurcated double ring.[26] A summary of the important structural parameters for the C14 conformer of $\gamma\gamma_{ACHC}$ is found in Table 5.1, while the corresponding data for representative members of each structural type are collected in Table 5.2.

It is important to note first that while both the C7/C7/C14 cycles and the C9/C14 bifurcated double rings incorporate a C14 ring, there are no C14 single-ring structures observed in γFA and γAF. The flexibility of the γ-peptide backbone in the absence of the cyclohexyl/ethyl constraints is sufficient to enable nearest-neighbor C7 or C9 H-bonds to occur in addition to the C14 H-bond. As we noted previously[26], the C7/C7/C14 cycles can be viewed as frustrated amide stacked structures in which formation of a C14 H-bond

(thereby closing the cycle) blocks collapse into amide stacking by either nearest-neighbor amide group. The end result is that amide groups reverse their orientation along the triamide backbone, with dihedral angles that are (g+, g-) rather than the (g+, g+) values that are locked in by the ACHC/ethyl group constraints.

The bifurcated C9/C14 double ring in γAF is also made possible by the increased flexibility afforded by the γ^2 Ala and Phe side chains relative to the ACHC/ethyl constraints. As is typical of bifurcated double rings, formation of two H-bonds to the same C=O group weakens both of them, with H-bond lengths ~0.08 Å longer than the C9 or C14 single-rings preferred by γ_{ACHC} and $\gamma\gamma_{ACHC}$, respectively. In γAF, the flexibility of the backbone enabled formation of these two H-bonds with minimal steric penalty, but in $\gamma\gamma_{ACHC}$, where steric constraints are significant, one optimal C14 H-bonded ring is preferred over the two weakened C9 and C14 rings.

5.4.2 Comparing the C14 Ring with the 14-Helix

Although this marks the first gas-phase study of γ_{ACHC} and $\gamma\gamma_{ACHC}$, Guo et al. have carried out a seminal study of the structures of longer oligomers of γ_{ACHC} in solution using NMR, and in crystalline form via X-ray crystallography.[14] These γ-peptides have been observed to form helical structures stabilized by C14 hydrogen bonds oriented from C- to N-terminus, termed the 2.6_{14}-helix since it incorporates 2.6 residues per turn.[7,8,14,15] A summary of the relevant structural parameters for the helix is shown in Table 2. Since all amide groups align with their dipoles nominally along the helical axis, a large macro-dipole moment is generated that points from N- to C-terminus (+ to -).[45]

A comparison of the relevant structural parameters (Tables 5.1 and 5.2) between the C14 of $\gamma\gamma_{ACHC}$, and the crystal and NMR structures, reveal striking similarities between the C14 hydrogen bonds that are formed. This remarkable similarity is demonstrated pictorially in Figure 5.10d which overlays the C14 structure of $\gamma\gamma_{ACHC}$ with a portion of the γ_{ACHC} pentamer crystal structure previously reported by Guo et al.[14] The remarkable similarity demonstrated by the dihedral angles and structural overlap, show that the C14 single ring is acting as if it were a part of a longer 2.6_{14}-helix even though it is the minimal length structure that can form a single C14 H-bond. It should be noted that, while the cyclically-constrained γ-peptide backbone is pre-organized in such a way that it robustly forms the 2.6_{14}-helix, γ-peptide oligomers with less stringent constraints also form this helix, sometimes with as few as 4 residues.[7,8] However, NMR data on these less-constrained γ-peptides point to some degree of heterogeneity in the conformational make-up in solution.

Computational studies of the folding propensities of γ-peptides have been reported previously by Baldauf *et. al.*, who have calculated the relative stabilities of various types of helices formed by γ-peptide hexamers.[17] Of these prototypical helix types, the helices composed of C9 (labeled as H_9^I) and C14 H-bonds (H_{14}^I, equivalent to 2.6_{14}) are of direct relevance to the present work. For the hexamer with hydrogen sidechains, the H_9^I and H_{14}^I helices are within a few kJ/mol at both the HF/6-31G* and DFT B3LYP/6-31G* levels of theory. A summary of the structural parameters of these helices is found in Table 5.2.

The H_9^I helix type is not experimentally observed, but the observed C9 conformer of γ_{ACHC} (Table 5.1) and the calculated C9/C9 sequential double ring of $\gamma\gamma_{ACHC}$ (Figure 5.9 and Table 5.2) have structural parameters that agree very nicely with that predicted for the H_9^I helix by Baldauf *et al,*[17] The calculated structure for the H_{14}^I helix structure has γ-peptide backbone and H-bond lengths very similar to those in the C14 conformer of $\gamma\gamma_{ACHC}$. The calculations predict that the macro-dipole grows to nearly 20 D in the hexamer. Comparisons with both experiment and theory provide convincing evidence to support the conclusion that the C14 of $\gamma\gamma_{ACHC}$ is indeed the early stages of the 2.6_{14}-helix.

In summary, although the C14 single-ring structure falls just short of forming a full turn in $\gamma\gamma_{ACHC}$, the present results indicate that the 14-helix is the preferred structure for the isolated molecule in incipient form as soon as a C14 hydrogen bond is possible. This result is striking because α-peptide oligomers often require ten or more residues before any regular secondary structure is expressed, while γ_{ACHC} residues begin forming a stable secondary structure in the length it takes to form a single turn!

5.4.3 Complexation of the H_2O to γ_{ACHC}: Effect on Conformational Preferences

While the steric constraints imposed by the cyclohexyl and ethyl groups has led to a strong tendency for C9 nearest-neighbor and C14 next-nearest-neighbor interactions in the isolated molecule, it is still an open question whether these same preferences will still hold in aqueous solution, where amide-H_2O H-bonds can compete with amide-amide bonds. In the present work, we have taken one small step towards probing the specific effects of complexation of a single H_2O molecule to γ_{ACHC}.

The presence of even a single H_2O molecule can have fascinating consequences on the conformational preferences of a molecule stabilized by intramolecular H-bonds. Since H_2O can both donate and accept H-bonds, it often forms H-bonded bridges between H-bond donor and acceptor sites in the molecule to which it is bound.[31] If the H_2O molecule can form a stable bridge with minimal perturbation to the intramolecular H-bonds already present in the monomer, the conformational preferences of the monomer may be retained in the complex.[27,42,46] Alternatively, the H_2O molecule may disturb this intramolecular H-bonding architecture, sometimes by forming a new bridge by insertion between intramolecular donor and acceptor sites. In the most extreme case, the H_2O molecule may re-shape the molecule to which it is bound, giving it a structure not among those present in the isolated molecule.[47,48]

In the γ_{ACHC}-H_2O complex, the γ_{ACHC} molecule retains its preference for a C9 H-bonded structure, forming a double-donor bridge between the free C=O group and the aromatic π cloud. As we have seen, the backbone dihedral angles hardly change (Table 5.1) between γ_{ACHC} and γ_{ACHC}-H_2O. The single largest indicator of the structural change imparted by the water is a lengthening of the C9 hydrogen bond from 1.93 to 1.96Å. This small change in H-bond length leads to a shift up in frequency of the C9 NH stretch by ~18 cm^{-1}.

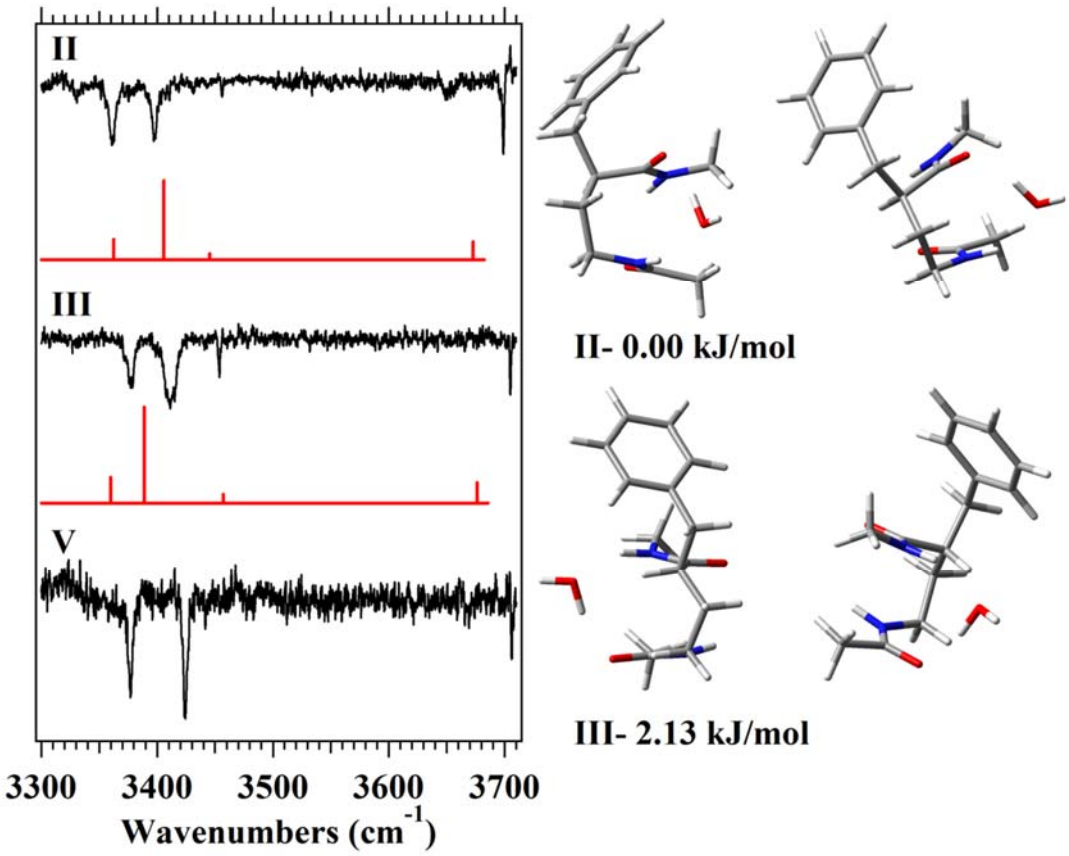


Figure 5.13. RIDIRS spectra in the NH and OH stretching region of the amide stacked γ-H_2O complexes along with the calculated spectra for the assigned conformations (red). To the right are the structures for the assigned conformations along with their relative energy.

Finally, the results on γ_{ACHC}-H_2O can be compared to the previous studies of its unconstrained analog, γ-H_2O.[27,49] Our intent in making this comparison is to see the ways in which the constrained backbone or differing phenyl ring positions in the two molecules affect the way in which H_2O binds to a C9 ring. Recall that γ (Ac-γ^2-hPhe-NHMe) possesses both C9 and amide stacked conformers in the absence of H_2O.[24] In the γ-H_2O complex, six conformational isomers of γ-H_2O are observed, three of which (conformers I, IV, and VI) retain a C9 H-bonded ring structure, with H_2O bound to the

free C=O group. The remaining three structures (conformers II, III, and V, Figure 5.13) have H_2O bound to the amide stacked structure, forming a bridge across the two amide planes.[27,49]

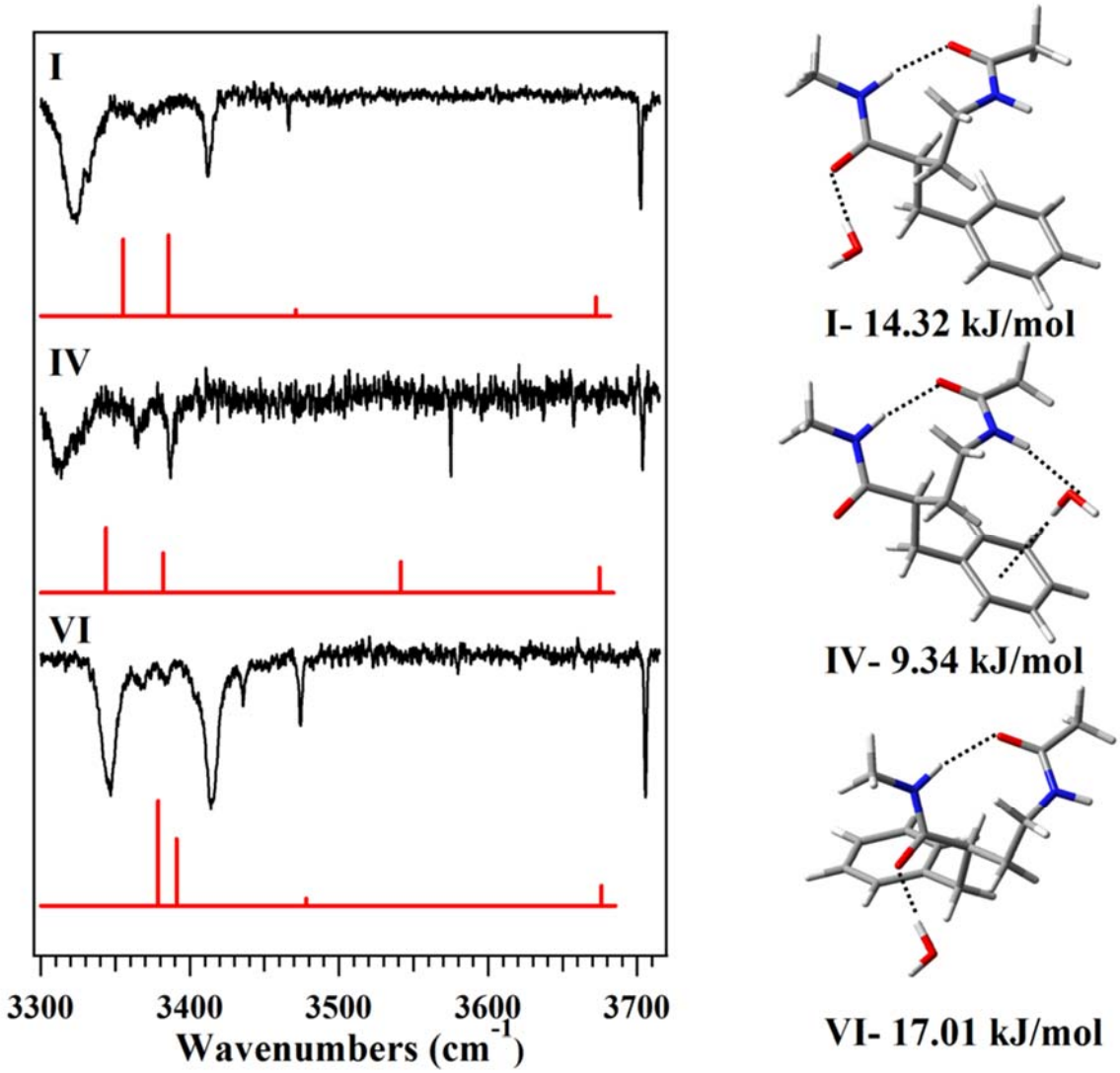


Figure 5.14. RIDIR spectra of the three C9 conformers of the unconstrained γ-H_2O complex taken in the hydride stretch region, compared with the best-fit spectra predicted by DFT M05-2X/6-31+G(d) level of theory. The structures assigned to these conformers are shown to the right.

Figure 5.14 presents RIDIR spectra of the three C=O bound C9 complexes of γ-H_2O in the NH/OH stretch regions, for comparison with that found in γ_{ACHC}-H_2O. Stick diagrams of the calculated NH/OH stretch harmonic frequencies and IR intensities are shown below each experimental spectrum, with the structures assigned to each spectrum shown to the right. Conformers I and VI incorporate the H_2O molecule as H-bond donor to the C-terminal C=O, differing only in the position of the phenyl ring. Conformer IV has the H_2O molecule acting as H-bond acceptor from the N-terminal NH, forming a bridge to the π cloud of the phenyl ring. By contrast, in γ_{ACHC}-H_2O, the H_2O molecule forms a double-donor bridge, with one OH bound to the C-terminal C=O and the other forming a π H-bond with the aromatic ring of the NH-benzyl cap. This unique binding motif is one that is probably dictated in large measure by the unique position of the phenyl ring in γ_{ACHC}, which has the phenyl ring a part of the C-terminal cap (Figure 5.1), and thus places the phenyl ring in an unusual position relative to the phenylalanine side chain in Ac-γ^2-hPhe-NHMe. It is also noteworthy that the single γ_{ACHC}-H_2O structure we observe for the constrained γ-peptide is also the global minimum, while in γ-H_2O, the lowest energy structures are amide-stacked, with the C9 structures calculated to be at least 9 kJ/mol less stable. In γ_{ACHC}-H_2O, the H_2O molecule binds to the single observed monomer structure, which also can accommodate a H_2O molecule in a bridge that makes it the global minimum. Thus, we see that in both γ-H_2O and γ_{ACHC}-H_2O complexes, the H_2O molecule takes up a range of positions that change the main C9 peptide backbone structure only slightly, indicating a robustness of the preference of the γ-peptide backbone for formation of C9 rings even in the presence of H_2O.

5.5 Conclusions

Conformational assignments of the gas-phase conformations of γ_{ACHC}, γ_{ACHC}-H_2O and $\gamma\gamma_{ACHC}$ were made by comparing the results of single-conformation IR spectra with the predictions of quantum chemical calculations. A single conformation of each of the molecules was observed in the cold, isolated environment of the molecular beam. γ_{ACHC} was found to fold exclusively into a C9 H-bonded ring conformation. This same C9 conformation is retained upon complexation to water, with the water molecule forming a bridge between the C-terminal carbonyl and the π-cloud, donating H-bonds to each.

The primary effect of the ACHC/ethyl constraint imposed on the γ-peptide backbone is to limit backbone structures to those having the g+/g+ orientation for θ and ζ. This backbone constraint effectively eliminates many of the conformations previously observed for the unconstrained γ-peptides from among the low-energy possibilities, including the C7 H-bonded ring and the amide stacked geometry. At the same time, the constrained backbone is near-optimal for formation of a strong C9 nearest-neighbor H-bond, leading to a C9 NH stretch fundamental in γ_{ACHC} that is lowered in frequency relative to its unconstrained counterpart.

In the capped triamide $\gamma\gamma_{ACHC}$ the single conformation observed was a C14 single-ring structure that closely mimics the first turn of a 14-helix. Once again, the (g+, g+) geometry built into the central dihedral angles by the ACHC/ethyl constraints enables formation, after adjustment of (ϕ,ψ), of a next-nearest-neighbor C14 interaction. This structure bears a close resemblance to the first turn of a 2.6_{14} helix identified by previous studies in non-polar solvent and in crystalline form. Indeed, the overlay of the two structures shown in Figure 5.10d is remarkable. Thus, we see that the constraints

built into the γ-peptide backbone by the combined presence of the cyclohexyl ring at Cβ–Cγ and ethyl group at Cα "lock in" the first stages of a 14-helix when the γ-peptide achieves the minimal structural capacity to do so. The effects of the constraints on the backbone is also seen in the computational studies that predict a sparse energy level structure in which C9 (in γ_{ACHC}) and C14 (in $\gamma\gamma_{ACHC}$) structures are significantly more stable than all others.

The present results point the way for future work along several directions. First, this study has probed only the effect of a single H_2O molecule on the smallest capped diamide in forming γ_{ACHC}-H_2O. It would be interesting to study the evolution of γ-peptide conformational preferences and the networks of H_2O-H_2O hydrogen bonds that form around it, as a function of number of H_2O molecules in γ_{ACHC}-$(H_2O)_n$ clusters. Extension to $\gamma\gamma_{ACHC}$ would also be useful in seeing whether the C14 single ring will be compromised or stabilized by addition of one or more H_2O molecules. Second, the recent studies of γ^4-substituted peptides[9-12] have raised the possibility that this simpler structural motif may also be preorganized for 14-helix formation, and may provide for reasonable mimicry of the α-helix when incorporated into α/β/γ-peptides.[13,50,51]

Finally, these constrained γ-peptide sub-units have also been incorporated into model α/γ- and β/γ-peptides whose conformational preferences are also being explored.[52] One intriguing outcome of this series of studies will be the further opportunity they provide for understanding the amide I/II regions of the infrared. By projecting the calculated normal modes onto a local mode basis, extraction of the local site frequencies and intermode couplings is possible.[37] These results can then also serve as exquisite tests

of current electrostatic models and their eventual extension beyond α-peptides to the full set of peptidomimetic synthetic foldamers.

5.6 References

(1) Gellman, S. H. *Acc. Chem Res.* **1998**, *31*, 173-180.

(2) Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. *Nat Chem Biol* **2007**, *3*, 252-262.

(3) Guichard, G.; Huc, I. *Chem. Commun.* **2011**, *47*, 5933-5941.

(4) Bautista, A. D.; Craig, C. J.; Harker, E. A.; Schepartz, A. *Curr. Opin. Chem. Bio.* **2007**, *11*, 685-692.

(5) Seebach, D.; Overhand, M.; Kühnle, F. N. M.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. *Helv. Chim. Acta* **1996**, *79*, 913-941.

(6) Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118*, 13071-13072.

(7) Hanessian, S.; Luo, X.; Schaum, R.; Michnick, S. *J. Am. Chem. Soc.* **1998**, *120*, 8569-8570.

(8) Hintermann, T.; Gademann, K.; Jaun, B.; Seebach, D. *Helv. Chim. Acta* **1998**, *81*, 983-1002.

(9) Bandyopadhyay, A.; Jadhav, S. V.; Gopi, H. N. *Chem. Commun.* **2012**, *48*, 7170-7172.

(10) Basuroy, K.; Dinesh, B.; Shamala, N.; Balaram, P. *Angew. Chem., Int. Ed.* **2012**, *51*, 8736-8739.

(11) Bandyopadhyay, A.; Gopi, H. N. *Org. Lett.* **2012**, *14*, 2770-2773.

(12) Jadhav, S. V.; Bandyopadhyay, A.; Gopi, H. N. *Org. Biomol. Chem.* **2013**, *11*, 509-514.

(13) Ananda, K.; Vasudev, P. G.; Sengupta, A.; Poopathi Raja, K. M.; Shamala, N.; Balaram, P. *J. Am. Chem. Soc.* **2005**, *127*, 16668-16674.

(14) Guo, L.; Zhang, W.; Reidenbach, A. G.; Giuliano, M. W.; Guzei, I. A.; Spencer, L. C.; Gellman, S. H. *Angew. Chem., Int. Ed.* **2011**, *50*, 5843-5846.

(15) Seebach, D.; Brenner, M.; Rueping, M.; Schweizer, B.; Jaun, B. *Chem. Commun.* **2001**, 207-208.

(16) Seebach, D.; Brenner, M.; Rueping, M.; Jaun, B. *Chem.--Euro. J.* **2002**, *8*, 573-584.

(17) Baldauf, C.; Günther, R.; Hofmann, H.-J. *Helv. Chim. Acta* **2003**, *86*, 2573-2588.

(18) Vasudev, P. G.; Chatterjee, S.; Shamala, N.; Balaram, P. *Acc. Chem Res.* **2009**, *42*, 1628-1639.

(19) Vasudev, P. G.; Chatterjee, S.; Shamala, N.; Balaram, P. *Chem. Rev.* **2010**, *111*, 657-687.

(20) Baquero, E. E.; James, W. H., III; Choi, S. H.; Gellman, S. H.; Zwier, T. S. *J. Am. Chem. Soc.* **2008**, *130*, 4784-4794.

(21) Baquero, E. E.; James, W. H., III; Choi, S. H.; Gellman, S. H.; Zwier, T. S. *J. Am. Chem. Soc.* **2008**, *130*, 4795-4807.

(22) James, W. H., III; Baquero, E. E.; Shubert, V. A.; Choi, S. H.; Gellman, S. H.; Zwier, T. S. *J. Am. Chem. Soc.* **2009**, *131*, 6574-6590.

(23) James, W. H., III; Baquero, E. E.; Choi, S. H.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *114*, 1581-1591.

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

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

(26) James, W. H., III; Buchanan, E. G.; Müller, C. W.; Dean, J. C.; Kosenkov, D.; Slipchenko, L. V.; Guo, L.; Reidenbach, A. G.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2011**, *115*, 13783-13798.

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

(28) Guo, L.; Chi, Y.; Almeida, A. M.; Guzei, I. A.; Parker, B. K.; Gellman, S. H. *J. Am. Chem. Soc.* **2009**, *131*, 16018-16020.

(29) Lubman, D. M. *Mass Spectrom. Rev.* **1988**, *7*, 559-592.

(30) Page, R. H.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988**, *88*, 4621-4636.

(31) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827-8839.

(32) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205-241.

(33) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440-467.

(34) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287-303.

(35) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289-300.

(36) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J.V.; Cioslowski, J.; Fox, D. J.; Revision C.01 ed.; Gaussian, Inc.: Wallingford CT, 2009.

(37) Buchanan, E. G.; James, W. H., III; Choi, S. H.; Guo, L.; Gellman, S. H.; Muller, C. W.; Zwier, T. S. *J. Chem. Phys.* **2012**, *137*, 094301-094316.

(38) Gotch, A. J.; Zwier, T. S. *J. Chem. Phys.* **1992**, *96*, 3388-3401.

(39) Pribble, R. N.; Zwier, T. S. *Science* **1994**, *265*, 75-79.

(40) Gierszal, K. P.; Davis, J. G.; Hands, M. D.; Wilcox, D. S.; Slipchenko, L. V.; Ben-Amotz, D. *J. Phys. Chem. Lett.* **2011**, *2*, 2930-2933.

(41) Frost, R. K.; Hagemeister, F. C.; Arrington, C. A.; Schleppenbach, D.; Zwier, T. S.; Jordan, K. D. *J. Chem. Phys.* **1996**, *105*, 2605-2617.

(42) Biswal, H. S.; Loquais, Y.; Tardivel, B.; Gloaguen, E.; Mons, M. *J. Am. Chem. Soc.* **2011**, *133*, 3931-3942.

(43) Shimanouchi, T. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg MD, 20899, 1972.

(44) Thomas, I. R.; Bruno, I. J.; Cole, J. C.; Macrae, C. F.; Pidcock, E.; Wood, P. A. *J. Appl. Crystallogr.* **2010**, *43*, 362-366.

(45) Seebach, D.; Hook, D. F.; Glättli, A. *Peptide Science* **2006**, *84*, 23-37.

(46) Cirtog, M.; Rijs, A. M.; Loquais, Y.; Brenner, V.; Tardivel, B.; Gloaguen, E.; Mons, M. *J. Phys. Chem. Lett.* **2012**, *3*, 3307-3311.

(47) Shubert, V. A.; James, W. H.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *113*, 8055-8066.

(48) Shubert, V. A.; Müller, C. W.; Zwier, T. S. *J. Phys. Chem. A* **2009**, *113*, 8067-8079.

(49) James, W. H., III, Purdue University, 2009.

(50) Shin, Y.-H.; Mortenson, D. E.; Satyshur, K. A.; Forest, K. T.; Gellman, S. H. *J. Am. Chem. Soc.* **2013**, *135*, 8149-8152.

(51) Sawada, T.; Gellman, S. H. *J. Am. Chem. Soc.* **2011**, *133*, 7336-7339.

(52) Kusaka, R.; Zhang, D.; Walsh, P. S.; Gord, J. R.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2013**.

CHAPTER 6. CONFORMATION-SPECIFIC SPECTROSCOPY OF CAPPED GLUTAMINE-CONTAINING PEPTIDES: ROLE OF A SINGLE GLUTAMINE RESIDUE ON PEPTIDE BACKBONE PREFERENCES

6.1 Introduction

Neurodegenerative diseases, such as Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD), are illnesses that ravage both mental acuity and motor function of affected individuals.[1-6] Although these illnesses involve different precursor proteins [β-amyloid (Aβ) and tau in AD,[3,4] α-synuclein in PD,[4,6] and huntingtin (htt) in HD],[5,7] resulting in different symptoms and pathogenesis, the structures of the toxic species appears to be related to one another across all of these diseases.[3,4,6]

This common taxonomy involves β-sheet oligomers that ultimately form insoluble amyloid fibrils.[8] These large β-sheet assemblies are held together by intermolecular H-bonds and by intra- and intermolecular sidechain-to-sidechain interactions. Seminal work by Eisenberg and co-workers,[9,10] and Bowers and co-workers[11] have provided detailed insight into both the structure and aggregation propensity of several short segments derived from amyloid-forming peptides (e.g., the hexamer Asn-Asn-Gln-Gln-Asn-Tyr from the yeast prion protein Sup35). Recently, 2D-IR experiments have been used to elucidate the structure of human islet amyloid polypeptide (hIAPP),[12-14] and γD-crystallin,[15] which both aggregate into fibril structures.

Glutamine is commonly found in many regions of proteins that form fibrilar structures, including ones responsible for neurodegenerative diseases. In Huntington's disease (HD), glutamine takes center stage. Occurrence of HD is associated with the presence of excess Gln in the protein huntingtin. This protein normally contains a Gln-repeat segment of 5 to 36 residues; increasing length of the poly-Gln segment is linked to disease pathogenesis.[4,5,7,16-18]

The flexible amide-containing side chain of glutamine (Gln, Q) is believed to be vitally important in forming the polar-zipper motif commonly associated with fibril structures.[18] Several computational and experimental studies have sought to understand the structural preferences of glutamine containing peptides and their aggregates.[18-25] Studies on larger glutamine repeat sequences point to formation of a β-hairpin structure comprised of β-sheet and β-turn regions.[26-28] In these situations the glutamine sidechain plays an integral role in aggregation by forming sidechain-to-backbone hydrogen bonds between successive β-hairpin structures, locking in the characteristic polar-zipper structure.[8-10,29]

In a fully-developed glutamine-rich peptide sequence, the presence of successive glutamine residues, each with a flexible amide-containing sidechain, creates many possibilities for sidechain-to-sidechain and sidechain-to-backbone hydrogen bonds. The presence of water further expands the types of H-bonded networks that can be formed. In order to achieve a molecular-scale understanding of the aggregation behavior of glutamine-rich peptides, one must develop quantitatively accurate molecular mechanics force fields. Such force fields would properly describe the delicate balance of attractive

forces and steric constraints that control the conformations, solubility and aggregation behavior of glutamine-rich peptides.

The development of accurate force fields benefits from experimental measurements that provide incisive tests of computational accuracy. Over the past several years, a number of groups world-wide have developed and utilized a powerful set of laser spectroscopy tools for studying the conformational preferences of short peptide sequences in the gas phase.[30-35] By studying the peptides in the absence of solvent, the inherent conformational preferences of the peptide are directly exposed, making a direct link with *ab initio* and force field calculations carried out on the isolated molecules. In the present paper, we focus on an important question that serves as a foundation for such theoretical developments; namely, 'What is the effect of a single glutamine residue on the inherent conformational preferences of a peptide backbone?'

We begin to address this question by recording the conformation-specific IR and UV spectra of several capped glutamine-containing peptides in the cold, collision-free environment of a molecular beam. Figure 6.1 shows the molecules which were studied, including three capped amino acids Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn) and one capped dipeptide (Ac-Ala-Gln-NHBn). These molecules are capped either at the N-terminus (with a benzyloxycarbonyl cap, labeled as 'Z'), or at the C-terminus (with an NH-benzyl cap, labeled as 'NHBn'), providing the aromatic chromophore needed for IR-UV double resonance techniques that enable the acquisition of IR and UV spectra of single conformations free from interference from other conformations present in the expansion. Conformation-specific spectroscopy has proven to be a powerful tool for revealing the inherent conformational preferences of both natural amino acids and

synthetic foldamers. In studies of this type to date, the spectroscopic signatures of incipient forms of several prototypical α-amino acid secondary structures have been obtained, including γ-turns, β-turns, β-hairpins, extended conformations, and helices.[36-41]

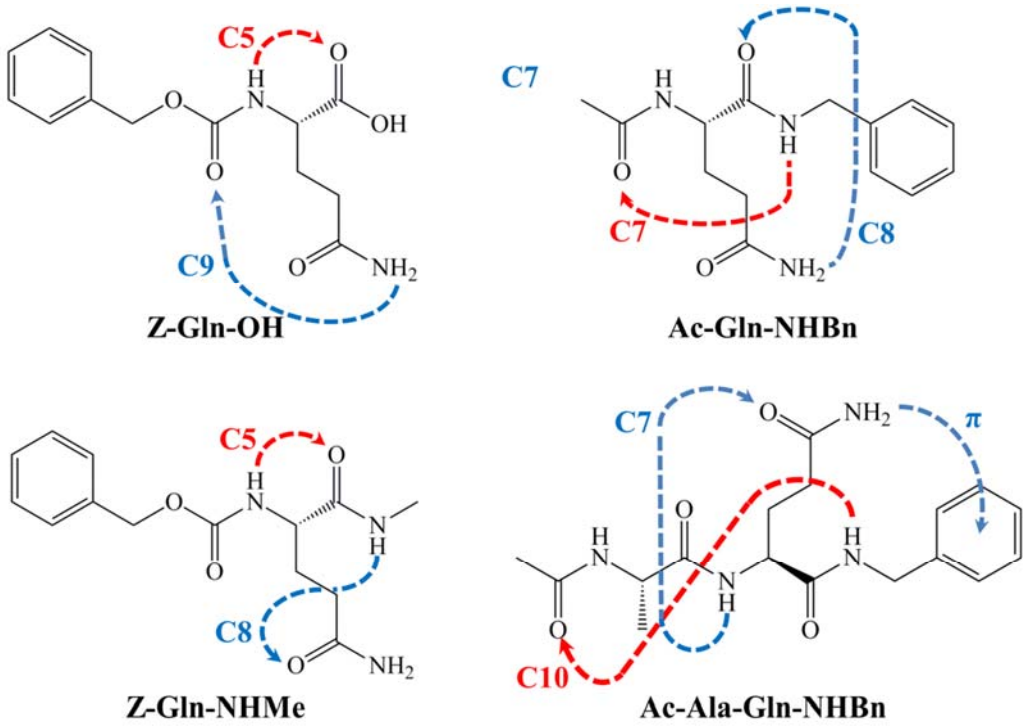


Figure 6.1. Chemical structures of the molecules studied in this work and example hydrogen bonds.

Analogous studies on synthetic foldamers (e.g., β-peptides, γ-peptides) have revealed the effects of increased backbone flexibility on conformational preferences, including the formation of unique secondary structures facilitated by the synthetic building-blocks of which they are composed.[42-46] The studies carried out by our group on

constrained and unconstrained γ-peptides are particularly relevant to Gln-containing peptides, because the spacing between backbone amides across a single γ-residue is identical to the spacing between the Gln sidechain amide group and the backbone amide involving the Gln α nitrogen.[47-51] In this sense, we can look to the results of previous studies on γ-peptides as a guide to some of the sidechain – backbone interactions we might anticipate within a single Gln residue. In particular, it is possible that the Gln sidechain may engage in amide stacking with the peptide backbone, given that a stacking interaction involving comparably spaced amide groups have been observed in small γ-peptides.[47,48]

In addition to amide stacking, there is a range of different sidechain-to-backbone and backbone-to-backbone hydrogen bonds that can be formed in short peptides that contain at least one Gln residue. These hydrogen bonds may compete with one another, as is shown in Figure 6.1. Our over-arching goal of understanding the effect of a single glutamine residue on a peptide's conformational preferences can thus be served by identifying the preferred H-bonding and/or stacking architectures adopted by short peptides that contain a Gln residue. As we shall see, the set of four molecules shown in Figure 6.1 exhibit conformational preferences that result in a rich variety of structures. While some of these observed structures stabilize the extended peptide backbone conformation that is characteristic of β-sheets, the Gln residue is also capable of stabilizing both γ- and β-turns, sometimes via

6.2 Experimental

6.2.1 Experimental Methods

6.2.1.1 Sample Handling

Benzyloxycarbonyl-Gln-OH (Z-Gln-OH) was purchased from Bachem and used without purification. Z-Gln-OH was converted to Z-Gln-NHMe using a published method[48]; this material was used for spectroscopic studies without purification. Ac-Gln-NHBn and Ac-Ala-Gln-NHBn were synthesized by the Gellman research group at the University of Wisconsin-Madison. To avoid sample degradation caused by thermal heating, we used a laser desorption source to bring the molecule of interest into the gas phase. A detailed description of the laser desorption source can be found in Chapter 2.[34,35,37,52-54]

6.2.1.2 Spectroscopic Techniques

The chamber and spectroscopic methods have been extensively detailed in Chapter 2. Only information specific to this set of experiments is provided in this section.

UV spectra were recorded for each sample using resonant two-photon ionization (R2PI).[53] The dye laser operated with Coumarin 540A (Exciton) laser dye to generate UV light in the origin region of the –NHBn chromophore (~265-270 nm). Conformation-specific IR spectra were recorded using resonant ion-dip infrared (RIDIR) spectroscopy.[34,35,51,54] The pump laser for the optical parametric converter could be

operated with, or without an injection seeder which influenced the final resolution of the IR light. This detail will be noted in the figure captions for each of the RIDIR spectra presented. RIDIR spectra were recorded in the hydride stretching region (~3000-3750 cm^{-1}) and the amide I/II region (~1400-1800 cm^{-1}). Conformation-specific UV spectra were recorded using infrared-ultraviolet holeburning (IR-UV HB) spectroscopy.[35]

6.2.2 Computational Methods

A set of starting structures was generated using the conformational search algorithm in the MACROMODEL suite of programs.[55] This conformational search used the Amber* force field[56] to find unique structures within a 50 kJ/mol energy window of the global minimum structure. For Z-Gln-OH separate conformational searches were performed on the *cis*- and *trans*-conformations of the –OH moiety in the carboxylic acid cap. From these structures the lowest ~30-100 conformations for each molecule were further optimized using density functional theory (DFT) calculations that employed the hybrid density functional M05-2X[57] with the 6-31+g(d) basis set. DFT calculations were performed using Gaussian09.[58] Harmonic vibrational frequencies were calculated for the optimized geometries. These frequencies were scaled by 0.9398 or 0.942 (Z-Gln-NHMe and all others, respectively), 0.951, and 0.960 for the NH-stretch, OH-stretch, and amide I/II regions, respectively, to account for anharmonicity in the harmonic frequencies.[33]

6.2.3 Structural Information and Nomenclature

The hydrogen bonds present in a given optimized structure are described using the common C*n* notation, where *n* denotes the number of atoms which compose the hydrogen bonded ring. In this work, we introduce a naming scheme in which backbone-to-backbone hydrogen bonds are listed first, separated by a single slash, with backbone H-bonds listed after a double slash in italics. For example, "C10//*C7/π*" denotes a structure in which a ten-membered hydrogen bonded ring is formed between two backbone amide groups, with the glutamine side chain involved in a double-bridge structure to the peptide backbone comprised of C7 and NH···π H-bonds. Crystal and NMR structures were taken from the RCSB Protein Data Bank[59] and the PDBeMotif[60] search tool from the European Bioinformatics institute.

6.3 Results and Analysis

6.3.1 Z-Glutamine-X (X= –OH and –NHMe)

6.3.1.1 UV Spectroscopy

R2PI and IR-UV holeburning spectra were recorded for both Z-Gln-NHMe and Z-Gln-OH in the parent mass channels (m/z= 293 and 280, respectively) over the S_0-S_1 origin region of the Z-cap (~37480-37665 cm^{-1}). These spectra are shown in Figure 6.2a and b, respectively. The R2PI and IR-UV holeburning spectra of both molecules show a series of sharp UV transitions that ride atop a broad background signal likely arising from

incomplete cooling. In Z-Gln-NHMe, the strong transition at 37602 cm^{-1} (marked with an asterisk) was used to record the RIDIR spectra which will be discussed in a later section (Section 6.3.1.3). The IR-UV holeburning spectrum recorded with IR holeburn laser fixed at 3560 cm^{-1} shows that all transitions and the broad background burn out together. Holeburning scans taken with hole-burn laser fixed on other IR transitions produced the same spectrum, indicating that a single conformation of Z-Gln-NHMe is present in the expansion. As a final check, RIDIR spectra were recorded on several different UV transitions, all of which showed the same infrared spectrum.

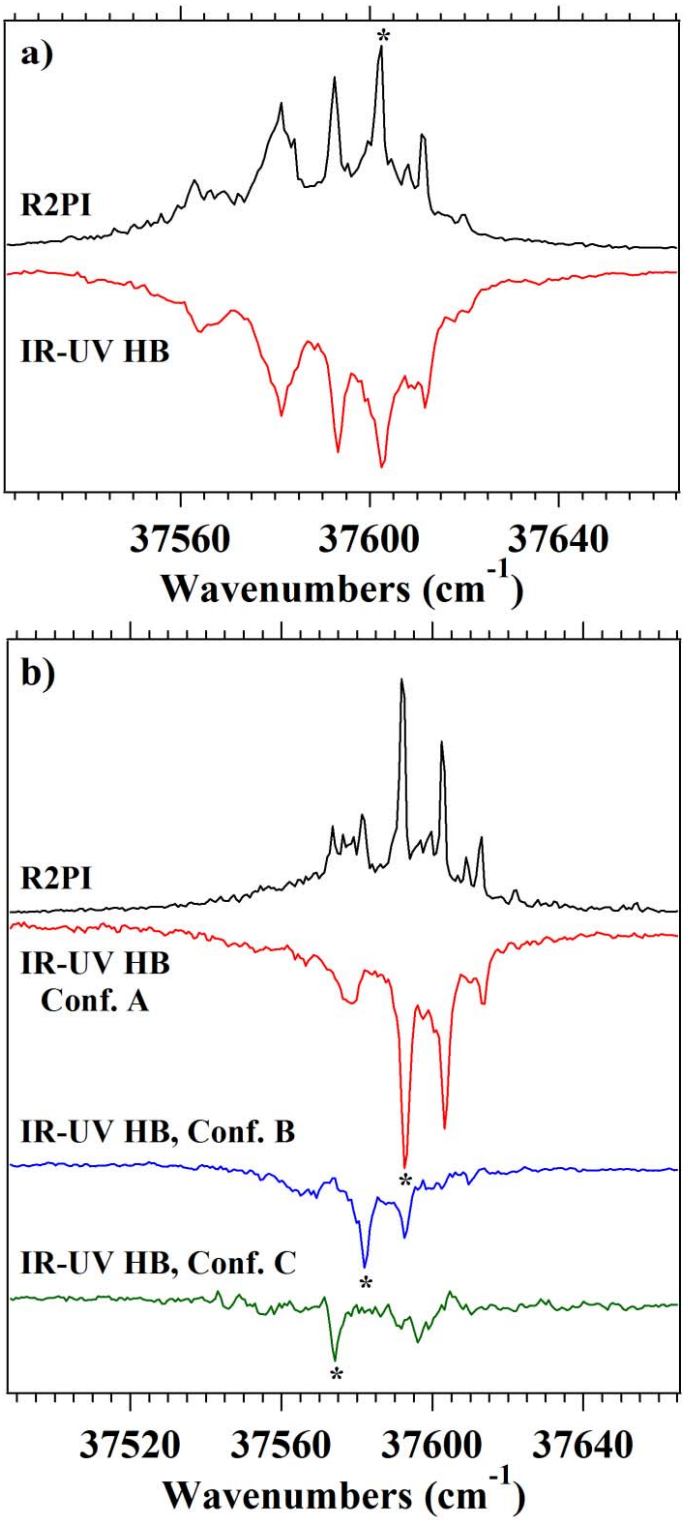


Figure 6.2. R2PI and IR-UV holeburning spectra for (a) Z-Gln-NHMe, and (b) Z-Gln-OH. Transitions marked with asterisks were used to record RIDIR spectra.

The R2PI (Figure 6.2b) spectrum of Z-Gln-OH shows a smaller background and narrower UV transitions indicative of better cooling of the laser-desorbed molecules in this case. IR-UV holeburning spectroscopy reveals the presence of three unique conformations (Figure 6.2b). The three strongest transitions at 37593, 37582 and 37574 cm^{-1} are assigned as the S_0-S_1 origin transitions for conformers A, B, and C, respectively, and were used to collect RIDIR spectra (Section 6.3.1.2). The broad transition at ~37578 cm^{-1} (Figure 6.2b) is believed to be a hot band of Conformer A, indicating the cooling is still not entirely complete. The proximity of the hot band due to A to the origin transitions of Conformers B and C will contribute to interferences in the RIDIR spectra recorded for these conformers. Minimal Franck-Condon activity for the three conformers indicates a minimal geometry change upon electronic excitation for all three conformers.

6.3.1.2 Z-Glutamine-OH: RIDIR Spectroscopy and Conformational Assignment

Figure 6.3 shows the RIDIR spectra recorded for the three assigned conformers of Z-Gln-OH in the NH/OH-stretch (Figure 6.3a,c,e) and amide I/II regions (Figure 6.3b,d, f). The red stick diagrams below the spectra are the predictions of calculations for the best-fit structures. A full comparison between observed and calculated IR transition frequencies for each of the conformers is included in Table 6.1. These stick diagrams not only help confirm the assignments but also provide insight to the groups responsible for the observed transitions, with labels given next to the calculated transitions describing the major carrier of a given IR transition. The hydride-stretch region for conformer A shows four distinct hydride-stretch fundamentals at 3562, 3443, 3430, and 3142 cm^{-1}. The bands

at 3562 and 3443 cm^{-1} are readily identified as the free antisymmetric and symmetric stretch fundamentals of the sidechain NH_2 moiety, respectively. The absence of the free carboxylic acid OH stretch (~3575 cm^{-1}) reveals that this group is involved in a hydrogen bond, most likely the band appearing at 3142 cm^{-1}. The final transition at 3430 cm^{-1} is shifted slightly lower in frequency than a free amide NH (~3450 cm^{-1}) indicative of a C5 hydrogen bond in an extended backbone geometry.

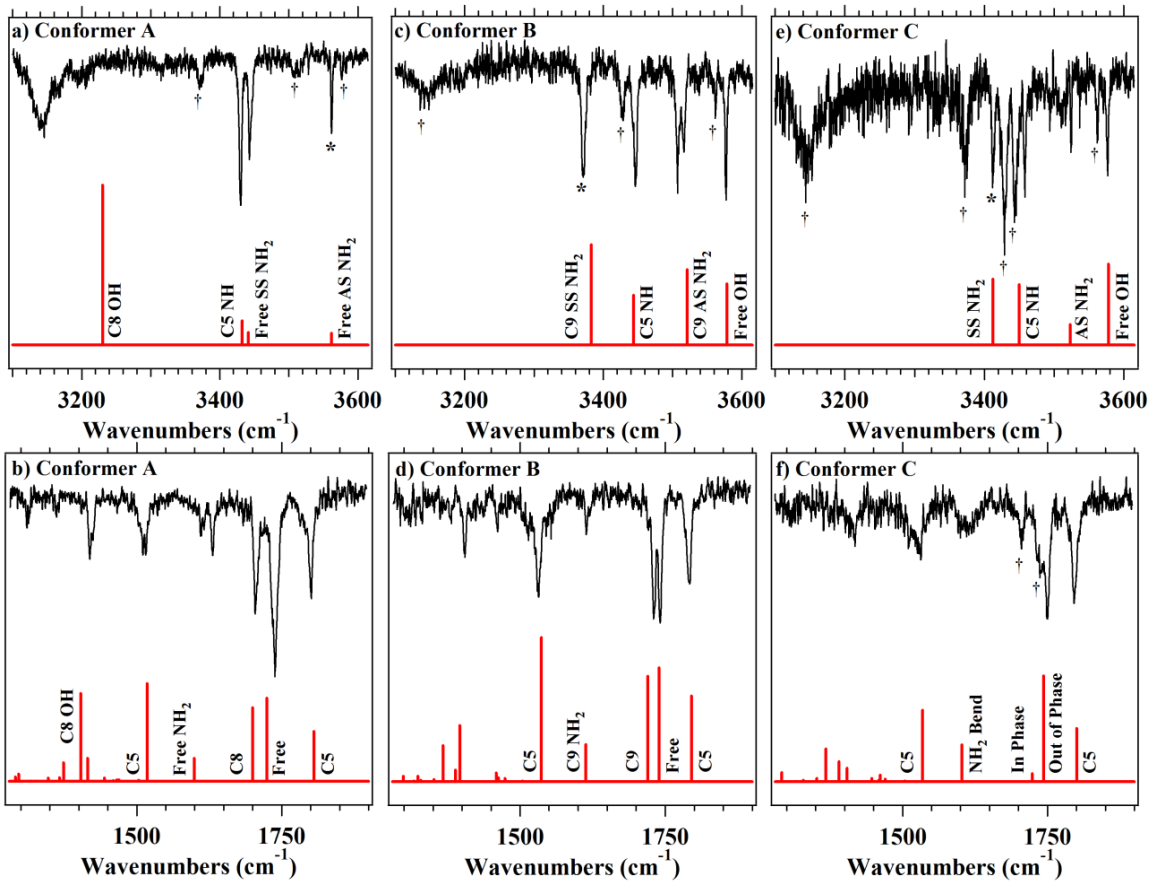


Figure 6.3. RIDIR spectra recorded for the three conformers of Z-Q-OH in the hydride-stretch region (top row, seeded) and amide I/II region (bottom row, unseeded): conformers A (a,b), B (c,d), and C (e,f). Stick-spectra below each experimental trace are scaled harmonic vibrational frequency calculations for the assigned conformations. Starred (*) transitions were used to record the IR-UV holeburning spectra shown in Figure 6.2. Transitions marked with a dagger (†) in the spectra of conformers B and C are spectroscopic interferences arising because the UV wavelength used for recording the RIDIR scans have some contribution from conformer A.

Based on comparison with calculation, we assign conformer A to a C5//*C8* (*trans*-OH) structure shown in Figure 6.4a, calculated to be 1.85 kJ/mol above the global minimum. In this structure the carboxylic acid OH has undergone an internal rotation into a *trans* configuration that facilitates formation of a very strong *C8* hydrogen bond of the C-terminal OH group with the sidechain carbonyl. The strength of this hydrogen is

evident from the magnitude of shift in the OH stretch (~430 cm^{-1}) from the nominally free position, reflecting its short H-bond length of 1.73 Å. In the absence of this stabilizing OH···O=C H-bond, the *trans*-OH minimum is calculated to be ~24 kJ/mol higher than *cis*-OH. Furthermore, there is a large barrier to internal rotation of this OH group (~55 kJ/mol in a simple glycine, Figure 6.5). Observation of such *trans* carboxylic acids have been seen in other uncapped gas-phase peptides when that isomerization can facilitate strong H-bond formation.[32]

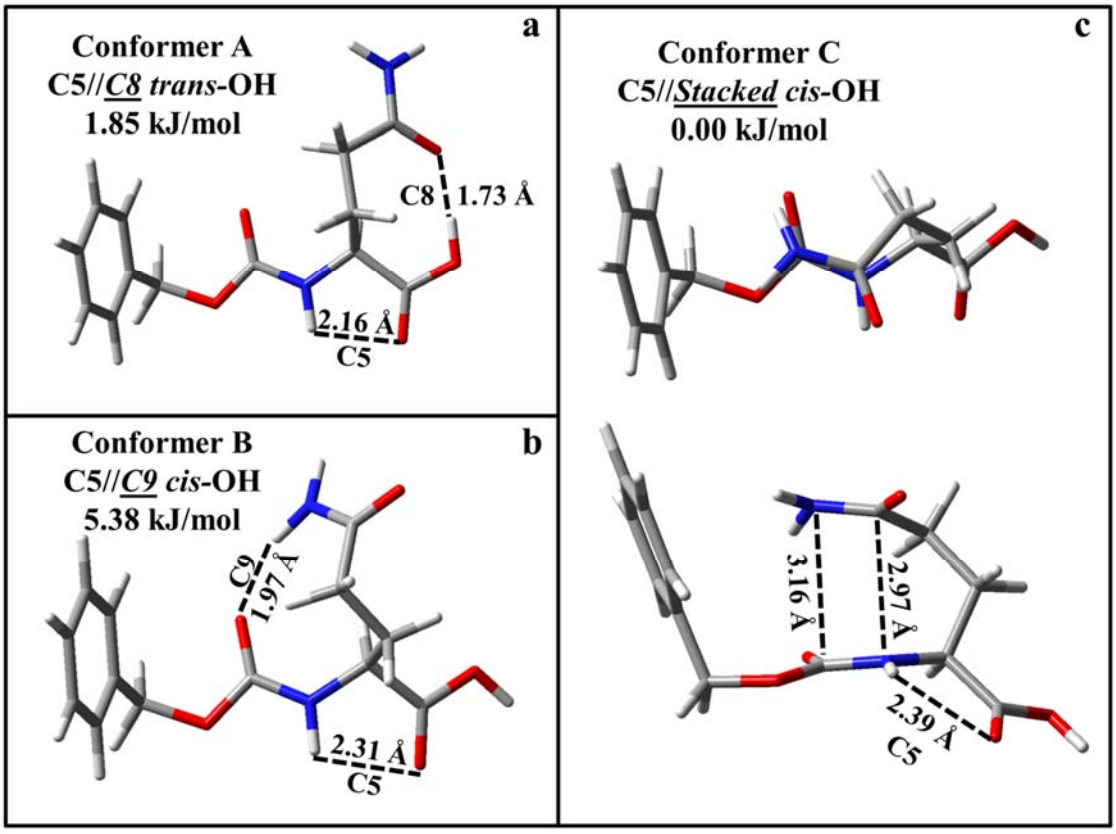


Figure 6.4. Assigned structures, relative energies, and hydrogen bond distances for conformers A-C (a-c, respectively) of Z-Gln-OH.

The corresponding comparison between experiment and calculation in the amide I/II region is shown in Figure 6.3b. Excellent agreement in both absolute frequency and intensity pattern confirm and strengthen the conformational assignment made from the hydride stretch region. This region, however, is less intuitively analyzed compared to the hydride stretching region. The amide I region shows three, clearly resolved transitions at 1801, 1739, and 1705 cm^{-1}. The highest frequency C=O stretch is assigned to the C5 hydrogen bonded carboxylic acid carbonyl stretch. The lowest frequency transition (1705 cm^{-1}) is assigned to the sidechain carbonyl involved in a *C8* hydrogen bond, which accounts for its shift to lower frequency. The free Z-cap carbonyl appears at 1739 cm^{-1}.

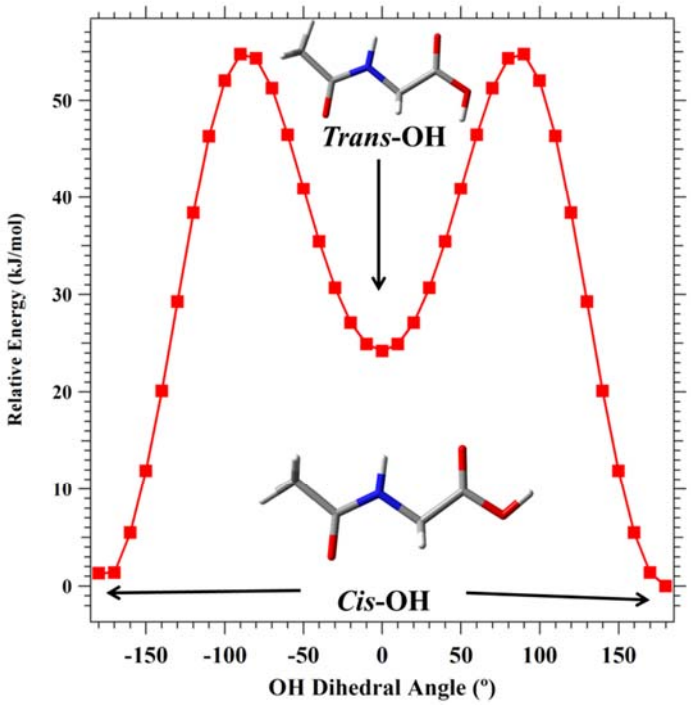


Figure 6.5. Relaxed potential energy scan of the hindered internal rotation of a carboxylic acid OH in a simply glycine residue showing the barrier height and relative energies of the *cis*- and *trans*-OH rotamers.

Calculation also aids in the assignment of the amide II region including the NH_2 bend, which appears at a much higher frequency (1612 cm^{-1}) than typical amide NH bending modes (~1600 cm^{-1}). The remaining strong transitions at 1513 and 1419 cm^{-1} are assigned to the C5 amide NH bend and the in-plane *C8* OH bend, respectively.

Analysis of conformers B and C is complicated by the presence of interfering transitions in the hydride stretching region, likely caused by overlapping transitions in the UV with not only one another, but also with the hot band of conformer A. Transitions that are caused by this spectral interference are labeled with daggers (†) in the spectra reported in Figure 6.3. Despite this complicating factor, it is still possible to make firm conformational assignments for these conformers.

Table 6.1. Experimental and calculated vibrational frequencies for the conformers of Z-Gln-OH, Z-Gln-NHMe, Ac-Gln-NHBn, and Ac-Ala-Gln-NHBn.

Structure Name	Hydrogen Bond Motif	OH Stretch		NH Stretch Frequency (cm^{-1})						C=O Stretch/ Amide I Frequency (cm^{-1})		NH- and OH-Bend/ Amide II Frequency (cm^{-1})					
				Asym. NH_2		Sym. NH_2		NH Stretch		C=O Stretch		NH Bend		NH_2 Bend		OH Bend	
		Exp.	Th.	Exp.	Th.	Exp.	Th.	Exp.	Th.	Exp.	Th.	Exp.	Th.	Exp.	Th.	Exp.	Th.
Z-Gln-NHMe																	
Conf.A	C5//*C8*			3560	3557	3442	3437	3402	3404	1740	1730	1584	1576	1605	1598		
								3351	3364	1721	1716	1510	1515				
										1709	1705						
Z-Gln-OH																	
Conf. A	C5//*C8* (*trans*-OH)	3142	3230	3562	3562	3443	3442	3430	3432	1801	1806	1513	1518	1613	1599	1419	1404
										1739	1725						
										1705	1700						
Conf. B	C5//*C9* (*cis*-OH)	3577	3578	3507/ 3517	3521	3371	3383	3446	3444	1792	1795	1533	1537	1614	1613		
										1741	1739						
										1731	1720						
Conf.C	C5//*Stacked* (*cis*-OH)	3577	3578	3524	3523	3412	3412	3458	3450	1796	1800	1532	1534	1609	1602		
										1750	1743						
										1737/ 1733	1724						
Ac-Q-NHBn																	
Conf. A	C5//*C8*			3562	3563	3443	3444	3402	3419	1720	1719	1577	1573	1603	1598		
								3310	3339	1703	1699	1509	1523				
										1695	1693						
Conf B	*C7/π*			3529	3535	3415	3416	3457	3468	1725	1728		1527		1595		
								3354	3376	1713	1708/ 1706	1521	1518				
Conf. C	C7//*C7/C8/π*			3509	3513	3400	3401	3343	3361	1706	1705	1558	1556/ 1552		1599		
								3270	3329	1682	1692/ 1680						
Ac-A-Q-NHBn																	
Conf. A	C10//*C7/π*			3520	3526	3408	3411	3470	3471	1718	1722	1546	1554	1601	1599		
								3361/ 3357	3380/ 3378	1706	1704	1531	1528				
										1694	1695	1508	1508				
										1681	1691						

Conformer B shows four new transitions in the hydride stretching region at 3577, 3507/3517, 3446, and 3371 cm^{-1} (Table 6.1, Figure 6.3c). The transition at 3577 cm^{-1} is ascribable to the free carboxylic acid OH stretch vibration,[37] which thereby retains its low energy *cis* rotamer configuration. The missing free anti-symmetric NH_2 stretch at ~3561 cm^{-1} indicates that the sidechain NH_2 is involved in a hydrogen bond. The transition at 3446 cm^{-1} falls into a range of frequencies ascribable either to free symmetric NH_2 stretches, or weakly bound amide NH groups involved in a C5 hydrogen bond. Since the NH_2 group is involved in a hydrogen bond, this transition is assigned to a C5 NH group. The remaining two transitions are assigned to the hydrogen bonded antisymmetric (3507/3517 cm^{-1}) and symmetric (3371 cm^{-1}) stretches of the NH_2 group.

Comparison with theory leads to the assignment of conformer B to an extended backbone C5//*C9* (*cis*-OH) structure that is calculated to be 5.38 kJ/mol above the global minimum (Figure 6.4b). This comparison yields excellent agreement between experiment and theory in both regions recorded experimentally. The sidechain NH_2 group is involved in a sidechain-to-backbone hydrogen bond (*C9*) towards the N-terminus. Based on the NH stretch frequency shifts of the NH_2 group (50-70 cm^{-1}), the *C9* H-bond is significantly weaker than its *C8* counterpart in conformer A involving the OH group (430 cm^{-1}). This weaker H-bond is reflected in the longer hydrogen bond length of the *C9* hydrogen bond (1.97 vs. 1.73 Å, Table 6.2).

Table 6.2. Hydrogen bond distances for hydrogen bonds found in the glutamines and similar hydrogen bonds found in γ-peptide foldamers.

Molecule	Hydrogen Bond Type	Hydrogen Bond Distance (Å)
Z-Gln-NHMe		
Conf. A	C5	2.12
(C5//*C8*)	*C8*	1.98
Z-Gln-OH		
Conf. A	C5	2.16
(C5//*C8*)	*C8*	1.73
Conf. B	C5	2.31
(C5//*C9*)	*C9*	1.97
Conf. C	C5	2.39
(C5//*Stacked)*	*Int. N→C*	2.97
	Ext. N→C	3.16
Ac-Gln-NHBn		
Conf. A	C5	2.11
(C5//*C8*)	*C8*	1.96
Conf. B	*C7*	1.98
(*C7/π*)	*π*	2.49
Conf. C	C7	1.92
(C7//*C7*/*C8*/*π*)	*C7*	1.99
	C8	2.24
	π	3.12, 2.98
Ac-Ala-Gln-NHBn		
Conf. A	C10	2.02
(C10//*C7*/*π*)	*C7*	1.97
	π	2.40
γ_{ACHC}[a]	C9	1.93
Ac-γ^2-hPhe-NHMe[b]		
C9	C9	2.00
Amide-Stacked	*Int. N→C*	2.90
	Ext. N→C	3.05
Gabapentin Derivative 1[b]		
C7	C7	2.04
Ac-Phe-NHMe[c]		
C7(a)	C7	2.17
C7(g+)	C7	2.01
Ac-Phe-Ala-NHMe[c]		
C5/C7	C7	2.07
Ac-Ala-Phe-NHMe[c]		
C7/C7(g-)	C7	2.05
	C7	2.04

[a] From ref. [51].
[b] From ref. [48].
[c] From ref. [33].

The amide I/II region provides additional evidence for the assignment (Figure 6.3c, d). The amide I region has three transitions centered at 1792, 1741, and 1731 cm^{-1}. The highest frequency transition is assigned to the backbone C5 hydrogen bonded carbonyl, consistent with other experiments on similar systems.[33] The transition at 1741 cm^{-1} is the free sidechain carbonyl stretch, and the transition at 1731 cm^{-1} results from the *C9* hydrogen bonded Z-cap carbonyl. Similarly good agreement is found in the amide II region where transitions at 1614 and 1533 cm^{-1} are assigned to the NH_2 and backbone C5 amide NH bends, respectively.

Conformer C is the most difficult to analyze because it is the most minor conformer, and as such, suffers from the largest interference caused by overlap in the UV transitions. However, since the transitions belonging to conformers A and B are known, identifying conformer C transitions is straight forward. In the hydride stretch region four transitions are identified at 3577, 3524, 3458, and 3412 cm^{-1} (Table 6.1, Figure 6.3c). The highest frequency transition is assigned to the free carboxylic acid OH stretch (3577 cm^{-1}).[37] Surprisingly, the other three transitions are all above 3400 cm^{-1}, indicating that there are no strong hydrogen bonds in the molecule. The lack of a free anti-symmetric stretch at ~3561 cm^{-1} also reveals that the NH_2 group is involved in at least a weak hydrogen bond.

Comparing these unique spectral attributes with the predictions of calculations, conformer C can be assigned to the extended backbone C5//*Amide-stacked* (*cis*-OH) geometry, calculated to be the global minimum structure (Figure 6.4c). On this basis, the transitions at 3524 and 3412 cm^{-1} are assigned as the antisymmetric and symmetric stretches of the NH_2 group, which is weakly hydrogen bonded to the π-cloud of the

Z-cap. The relatively small frequency shifts of these transitions from the nominally free positions (by -30 and -35 cm^{-1}, respectively) are consistent with the weak π H-bond involved. The remaining transition (3458 cm^{-1}) is assigned to a C5 NH group, occurring near the anticipated position based on the C5 rings in A and B.

Once again, spectra in the amide I/II regions confirm the structural assignment (Figure 6.3f). The transition at 1796 cm^{-1} is due to the C5 C=O stretch of the COOH group. The transitions at 1737 and 1705 cm^{-1} are interference from the two major transitions of conformer A, leaving the pair of transitions at 1750 and 1733 cm^{-1} yet to assign. The unique geometry of the *amide-stacked* planes produces an interesting intensity pattern in the other two carbonyl stretches, both of which are involved in the stack. These two C=O groups are nearly anti-parallel to one another, and are in similar environments, so that coupling between them produces two delocalized C=O stretch modes that are best described as in- and out-of-phase linear combinations of the two C=O stretch motions. The dipole derivatives nearly cancel in the mode in which the two C=O groups stretch in-phase, but constructively interfere in the out-of-phase motion, leading to a large asymmetry in calculated intensity. On this basis, we assign the strong and weak transitions at 1750 and 1733 cm^{-1} to the out-of-phase and in-phase doublet of amide stacked C=O stretch modes. The amide II region has two transitions at 1609 and 1532 cm^{-1}, which are assigned to the NH_2 and C5 amide NH bending fundamentals, respectively.

Several other interesting or low-energy structures have been calculated for Z-Gln-OH and are shown in Figure 6.6. The structures are not experimentally assigned to a conformation present in the molecular beam, but they do represent interesting structures that are representative of similar geometries assigned for the other glutamine containing peptides. These structures include a C7//*C7/C8* *trans*-OH structure (Figure 6.6a), a *C7/amide-stacked* *trans*-OH structure (Figure 6.6b), and a C7//*C7* *trans*-OH structure (Figure 6.6c).

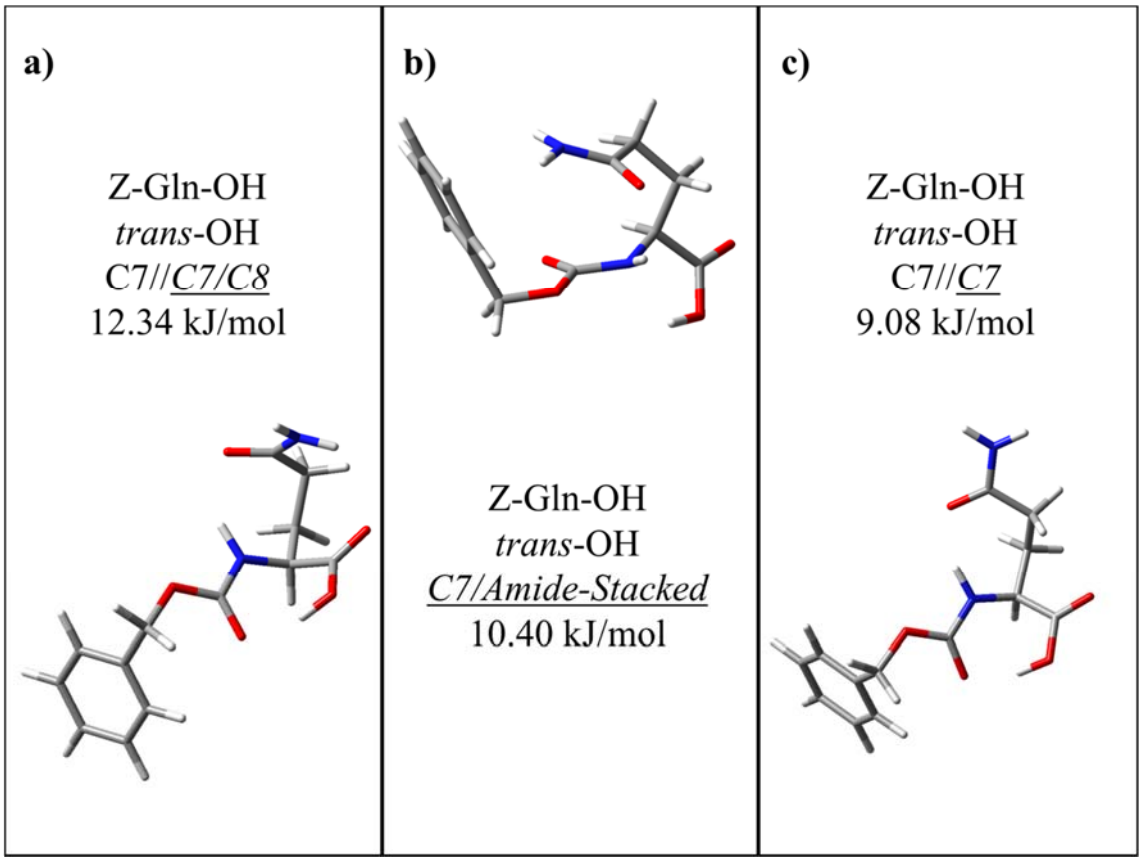


Figure 6.6. Low-energy and other structures of Z-Gln-OH which were not experimentally assigned: (a) a C7//*C7/C8* structure, (b) a *C7/amide-stacked* structure, and (c) a C7//*C7* structure.

6.3.1.3 Z-Glutamine-NHMe: IR Spectroscopy and Conformational Assignment

Figure 6.7a and b show the RIDIR spectra recorded for the sole conformer found experimentally for Z-Gln-NHMe (conformer A). The NH stretch spectrum has four clear transitions at 3560, 3442, 3402, and 3351 cm^{-1} (Table 6.1, Figure 6.7a). As was the case in the Z-Gln-OH conformers, the transitions at 3560 and 3442 cm^{-1} are readily assigned to the free antisymmetric and symmetric stretches of the NH_2 group, pointing towards a structure with a free NH_2 group. The remaining transitions arise from the backbone amide NH's which are involved in one weak and one strong hydrogen bond, strongly reminiscent of the C5//*C8* motif of Z-Gln-OH conformer A.

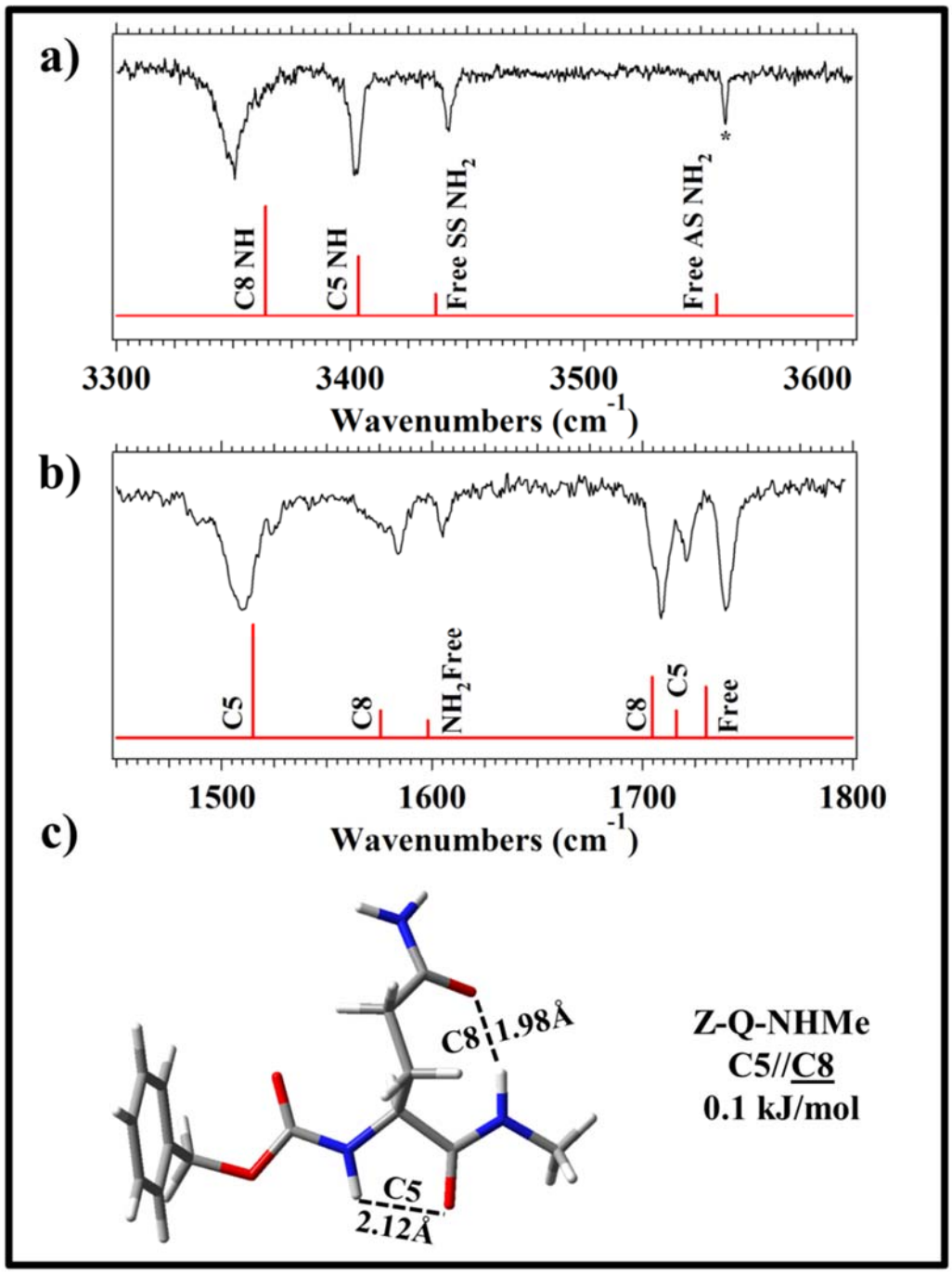


Figure 6.7. RIDIR spectra of the single observed conformation of Z-Gln-NHMe in the NH-stretch region (a, seeded) and amide I/II region (b, seeded). Stick spectra below each experimental trace are scaled harmonic vibrational frequency calculations for the assigned C5//*C8* (c) conformation. Starred (*) transition was used to record the IR-UV holeburning spectrum shown in Figure 6.2.

Comparing experiment with the predictions of calculations (Figure 6.7) enables the assignment of conformer A of Z-Gln-NHMe to an extended backbone C5//*C8* structure, where the sidechain forms a hydrogen bond with the backbone towards the C-terminus. The assigned structure, shown in Figure 6.7c, is 0.1 kJ/mol above the calculated global minimum. The final two transitions can now be assigned to the backbone C5 (3402 cm^{-1}) and the *C8* (3351 cm^{-1}) backbone-to-sidechain amide NH stretch fundamentals.

The amide I region comprises three well-resolved transitions at 1740, 1721, and 1709 cm^{-1}, assignable based on the calculations to the (nominally) free, C5, and *C8* C=O stretch fundamentals, respectively (Figure 6.7b). The amide II region is similarly made up of three transitions at 1605, 1584, and 1510 cm^{-1} due to the free NH_2, *C8* amide NH, and C5 amide NH bend fundamentals, respectively. The large intensity of the C5 NH bend and the intensity asymmetry with its *C8* counterpart arise from two effects. First, the C5 NH bend oscillates toward and away from the C=O group involved in the C5 H-bond. This interaction enhances the dipole derivative, increasing the intensity of the C5 NH bend.[61] Second, the anti-parallel arrangement of successive amide groups linked by a C5 H-bond enhances coupling between the two NH groups (despite their different H-bonding environments), producing partially delocalized NH bend modes, with the in-phase oscillation gaining in intensity, much as it does in β-strand and β-sheet structures.[33,61]

It is worth noting that, although the structural families of conformer A of Z-Gln-OH and Z-Gln-NHMe are both C5//*C8*, the former molecule forms a *C8* using its C-terminal OH rather than an amide NH. The OH···O=C *C8* is much shorter than

NH···O=C (1.72 Å vs. 1.98 Å), producing a much larger shift to lower frequency in Z-Gln-OH (-430 cm^{-1}) than in Z-Gln-NHMe (-120 cm^{-1}, Table 6.1 and Table 6.2).

While only a single conformer was characterized experimentally, several other low energy structures were calculated for Z-Gln-NHMe including *Amide-Stacked* (0.00 kJ/mol), C5//*C9/C8* with an extended backbone (1.17 kJ/mol), C7//*C7*/*C8* with a γ-turn backbone (8.38 kJ/mol), and *C7* with a turned backbone (8.19 kJ/mol). Structures for these unobserved structures are shown in Figure 6.8. While it is possible that minor conformers could be present and were obscured by overlap with bands from conformer A (Figure 6.2a), the single dominant conformer is the C5//*C8* conformer just described.

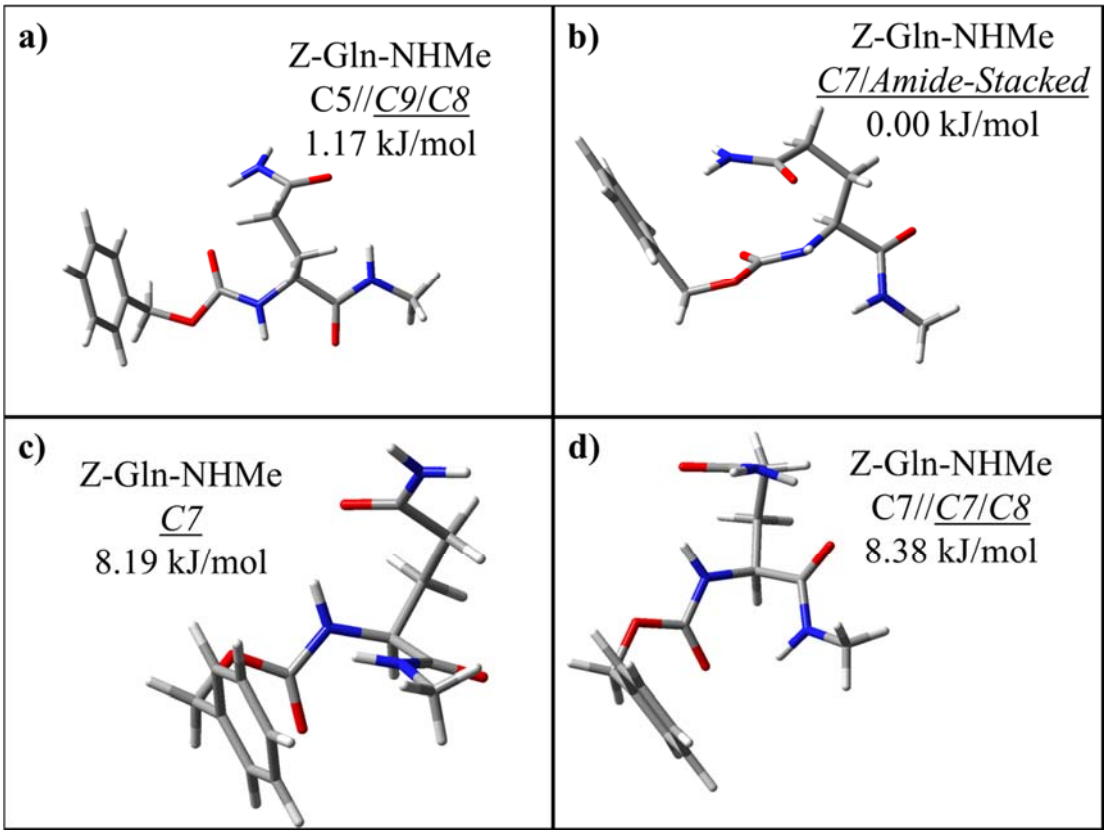


Figure 6.8. Low-energy and other structures of Z-Gln-NHMe which were not experimentally assigned: (a) a C5//*C9*/*C8* double bridge structure, (b) a *C7*/*amide-stacked* structure, (c) a *C7* structure, and (d) a C7//*C7*/*C8* structure.

6.3.2 Ac-Glutamine-NHBn

6.3.2.1 UV Spectroscopy

Figure 6.9a shows the R2PI spectrum recorded for the Ac-Gln-NHBn molecule. The spectrum was recorded by monitoring ion yield in the parent mass channel (m/z=272) as the laser was tuned through the –NHBn S_0-S_1 origin region (~37300-37740 cm^{-1}). The resulting spectrum shows a very dense region with sharp UV transitions riding on a broad background likely caused by incomplete cooling of a portion of the

desorbed sample. UV transitions due to three unique conformations were observed in the molecular beam as identified by RIDIR spectroscopy and IR-UV holeburning, with their origin transitions at 37548 (A), 37619 (B), and 37554 (C) cm^{-1}, labeled in Figure 6.9a. IR-UV holeburning (Figure 6.9b) was performed on conformer A using a unique IR transition at 3402 cm^{-1}. Much of the dense vibronic activity in the R2PI spectrum is associated with conformer A, as is the broad background. The IR-UV holeburning scan is missing the transitions ascribed to conformers B and C, identifying them for RIDIR scans which proved that they were due to different conformational isomers. Holeburning scans were not performed on these conformers due to their minor population and our limited sample size. The transition at 37425 cm^{-1} was investigated minimally, but its carrier could not be identified. It is likely a water cluster built off of conformer A based on cursory investigation in the IR.

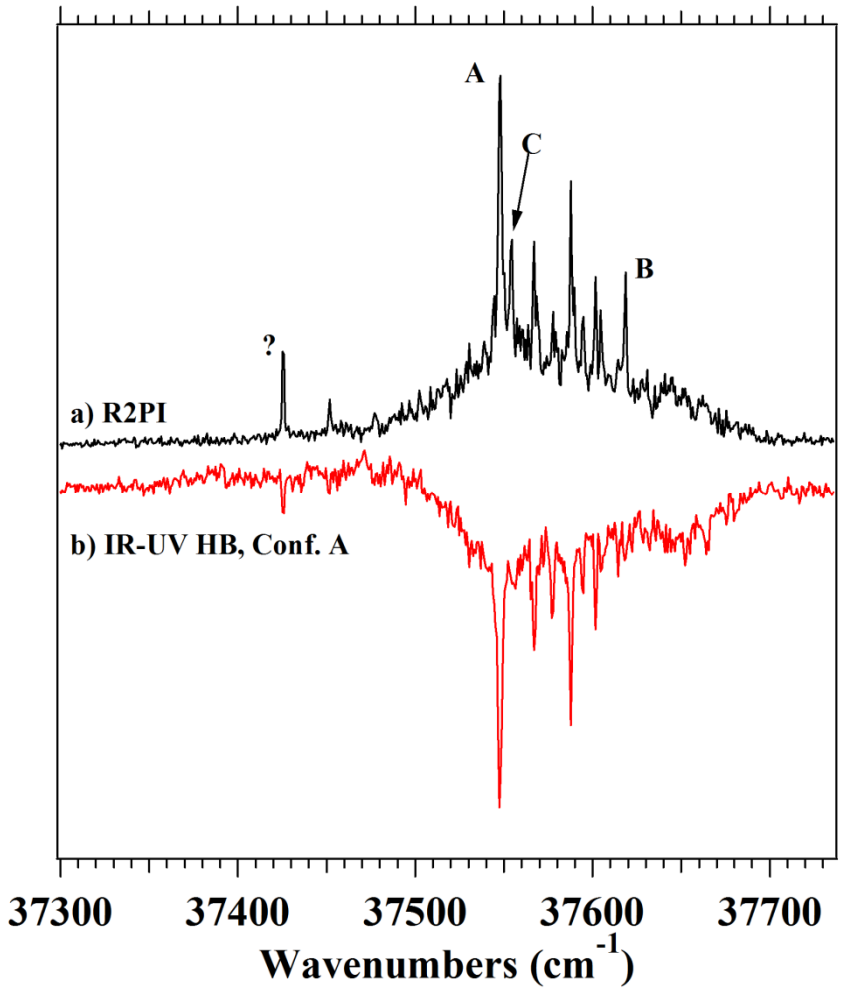


Figure 6.9. R2PI (a) of Ac-Gln-NHBn. IR-UV holeburning spectrum (b) for conformer A. Transitions used to record RIDIR spectra for conformers A, B, and C are labeled with A, B, and C, respectively.

6.3.2.2 IR Spectroscopy and Conformational Assignments

Figure 6.10 presents the NH stretch (upper) and amide I/II (lower) RIDIR spectra recorded for the three conformers of Ac-Gln-NHBn. Figure 6.10a shows the experimental spectrum recorded in the NH stretch region, with four transitions at 3562, 3443, 3402, and 3310 cm^{-1} (Table 6.1). This spectrum is nearly identical to the spectrum recorded for conformer A of Z-Gln-NHMe (C5//*C8*), with the three higher-frequency IR transitions

appearing within 1 cm^{-1} of each other. Calculations confirm that conformer A of Ac-Gln-NHBn is also a C5//*C8* structure. This structure, shown in Figure 6.11a, is calculated to be just 1.54 kJ/mol above the global minimum. The primary difference between the C5//*C8* conformers of Z-Gln-NHMe and Ac-Gln-NHBn is the length of the *C8* hydrogen bond formed between the sidechain and the backbone, which is 1.98 Å in Z-Gln-NHMe, but shrinks to 1.96 Å in Ac-Gln-NHBn (Table 6.2). The shorter H-bond in Ac-Gln-NHBn leads to a larger shift in frequency of the *C8* NH stretch from its free position down to 3310 cm^{-1}, compared to 3364 cm^{-1} in Z-Gln-NHMe.

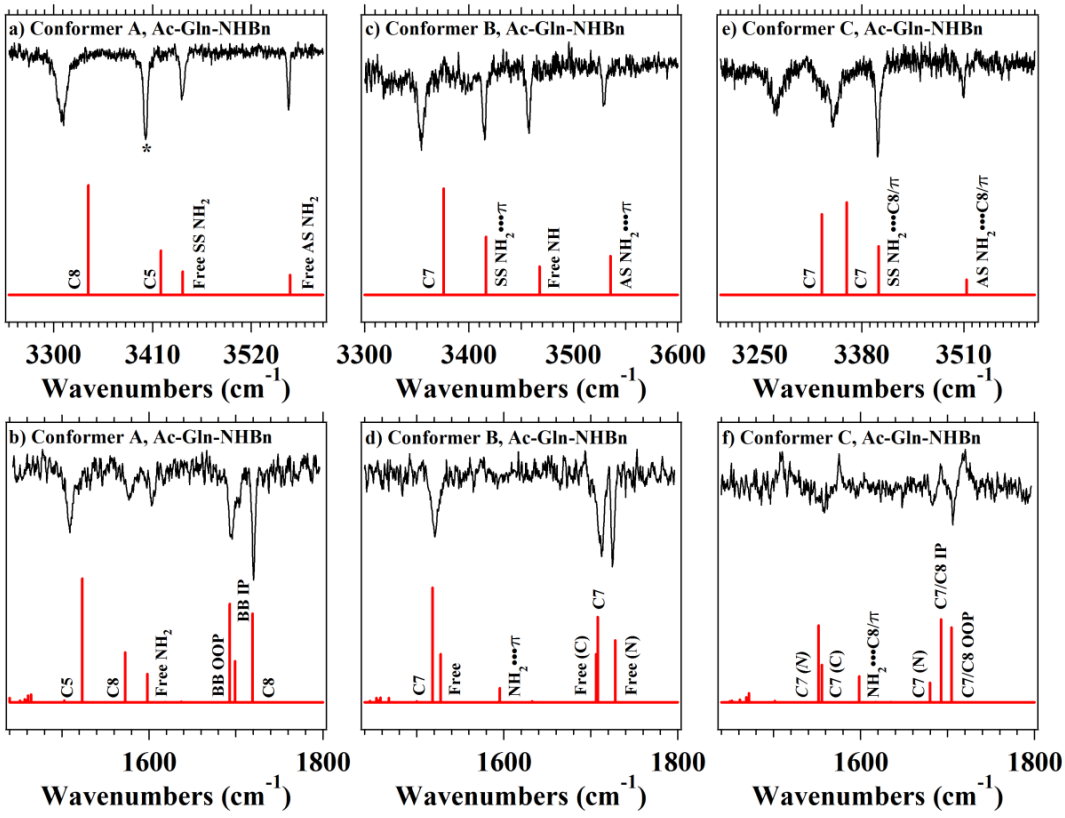


Figure 6.10. RIDIR spectra recorded for the three conformers of Ac-Gln-NHBn in the NH-stretch region (top row, unseeded) and amide I/II region (bottom row, unseeded): conformers A (a,b), B (b,d), and C (e,f). Stick-spectra below each experimental trace are scaled harmonic vibrational frequency calculations for the assigned conformations. The starred (*) transition was used to record the IR-UV holeburning spectrum shown in Figure 6.9.

Interestingly, while the C5//*C8* conformers of Ac-Gln-NHBn (Figure 6.10b) and Z-Gln-NHMe (Figure 6.7b) have amide I spectra that are very similar in appearance, the difference in strength of the *C8* H-bond and the differing positions of the aromatic caps shifts the absolute frequencies, re-shuffles the order, and changes the coupling between the three C=O stretch transitions. The stronger *C8* H-bond involving the side-chain C=O group shifts the *C8* C=O stretch up in frequency from 1709 to 1720 cm^{-1}, and localizes its motion on the *C8* C=O oscillator. The nominally "free" and C5 amide C=O groups in the

extended backbone are nearly degenerate with, and anti-parallel to, one another. Coupling between them produces in-phase and out-of-phase modes delocalized over the pair of C=O groups, with frequencies of 1695 and 1703 cm^{-1} for the out-of-phase (oop) and in-phase (ip) combinations, respectively, as labeled in the figure. This coupling is reflected in the asymmetric intensity distribution between the pair, with the out-of-phase combination, with its constructive interference, gaining intensity at the expense of the in-phase combination.

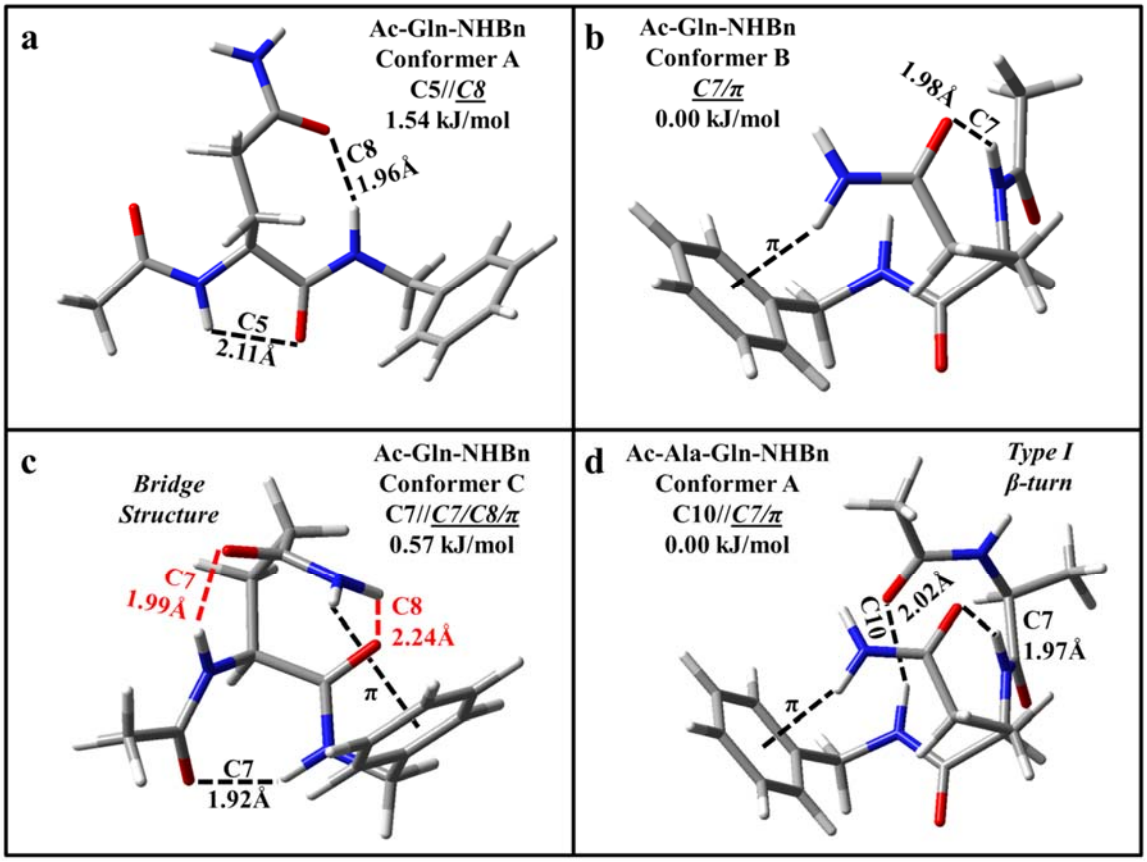


Figure 6.11. Three assigned structures for the three conformers of Ac-Gln-NHBn (a, b, c) and the single conformer of Ac-Ala-Gln-NHBn (d). Also shown for each of the assigned conformers are their relative energies, hydrogen bonding pattern and bond lengths. The hydrogen bonds labeled in red indicate a hydrogen bond formed when glutamine forms a hydrogen bonded "bridge" motif.

The amide II region of conformer A is also very similar in appearance to its C5//*C8* counterpart in Z-Q-NHMe, with the C5 NH bend at low frequency (1509 cm^{-1}) and possessing its characteristic large intensity.

Conformer B RIDIR spectra are shown in Figures 6.10c and d. The NH stretch spectrum bears little resemblance to the spectra recorded for the other molecules, indicating that it arises from a new structural family. The NH stretch fundamentals appear at 3529, 3457, 3415, and 3354 cm^{-1} (Table 6.1). The NH_2 symmetric and antisymmetric stretch transitions both appear shifted down in frequency slightly from their nominally free value, consistent with the presence of a weak H-bond involving the aromatic π cloud. The remaining transitions at 3354 and 3457 cm^{-1} are representative of a strongly hydrogen bonded backbone NH and a nearly free amide NH, respectively.

Comparison of the NH stretch spectra with theory produces an assignment of a *C7/π* structure in which the only H-bonds are those between the Gln side chain and the peptide backbone (*C7*, N-term NH to Gln C=O) or aromatic cap (Gln NH_2 to Bn π), with no backbone-to-backbone hydrogen bonds. The assigned structure is the calculated global minimum (Figure 6.11b), with calculated IR transitions that provide an excellent match in both the NH stretch and amide I/II regions. As anticipated the single strong H-bond is the *C7* H-bond between the N-terminal NH and the Gln C=O group.

The amide I and II regions provide corroborating support for the assignment, but due to the small signal size and lower laser power several of the lower intensity transitions are not resolved experimentally. The amide I region shows two clearly resolved transitions at 1725 and 1713 cm^{-1}, while the amide II region displays only a single broad transition at 1521 cm^{-1} (Table 6.1, Figure 6.10d). Comparison with

calculations makes it possible to assign these transitions and partially explain the reason for the missing transitions. In the amide I region, the transition at 1725 cm^{-1} is the free C=O stretch of the N-terminal carbonyl. The broader peak at 1713 cm^{-1} likely contains two nearly degenerate C=O stretch transitions of the sidechain *C7* and C-terminal free carbonyls. The predicted splitting of 2 cm^{-1} is likely unresolvable in the experimental spectrum. The amide II region shows a similar near degeneracy between the free amide NH and *C7* amide NH, which are predicted to be split by 9 cm^{-1}, and are experimentally unresolved, producing a single broadened transition at 1521 cm^{-1}.

Conformer C is the final minor conformer of Ac-Gln-NHBn, with RIDIR spectra shown in Figure 6.10e and f. The NH stretch spectrum of this conformer is also unique compared to those so far observed. All four NH stretch transitions (3509, 3400, 3343, and 3270 cm^{-1} (Table 6.1), are shifted down in frequency from their nominally free values (3560/3445 for NH_2 and 3450-3480 for amide NH), suggesting that the NH_2 and two amide NH groups are all involved in H-bonds.

Comparison with theory leads to an assignment for conformer C to the C7//*C7*/*C8*/π structure shown pictorially in Figure 6.11c. This structure is only 0.57 kJ/mol above the global minimum, and is a highly compact structure in which every group is involved in a hydrogen bond. The backbone C7 ring is locked in place by a set of two sidechain-to-backbone H-bonds, with the side-chain C=O forming a *C7* ring with the N-terminal NH, and the NH_2 group engaged in a *C8* H-bond with the C-terminal C=O. This double-bridge capping arrangement of the Gln side-chain stabilizes the C7 peptide backbone, which forms what is formally an inverse γ-turn, with Ramachandran angles of (-82.4, 56.3)°. Such structures have been previously identified in the literature

by several groups.[33,39,62] The two hydrogen bonded transitions at 3343 and 3270 cm^{-1} are now identified as arising from the *C7* sidechain-to-backbone hydrogen bond and the C7 hydrogen bond formed within the backbone, respectively. The predicted NH stretch spectrum shown in Figure 6.10e shows an excellent match to the experimental spectrum.

The recorded amide I/II spectrum (Figure 6.10f) is complicated by the small signal levels, and the proximity of the S_0-S_1 origin transition (37554 cm^{-1}) to the origin transition of conformer A (37548 cm^{-1}). The depletion signals for conformer C are observed amidst gains in ion intensity caused by the neighboring major conformer. This phenomenon has been observed previously and complicates the experimental spectrum.[63] Still, two identifiable transitions are present in the amide I region at 1706 and 1682 cm^{-1}. The lower frequency transition is assigned to the C=O stretch involved in the backbone C7. The amide II transitions are difficult to observe in the midst of the gains from conformer A. The only identifiable transition is a weak, broad absorption at 1558 cm^{-1}. This transition likely contains both the amide NH (C-terminal) C7 and the *C7* amide NH (N-terminal) which are predicted to be nearly degenerate at 1556 and 1552 cm^{-1}, respectively.

The backbone C7 in conformer C produces the lowest frequency NH stretch observed so far (3270 cm^{-1}), providing evidence that this H-bond is particularly strong. This backbone C7 is also substantially stronger than the *C7* hydrogen bonds formed between the sidechain and the backbone, which show shifts of only ~120 cm^{-1} from their nominally free position. This strength is borne out by the calculations, which predict a C7 H-bond length of 1.92 Å, compared to 1.98/1.99 Å for the sidechain-to-backbone *C7*'s in conformers B and C respectively (Table 6.2).

Figure 6.12 shows a select number of low-energy or otherwise interesting structures of Ac-Gln-NHBn. The structures include a C5//*C9/C8* structure (Figure 6.12a), a C7//*C7* structure (Figure 6.12b), and a C5//*C9/Amide-Stacked* structure (Figure 6.12c). These structures, and their importance, will be taken up more thoroughly in a later section (Section 6.4).z

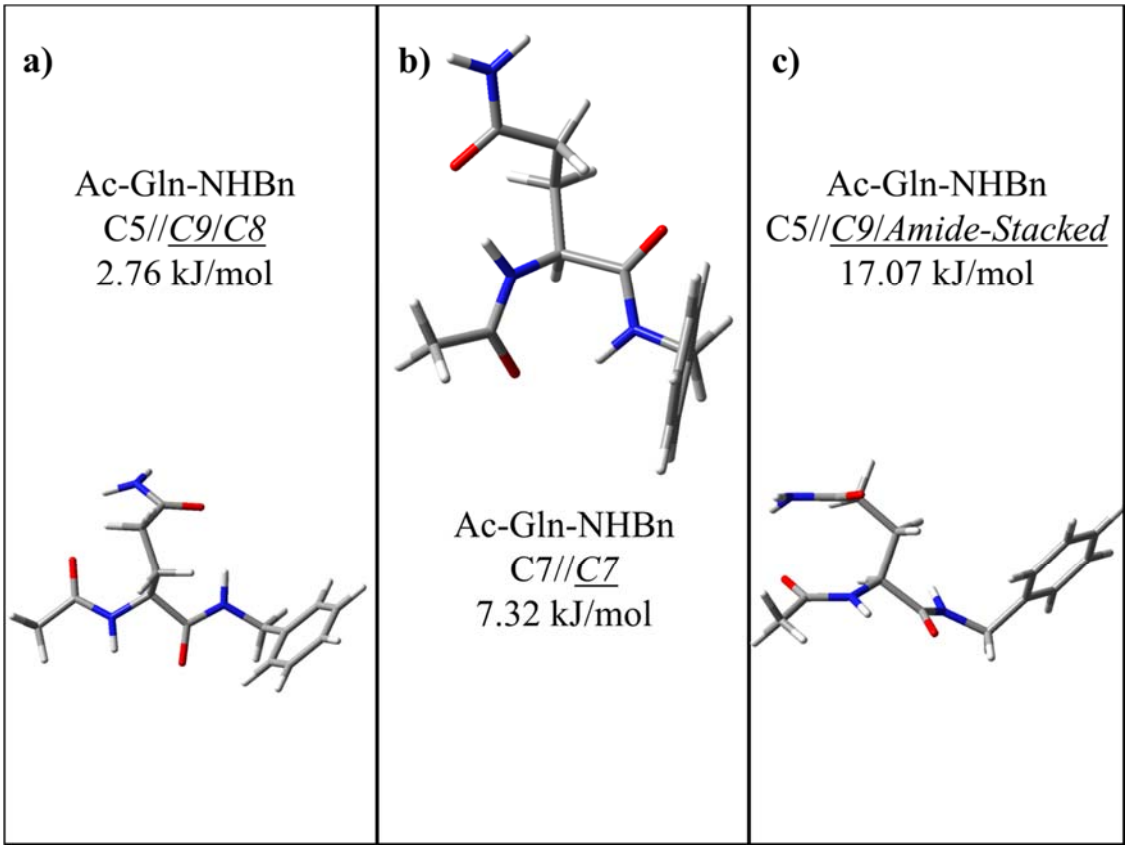


Figure 6.12. Low-energy and other structures of Ac-Gln-NHBn which were not experimentally assigned: (a) a C5//*C9/C8* structure, (b) a C7//*C7* structure, and (c) a C5//*C9/amide-stacked* structure.

6.3.3 Ac-Alanine-Glutamine-NHBn

6.3.3.1 UV Spectroscopy

Figure 6.13a shows the R2PI and IR-UV holeburning spectra for Ac-Ala-Gln-NHBn. The R2PI spectrum (black) shows very little background and is relatively uncongested. IR-UV holeburning (red) proves that all transitions belong to a single conformation with an S_0-S_1 origin transition at 37621 cm^{-1}. The spectrum shows a short Franck-Condon progression involving a 19 cm^{-1} vibration, presumably arising from a low frequency mode involving the backbone and sidechain groups vibrating against the π-cloud of the UV chromophore.

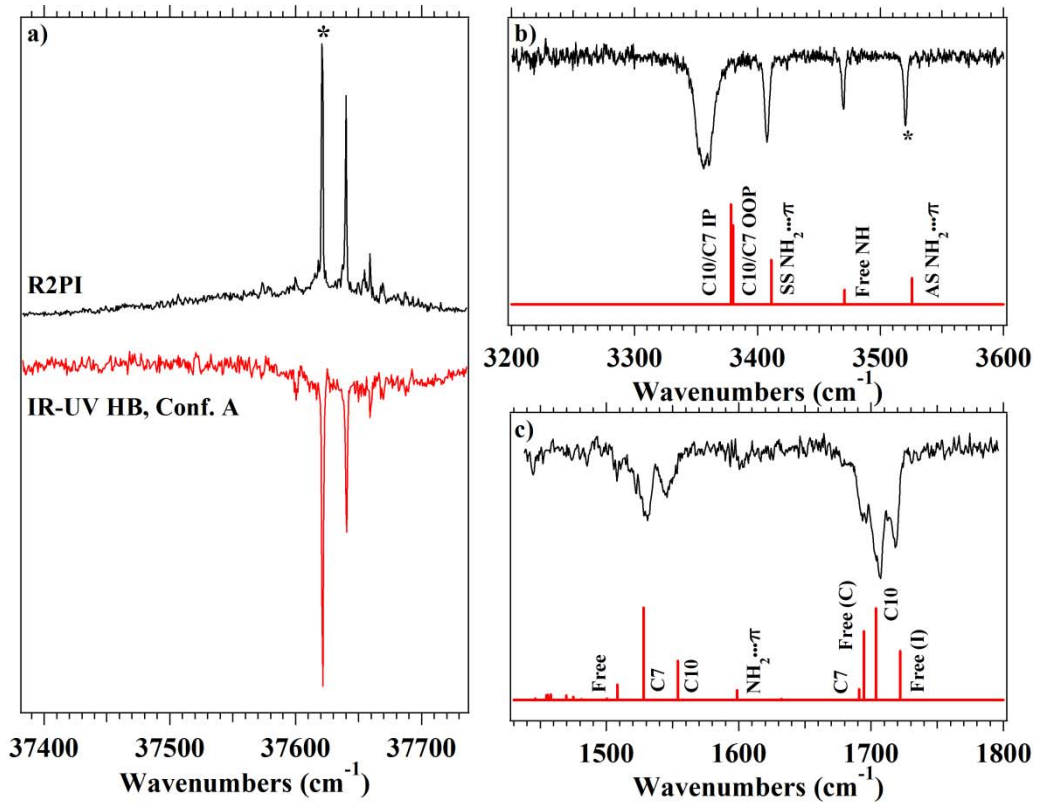


Figure 6.13. Experimental spectra recorded for Ac-Ala-Gln-NHBn. R2PI (a) and IR-UV holeburning (a) for the sole observed conformation. Experimental RIDIR spectra recorded in the NH-stretch (b, unseeded) and amide I/II (c, unseeded) regions shown above the labeled red stick spectra from the assigned structure's harmonic frequency calculation. Starred transitions (*) indicated transitions used to record RIDIR and IR-UV holeburning spectra.

6.3.3.2 IR Spectroscopy and Conformational Assignments

Figures 6.13b and c present the RIDIR spectra recorded for the single observed conformer of Ac-Ala-Gln-NHBn in the NH stretch and amide I/II regions, respectively. The NH stretch spectrum shows three sharp transitions at 3520, 3470, and 3408 cm^{-1} and a broad band with a barely-resolved splitting at 3361/3357 cm^{-1} (Table 6.1). The overall spectrum is similar in appearance to that of conformer B of Ac-Gln-NHBn. Following that lead, the transitions at 3520 and 3408 cm^{-1} can be tentatively assigned as symmetric

and antisymmetric stretch fundamentals of the NH_2 group, slightly shifted from their free positions by virtue of a π H-bond. The structure also possesses a completely free amide NH at 3470 cm^{-1}. The broadened transition at 3361/3357 cm^{-1} shows a slight splitting which indicates that two nearly degenerate transitions are present at a frequency typical of both C7 and C10 hydrogen bonds.[33,38,39,62,64]

The stick spectra shown below the experimental spectra in Figure 6.13b present an excellent fit, and confirm the tentative assignments. We assign this conformer to the C10//*C7*/*π* structure shown in Figure 6.11d, which is calculated to be the global minimum. This structure has a backbone C10 hydrogen bond that links the N-terminal C=O and the C-terminal amide NH, forming a type I β-turn, the most common β-turn in biological settings.[65] The backbone dihedrals for this conformer are (-70, -13.3)° and (-73, -11)° for *i*+1 and *i*+2 Ramachandran angles, respectively (Table 6.3), close to those for an idealized type I β-turn [(-60, -30)° and (-90, 0)°].[65] The glutamine residue here behaves in a similar fashion to conformer B of Ac-Gln-NHBn, with the sidechain forming a *C7* hydrogen bond towards the N-terminus nearest-neighbor, and the NH_2 forming a hydrogen bond with the π-cloud of the NHBn cap. The C10 and *C7* NH stretch transitions are nearly degenerate, and highly coupled, leading to the broad, slightly split pair of transitions at 3361/3357 cm^{-1}. A detailed discussion of the similarities between the gas-phase and crystal/NMR β-turn structures will be taken up in Section 6.4.1.3.

Table 6.3. Ramachandran angle comparison of type-I β-turns with sequence X(AQ)X.

Structure	Turn Sequence/ Residue #	Turn Residues *(i+1)* (ϕ,ψ)°	*(i+2)* (ϕ,ψ)°
Ac-Q-NHBn, Conf. B[a]	X(XQ)X		(-87.2, -9.1)
Ac-AQ-NHBn[a]	X(AQ)X	(-70, -13.3)	(-73, -11)
Complex, Hydrolase Product[b]	F(AQ)L 9-12	(-57.0, -21.7)	(-98.4, 1.67)
Nuclear Protein[c]	K(AQ)F 92-95	(-57.1, -34.1)	(-64.0, -23.7)
Signaling Protein, Cell Cycle[d]	E(AQ)G 23-26	(-66.6, -19.1)	(-91.8, -2.8)
Viral Protein, Anti-Viral Protein[e]	E(AQ)Q 560-563	(-65.5, -43.1)	(-64.4, -31.3)
Xylanase[f]	A(AQ)N 69-72	(-56.8, -33.3)	(-84.5, -14.1)
Electron Transport (Heme Protein)[g]	C(AQ)C 14-17	(-63.3, -15.0)	(-76.6, -8.8)
Toxoflavin Binding Protein[h]	G(AQ)H 26-29	(-52.9, -27.8)	(-68.8, -17.8)
Histocompatibility Anitgen i-ak[i]	F(AQ)L 48-51	(-65.2, -6.1)	(-93.9, -21.3)
Iron-Sulfur-Binding Protein[j]	I(AQ)G 69-72	(-64.3, -27.5)	(-90.2, 2.0)
Hydrolase, Ligase/Inhibitor[k]	V(AQ)E 146-149	(-54.5, -39.4)	(-78.4, -7.7)
Anti-Oncogene Protein[l]	D(AQ)A 352-355	(-67.6, -23.0)	(-83.5, -10.6)
Viral Protein, Metal Binding Protein[m]	F(AQ)C 34-37	(-58.9, -30.9)	(-78.8, -40.3)
Structural Genomics, Signaling Protein[n]	L(AQ)A 37-40	(-48.7, -38.4)	(-73.1, -35.8)
Average of Crystal Structures	X(AQ)X	(-59.9, -27.7)	(-80.5, -16.2)

[a] This work
[b] (PDB ID: 1af2).[66]
[c] (PDB ID: 3g36).[67]
[d] (PDB ID: 1mb3).[68]
[e] (PDB ID: 4i2L).[69]
[f] (PDB ID: 1w3h).[70]
[g] (PDB ID: 5cyt).[71]
[h] (PDB ID: 3ouL).[72]
[i] (PDB ID: 1iak).[73]
[j] (PDB ID: 3zyy).[74]
[k] (PDB ID: 4dm9).[75]
[l] (PDB ID: 1oLh).[76]
[m] (PDB ID: 2a51).[77]
[n] (PDB ID: 2cs4).[78]

The amide I/II region (Figure 6.13c) also shows an excellent match to theory in both absolute frequency and intensity pattern. The amide I region has transitions at 1718, 1706 and 1694 cm^{-1} with a weak shoulder at ~1680 cm^{-1}. The highest frequency transition at 1718 cm^{-1} is assigned to the free C=O stretch of the interior amide group, while the transition at 1694 cm^{-1} is from the free C-terminal C=O stretch. The remaining two C=O stretch modes at 1706 (high intensity) and 1681 cm^{-1} (very weak) show a large intensity asymmetry resulting from the nearly anti-parallel orientation of the Ac and Gln C=O groups in this structure, which are engaged in C10 and *C7* H-bonds, respectively. The normal modes are nearly equal mixes of the local mode oscillations, with the in-phase combination at 1681 cm^{-1} near zero intensity, and the out-of-phase counterpart at 1706 cm^{-1} with unusual intensity. The amide II region has four well resolved transitions at 1601, 1546, 1531, and 1508 cm^{-1} which are assigned to the NH_2 bend, the C10, the *C7*, and free amide NH bending vibrations, respectively.

The structures in Figure 6.14 are calculated structures of the higher energy minima on the potential energy surface. These structures are not assigned experimentally and are mere calculated high energetic minima on the potential energy surface. They represent both turn (Figure 6.14a and c) and fully extended (Figure 6.14c and d) backbone conformational families

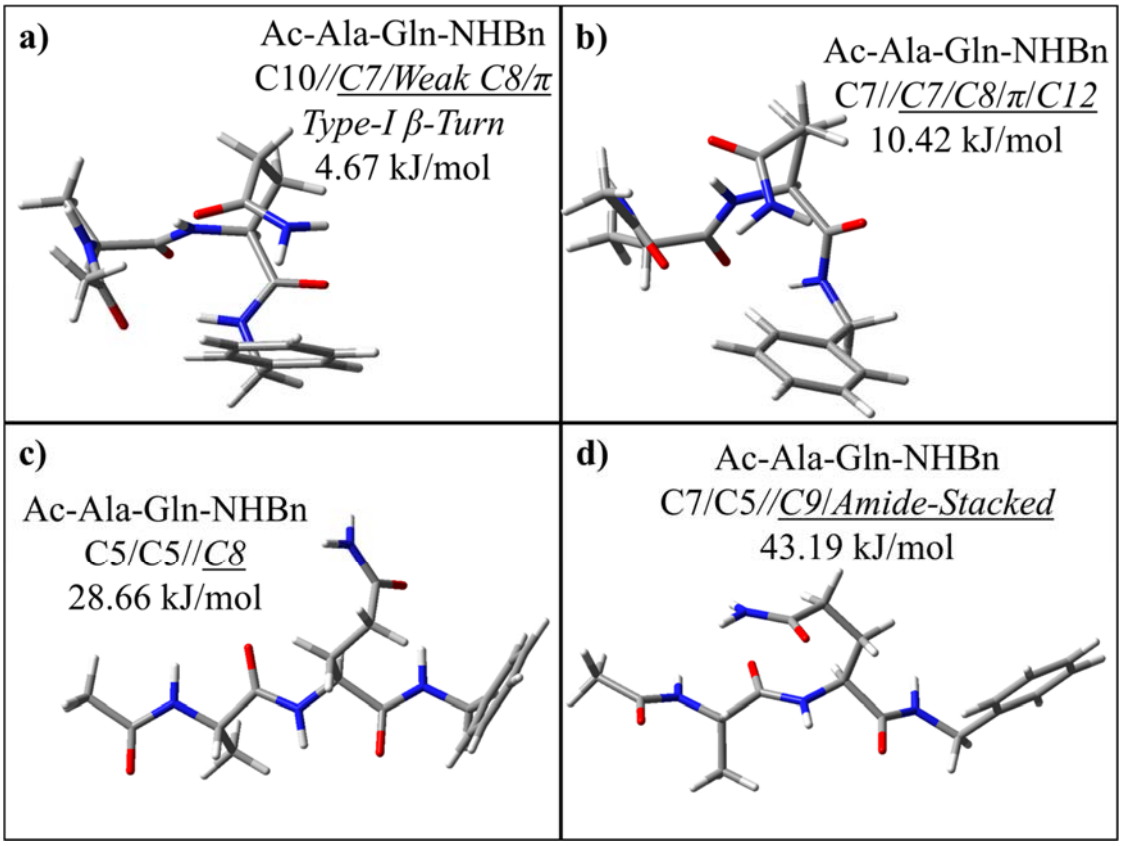


Figure 6.14. Low-energy and other structures of Ac-Ala-Gln-NHBn which were not experimentally assigned: (a) a C10//*C7/ Weak C8/π* type-I β-turn, (b) a C7//*C7/C8/π/C12* structure, (c) a C5/C5//*C8* structure, and (d) a C5/C5//*Amide-Stacked* structure.

6.4 Discussion

Using single-conformation spectroscopy, we have recorded infrared spectra of a total of eight conformations of four short Gln-containing peptides of different lengths and capping arrangements: three in Z-Gln-OH, one in Z-Gln-NHMe, three in Ac-Q-NHBn, and a single conformer in Ac-AQ-NHBn. The jet-cooled gas-phase data includes representative examples of five unique H-bonding architectures: three C5//*C8* conformers, one C5//*C9*, one amide-stacked structure, one C7//*C7/C8/π* double-bridge cap, and a C10//*C7/π* type I β-turn. Representative examples of each of these structural

types are shown in Figure 6.15. A sixth type, the *C7/π* conformer of Ac-Gln-NHBn, shares the same stabilizing *C7* Gln H-bond and backbone bend as in the second half of the type I β-turn, and therefore is not shown separately in the figure (Figure 6.16).

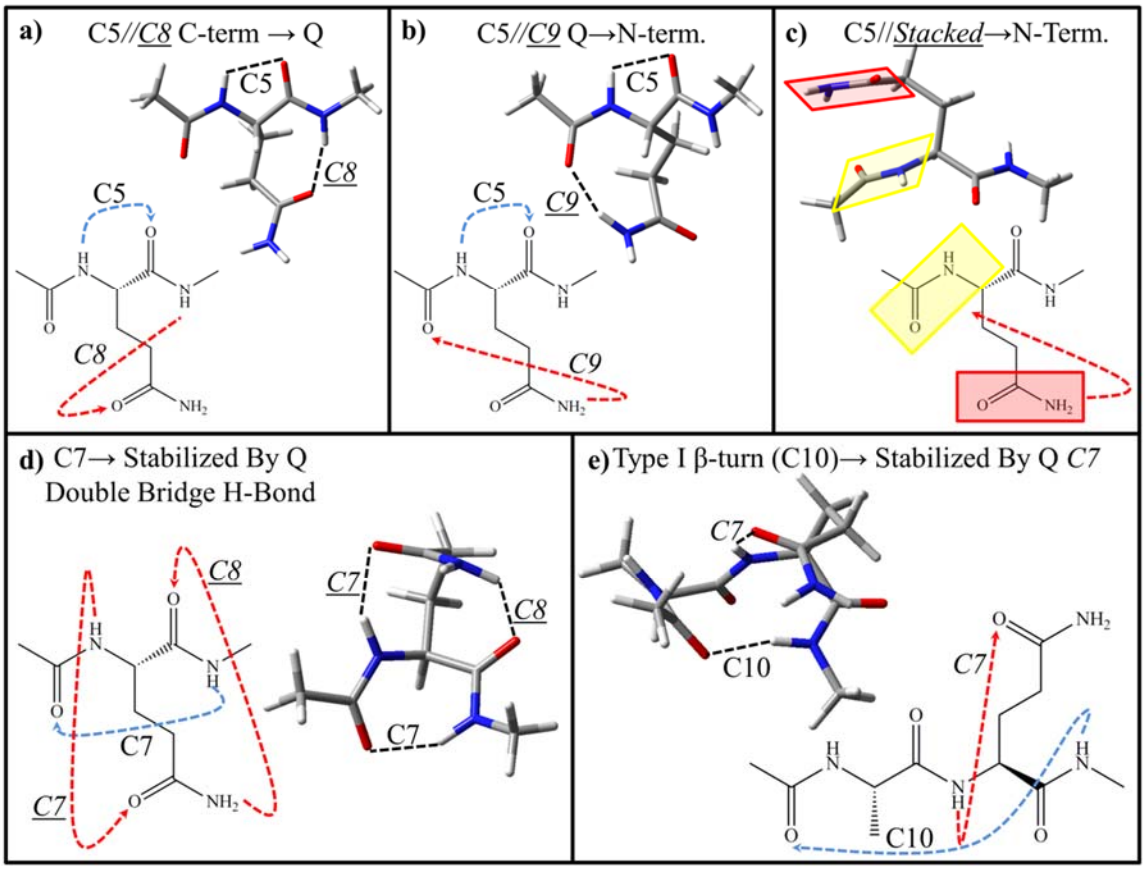


Figure 6.15. Hydrogen bonding motifs of the glutamine sidechain: (a) *C8* sidechain-to-backbone hydrogen bond towards the N-terminus, (b) *C9* sidechain-to-backbone hydrogen bond towards the C-terminus, (c) sidechain amide stacking towards the N-terminus, (d) backbone C7 hydrogen bond stabilized by a pair of Gln side-chain hydrogen bonds (*C7/C8*) that bridge adjacent amide groups in the peptide backbone , and (d) a type-I β-turn (C10) stabilized through a *C7* sidechain-to-backbone hydrogen bond towards the N-terminus.

The active role played by the Gln side chain in dictating structure, with its amide group capable of forming H-bonds to the peptide backbone amides, is clear from the figure. In each structural motif, the Gln sidechain is involved in hydrogen bond(s) that are either the primary H-bonding interaction stabilizing the structure or that enhance the stability of an otherwise common peptide backbone structure.

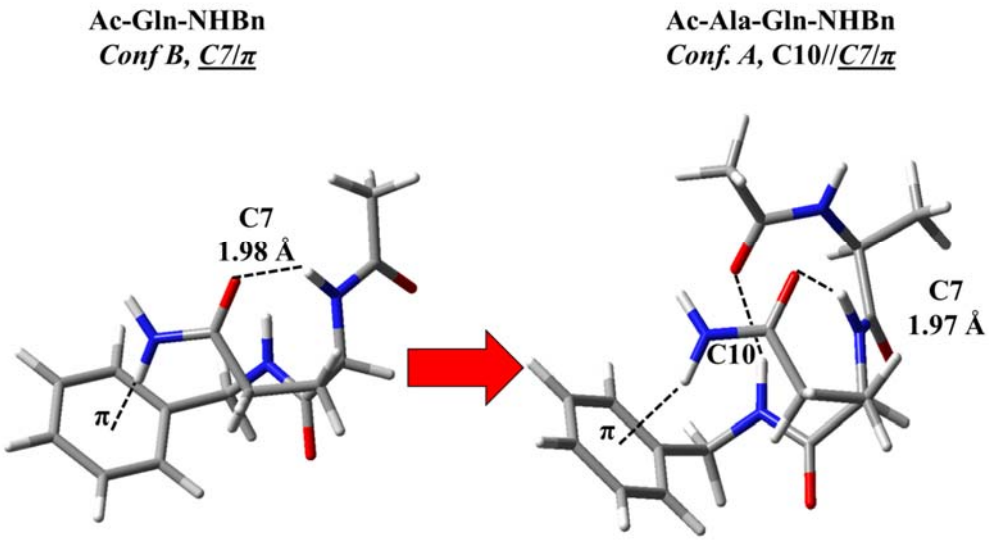


Figure 6.16. Comparison of the *C7/π* structure of glutamine which is found in both Ac-Gln-NHBn conformer B and Ac-Ala-Gln-NHBn conformer A.

Figure 6.17 presents a series of energy level diagrams that compare the calculated relative energies of the conformational minima within 15 kJ/mol of the global minimum for each of the capped amino acids (Z-Gln-OH, Z-Gln-NHMe, Ac-Gln-NHBn) and the sole capped dipeptide studied here (Ac-Ala-Gln-NHBn). The figure also includes calculations in which the cap containing the aromatic ring is removed, in order to assess

the effects of the aromatic ring on the stabilities of the structures. In Section 6.4.1 we use Figures 6.15 and 6.17 to assess the ways in which a single glutamine side chain interacts with, and stabilizes various local peptide backbone architectures, including the β-turn structure found in longer XAQY sequences in proteins. In Section 6.4.2, we will take a wider view of the potential energy surfaces, comparing the observed structures with other competing possibilities, and assessing the role played by the Z- and -NHBn caps.

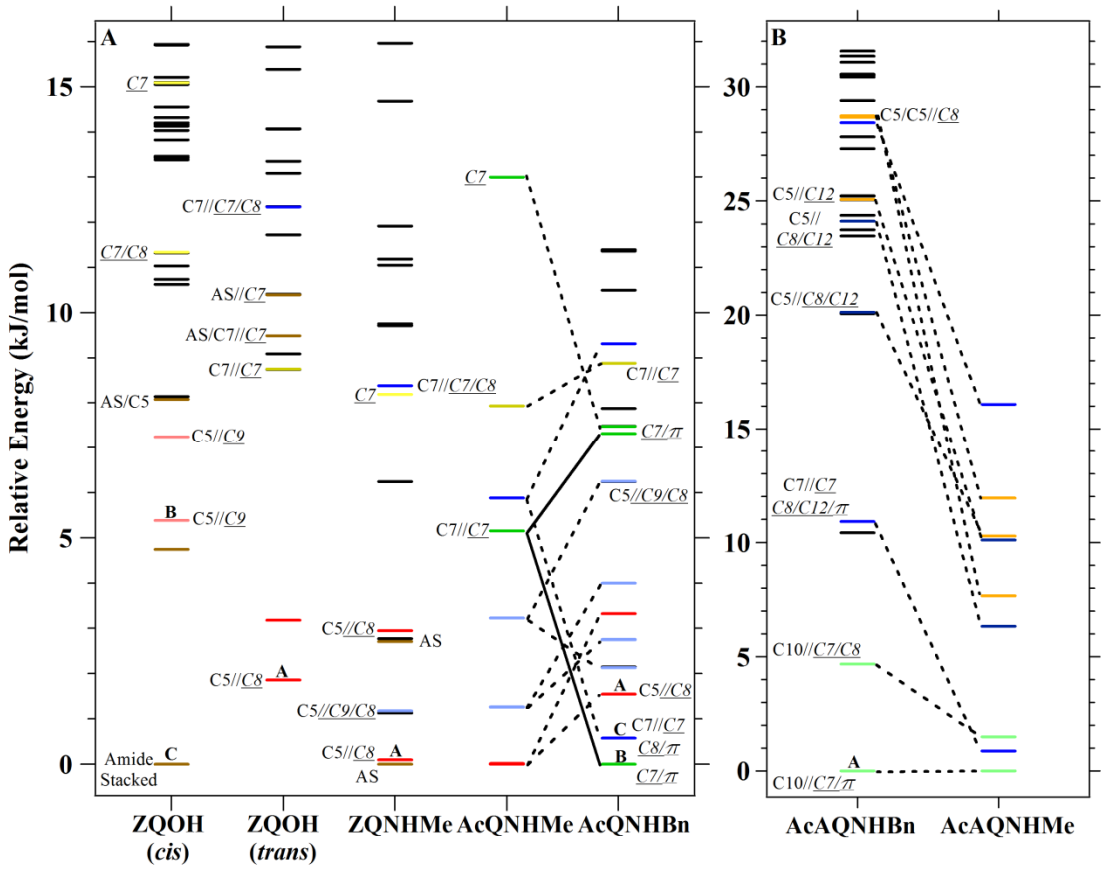


Figure 6.17. Energy level diagram for (a) the capped amino acids Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn and (b) capped dipeptide Ac-Ala-Gln-NHBn. The relative zero-point corrected energies of conformational minima are calculated at the DFT M05-2X/6-31+G(d) level of theory. The molecules observed experimentally are compared with Ac-Gln-NHMe and Ac-Ala-Gln-NHMe in the figure in order to provide a point of comparison of the effects of the aromatic cap at the N- and C-terminus. Levels connected with a dotted line retain the same geometry in the –NHMe capped structure as the -NHBn structures, solid lines indicate a change in geometry. Selected levels are labeled with their hydrogen bonding motif. The energy levels are color-coded by structural family: nominally amide-stacked (brown), C5//*C9* (pink), C5//*C8* (red), C5/C5 containing and C5//*C12* (orange), general glutamine bridging geometries (Navy blue), C7//*C7/C8* containing (blue), C5//*C9/C8* (light blue), general turned structures (yellow), *C7/π* (green), C7//*C7* (gold), β-turns (light green), general extended structures (maroon), and unclassified (black). Energy levels of assigned conformers are indicated by their letter labels.

6.4.1 Sidechain-to-Backbone Hydrogen Bonds and Local Peptide Backbone Structures

6.4.1.1 Extended Backbone Structures

Five of seven assigned conformations of the capped Gln-containing amino acids (Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn) contain an extended peptide backbone, with its characteristic nearest-neighbor C5 hydrogen bond. When the nearest-neighbor amide NH on the C-terminal side of the Gln residue forms an H-bond with the Gln C=O group, a *C8* H-bond is formed, as shown in Figure 6.15a. The flexibility and length of the Gln side chain is such that this *C8* H-bond is formed with good strength while maintaining an extended peptide backbone. C5//*C8* structures are formed in Z-Gln-NHMe and Ac-Gln-NHBn, while its OH···O=C analog is formed in Z-Gln-OH. In each case, these structures are among the calculated lowest-energy structures (Figure 6.17a).

Alternatively, when the Gln NH_2 group acts as H-bond donor to the amide C=O group on the N-terminal side, a *C9* H-bond is formed (Figure 6.15b). This side-chain to backbone H-bond is also best formed with the peptide backbone in an extended C5 conformation. Interestingly, the *C9* H-bonded ring is similar to that formed in γ-peptides studied previously by our group (Ac-γ^2-hPhe-NHMe and its cyclically-constrained counterpart),[47-51] which share the same three-carbon alkyl group spacing between amide groups with the same relative directions of the two amide groups necessary for C9 H-bond formation (-CONH-C-C-C-$CONH_2$). Indeed, the calculated H-bond lengths (Table 6.2) and dihedral angles are very similar to those in the γ-peptides, as is the frequency of the C9 NH stretch fundamental.

The C5//*Stacked* structure of Z-Gln-OH conformer C is the first amide-stacked structure experimentally observed in a natural amino acid. This structure is shown in generic form in Figure 6.15c, where the two amide stacking planes are identified. Like the *C9* H-bond just considered, amide stacking was first discovered and characterized in a synthetic γ-peptide foldamer.[47-50] It is an intriguing structure because it contains no amide-amide H-bonds, but is held together by electrostatic and dispersive interactions between the anti-parallel planes of the two amide groups. These amide planes are 0.07-0.10 Å further separated than those in the γ-peptide structure of Ac-γ^2-hPhe-NHMe (Table 6.2), with dihedral angles along the 3-carbon spacer that differ by no more than a few degrees.[47-50] Although this is the first identification of an amide-stacked geometry in a natural amino-acid, the structure receives additional stabilization beyond the interactions present in the amide stack from an $NH_2\cdots\pi$ interaction with the Z-cap aromatic ring, which would not be present when this cap is removed. As we shall see, this energetic stabilization from the cap makes it less likely that single Gln residues would take up this configuration in natural environments without additional stabilizing interactions (Figure 6.12). As Figure 6.17a shows, an amide-stacked structure is also calculated to be among the lowest energy structures of Z-Gln-NHMe, but is not observed experimentally. Instead, a single C5//*C8* structure is found, suggesting that the C5//*C8* structure is the global minimum, and that interconversion of the amide stacked conformer to the C5//*C8* minimum may occur during the cooling process.

The C5//*C8*, C5//*C9* and C5//*Stacked* structures all utilize sidechain to backbone interactions to stabilize the extended (C5) peptide backbone geometry characteristic of a β-sheet structure. It is possible that these local interactions of the Gln side chain with the

peptide backbone could play a role in enabling β-sheet formation by stabilizing the extended backbone structure locally so that H-bonding sites along the extended backbone on the opposite side from the Gln side chain(s) are free to engage in H-bonds that lead to β-sheet formation. If this stabilization of extended structures is so prevalent in Gln residues, it may be one contributing factor to glutamine's prevalence in fibril forming regions of peptides.

6.4.1.2 Turned Backbone Structures

While the majority of observed structures in our X-Gln-Y series (X=Z, Ac; Y=OH, NHMe, NHBn) incorporate the extended C5 peptide backbone, the single Gln side chain can also interact with the peptide amide groups to support other local peptide backbone architectures. Indeed, two of the three structures identified in Ac-Gln-NHBn involve a turned peptide backbone that is stabilized by strong sidechain-to-backbone hydrogen bonds: the inverse γ-turn C7//*C7*/*C8*/*π* geometry taken up by conformer C (Figure 6.11c) and the *C7*/*π* geometry of conformer B, (Figure 6.11b), both of which are calculated to be slightly lower in energy than the C5//*C8* conformer A (Figure 6.17a and 6.11a).

The C7//*C7*/*C8* structure, shown in schematic form in Figure 6.15d, involves the peptide backbone in an inverse γ-turn with its characteristic C7 H-bond. This C7 hydrogen bond is stabilized and strengthened by the sidechain forming a doubly hydrogen bonded bridge in which the Gln side-chain C=O and NH groups form *C7* and *C8* H-bonds, respectively, to the NH and C=O groups on adjacent amide groups in the peptide backbone. In order to accommodate this pair of additional H-bonds, the peptide

backbone must tip these groups out-of-plane towards the bridging Gln side chain, thereby enforcing formation of a peptide C7 H-bond. In so doing, the Gln side chain completes a hydrogen-bonded cycle involving three H-bonds that have the potential for cooperative strengthening, as each amide group is involved in H-bonds both as donor and acceptor.

In Ac-Gln-NHBn, the structure gains additional stabilization from the NHBn cap, which forms a π H-bond with the remaining NH group of the Gln NH_2 (not shown in Figure 6.15d). When the -NHBn group is removed to form Ac-Gln-NHMe (Figure 6.17a), the corresponding structure is about 5 kJ/mol higher in energy than the C5//*C8* structure (Table 6.4), which becomes the global minimum.

Table 6.4. Calculated energy comparison of Ac-Gln-NHBn and Ac-Gln-NHMe.

Hydrogen Bonding Pattern	Relative Energy (kJ/mol)	
	NHBn Cap	*NHMe Cap*
C7//*C7/C8* *Conf. C*	0.57	5.88
C7//*C7*	8.88	7.93
C5//*C9/C8*	6.25	3.23
C5//*C9/C8*	2.13	3.23
C5//*C9/C8*	2.76	1.27
C7//*C7/C8*	9.31	5.89
C5//*C9/C8*	3.99	1.26
C7/π → C7//*C7* *Conf. B*	0.00	5.15
C7//*C7*	7.32	5.15
C5//*C9/C8*	2.75	1.27
C7/π → *C7*	7.47	12.99
C5//*C8* *Conf. A*	1.54	0.01
C5//*C8*	3.32	0.00

With or without the NHBn cap, the inverse γ-turn found here has an unusually short H-bond, with an NH···O=C H-bond length (1.92 Å) about 0.1 Å smaller than C7 H-bonds in other natural peptides or synthetic foldamers studied to date, as summarized in Table 6.2. This is mirrored in its large frequency shift of the NH stretch fundamental, down to 3270 cm^{-1}, almost 200 cm^{-1} below its free value. The glutamine sidechain of this inverse γ-turn structure also forms a *C7* H-bond which is quite strong, with an NH···O=C H-bond length of 1.99 Å, while the *C8* H-bond is quite weak (2.24 Å), in part because the π H-bond is also present to the same NH_2 group.

The *C7/π* conformation has a backbone bend which contains no standard hydrogen bonds and Ramachandran angles of (-87.2, -9.1)° (Table 6.3). This unusual backbone geometry is stabilized by a sidechain-to-backbone *C7* with a hydrogen bond length of 1.98 Å. This hydrogen bond is very similar to the one formed in the inverse γ-turn structure, with an NH stretch frequency (3354 cm^{-1} vs. 3343 cm^{-1}) that reflects that similarity. Once again, the π H-bond between the other NH bond in the NH_2 group and the phenyl ring of the NHBn cap provides additional stabilization. When the phenyl ring is replaced by a methyl group to form Ac-Gln-NHMe, calculations predict a destabilization of about 5 kJ/mol, as shown in the energy level diagrams of Figure 6.17a (Table 6.4). Interestingly, the unusual backbone turn reconfigures to form a C7 H-bond in the process. As we shall see in the next section, the peptide backbone makes use of an unusual turn in forming the *C7/π* structure that reappears as an element of the energetically preferred global minimum in Ac-Ala-Gln-NHBn and its simpler parent Ac-Ala-Gln-NHMe, marking it as an important structural element in Gln's interaction with longer peptide sequences (Figure 6.16). To this we now turn.

6.4.1.3 Type I β-Turn

The results on Ac-Ala-Gln-NHBn are especially important, because they offer a first glimpse at the conformational preferences imposed by a single Gln residue in a longer peptide sequence, here extended by the inclusion of an Ala residue on the N-terminal side. As we have seen, the entire population of Ac-Ala-Gln-NHBn is funneled into a single structure, the C10//*C7/π* conformer shown in schematic form (minus the Bn cap) in Figure 6.15e.

As we have already noted, the C10 H-bond constitutes a type I β-turn, an important secondary structural element that is made all the more relevant by its role in linking successive anti-parallel β-sheet segments in the fibrils formed in neurodegenerative diseases related to poly-glutamine repeat disorders including Huntington's disease.[26,27,79] In many peptide sequences, formation of a β-turn with its C10 H-bond leaves the central amide group of the turn without an amide-amide H-bond. However, the C=O group of the Gln side chain forms a strong *C7* H-bond to this central amide NH, the same C7 H-bond formed in the double-bridge C7//*C7/C8/π* and *C7/π* structures of Ac-Gln-NHBn. This additional H-bond stabilizes the C10//*C7/π* β-turn structure, making it the global minimum in Ac-Ala-Gln-NHBn with only one other structure, also a β-turn, within 11 kJ/mol, as shown in Figure 6.17b. The observed structure remains the global minimum when the phenyl cap is replaced by methyl (Figure 6.17b), confirming that the Gln side chain adds to the stability of the type I β-turn principally through the Gln side chain H-bond to the interior amide NH rather than the π-hydrogen bond (Table 6.5).

Table 6.5. Calculated energy comparison of Ac-Ala-Gln-NHBn and Ac-Ala-Gln-NHMe.

Hydrogen Bonding Pattern	Relative Energy (kJ/mol)	
	NHBn Cap	*NHMe Cap*
C5/C5//*C8*	28.66	11.94
C7//*C7/C8/C12*	10.42	0.89
C5//*C12*	25.07	10.27
C5/C5//*C8*	28.73	7.66
C5/C7//*C7/C8*	28.43	16.08
C5//*C8/C12*	24.13	6.33
C7//*C12/C8*	20.13	10.10
C7//*C7/C8/C12*	10.91	0.88
C10//*C7*, Assigned Conformer	0.00	0.00
C5/C5//*Stacked*	43.19	23.23
C10//*C7/C8*	4.67	1.50

The assigned C10//*C7/π* type-I β-turn structure of Ac-Ala-Gln-NHBn can be compared against similar β-turns found in protein crystal and solution NMR structures with Ala and Gln as the *i*+1 and *i*+2 residues, respectively. One such comparison of this type is shown in Figure 6.18a and b. In these overlays the backbone geometry of Ac-Ala-Gln-NHBn is compared with the backbone of a type-I β-turn from the Myoglobin and Cytochrome C crystal structure (PDB ID:5cyt, C(AQ)C sequence).[71] The type-I β-turn from the crystal structure also contains a similar *C7* hydrogen bond between the interior amide NH and the Gln sidechain. These overlays demonstrate a striking agreement of the gas phase type-I β-turn with both the idealized type-I β-turn motif of the backbone (Figure 6.18a) but also representative type-I β-turn crystal structures of the same sequence (Figure 6.18b). This excellent agreement extends to several XAQY type-I β-turns found in the X-ray crystal and solution NMR structures of a number of proteins,

as summarized in Table 2. The average *i*+1 and *i*+2 Ramachandran angles of the crystal and solution NMR structures are (-59.9, -27.7)° and (-80.5, -16.2)°, in close correspondence with the gas-phase values of (-70, -13.3)° and (-73, -11)° (literature ranges for type-I β-turns are (-60, -30)° and (-90, 0)°).[65] At the same time, the majority of the PDB structures in Table 6.3 show the Gln amide group forming hydrogen bond(s) with other nearby parts of the protein or with its solvent environment. Thus, while the Gln sidechain H-bond to the center of the β-turn is stabilizing, it is not normative in solution or crystalline environments.

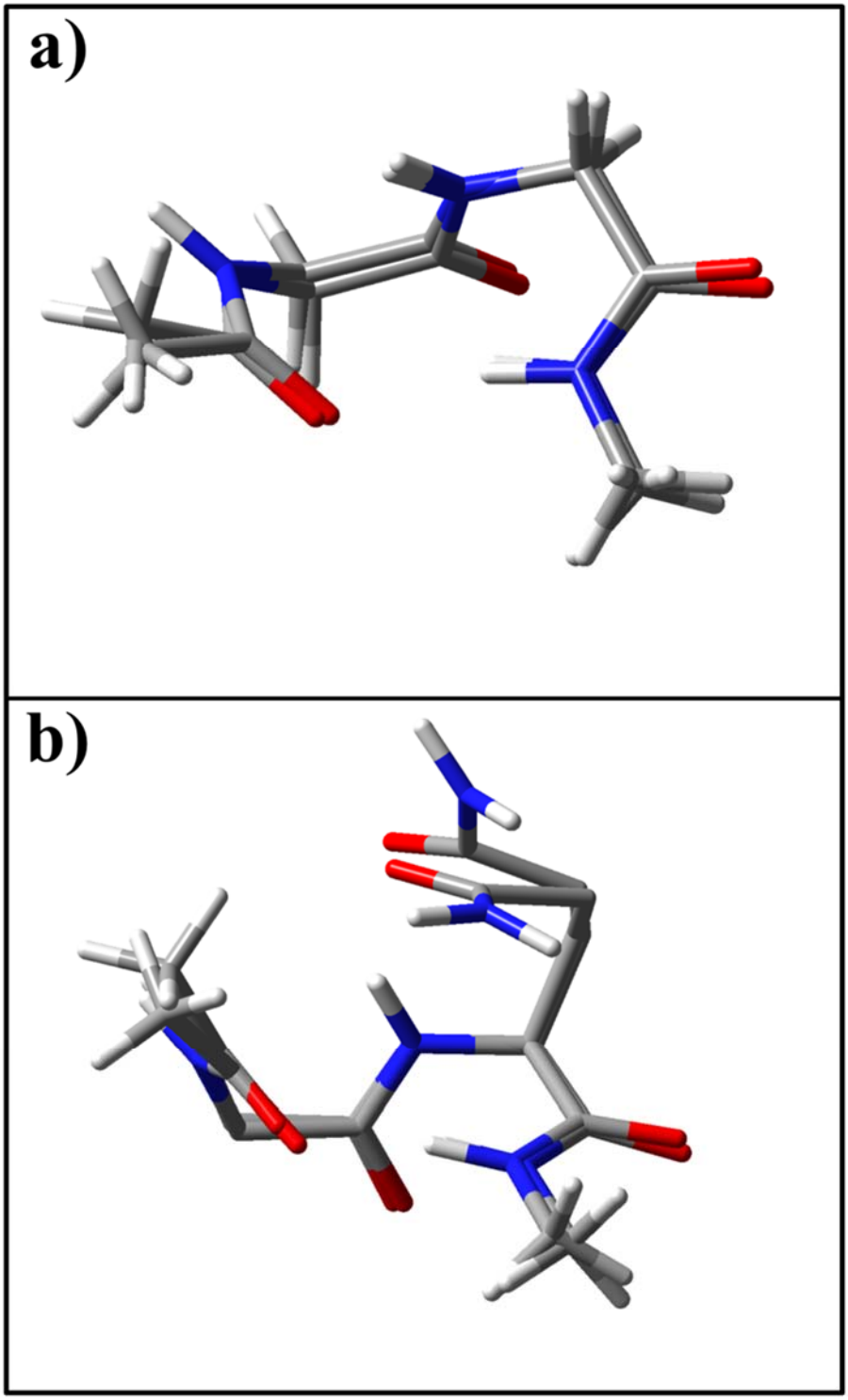


Figure 6.18. Two views (a,b) of the assigned gas-phase type-I β-turn of Ac-Ala-Gln-NHBn overlaid on top of the type-1 β-turn (residues 14-17) of Myoglobin and Cytochrome C (PDB ID:5cyt):[71] (a) View of the peptide backbones without the Gln sidechain, and (b) including the Gln side chains, showing their shared C10//*C7* motif.

The similarities in backbone structures between gas-phase and β-turns in proteins indicate that the inherent conformational preferences uncovered by the present gas-phase studies are robust to at least some of the environmental perturbations that arise from the protein environment and crystal formation. The present results provide IR spectra of this Gln-containing β-turn, which may facilitate its identification in solution, and contribute to a refinement of the theoretical models of the amide I/II regions on which certain condensed phase structural deductions are built (for examples see ref. [33] and references found within).

6.4.2 Exploring the Potential Energy Surface

The energy level diagrams in Figure 6.17 have provided firm evidence that the observed conformations are in every case among the lowest energy structures of each molecule calculated at the DFT M05-2X/6-31+G(d) level of theory. At the same time, in cases where more than one conformation is observed (Z-Gln-OH and Ac-Gln-NHBn), the observed conformations represent lowest-energy examples of different H-bonding architectures, consistent with a rugged potential energy landscape in which population is trapped, at least in part, behind barriers that prevent isomerization between certain families during the collisional cooling process. Assured by the reasonable correspondence between observation and calculation, in this section we take a more global view of the molecules' potential energy landscapes afforded by the near-exhaustive set of conformational minima calculated at this level of theory.

While the observed conformations are always among the lowest-energy computed structures, there are other structures that are also low in energy, yet are not observed. As others have noted, under thermalized conditions, it is the free energy landscape that is of primary importance for determining population abundances.[80,81] In the initial desorption plume, tightly-folded minima with stiffer vibrational frequencies will be discriminated against on entropic grounds relative to more extended structures. In addition, it is possible that population initially in one conformational minimum will be transferred into a lower-energy minimum during the collisional cooling process.[82]

Despite not being experimentally observed, it is worth commenting briefly on the other types of low-energy hydrogen-bonded families identified by the calculations, since they add to the rich variety of possibilities afforded by a single glutamine residue. Figures showing representative structures in these families are found in previous sections (Figures 6.6, 6.8, 6.12, and 6.14).

In the triad of capped amino acids (Z-Gln-NHMe, Ac-Gln-NHMe, and Ac-Gln-NHBn), C5//*C9/C8* conformations exist (labeled in light blue) with calculated energies within 2-3 kJ/mol of the global minimum. Conformers in this family engage the Gln amide group as a doubly H-bonded bridge to the C5 peptide backbone, with the C=O on the N-terminus and NH on the C-terminus forming *C9* and *C8* H-bonds, respectively, offering yet a fourth means by which a single Gln residue can stabilize an extended peptide backbone.

In Ac-Gln-NHBn, the *C7/π* structure is calculated to be the global minimum. As we have noted, this structure is highly unusual in that it has a turned backbone that has no backbone H-bond (i.e., neither C5 nor C7), and is stabilized by the bridging *C7* and π

H-bonds formed by the Gln side chain. When the benzyl cap is removed to form Ac-Gln-NHMe, the backbone reconfigures to form a C7//*C7* structure, which less stable than the C5//*C8* global minimum by 5 kJ/mol (Figure 6.17a, Table 6.4). This structure uses a different Gln side chain configuration in forming the C7 backbone-to-sidechain H-bond (Figure 6.12), than its C7//*C7/C8/π* counterpart, and the two possibilities are close in energy in the absence of an aromatic cap.

Finally, in Ac-Ala-Gln-NHBn, a second tightly folded structure, labeled as C7//*C7/C8/C12/π* drops in energy to within 2 kJ/mol of the C10//*C7/π* global minimum when the NHBn cap is removed in this larger peptide (Figure 6.17b, Table 6.5). This structure, shown in Figure 6.14b, engages the Gln side chain in three H-bonds involving both hydrogens of the NH_2 group and the C=O group. This illustrates the expanding possibilities that exist when the peptide backbone is lengthened.

6.5 Conclusion

The central issue we have sought to address in this work is the way in which the sidechain of a glutamine residue interacts with nearby amide groups, and the impact of such sidechain-to-backbone interactions on local conformational preferences. The structural families observed demonstrate the unique ability of the glutamine sidechain to form a variety of hydrogen bonds with the peptide backbone. The molecules that contain only a single Gln residue probe the preferences of the backbone immediately adjacent to the Gln sidechain, while Ac-Ala-Gln-NHBn reveals possibilities for longer range interactions.

A total of seven conformers of Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn have been experimentally assigned, including (i) C5//*C9*, C5//*C8*, and C5//*amide-stacked* structures that incorporate an extended peptide backbone, (ii) the *C7/π* turn in which the H-bond formed between the Gln side chain and the peptide backbone is the only one present, and (iii) the C7//*C7/C8/π* inverse γ-turn with its Gln doubly H-bonded bridge. In the larger Ac-Ala-Gln-NHBn molecule, all the population resides in a single type-I β-turn (C10//*C7/π*) conformation in which the Gln residue stabilizes the structure by forming a *C7* H-bond with the central amide group in the β-turn, which otherwise would be without a hydrogen-bonding partner. With the single exception of the *C7/π* conformer, in every case the sidechain-to-backbone hydrogen bonds are either the primary interaction stabilizing the molecule, or provide additional stabilization to a structure already prevalent in the absence of glutamine.

In this bottom-up approach to studying peptide structure, we also begin to address whether the structures preferred in the smaller peptides carry over to larger ones, and if so, in what way. Comparison of the *C7/π* structure of Ac-Gln-NHBn (Figure 6.11b) with the C10//*C7/π* conformer of Ac-Ala-Gln-NHBn (Figure 6.11d) displays the structural 'inheritance' between the two structures (Figure 6.16). The glutamine side-chain in two structures adopts the same *C7/π* geometry, with backbone Ramachandran angles about the Gln residues [(-87.2, -9.1)° cf. (-73, -11°)] and H-bond lengths [1.98 cf. 197 Å] that are strikingly similar. Furthermore, comparison of the backbone of the β-turn from Ac-Ala-Gln-NHBn with the backbones of a series of X-AQ-Y β-turns found in the protein data bank reveal them to be highly conserved, proving that the β-turn formed in

the isolated molecule is robustly retained in the crystalline protein environment, despite the presence of additional intermolecular interactions in the latter case.

Although this initial work on Gln-containing peptides has uncovered certain aspects of the primary conformational influences of a single Gln residue, several important interactions that are significant in fibril formation have not been explored. Sidechain-to-sidechain interactions between glutamine residues will undoubtedly play an important structure-determining role not only in monomers of glutamine rich peptides, but also in the aggregation of these peptides into fibrils. Studies of the present type on model Gln-rich peptides and their aggregates are needed to more directly probe the pivotal role played by Gln-Gln interactions, including those that stabilize the fibril structures via polar zippers.[8-10] Gln-Gln H-bonds are likely also to play a critical role in the progenitor species thought to be the key toxic species neurodegenerative diseases, motivating future work in this direction.

6.6 References

(1) Roberson, E. D.; Mucke, L. *Science* **2006**, *314*, 781-784.

(2) Goedert, M.; Spillantini, M. G. *Science* **2006**, *314*, 777-781.

(3) Goedert, M. *Science* **2015**, *349*.

(4) Soto, C. *Nat. Rev. Neurosci.* **2003**, *4*, 49.

(5) Ross, C. A. *Neuron* **2002**, *35*, 819-822.

(6) Ross, C. A.; Poirier, M. A. *Nature Medicine* **2004**, *10*, S10-S17.

(7) Mangiarini, L.; Sathasivam, K.; Seller, M.; Cozens, B.; Harper, A.; Hetherington, C.; Lawton, M.; Trottier, Y.; Lehrach, H.; Davies, S. W.; Bates, G. P. *Cell* **1996**, *87*, 493-506.

(8) Perutz, M. *Protein Science* **1994**, *3*, 1629-1637.

(9) Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D. *Nature* **2005**, *435*, 773-778.

(10) Sawaya, M. R.; Sambashivan, S.; Nelson, R.; Ivanova, M. I.; Sievers, S. A.; Apostol, M. I.; Thompson, M. J.; Balbirnie, M.; Wiltzius, J. J. W.; McFarlane, H. T.; Madsen, A. Ø.; Riekel, C.; Eisenberg, D. *Nature* **2007**, *447*, 453-457.

(11) Bleiholder, C.; Dupuis, N. F.; Wyttenbach, T.; Bowers, M. T. *Nature Chemistry* **2011**, *3*, 172-177.

(12) Shim, S.-H.; Gupta, R.; Ling, Y. L.; Strasfeld, D. B.; Raleigh, D. P.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6614-6619.

(13) Wang, L.; Middleton, C. T.; Singh, S.; Reddy, A. S.; Woys, A. M.; Strasfeld, D. B.; Marek, P.; Raleigh, D. P.; de Pablo, J. J.; Zanni, M. T.; Skinner, J. L. *J. Am. Chem. Soc.* **2011**, *133*, 16062-16071.

(14) Middleton, C. T.; Marek, P.; Cao, P.; Chiu, C.-c.; Singh, S.; Woys, A. M.; de Pablo, J. J.; Raleigh, D. P.; Zanni, M. T. *Nat Chem* **2012**, *4*, 355-360.

(15) Moran, S. D.; Woys, A. M.; Buchanan, L. E.; Bixby, E.; Decatur, S. M.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 3329-3334.

(16) Ross, C. A.; Poirier, M. A.; Wanker, E. E.; Amzel, M. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1-3.

(17) Zoghbi, H. Y.; Orr, H. T. *Annual Review of Neuroscience* **2000**, *23*, 217.

(18) Perutz, M. F.; Johnson, T.; Suzuki, M.; Finch, J. T. *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5355-5358.

(19) Bauer, M. T.; Gilmore, K. A.; Petty, S. A. *Biochemical and Biophysical Research Communications* **2011**, *406*, 348-352.

(20) Kecel, S.; Ozel, A. E.; Akyuz, S.; Celik, S. *Spectroscopy: An International Journal* **2010**, *24*, 219-232.

(21) Vaden, T. D.; Gowers, S. A. N.; de Boer, T. S. J. A.; Steill, J. D.; Oomens, J.; Snoek, L. C. *J. Am. Chem. Soc.* **2008**, *130*, 14640-14650.

(22) Tarditi, A. M.; Klipfel, M. W.; Rodriguez, A. M.; Suvire, F. D.; Chasse, G. A.; Farkas, O.; Perczel, A.; Enriz, R. D. *Journal of Molecular Structure: THEOCHEM* **2001**, *545*, 29-47.

(23) Xiong, K.; Punihaole, D.; Asher, S. A. *Biochemistry* **2012**, *51*, 5822-5830.

(24) Gsponer, J.; Haberthür, U.; Caflisch, A. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 5154-5159.

(25) Marchut, A. J.; Hall, C. K. *Biophysical Journal* **2006**, *90*, 4574-4584.

(26) Buchanan, L. E.; Carr, J. K.; Fluitt, A. M.; Hoganson, A. J.; Moran, S. D.; de Pablo, J. J.; Skinner, J. L.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5796-5801.

(27) Kar, K.; Hoop, C. L.; Drombosky, K. W.; Baker, M. A.; Kodali, R.; Arduini, I.; van der Wel, P. C. A.; Horne, W. S.; Wetzel, R. *Journal of Molecular Biology* **2013**, *425*, 1183-1197.

(28) Thakur, A. K.; Wetzel, R. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 17014-17019.

(29) Schneider, R.; Schumacher, M. C.; Mueller, H.; Nand, D.; Klaukien, V.; Heise, H.; Riedel, D.; Wolf, G.; Behrmann, E.; Raunser, S.; Seidel, R.; Engelhard, M.; Baldus, M. *Journal of Molecular Biology* **2011**, *412*, 121-136.

(30) Loquais, Y.; Gloaguen, E.; Habka, S.; Vaquero-Vara, V.; Brenner, V.; Tardivel, B.; Mons, M. *J. Phys. Chem. A* **2015**, *119*, 5932-5941.

(31) Gloaguen, E.; Loquais, Y.; Thomas, J. A.; Pratt, D. W.; Mons, M. *J. Phys. Chem. B* **2013**, *117*, 4945-4955.

(32) Jaeqx, S.; Du, W.; Meijer, E. J.; Oomens, J.; Rijs, A. M. *J. Phys. Chem. A* **2013**, *117*, 1216-1227.

(33) Buchanan, E. G.; James, W. H., III; Choi, S. H.; Guo, L.; Gellman, S. H.; Muller, C. W.; Zwier, T. S. *J. Chem. Phys.* **2012**, *137*, 094301-094316.

(34) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205-241.

(35) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827-8839.

(36) Burke, N. L.; Redwine, J. G.; Dean, J. C.; McLuckey, S. A.; Zwier, T. S. *International Journal of Mass Spectrometry* **2015**, *378*, 196-205.

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

(38) Gloaguen, E.; Pollet, R.; Piuzzi, F.; Tardivel, B.; Mons, M. *Phys. Chem. Chem. Phys.* **2009**, *11*, 11385-11388.

(39) Chin, W.; Piuzzi, F.; Dognon, J. P.; Dimicoli, I.; Mons, M. *J. Chem. Phys.* **2005**, *123*, 084301.

(40) Stearns, J. A.; Boyarkin, O. V.; Rizzo, T. R. *J. Am. Chem. Soc.* **2007**, *129*, 13820-13821.

(41) Brenner, V.; Piuzzi, F.; Dimicoli, I.; Tardivel, B.; Mons, M. *Angewandte Chemie-International Edition* **2007**, *46*, 2463-2466.

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

(43) Kusaka, R.; Zhang, D.; Walsh, P. S.; Gord, J. R.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2013**.

(44) Alauddin, M.; Gloaguen, E.; Brenner, V.; Tardivel, B.; Mons, M.; Zehnacker-Rentien, A.; Declerck, V.; Aitken, D. J. *Chem.--Euro. J.* **2015**, *21*, 16479-16493.

(45) James, W. H., III; Baquero, E. E.; Shubert, V. A.; Choi, S. H.; Gellman, S. H.; Zwier, T. S. *J. Am. Chem. Soc.* **2009**, *131*, 6574-6590.

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

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

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

(49) James, W. H., III; Buchanan, E. G.; Müller, C. W.; Dean, J. C.; Kosenkov, D.; Slipchenko, L. V.; Guo, L.; Reidenbach, A. G.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2011**, *115*, 13783-13798.

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

(51) Walsh, P. S.; Kusaka, R.; Buchanan, E. G.; James, W. H.; Fisher, B. F.; Gellman, S. H.; Zwier, T. S. *J. Phys. Chem. A* **2013**, *117*, 12350-12362.

(52) Dean, J. C.; Kusaka, R.; Walsh, P. S.; Allais, F.; Zwier, T. S. *J. Am. Chem. Soc.* **2014**, *136*, 14780-14795.

(53) Lubman, D. M. *Mass Spectrom. Rev.* **1988**, *7*, 559-592.

(54) Page, R. H.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988**, *88*, 4621-4636.

(55) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440-467.

(56) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287-303.

(57) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289-300.

(58) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J.V.; Cioslowski, J.; Fox, D. J.; Revision C.01 ed.; Gaussian, Inc.: Wallingford CT, 2009.

(59) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Research* **2000**, *28*, 235-242.

(60) European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe-site/pdbemotif/).

(61) Torii, H. *J. Phys. Chem. Lett.* **2012**, *3*, 112-116.

(62) Gloaguen, E.; Pagliarulo, F.; Brenner, V.; Chin, W.; Piuzzi, F.; Tardivel, B.; Mons, M. *Phys. Chem. Chem. Phys.* **2007**, *9*, 4491-4497.

(63) Dean, J. C.; Buchanan, E. G.; James, W. H., III; Gutberlet, A.; Biswas, B.; Ramachandran, P. V.; Zwier, T. S. *J. Phys. Chem. A* **2011**, *115*, 8464-8478.

(64) Chin, W.; Compagnon, I.; Dognon, J. P.; Canuel, C.; Piuzzi, F.; Dimicoli, I.; von Helden, G.; Meijer, G.; Mons, M. *J. Am. Chem. Soc.* **2005**, *127*, 1388-1389.

(65) Lewis, P. N.; Momany, F. A.; Scheraga, H. A. *Biochimica et Biophysica Acta (BBA) - Protein Structure* **1973**, *303*, 211-229.

(66) Xiang, S.; Short, S. A.; Wolfenden, R.; Carter, C. W. *Biochemistry* **1997**, *36*, 4768-4774.

(67) Wang, X.; Lou, Z.; Dong, X.; Yang, W.; Peng, Y.; Yin, B.; Gong, Y.; Yuan, J.; Zhou, W.; Bartlam, M.; Peng, X.; Rao, Z. *Journal of Molecular Biology* **2009**, *390*, 530-537.

(68) Guillet, V.; Ohta, N.; Cabantous, S.; Newton, A.; Samama, J.-P. *J. Biol. Chem.* **2002**, *277*, 42003-42010.

(69) Yao, X.; Chong, H. H.; Waltersperger, S.; Wang, M. T.; He, Y. X.; Cui, S. *New HIV entry inhibitor MTSFT/T23 complex (DOI: 10.2210/pdb4i2l/pdb)* **2014**.

(70) Andrews, S. R.; Taylor, E. J.; Pell, G.; Vincent, F.; Ducros, V. M.-A.; Davies, G. J.; Lakey, J. H.; Gilbert, H. J. *J. Biol. Chem.* **2004**, *279*, 54369-54379.

(71) Takano, T. *Refinement of Myoglobin and Cytochrome C (DOI: 10.2210/pdb5cyt/pdb)* **1989**.

(72) Jung, W.-S.; Lee, J.; Kim, M.-I.; Ma, J.; Nagamatsu, T.; Goo, E.; Kim, H.; Hwang, I.; Han, J.; Rhee, S. *PLoS ONE* **2011**, *6*, e22443.

(73) Fremont, D. H.; Monnaie, D.; Nelson, C. A.; Hendrickson, W. A.; Unanue, E. R. *Immunity* **1998**, *8*, 305-317.

(74) Hennig, S. E.; Jeoung, J.-H.; Goetzl, S.; Dobbek, H. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 5235-5240.

(75) Davies, C. W.; Chaney, J.; Korbel, G.; Ringe, D.; Petsko, G. A.; Ploegh, H.; Das, C. *Bioorganic & Medicinal Chemistry Letters* **2012**, *22*, 3900-3904.

(76) Clore, G. M.; Omichinski, J. G.; Sakaguchi, K.; Zambrano, N.; Sakamoto, H.; Appella, E.; Gronenborn, A. M. *Science* **1994**, *265*, 386-391.

(77) Morellet, N.; Meudal, H.; Bouaziz, S.; Roques, Bernard P. *Biochemical Journal* **2006**, *393*, 725-732.

(78) Nagashima, T.; Hayashi, F.; Yokoyama, S. *Solution structure of N-terminal domain of chromosome 12 open reading frame 2 (DOI: 10.2210/pdb2cs4/pdb)* **2005**.

(79) Hoop, C. L.; Lin, H.-K.; Kar, K.; Magyarfalvi, G.; Lamley, J. M.; Boatz, J. C.; Mandal, A.; Lewandowski, J. R.; Wetzel, R.; van der Wel, P. C. A. *Proc. Natl. Acad. Sci. USA* **2016**.

(80) Shubert, V. A.; Baquero, E. E.; Clarkson, J. R.; James, W. H.; Turk, J. A.; Hare, A. A.; Worrel, K.; Lipton, M. A.; Schofield, D. P.; Jordan, K. D.; Zwier, T. S. *J. Chem. Phys.* **2007**, *127*.

(81) Gloaguen, E.; de Courcy, B.; Piquemal, J. P.; Pilme, J.; Parisel, O.; Pollet, R.; Biswal, H. S.; Piuzzi, F.; Tardivel, B.; Broquier, M.; Mons, M. *J. Am. Chem. Soc.* **2010**, *132*, 11860-11863.

(82) Ruoff, R. S.; Klots, T. D.; Emilsson, T.; Gutowsky, H. S. *J. Chem. Phys.* **1990**, *93*, 3142.

CHAPTER 7. INHERENT CONFORMATIONAL PREFERENCES OF AC-GLN-GLN-NHBN: SIDECHAIN HYDROGEN BONDING SUPPORTS A β-TURN THAT MAY BE RELEVANT TO POLYGLUTAMINE SEGMENT MISFOLDING

7.1 Introduction

Neurodegenerative diseases such as Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD) cause significant degradation of physical motor function and mental acuity. Although these diseases affect different regions of the brain, and involve different precursor proteins, pathogenesis in each case seems to involve the misfolding and subsequent aggregation of precursor proteins into characteristic β-sheet rich fibrils.[1,2]

The CAG codon repeat, or glutamine repeat, neurodegenerative disorders are triggered by the extension of a polyglutamine (poly-Gln) segment within a precursor protein beyond a threshold value, which for Huntington's disease is 36 glutamine residues.[3,4 5] Long poly-Gln segments are highly prone to aggregation. The fundamental studies reported here are motivated by the hypothesis that elucidating the intrinsic conformational propensities of glutamine-rich peptides will yield a deeper understanding of poly-Gln-driven disease pathogenesis.[2,4,6]

Recent studies have identified the formation of β-turns and β-hairpins in poly-Gln segments as critical to fibril formation. Wetzel and co-workers showed that placement of β-hairpin-promoting moieties at the center of a polyglutamine tract greatly accelerates

aggregation and reduces the number of molecules required to nucleate aggregation.[7,8] These studies, however, involved the use of glutamine-free turn-forming units; therefore, the contribution of local β-turn formation to fibril formation by pure poly-Gln segments remains uncertain. Even in well-formed poly-Gln fibrils, spectroscopic signals associated with a turn would be difficult to detect in the context of much stronger signals from β-sheet regions.[9,10] The potential for β-turn/β-hairpin formation in glutamine-rich segments is relevant not only to fibrillar assembly but also to the formation of oligomeric species that may be critical in terms of poly-Gln-associated neurotoxicity.[11-15]

Here we describe an unusual experimental strategy for assessing local folding preferences in a glutamine-rich peptide. Using single-conformation laser spectroscopy, we have examined the inherent conformational preferences of Ac-Gln-Gln-NHBn (Figure 7.1a), which is the smallest glutamine-rich molecule competent to form a β-turn. These studies have been conducted in the gas phase, which can be regarded as the ultimate nonpolar environment, in which only intramolecular noncovalent interactions are possible.[16-18] Our results demonstrate that, in the absence of intermolecular interactions with water or other peptides, the dominant observed conformer of Ac-Gln-Gln-NHBn contains a type-I β-turn, with each glutamine side chain carbonyl forming a hydrogen bond to the nearest NH in the backbone (Figure 7.1b). If β-turn formation within a poly-Gln segment is a rare event that initiates the misfolding pathway,[7] our work suggests that a nonpolar environment may encourage misfolding.

Related previous studies revealed sidechain-to-backbone and backbone-to-backbone hydrogen bonding patterns that are most prevalent for an isolated glutamine residue.[19] The current study of Ac-Gln-Gln-NHBn was aimed at revealing the more

complex interplay among internal interactions when sidechain-to-sidechain hydrogen bonding is possible.

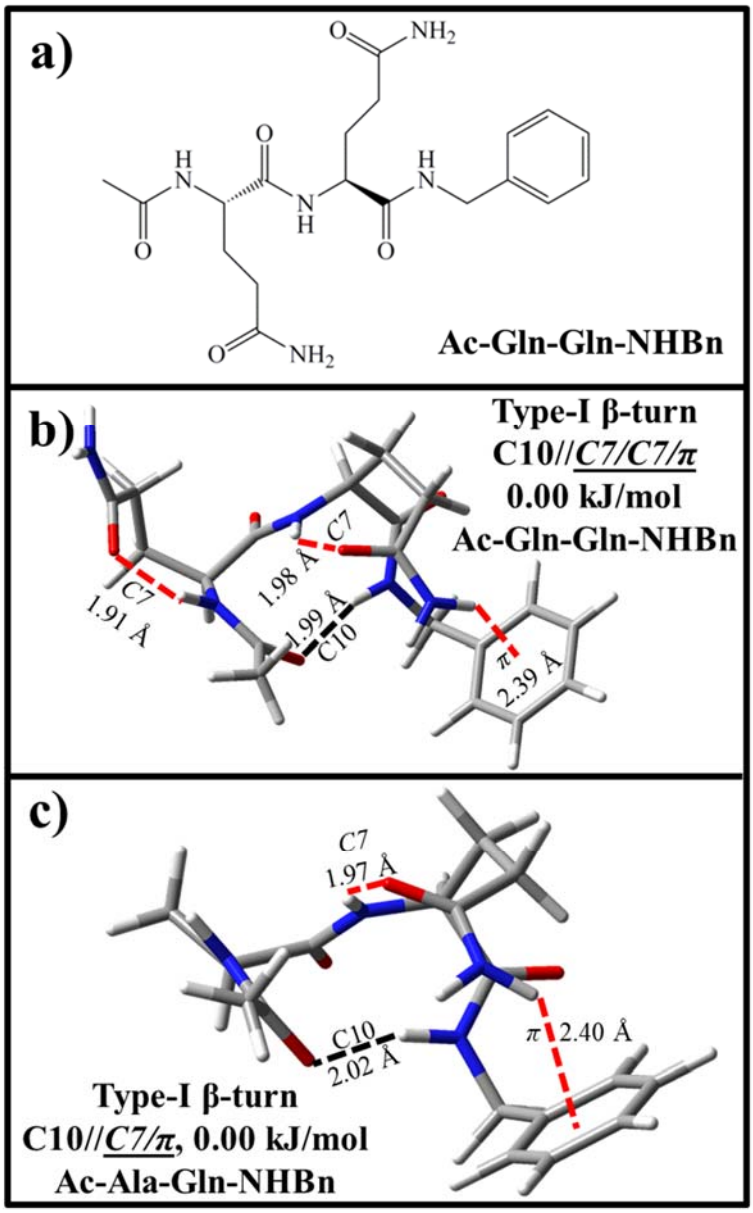


Figure 7.1. (a) The chemical structure and (b) experimentally assigned type-I β-turn (C10//*C7/C7/π*) structure of Ac-Gln-Gln-NHBn. (c) For comparison, the type-I β-turn (C10//*C7/π*) assigned to the sole experimentally identified structure of Ac-Ala-Gln-NHBn. Structures were calculated at the M05-2x/6-31+g(d) level of theory.[19]

7.2 Experimental

A full description of the experimental and theoretical methods used in this chapter can be found in Chapter 2. Sample was synthesized at the University of Wisconsin-Madison and used without further purification. Laser desorption was used to bring the peptide into the gas phase, where it is entrained in a supersonic expansion and cooled into the zero-point levels of low-lying conformational minima. The electronic spectrum due to the NHBn cap was recorded using resonant two-photon ionization (R2PI) while the single-conformation IR spectrum was obtained using resonant ion-dip infrared (RIDIR) spectroscopy.[16-18,20,21]

Generating a set of structures for Ac-Gln-Gln-NHBn is a multi-step process involving several different computational methods. An initial set of starting structures is generated from a conformational search using the Amber* force field[22] in the MACROMODEL suite of programs.[23] From this set of structures calculated using the force field, a subset of the low-energy structures (~90 structures) are fully optimized using Truhlar's hybrid density functional M05-2x with the 6-31+g(d) basis set.[24,25] As described in this report, the similarities between the spectra of Ac-Ala-Gln-NHBn and Ac-Gln-Gln-NHBn points towards a retained conformational preference. With this in mind, the AQ structure was altered to be QQ by replacing the sidechain of the first residue, while retaining the C10//*C7/π* structure that was low energy from the set of conformations optimized from the conformational search performed on the Ac-Ala-Gln-NHBn. Two different *C7* hydrogen bonds are possible for the *i+1* Gln residue whilst still retaining the C10//*C7/π* architecture. The *C7*'s differ in the approach of the sidechain with one formed "in-plane" and one formed "out-of-plane". Both

structures will be shown and discussed in this chapter. These altered structures were then optimized using the standard DFT method and were calculated to be the two lowest energy minima. Unfortunately these *exact* structures are missing from the force field conformational search; the C10//*C7/C7/π* family is found by the search. Harmonic vibrational frequencies were calculated for each of the optimized structures to allow for comparison with experimental IR spectra, and zero-point energy correcting the energies of optimized structures. Harmonic frequencies were scaled by 0.9406 in the NH-stretching region to account for anharmonicity in the calculated normal modes.

The naming scheme used in this work is retained from the earlier study on Ac-A-Q-NHBn, by listing the hydrogen bonds by type ('Cn' = an n-membered H-bonded ring), ordered from N- to C-terminus as follows: backbone-backbone//*sidechain-backbone*//**sidechain-sidechain**.[19]

7.3 Results and Discussion

Figure 7.2a shows the R2PI spectrum of Ac-Gln-Gln-NHBn recorded in the S_0-S_1 origin region of the –NHBn UV chromophore (~37450-37800 cm^{-1}). The spectrum shows several sharp transitions that appear on a background that likely arises due to incomplete cooling of a fraction of the laser desorbed molecules in the molecular beam. The transition at 37628 cm^{-1} is the S_0-S_1 origin transition for the single identified conformer for Ac-Gln-Gln-NHBn. Figure 7.2b shows the experimental RIDIR spectrum in the NH-stretch region. Five well resolved transitions are evident at 3562, 3522, 3445, 3404, and 3345 cm^{-1}, while a sixth asymmetrically broadened transition at 3324 cm^{-1} shows a shoulder centered around 3334 cm^{-1} (Table 7.1). Comparison of the experimental

spectrum with the calculated frequencies and infrared intensities for the global minimum structure—a type-I β-turn C10//*C7/C7/π* conformation—shows excellent agreement in both absolute frequency and intensity pattern (Table 7.1 and Figure 7.2b).

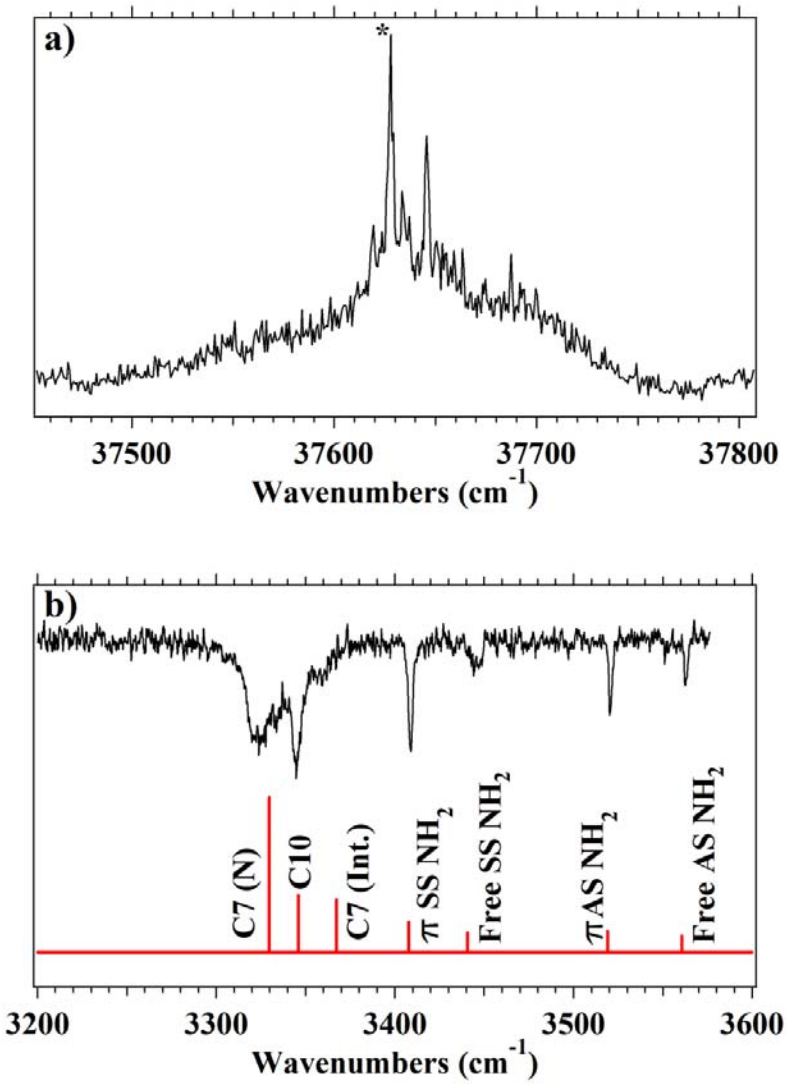


Figure 7.2. (a) R2PI spectrum of Ac-Gln-Gln-NHBn recorded in the monomer mass channel (m/z= 405) in the –NHBn S_0-S_1 origin region (~37450-37800 cm-1). (b) RIDIR spectrum recorded in the NH-stretch region for the major conformer. The labeled stick spectra below the experimental trace display the scaled NH stretch harmonic vibrational frequencies for the assigned C10//*C7/C7/π* type-I β-turn conformer. The starred transition (*) was used to record the RIDIR spectrum.

Table 7.1. Experimental and Theoretical Vibrational Frequencies for the Ac-Ala-Gln-NHBn and Ac-Gln-Gln-NHBn β-turns.

Structure Name	Hydrogen Bond Motif	NH Stretch Frequency (cm^{-1})								
		Asym. NH_2			Sym. NH_2			NH Stretch		
		Exp.	Th.	H-Bond	Exp.	Th.	H-Bond	Exp.	Th.	H-Bond
Ac-Ala-Gln-NHBn[a]										
Conf. A	C10//*C7*/π	3520	3526	π	3408	3411	π	3470	3471	Free
								3361/ 3357	3380/ 3378	C10/*C7*
Ac-Gln-Gln-NHBn										
Conf. A	C10//*C7*/*C7*/π//	3562	3561	Free	3445	3441	Free	3345	3367	*C7* *(Int.)*
		3521	3519	π	3409	3408	π	3334	3346	C10
								3324	3330	*C7* *(N-term.)*

[a]Ref [19].

Note that both Gln side chains engage in C7 hydrogen bonds with the peptide backbone, thereby forming a tightly held and highly stable β-turn. The close agreement between the experimental and calculated IR spectra allows the transitions to be readily assigned. The transitions at 3562 and 3445 cm^{-1} are the free antisymmetric and symmetric stretches of the *i+1* Gln $–NH_2$ moiety, respectively. The π bound antisymmetric and symmetric $–NH_2$ stretch fundamentals of the *i+2* Gln are shifted slightly to lower frequency from their nominally free positions, to 3521 and 3409 cm^{-1}, respectively due to formation of a π H-bond. The final closely-spaced set of broadened transitions arise from the strongly hydrogen bonded backbone amide NH groups. The transitions at 3345, 3334, and 3324 cm^{-1} are assigned to the *C7* ring formed between the interior NH and the *i+2*

Gln sidechain, the backbone-backbone C10, and the *C7* formed by the *i+1* Gln sidechain and the N-terminal NH, respectively (Table 7.1 and Figure 7.2b).

These three hydrogen bonds work in concert to stabilize the structure. Their NH stretch fundamentals are shifted down in frequency by about 20-40 cm^{-1} relative to the corresponding C10/C7 pair in Ac-Ala-Gln-NHBn (Figure 7.3).

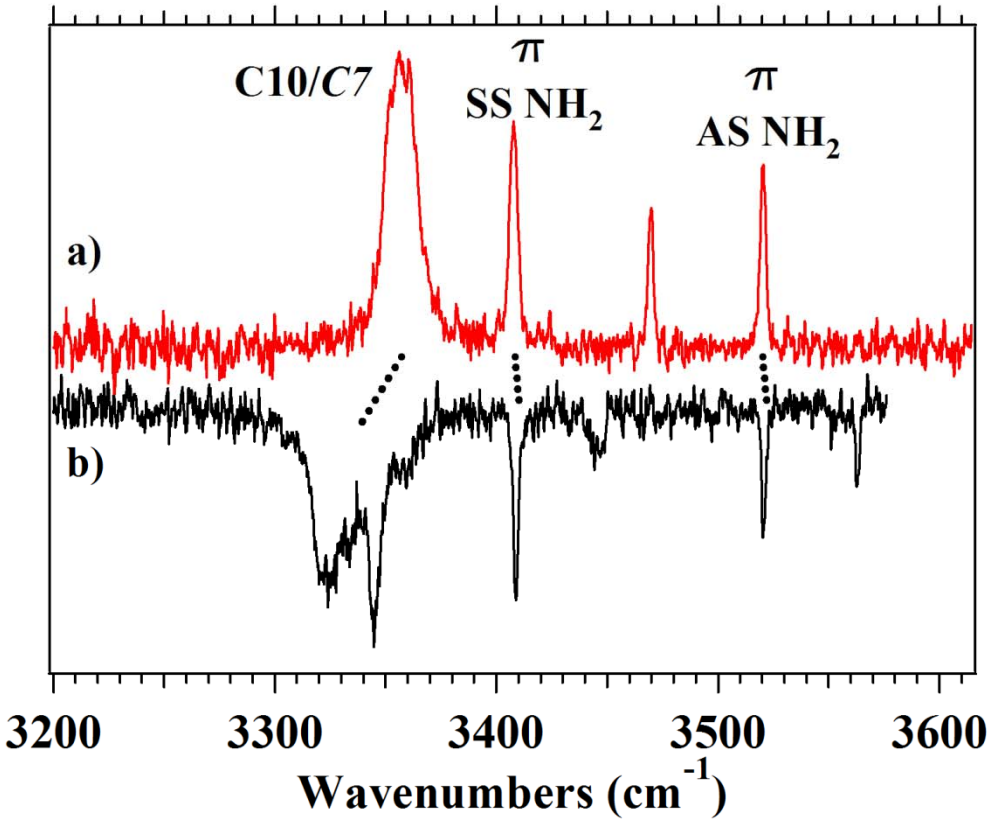


Figure 7.3. Experimental NH-stretch RIDIR spectra for the type-I β-turn conformers of (a) Ac-Ala-Gln-NHBn and (b) Ac-Gln-Gln-NHBn. The spectral signatures of the *i+2* Gln residue are highlighted.[19]

The type-I β-turns formed by Ac-Ala-Gln-NHBn and Ac-Gln-Gln-NHBn are structurally similar (Figure 7.1 b and c, Figure 7.3, Table 7.2), but there is both

spectroscopic and computational evidence for a synergistic stabilization induced by the pair of C7 H-bonds between the two Gln side chains and the peptide backbone. The *C7* H-bond formed by the *i+1* Gln C=O group to the N-terminal NH is calculated to be unusually short (1.91 Å), and this is reflected in a lowering in the frequency of its NH stretch fundamental by about 30 cm^{-1} relative to the corresponding C7 H-bond in Ac-Ala-Gln-NHBn (Figure 7.3, Tables 7.2 and 7.3). A similar shift is observed in the C10 H-bond that characterizes the β-turn, with calculated reduction in length from 2.02 to 1.99 Å. These transitions are also broadened significantly compared to non-hydrogen bonded NH stretch fundamentals as a result of the strong hydrogen bonding interaction.

Table 7.2. Hydrogen bond distances for hydrogen bonds found in Ac-Ala-Gln-NHBn and Ac-Gln-Gln-NHBn.

Molecule	Hydrogen Bond	
	Type	Distance (Å)
Ac-Ala-Gln-NHBn[a,b]		
Conf. A	C10	2.02
(C10//*C7*/*π*)	*C7*	1.97
	π	2.40
Ac-Gln-Gln-NHBn		
Conf. A	C10	1.99
C10//*C7*/*C7*/*π*//	*C7 (Int.)*	1.98
	C7 (N-term.)	1.91
	π	2.39

[a] From ref. [19].

[b] The C10 hydrogen bond length was inaccurately reported in ref. [19] and the correct value is reported here.

Table 7.3. Ramachandran angle comparison of type-I β-turns with sequence X(QQ)X.

Structure	Turn Sequence/ Residue #	Turn Residues *(i+1)* (ϕ,ψ)°	Turn Residues *(i+2)* (ϕ,ψ)°
Ac-AQ-NHBn[a]	X(AQ)X	(-70, -13.3)	(-73, -11)
Ac-QQ-NHBn	X(QQ)X	(-72.5, -8.0)	(-71.2, -12.5)
Antimicrobial Protein, Viral Protein[b]	F(QQ)I 48-51	(-63.7, -14.2)	(-75.4, -24.7)
Ribosome Inhibitor, Hydrolase[c]	D(QQ)T 9-12	(-67.7, -6.2)	(-102.7, -8.4)
Pyrophosphorylase[d]	L(QQ)G 78-81	(-59.2, -34.6)	(-94.4, 21.0)
Serum Protein[e]	N(QQ)N 141-144	(-48.1, -39.8)	(-101.7, 8.7)
Sugar Binding Protein, Oxidoreductase[f]	E(QQ)V 55-58	(-61.7, -27.8)	(-87.0, -11.2)
Transferase/ Immune System[g]	H(QQ)W 208-211	(-63.8, -30.6)	(-99.9, 8.9)
Viral Protein Replication[h]	L(QQ)L 22-25	(-63.7, -41.4)	(-60.4, -20.3)
Membrane Protein, heme Binding Protein[i]	S(QQ)D 154-157	(-63.4, -7.5)	(-91.2, -21.4)
Average of Crystal Structures	X(QQ)X	(-61.4, -25.3)	(-89.1, -5.9)

[a] From ref. [19].
[b] From ref. [26] (PDB ID: 4f87).
[c] From ref. [27] (PDB ID: 1m2t).
[d] From ref. [28] (PDB ID: 1dy3).
[e] From ref. [29] (PDB ID: 1c28).
[f] From ref. [30] (PDB ID: 3e7j).
[g] From ref. [31] (PDB ID: 3grw).
[h] From ref. [32] (PDB ID: 2ahm).
[i] From ref. [33] (PDB ID: 3csL).

The observed gas-phase structure can be directly compared with naturally occurring type-I β-turns of sequence Xaa-Gln-Gln-Yaa from crystal and NMR structures of proteins.[34] Table 7.3 provides the dihedral angles for the *i+1* and *i+2* residues for eight such structures. It is immediately clear from this comparison that the gas phase and condensed phase structures share the same backbone arrangement, confirming the gas phase structure as an accurate representation of folding patterns common to Gln-Gln segments in condensed phases. The crystal and NMR structures collectively manifest a degree of inhomogeneity in the β-turns that is not observed in the Ac-Ala-Gln-NHBn or Ac-Gln-Gln-NHBn gas phase structures, due to interactions of Gln side chains with their surroundings in crystals or solution, a factor that is irrelevant in the gas phase. Our observation of a specific pattern of Gln sidechain-to-backbone hydrogen bonding in the gas phase, and the evidence that such peripheral hydrogen bonds stabilize the β-turn, raises the possibility that nonpolar environments promote turn formation and ultimately fibril formation by glutamine-rich segments of proteins.

In order to better understand the role of β-turn structures in the larger context of other possible structures for the isolated molecule, the energy level diagram shown in Figure 7.4 was constructed that summarizes the relative energies of all structures within 37 kJ/mol of the global minimum calculated at the DFT M05-2X/6-31+g(d) level of theory. This diagram is surprisingly sparse in its low energy region, with only 8 structures observed in the first 15 kJ/mol. Even more striking, of these eight, four are type-I β-turns, as are three of the lowest four, including the global minimum, which is the single observed conformation. It is clear on this basis that the β-turn is preferred over other structures for Ac-Gln-Gln-NHBn in non-polar environments. The –NHBn cap used in

this molecule contributes modestly to stabilizing the observed conformation through the NH···π hydrogen bond. As is shown in Figure 7.4b, when the low energy structures are re-optimized with the NHBn cap replaced by NHMe, the assigned conformation is not the global minimum, but remains amongst the lowest energy structures.

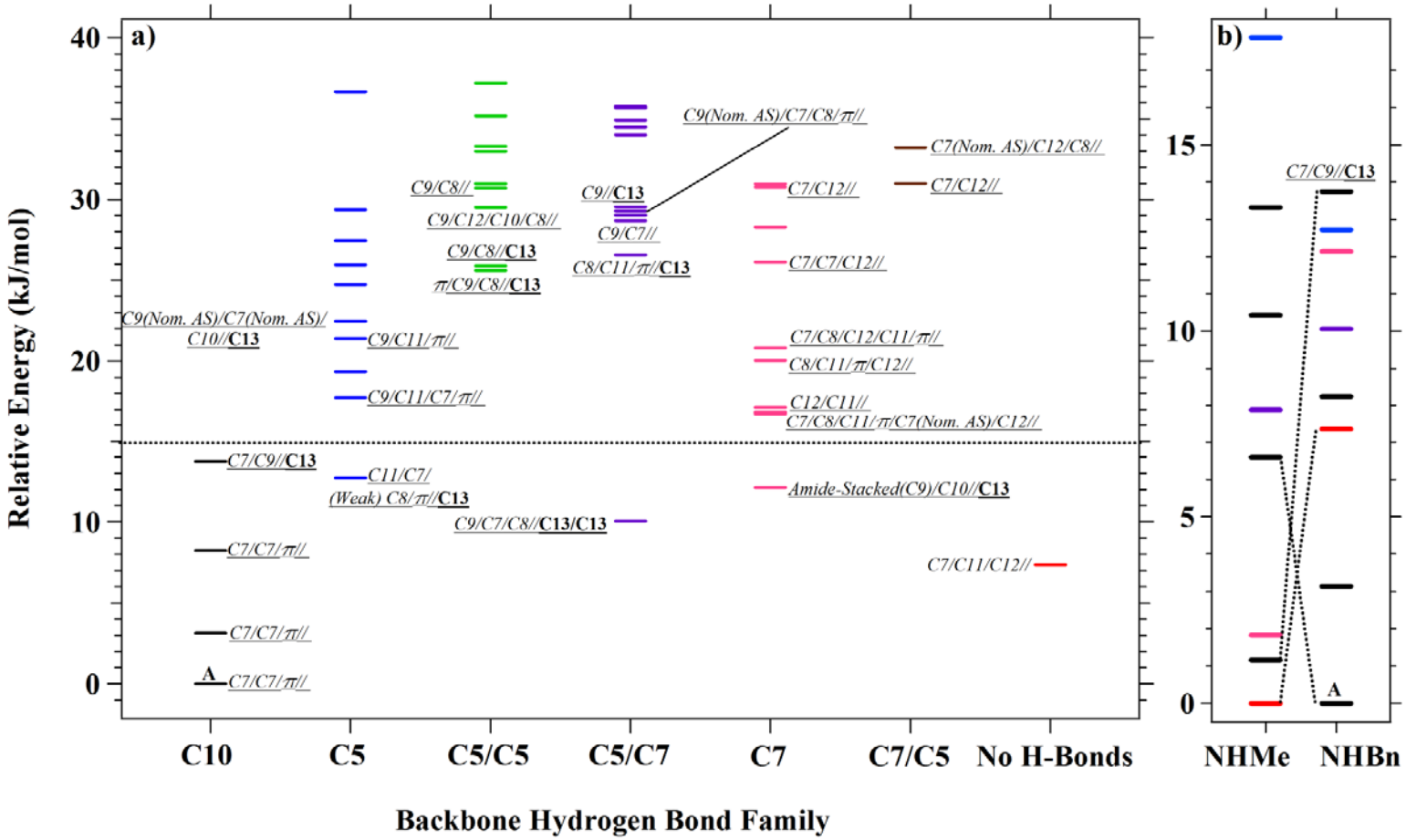


Figure 7.4. (a) Energy level diagram containing all calculated structures within 37 kJ/mol of the calculated global minimum at M05-2x/6-31+g(d). Structures are grouped by their peptide backbone architecture involving the backbone-backbone hydrogen bonds. Labels on individual structures list the sidechain-backbone (first, in italics) and sidechain-sidechain (second, in bold) hydrogen bonds, separated by a double-slash (//). (b) Comparison of the energies of low-energy (below dashed line in (a)) conformations calculated with the –NHBn or –NHMe caps. Backbone-backbone hydrogen bond families are color coded as in part (a).

The preference for β-turn formation in Ac-Gln-Gln-NHBn is all the more notable when compared with the preferences observed in other capped dipeptides studied to date that don't contain Gln residues.[35,36] In these studies similar β-turn motifs (C10) are observed experimentally in some cases, but the C10 structures are often minor conformers that are in competition with C5/C5 and C5/C7 hydrogen bonded families. The presence of β-turn motifs is not surprising in small peptides because of the limited number of hydrogen bonding patterns available to peptides this size. However, the studies of Ac-Ala-Gln-NHBn and Ac-Gln-Gln-NHBn mark the first study on naturally occurring amino acids in which only a β-turn motif is observed.[35,36] We surmise on this basis that the glutamine sidechains are "locking-in" the type-I β-turn.

Although the assigned global minimum structure possesses no sidechain-to-sidechain hydrogen bonds, Ac-Gln-Gln-NHBn is the smallest molecule in which such interactions are possible. Indeed, such sidechain-to-sidechain hydrogen bonds are prevalent among the higher energy structures (Figure 7.4, 7.5, and 7.6). The most common calculated sidechain-sidechain interaction is the **C13** hydrogen bond formed between the $-NH_2$ and C=O groups of the two adjacent Gln sidechains. Close inspection of Figure 7.4 reveals that this type of interaction is possible in nearly all families of backbone geometries (C10, C5, C5/C5, C7, and C5/C7). In this sense, it appears that these sidechain-sidechain interactions are typically hydrogen bonds of convenience rather than interactions that uniquely stabilize a particular backbone conformation.

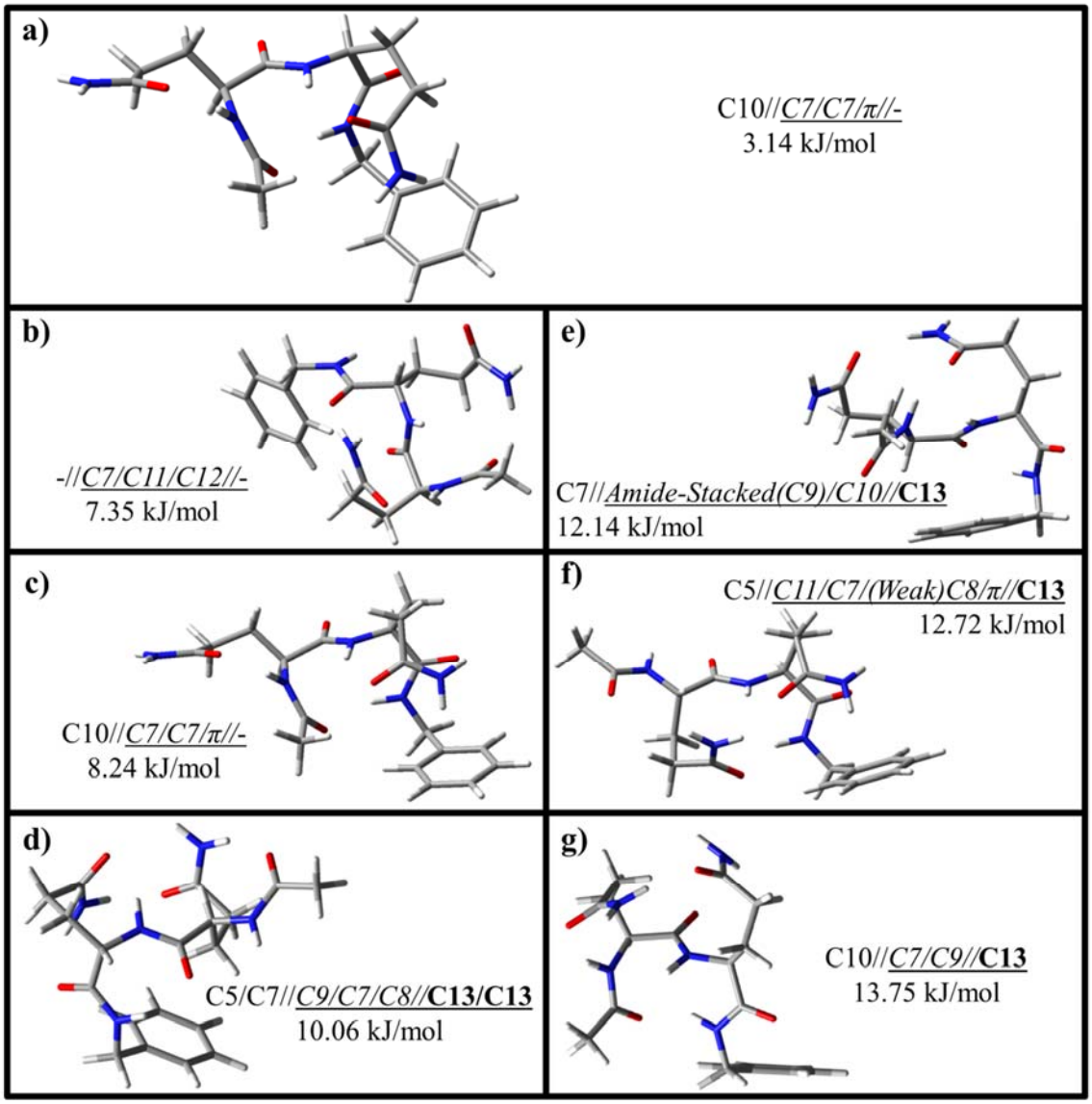


Figure 7.5. Calculated low-energy structures that are not assigned to the experimentally observed conformation. Structures are labeled with their hydrogen bonding motif and zero-point corrected energies relative to the global minimum. All structures were optimized using M05-2x/6-31+g(d).

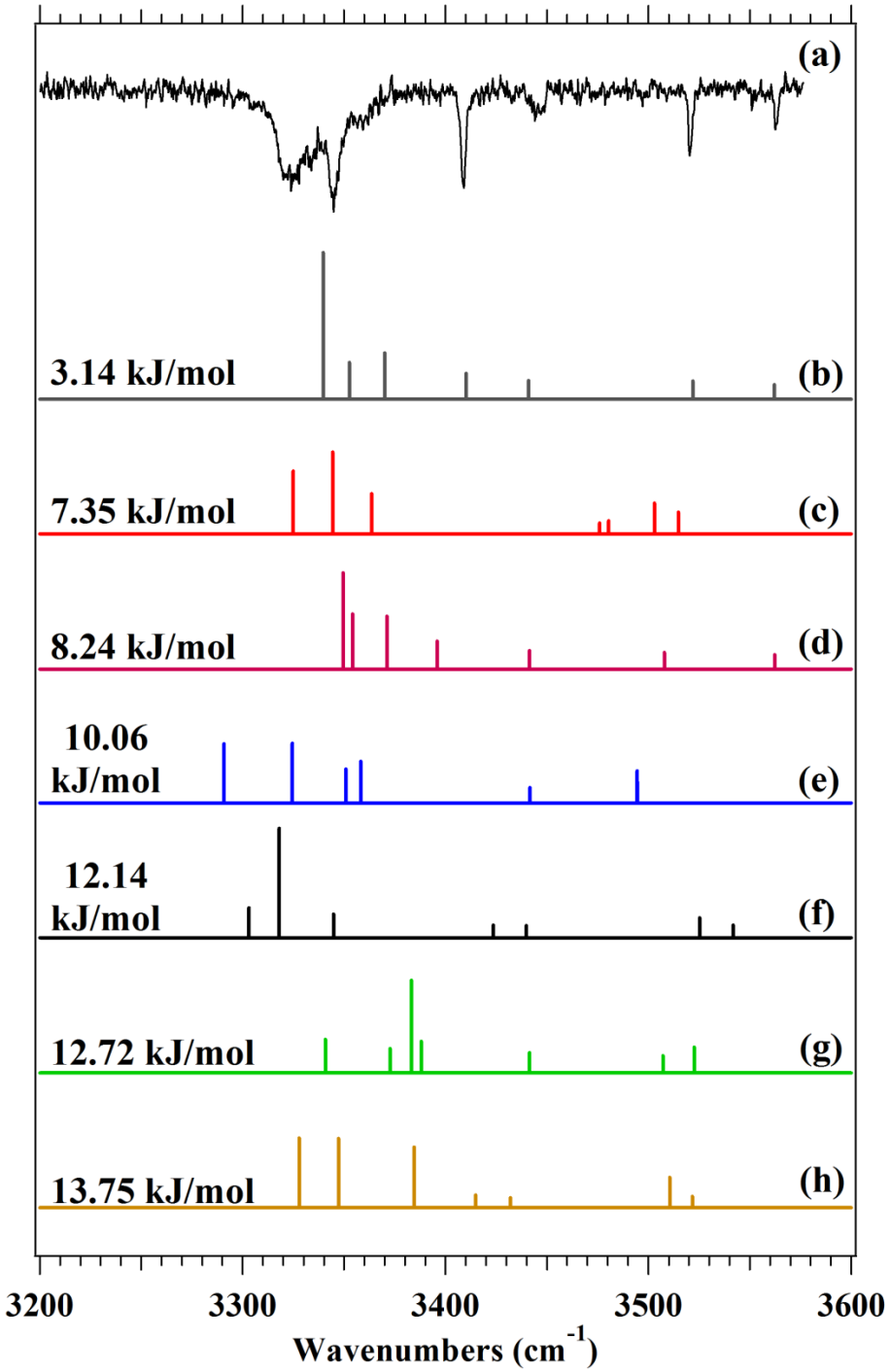


Figure 7.6. Experimental NH-stretch spectrum (a), along with the predicted spectra for several low-energy, non-assigned structures: a (b) C10//*C7/C7/π*//- β-turn structure, a (c) -//*C7/C11/12*//- structure, a (d) C10//*C7/C7/π*//- β-turn structure, a (e) C5/C7//*C9/C7/C8*//**C13/C13** structure, a (f) C7//*Amide-Stacked (C9)/ C10*//**C13** structure, a (g) C5//*C11/C7/(Weak)C8/π*//**C13** structure, and a (h) C10//*C7/C9*//**C13** β-turn structure. Relative energies of each are also provided in the figure.

Although sidechain-to-sidechain interactions are important features of the aggregated fibril structures, the gas phase results point towards a diminished role for such interactions in non-aggregated poly-Gln segments. The low-energy region (< ~10 kJ/mol) of the calculated potential energy diagram (Figure 7.4) shows that the four lowest-energy structures (<10 kJ/mol) are devoid of any sidechain-to-sidechain interactions and instead contain exclusively backbone-backbone and side-chain-backbone H-bonds. The first calculated structure that contains a sidechain-sidechain hydrogen bond is a C5/C7//*C9/C7/C8*//**C13/C13** (Figure 7.5) structure calculated to be ~10 kJ/mol above the global minimum.

It is notable that structural preferences manifested by isolated Ac-Gln-Gln-NHBn are evident to some degree also in the single-glutamine analogs Ac-Ala-Gln-NHBn and Ac-Gln-NHBn (Figure 7.1, Table 7.3). In all three molecules, low-energy structures involve the Gln sidechain C=O in a C7 H-bond with the backbone NH of that residue, an interaction that induces a turn in the backbone.[19]

7.4 Conculsion

he present results strongly suggest that a -Gln-Gln- segment in a nonpolar environment will readily and preferentially adopt a sidechain-stabilized β-turn. This folding behavior could be important for disease pathogenesis. Experimental efforts to characterize the forms of monomeric or oligomeric poly-Gln segments that are proposed to exert toxic effects are complicated by the transience of these hypothetical species. Moreover, such species are likely to occur at low abundance in the conformational equilibria available to poly-Gln segments in solution. For segments with a Gln repeat

length beyond the threshold for disease onset, the "misfolded" monomers aggregate quickly to insoluble states, so that little of the proposed toxic forms would be present.[7,11-15] Available data for poly-Gln segments in solution suggest a predominance of random coil-like structures.[15] Since recent studies have identified formation of β-turn/β-hairpin structures as the bottle neck along the pathogenic pathway[7-15], the present results suggest that transient entry of a short Gln-rich portion into a nonpolar environment, perhaps within a lipid bilayer, might lead to local β-turn formation and thereby promote intermolecular association of poly-Gln segments.

7.5 References

(1) Soto, C. *Nat. Rev. Neurosci.* **2003**, *4*, 49.

(2) Ross, C. A.; Poirier, M. A. *Nature Medicine* **2004**, *10*, S10-S17.

(3) Zoghbi, H. Y.; Orr, H. T. *Annual Review of Neuroscience* **2000**, *23*, 217.

(4) Ross, C. A. *Neuron* **2002**, *35*, 819-822.

(5) Mangiarini, L.; Sathasivam, K.; Seller, M.; Cozens, B.; Harper, A.; Hetherington, C.; Lawton, M.; Trottier, Y.; Lehrach, H.; Davies, S. W.; Bates, G. P. *Cell* **1996**, *87*, 493-506.

(6) Ross, C. A.; Poirier, M. A.; Wanker, E. E.; Amzel, M. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 1-3.

(7) Kar, K.; Hoop, C. L.; Drombosky, K. W.; Baker, M. A.; Kodali, R.; Arduini, I.; van der Wel, P. C. A.; Horne, W. S.; Wetzel, R. *Journal of Molecular Biology* **2013**, *425*, 1183-1197.

(8) Thakur, A. K.; Wetzel, R. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 17014-17019.

(9) Buchanan, L. E.; Carr, J. K.; Fluitt, A. M.; Hoganson, A. J.; Moran, S. D.; de Pablo, J. J.; Skinner, J. L.; Zanni, M. T. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 5796-5801.

(10) Hoop, C. L.; Lin, H.-K.; Kar, K.; Magyarfalvi, G.; Lamley, J. M.; Boatz, J. C.; Mandal, A.; Lewandowski, J. R.; Wetzel, R.; van der Wel, P. C. A. *Proc. Natl. Acad. Sci. USA* **2016**.

(11) Miller, J.; Arrasate, M.; Brooks, E.; Libeu, C. P.; Legleiter, J.; Hatters, D.; Curtis, J.; Cheung, K.; Krishnan, P.; Mitra, S.; Widjaja, K.; Shaby, B. A.; Lotz, G. P.; Newhouse, Y.; Mitchell, E. J.; Osmand, A.; Gray, M.; Thulasiramin, V.; Saudou, F.; Segal, M.; Yang, X. W.; Masliah, E.; Thompson, L. M.; Muchowski, P. J.; Weisgraber, K. H.; Finkbeiner, S. *Nat Chem Biol* **2011**, *7*, 925-934.

(12) Zhang, Q. C.; Yeh, T.-l.; Leyva, A.; Frank, L. G.; Miller, J.; Kim, Y. E.; Langen, R.; Finkbeiner, S.; Amzel, M. L.; Ross, C. A.; Poirier, M. A. *J. Biol. Chem.* **2011**, *286*, 8188-8196.

(13) Nucifora, L. G.; Burke, K. A.; Feng, X.; Arbez, N.; Zhu, S.; Miller, J.; Yang, G.; Ratovitski, T.; Delannoy, M.; Muchowski, P. J.; Finkbeiner, S.; Legleiter, J.; Ross, C. A.; Poirier, M. A. *J. Biol. Chem.* **2012**, *287*, 16017-16028.

(14) Peters-Libeu, C.; Miller, J.; Rutenber, E.; Newhouse, Y.; Krishnan, P.; Cheung, K.; Hatters, D.; Brooks, E.; Widjaja, K.; Tran, T.; Mitra, S.; Arrasate, M.; Mosquera, L. A.; Taylor, D.; Weisgraber, K. H.; Finkbeiner, S. *Journal of Molecular Biology* **2012**, *421*, 587-600.

(15) Kar, K.; Jayaraman, M.; Sahoo, B.; Kodali, R.; Wetzel, R. *Nature Structural & Molecular Biology* **2011**, *18*, 328-336.

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

(17) Lubman, D. M. *Mass Spectrom. Rev.* **1988**, *7*, 559-592.

(18) Page, R. H.; Shen, Y. R.; Lee, Y. T. *J. Chem. Phys.* **1988**, *88*, 4621-4636.

(19) Walsh, P. S.; Dean, J. C.; McBurney, C.; Kang, H.; Gellman, S. H.; Zwier, T. S. *Phys. Chem. Chem. Phys.* **2016**, *18*, 11306-11322.

(20) Zwier, T. S. *Annu. Rev. Phys. Chem.* **1996**, *47*, 205-241.

(21) Zwier, T. S. *J. Phys. Chem. A* **2001**, *105*, 8827-8839.

(22) Weiner, P. K.; Kollman, P. A. *J. Comput. Chem.* **1981**, *2*, 287-303.

(23) Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caufield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440-467.

(24) Zhao, Y.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, *3*, 289-300.

(25) Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V.N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J.V.; Cioslowski, J.; Fox, D. J.; Revision C.01 ed.; Gaussian, Inc.: Wallingford CT, 2009.

(26) McGowan, S.; Buckle, A. M.; Mitchell, M. S.; Hoopes, J. T.; Gallagher, D. T.; Heselpoth, R. D.; Shen, Y.; Reboul, C. F.; Law, R. H. P.; Fischetti, V. A.; Whisstock, J. C.; Nelson, D. C. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12752-12757.

(27) Krauspenhaar, R.; Rypniewski, W.; Kalkura, N.; Moore, K.; DeLucas, L.; Stoeva, S.; Mikhailov, A.; Voelter, W.; Betzel, C. *Acta Crystallographica Section D* **2002**, *58*, 1704-1707.

(28) Stammers, D. K.; Achari, A.; Somers, D. O. N.; Bryant, P. K.; Rosemond, J.; Scott, D. L.; Champness, J. N. *FEBS Letters* **1999**, *456*, 49-53.

(29) Shapiro, L.; Scherer, P. E. *Current Biology* **1998**, *8*, 335-340.

(30) Shaya, D.; Zhao, W.; Garron, M.-L.; Xiao, Z.; Cui, Q.; Zhang, Z.; Sulea, T.; Linhardt, R. J.; Cygler, M. *J. Biol. Chem.* **2010**, *285*, 20051-20061.

(31) Qing, J.; Du, X.; Chen, Y.; Chan, P.; Li, H.; Wu, P.; Marsters, S.; Stawicki, S.; Tien, J.; Totpal, K.; Ross, S.; Stinson, S.; Dornan, D.; French, D.; Wang, Q.-R.; Stephan, J.-P.; Wu, Y.; Wiesmann, C.; Ashkenazi, A. *The Journal of Clinical Investigation* **2009**, *119*, 1216-1229.

(32) Zhai, Y.; Sun, F.; Li, X.; Pang, H.; Xu, X.; Bartlam, M.; Rao, Z. *Nat Struct Mol Biol* **2005**, *12*, 980-986.

(33) Krieg, S.; Huché, F.; Diederichs, K.; Izadi-Pruneyre, N.; Lecroisey, A.; Wandersman, C.; Delepelaire, P.; Welte, W. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 1045-1050.

(34) European Bioinformatics Institute (http://www.ebi.ac.uk/pdbe-site/pdbemotif/).

(35) Schwing, K.; Fricke, H.; Bartl, K.; Polkowska, J.; Schrader, T.; Gerhards, M. *ChemPhysChem* **2012**, *13*, 1576-1582.

(36) Chin, W.; Piuzzi, F.; Dimicoli, I.; Mons, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1033-1048.

VITA

VITA

Patrick S. Walsh was born in Bristol, PA on May 23, 1989 to Beth G. and Mark T. Walsh. He has an older brother, Michael E., who resides in Holland, PA. Patrick grew up in Richboro, PA graduating from Council Rock High School South in 2007. After graduation he enrolled at the University of Pittsburgh in Pittsburgh, PA where he majored in Chemistry and minored in German Studies. Patrick graduated with a B.S., *cum laude* with Program Honors in Chemistry, in 2011. In June of 2011, Patrick began graduate school at Purdue University and joined Prof. Timothy S. Zwier's research group. During his time in the group Patrick used single-conformation spectroscopy to study hydrogen bonding networks in water clusters and peptides. As a graduate student, Patrick played way too much Corn-Hole at ISMS and probably had more fun than he should have. Patrick received his Ph.D. in Physical Chemistry in 2016, after which he joined Prof. Feng Gai's research group at the University of Pennsylvania as a postdoctoral researcher. Patrick enjoys cheesesteaks, Flyers hockey, and go-karting. Patrick hates the Pittsburgh Penguins and the Dallas Cowboys, they are the worst.

PUBLICATIONS

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 11306

Conformation-specific spectroscopy of capped glutamine-containing peptides: role of a single glutamine residue on peptide backbone preferences†

Patrick S. Walsh,[a] Jacob C. Dean,‡[a] Carl McBurney,§[b] Hyuk Kang,[c] Samuel H. Gellman[b] and Timothy S. Zwier*[a]

Received 16th February 2016,
Accepted 13th March 2016

DOI: 10.1039/c6cp01062h

www.rsc.org/pccp

The conformational preferences of a series of short, aromatic-capped, glutamine-containing peptides have been studied under jet-cooled conditions in the gas phase. This work seeks a bottom-up understanding of the role played by glutamine residues in directing peptide structures that lead to neurodegenerative diseases. Resonant ion-dip infrared (RIDIR) spectroscopy is used to record single-conformation infrared spectra in the NH stretch, amide I and amide II regions. Comparison of the experimental spectra with the predictions of calculations carried out at the DFT M05-2X/6-31+G(d) level of theory lead to firm assignments for the H-bonding architectures of a total of eight conformers of four molecules, including three in Z-Gln-OH, one in Z-Gln-NHMe, three in Ac-Gln-NHBn, and one in Ac-Ala-Gln-NHBn. The Gln side chain engages actively in forming H-bonds with nearest-neighbor amide groups, forming C8 H-bonds to the C-terminal side, C9 H-bonds to the N-terminal side, and an amide-stacked geometry, all with an extended (C5) peptide backbone about the Gln residue. The Gln side chain also stabilizes an inverse γ-turn in the peptide backbone by forming a pair of H-bonds that bridge the γ-turn and stabilize it. Finally, the entire conformer population of Ac-Ala-Gln-NHBn is funneled into a single structure that incorporates the peptide backbone in a type I β-turn, stabilized by the Gln side chain forming a C7 H-bond to the central amide group in the β-turn not otherwise involved in a hydrogen bond. This β-turn backbone structure is nearly identical to that observed in a series of X-(AQ)-Y β-turns in the protein data bank, demonstrating that the gas-phase structure is robust to perturbations imposed by the crystalline protein environment.

1. Introduction

Neurodegenerative diseases, such as Alzheimer's disease (AD), Parkinson's disease (PD), and Huntington's disease (HD), are illnesses that ravage both mental acuity and motor function of affected individuals.[1–6] Although these illnesses involve different precursor proteins [β-amyloid (Aβ) and tau in AD,[3,4] α-synuclein in PD,[4,6] and huntingtin (htt) in HD],[5,7] resulting in different symptoms and pathogenesis, the structures of the toxic species appear to be related to one another across all of these diseases.[3,4,6]

This common taxonomy involves β-sheet oligomers that ultimately form insoluble amyloid fibrils.[8] These large β-sheet assemblies are held together by intermolecular H-bonds and by intra- and intermolecular sidechain-to-sidechain interactions. Seminal work by Eisenberg and co-workers,[9,10] and Bowers and co-workers[11] have provided detailed insight into both the structure and aggregation propensity of several short segments derived from amyloid-forming peptides (*e.g.*, the hexamer Asn-Asn-Gln-Gln-Asn-Tyr from the yeast prion protein Sup35). Recently, 2D-IR experiments have been used to elucidate the structure of human islet amyloid

[a] *Department of Chemistry, Purdue University, West Lafayette, Indiana 47901, USA. E-mail: zwier@purdue.edu*
[b] *Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706, USA*
[c] *Department of Chemistry, Ajou University, Republic of Korea*

† Electronic supplementary information (ESI) available: Detailed information regarding the synthetic procedures for Ac-Gln-NHBn and Ac-Ala-Gln-NHBn, potential energy scan of the carboxylic acid OH rotation, additional calculated non-assigned structures of Z-Gln-OH, Z-Gln-NHMe, Ac-Gln-NHBn, and Ac-Ala-Gln-NHBn, comparison of NHBn *versus* NHMe structure of double bridge structure calculated for Ala-Gln, a comparison between Ac-Gln-NHBn conformer B and Ac-Ala-Gln-NHBn conformer A, a table summarizing the experimental and calculated vibrational frequencies, and a table summarizing the energies of Ac-Gln-X and Ac-Ala-Gln-X molecules (X = –NHBn and –NHMe). See DOI: 10.1039/c6cp01062h

‡ Present address: Department of Chemistry, Princeton University, Princeton, NJ 08540, USA.
§ Present address: Pharmaceutical Product Development, LLC (PPD), Middleton, WI 53562, USA.

polypeptide (hIAPP),[12–14] and γD-crystallin,[15] which both aggregate into fibril structures.

Glutamine is commonly found in many regions of proteins that form fibrillar structures, including ones responsible for neurodegenerative diseases. In Huntington's disease (HD), glutamine takes center stage. Occurrence of HD is associated with the presence of excess Gln in the protein huntingtin. This protein normally contains a Gln-repeat segment of 5 to 36 residues; increasing length of the poly-Gln segment is linked to disease pathogenesis.[4,5,7,16–18]

The flexible amide-containing side chain of glutamine (Gln, Q) is believed to be vitally important in forming the polar-zipper motif commonly associated with fibril structures.[18] Several computational and experimental studies have sought to understand the structural preferences of glutamine containing peptides and their aggregates.[18–25] Studies on larger glutamine repeat sequences point to formation of a β-hairpin structure comprised of β-sheet and β-turn regions.[26–28] In these situations the glutamine sidechain plays an integral role in aggregation by forming sidechain-to-backbone hydrogen bonds between successive β-hairpin structures, locking in the characteristic polar-zipper structure.[8–10,29]

In a fully-developed glutamine-rich peptide sequence, the presence of successive glutamine residues, each with a flexible amide-containing sidechain, creates many possibilities for sidechain-to-sidechain and sidechain-to-backbone hydrogen bonds. The presence of water further expands the types of H-bonded networks that can be formed. In order to achieve a molecular-scale understanding of the aggregation behavior of glutamine-rich peptides, one must develop quantitatively accurate molecular mechanics force fields. Such force fields would properly describe the delicate balance of attractive forces and steric constraints that control the conformations, solubility and aggregation behavior of glutamine-rich peptides.

The development of accurate force fields benefits from experimental measurements that provide incisive tests of computational accuracy. Over the past several years, a number of groups world-wide have developed and utilized a powerful set of laser spectroscopy tools for studying the conformational preferences of short peptide sequences in the gas phase.[30–35] By studying the peptides in the absence of solvent, the inherent conformational preferences of the peptide are directly exposed, making a direct link with *ab initio* and force field calculations carried out on the isolated molecules. In the present paper, we focus on an important question that serves as a foundation for such theoretical developments; namely, 'What is the effect of a single glutamine residue on the inherent conformational preferences of a peptide backbone?'

We begin to address this question by recording the conformation-specific IR and UV spectra of several capped glutamine-containing peptides in the cold, collision-free environment of a molecular beam. Fig. 1 shows the molecules which were studied, including three capped amino acids Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn) and one capped dipeptide (Ac-Ala-Gln-NHBn). These molecules are capped either at the N-terminus (with a benzyloxycarbonyl cap, labeled as 'Z'), or at the C-terminus

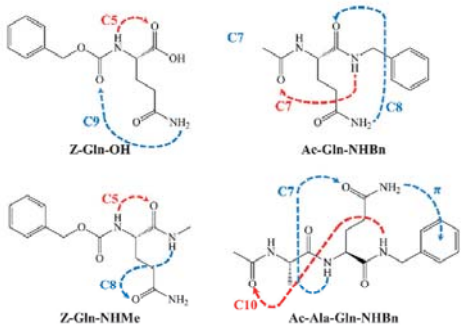


Fig. 1 Chemical structures of the molecules studied in this work.

(with an NH-benzyl cap, labeled as 'NHBn'), providing the aromatic chromophore needed for IR-UV double resonance techniques that enable the acquisition of IR and UV spectra of single conformations free from interference from other conformations present in the expansion. Conformation-specific spectroscopy has proven to be a powerful tool for revealing the inherent conformational preferences of both natural amino acids and synthetic foldamers. In studies of this type to date, the spectroscopic signatures of incipient forms of several prototypical α-amino acid secondary structures have been obtained, including γ-turns, β-turns, β-hairpins, extended conformations, and helices.[36–41]

Analogous studies on synthetic foldamers (e.g., β-peptides, γ-peptides) have revealed the effects of increased backbone flexibility on conformational preferences, including the formation of unique secondary structures facilitated by the synthetic building-blocks of which they are composed.[42–46] The studies carried out by our group on constrained and unconstrained γ-peptides are particularly relevant to Gln-containing peptides, because the spacing between backbone amides across a single γ-residue is identical to the spacing between the Gln sidechain amide group and the backbone amide involving the Gln α nitrogen.[47–51] In this sense, we can look to the results of previous studies on γ-peptides as a guide to some of the sidechain ↔ backbone interactions we might anticipate within a single Gln residue. In particular, it is possible that the Gln sidechain may engage in amide stacking with the peptide backbone, given that a stacking interaction involving comparably spaced amide groups have been observed in small γ-peptides.[47,48]

In addition to amide stacking, there is a range of different sidechain-to-backbone and backbone-to-backbone hydrogen bonds that can be formed in short peptides that contain at least one Gln residue. These hydrogen bonds may compete with one another, as is shown in Fig. 1. Our over-arching goal of understanding the effect of a single glutamine residue on a peptide's conformational preferences can thus be served by identifying the preferred H-bonding and/or stacking architectures adopted by short peptides that contain a Gln residue. As we shall see, the set of four molecules shown in Fig. 1 exhibit conformational preferences that result in a rich variety of structures. While some of

these observed structures stabilize the extended peptide backbone conformation that is characteristic of β-sheets, the Gln residue is also capable of stabilizing both γ- and β-turns, sometimes *via* surprising mechanisms.

2. Experimental

2.1. Experimental methods

2.1.1. Sample handling. Benzyloxycarbonyl-Gln-OH (Z-Gln-OH) was purchased from Bachem and used without purification. Z-Gln-OH was converted to Z-Gln-NHMe using a published method;[48] this material was used for spectroscopic studies without purification. Ac-Gln-NHBn and Ac-Ala-Gln-NHBn were synthesized as described in the ESI.† To avoid sample degradation caused by thermal heating, we used a laser desorption source to bring the molecule of interest into the gas phase. The laser desorption source and experimental apparatus have been described previously.[34,35,37,52–54]

2.1.2. Spectroscopy techniques. Ultraviolet (UV) light was generated using the doubled output of a Nd:YAG (355 nm, Continuum SureLite II) pumped dye laser (Radiant Dyes Narrow-Scan) operating at 20 Hz. Typical powers are kept below ~0.5 mJ per pulse to avoid saturation. Infrared (IR) light was generated using an optical parametric converter (Laservision) pumped by a Q-switched Nd:YAG laser (Continuum SureLite III/EX) operating at 10 Hz. The pump laser could be operated in both seeded and unseeded modes, providing better spectral resolution when required (specified in the figures). IR light could be generated in both the hydride stretch (~3100–3750 cm^{-1}, 10–15 mJ per pulse) and the amide I and II (~1300–1900 cm^{-1}, ~500–750 μJ per pulse) regions to provide sufficient spectral content to make firm assignments.

UV spectra were recorded for each sample using resonant two-photon ionization (R2PI).[53] R2PI spectra were recorded by scanning the UV laser (~265–270 nm) through the S_0–S_1 origin region of the molecules of interest and plotting the mass-selective ion signal *versus* the excitation wavelength.

Conformation-specific IR spectra were recorded using resonant ion-dip infrared (RIDIR) spectroscopy.[34,35,54] In this technique a 20 Hz UV and 10 Hz IR laser are spatially overlapped with the IR temporally preceding the UV by ~200 ns. The UV laser has its wavelength fixed on an origin transition of a single conformation while the IR laser is scanned through the region of interest. When the IR laser is resonant with a transition which shares the same ground state as the UV transition, a depletion of ion signal is observed. The repetition mismatch between the UV and IR lasers allows for this depletion to be monitored in real time in an "on-off" fashion using active baseline subtraction in a gated integrator (Stanford Research). An infrared spectrum is recorded by plotting this depletion signal as a function of IR laser wavelength.[34,35,51,54]

Conformation-specific UV spectra were recorded using infrared-ultraviolet holeburning (IR-UV HB) spectroscopy.[35] The experimental setup of the IR-UV HB is the same as in RIDIR spectroscopy with one key difference; in IR-UV HB the IR laser is fixed to a unique IR transition of a particular conformation, while the UV laser is scanned through the electronic origin region. Once again, the depleted ion signal is measured using active baseline subtraction and plotted as a function of the UV excitation wavelength.

2.2. Computational methods

A set of starting structures was generated using the conformational search function in the MACROMODEL suite of programs.[55] This conformational search used the Amber* force field[56] to find all unique structures within a 50 kJ mol^{-1} energy window of the global minimum structure. For Z-Gln-OH separate conformational searches were performed on the *cis* and *trans* conformations of the –OH moiety in the carboxylic acid cap. From these structures the lowest ~30–100 conformations for each molecule were further optimized using density functional theory (DFT) calculations that employed the hybrid density functional M05-2X[57] with the 6-31+g(d) basis set. DFT calculations were performed using Gaussian09.[58] Harmonic vibrational frequencies were calculated for the optimized geometries. These frequencies were scaled by 0.9398 or 0.942 (Z-Gln-NHMe and all others, respectively), 0.951, and 0.960 for the NH-stretch, OH-stretch, and amide I/II regions respectively to account for anharmonicity in the harmonic frequencies.[33]

2.3. Structural information and nomenclature

The hydrogen bonds present in a given optimized structure are described using the common C_n notation, where n denotes the number of atoms which compose the hydrogen bonded ring. In this work, we introduce a naming scheme in which backbone-to-backbone hydrogen bonds are listed first, separated by a single slash, with backbone H-bonds listed after a double slash in italics. For example, "C10//*C7/π*" denotes a structure in which a ten-membered hydrogen bonded ring is formed between two backbone amide groups, with the glutamine side chain involved in a double-bridge structure to the peptide backbone comprised of C7 and NH···π H-bonds. Crystal and NMR structures were taken from the RCSB Protein Data Bank[59] and the PDBeMotif[60] search tool from the European Bioinformatics institute.

3. Results and analysis

3.1. Z-glutamine-X (X = –OH or –NHMe)

3.1.1. UV spectroscopy. R2PI and IR-UV holeburning spectra were recorded for both Z-Gln-NHMe and Z-Gln-OH in the parent mass channels (m/z = 293 and 280, respectively) over the S_0–S_1 origin region of the Z-cap (~37 480–37 665 cm^{-1}). These spectra are shown in Fig. 2a and b, respectively. The R2PI and IR-UV holeburning spectra of both molecules show a series of sharp UV transitions that ride atop a broad background signal likely arising from incomplete cooling. In Z-Gln-NHMe, the strong transition at 37 602 cm^{-1} (marked with an asterisk) was used to record the RIDIR spectra which will be discussed in a later section (Section 3.1.3). The IR-UV holeburning spectrum recorded with IR holeburn laser

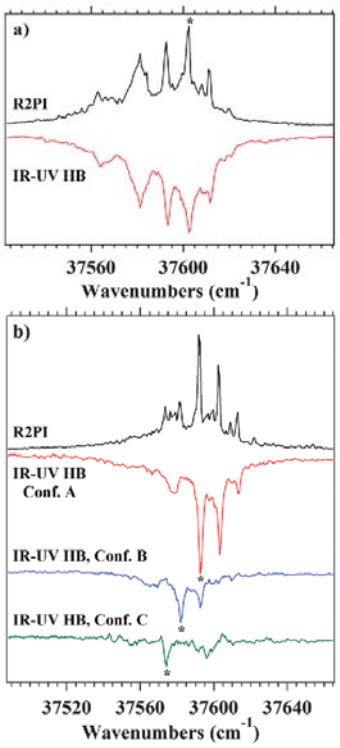


Fig. 2 R2PI and IR-UV holeburning spectra for (a) Z-Gln-NHMe, and (b) Z-Gln-OH. Transitions marked with asterisks were used to record RIDIR spectra.

fixed at 3560 cm $^{-1}$ shows that all transitions and the broad backbround burn out together. Holeburning scans taken with hole-burn laser fixed on other IR transitions, produced the same spectrum, indicating that a single conformation of Z-Gln-NHMe is present in the expansion. As a final check, RIDIR spectra were recorded on several different UV transitions, all of which showed the same infrared spectrum.

The R2PI (Fig. 2b) spectrum of Z-Gln-OH shows a smaller background and narrower UV transitions indicative of better cooling of the laser-desorbed molecules in this case. IR-UV holeburning spectroscopy reveals the presence of three unique conformations (Fig. 2b). The three strongest transitions at 37 593, 37 582 and 37 574 cm $^{-1}$ are assigned as the S_0–S_1 origin transitions for conformers A, B, and C, respectively, and were used to collect RIDIR spectra (Section 3.1.2). The broad transition at ~37 578 cm $^{-1}$ (Fig. 2b) is believed to be a hot band of conformer A, indicating the cooling is still not entirely complete. The proximity of the hot band due to A to the origin transitions of conformers B and C will contribute to interferences in the RIDIR spectra recorded for these conformers. Minimal Franck–Condon activity for the three conformers indicates a minimal geometry change upon electronic excitation for all three conformers.

3.1.2. Z-glutamine-OH: RIDIR spectroscopy and conformational assignments. Fig. 3 shows the RIDIR spectra recorded for the three assigned conformers of Z-Gln-OH in the NH/OH-stretch (Fig. 3a, c and e) and amide I/II regions (Fig. 3b, d and f). The red stick diagrams below the spectra are the predictions of calculations for the best-fit structures. A full comparison between observed and calculated IR transition frequencies for each of the conformers is included in ESI,† (Table S1). These stick diagrams not only help confirm the assignments but also provide insight to the groups responsible for the observed transitions, with labels given next to the calculated transitions describing the major carrier of a given IR transition. The hydride-stretch region for conformer A shows four distinct hydride-stretch fundamentals at 3562, 3443, 3430, and 3142 cm $^{-1}$. The bands at 3562 and 3443 cm $^{-1}$ are readily identified as the free antisymmetric and symmetric stretch fundamentals of the sidechain NH_2 moiety, respectively. The absence of the free carboxylic acid OH stretch (~3575 cm $^{-1}$) reveals that this group is involved in a hydrogen bond, most likely the band appearing at 3142 cm $^{-1}$. The final transition at 3430 cm $^{-1}$ is shifted slightly lower in frequency than a free amide NH (~3450 cm $^{-1}$) indicative of a C5 hydrogen bond in an extended backbone geometry.

Based on comparison with calculation, we assign conformer A to a C5//*C8* (*trans*-OH) structure shown in Fig. 4a, calculated to be 1.85 kJ mol $^{-1}$ above the global minimum. In this structure the carboxylic acid OH has undergone an internal rotation into a *trans* configuration that facilitates formation of a very strong *C8* hydrogen bond of the C-terminal OH group with the side-chain carbonyl. The strength of this hydrogen is evident from the magnitude of shift in the OH stretch (~430 cm $^{-1}$) from the nominally free position, reflecting its short H-bond length of 1.73 Å. In the absence of this stabilizing OH···O=C H-bond, the *trans*-OH minimum is calculated to be ~24 kJ mol $^{-1}$ higher than *cis*-OH. Furthermore, there is a large barrier to internal rotation of this OH group (~55 kJ mol $^{-1}$ in a simple glycine, ESI,† Fig. S1). Observation of such *trans* carboxylic acids have been seen in other uncapped gas-phase peptides when that isomerization can facilitate strong H-bond formation.[32]

The corresponding comparison between experiment and calculation in the amide I/II region is shown in Fig. 3b. Excellent agreement in both absolute frequency and intensity pattern confirm and strengthen the conformational assignment made from the hydride stretch region. This region, however, is less intuitively analyzed compared to the hydride stretching region. The amide I region shows three, clearly resolved transitions at 1801, 1739, and 1705 cm $^{-1}$. The highest frequency C=O stretch is assigned to the C5 hydrogen bonded carboxylic acid carbonyl stretch. The lowest frequency transition (1705 cm $^{-1}$) is assigned to the sidechain carbonyl involved in a *C8* hydrogen bond, which accounts for its shift to lower frequency. The free Z-cap carbonyl appears at 1739 cm $^{-1}$.

Calculation also aids in the assignment of the amide II region including the NH_2 bend, which appears at a much higher frequency

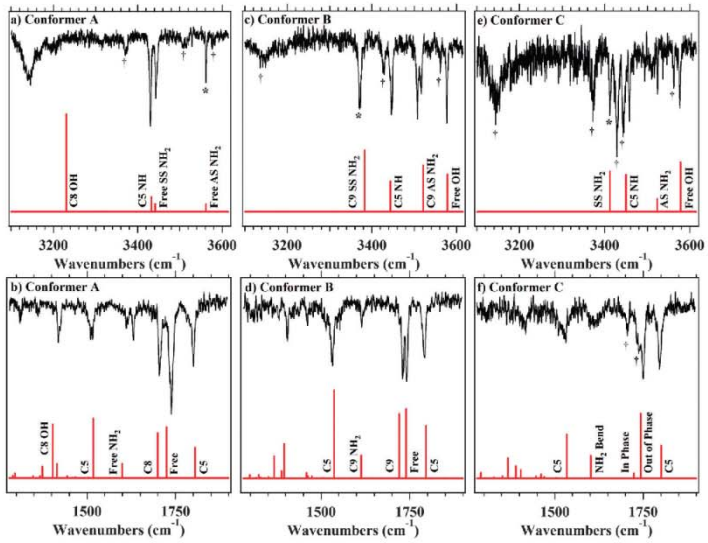


Fig. 3 RIDIR spectra recorded for the three conformers of Z-Q-OH in the hydride-stretch region (top row, seeded) and amide I/II region (bottom row, unseeded): conformers A (a and b), B (c and d), and C (e and f). Stick-spectra below each experimental trace are scaled harmonic vibrational frequency calculations for the assigned conformations. Starred (*) transitions were used to record the IR-UV holeburning spectra shown in Fig. 2. Transitions marked with a dagger (†) in the spectra of conformers B and C are spectroscopic interferences arising because the UV wavelength used for recording the RIDIR scans have some contribution from conformer A.

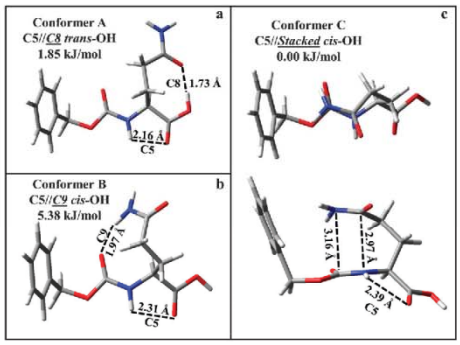


Fig. 4 Assigned structures, relative energies, and hydrogen bond distances for conformers A–C (a–c, respectively) of Z-Gln-OH.

(1612 cm $^{-1}$) than typical amide NH bending modes (~1600 cm $^{-1}$). The remaining strong transitions at 1513 and 1419 cm $^{-1}$ are assigned to the C5 amide NH bend and the in-plane C8 OH bend, respectively.

Analysis of conformers B and C is complicated by the presence of interfering transitions in the hydride stretching region, likely caused by overlapping transitions in the UV with not only one another, but also with the hot band of conformer A. Transitions that are caused by this spectral interference are labeled with daggers (†) in the spectra reported in Fig. 3. Despite this complicating factor, it is still possible to make firm conformational assignments for these conformers.

Conformer B shows four new transitions in the hydride stretching region at 3577, 3507/3517, 3446, and 3371 cm $^{-1}$ (Table S1, ESI,† Fig. 3c). The transition at 3577 cm $^{-1}$ is ascribable to the free carboxylic acid OH stretch vibration,[37] which thereby retains its low energy *cis* rotamer configuration. The missing free anti-symmetric NH_2 stretch at ~3561 cm $^{-1}$ indicates that the sidechain NH_2 is involved in a hydrogen bond. The transition at 3446 cm $^{-1}$ falls into a range of frequencies ascribable either to free symmetric NH_2 stretches, or weakly bound amide NH groups involved in a C5 hydrogen bond. Since the NH_2 group is involved in a hydrogen bond, this transition is assigned to a C5 NH group. The remaining two transitions are assigned to the hydrogen bonded antisymmetric (3507/3517 cm $^{-1}$) and symmetric (3371 cm $^{-1}$) stretches of the NH_2 group.

Comparison with theory leads to the assignment of conformer B to an extended backbone C5//C9 (*cis*-OH) structure that is calculated to be 5.38 kJ mol $^{-1}$ above the global minimum (Fig. 4b). This comparison yields excellent agreement between experiment and theory in both regions recorded experimentally. The sidechain NH_2 group is involved in a sidechain-to-backbone hydrogen bond (C9) towards the N-terminus. Based on the NH stretch frequency shifts of the NH_2 group (50–70 cm $^{-1}$), the C9

H-bond is significantly weaker than its *C8* counterpart in conformer A involving the OH group (430 cm $^{-1}$). This weaker H-bond is reflected in the longer hydrogen bond length of the *C9* hydrogen bond (1.97 *vs.* 1.73 Å, Table 1).

The amide I/II region provides additional evidence for the assignment (Fig. 3c and d). The amide I region has three transitions centered at 1792, 1741, and 1731 cm $^{-1}$. The highest frequency transition is assigned to the backbone C5 hydrogen bonded carbonyl, consistent with other experiments on similar systems.[33] The transition at 1741 cm $^{-1}$ is the free sidechain carbonyl stretch, and the transition at 1731 cm $^{-1}$ results from the *C9* hydrogen bonded Z-cap carbonyl. Similarly good agreement is found in the amide II region where transitions at 1614 and 1533 cm $^{-1}$ are assigned to the NH_2 and backbone C5 amide NH bends, respectively.

Table 1 Hydrogen bond distances for hydrogen bonds found in the glutamines and similar hydrogen bonds found in γ-peptide foldamers

	Hydrogen bond	
Molecule	Type	Distance (Å)
Z-Gln-NHMe		
Conf. A	C5	2.12
(C5//*C8*)	*C8*	1.98
Z-Gln-OH		
Conf. A	C5	2.16
(C5//*C8*)	*C8*	1.73
Conf. B	C5	2.31
(C5//*C9*)	*C9*	1.97
Conf. C	C5	2.39
(C5//*stacked*)	*Int. N → C*	2.97
	Ext. N → C	3.16
Ac-Gln-NHBn		
Conf. A	C5	2.11
(C5//*C8*)	*C8*	1.96
Conf. B	*C7*	1.98
(*C7*/*π*)	*π*	2.49
Conf. C	C7	1.92
(C7//*C7*/*C8*/*π*)	*C7*	1.99
	C8	2.24
	π	3.12, 2.98
Ac-Ala-Gln-NHBn		
Conf. A	C10	2.24
(C10//*C7*/*π*)	*C7*	1.97
	π	2.40
γ_{ACHC} [a]	C9	1.93
Ac-γ^2-hPhe-NHMe[b]		
C9	C9	2.00
Amide-stacked	*Int. N → C*	2.90
	Ext. N → C	3.05
Gabapentin derivative 1[b]		
C7	C7	2.04
Ac-Phe-NHMe[c]		
C7(a)	C7	2.17
C7(g+)	C7	2.01
Ac-Phe-Ala-NHMe[c]		
C5/C7	C7	2.07
Ac-Ala-Phe-NHMe[c]		
C7/C7(g−)	C7	2.05
	C7	2.04

[a] From ref. 51. [b] From ref. 48. [c] From ref. 33.

Conformer C is the most difficult to analyze because it is the most minor conformer, and as such, suffers from the largest interference caused by overlap in the UV transitions. However, since the transitions belonging to conformers A and B are known, identifying conformer C transitions is straight forward. In the hydride stretch region four transitions are identified at 3577, 3524, 3458, and 3412 cm $^{-1}$ (Table S1, ESI,† Fig. 3c). The highest frequency transition is assigned to the free carboxylic acid OH stretch (3577 cm $^{-1}$).[37] Surprisingly, the other three transitions are all above 3400 cm $^{-1}$, indicating that there are no strong hydrogen bonds in the molecule. The lack of a free anti-symmetric stretch at ~3561 cm $^{-1}$ also reveals that the NH_2 group is involved in at least a weak hydrogen bond.

Comparing these unique spectral attributes with the predictions of calculations, conformer C can be assigned to the extended backbone C5//*amide-stacked* (*cis*-OH) geometry, calculated to be the global minimum structure (Fig. 4c). On this basis, the transitions at 3524 and 3412 cm $^{-1}$ are assigned as the antisymmetric and symmetric stretches of the NH_2 group, which is weakly hydrogen bonded to the π-cloud of the Z-cap. The relatively small frequency shifts of these transitions from the nominally free positions (by −30 and −35 cm $^{-1}$, respectively) is consistent with the weak π H-bond involved. The remaining transition (3458 cm $^{-1}$) is assigned to a C5 NH group, occurring near the anticipated position based on the C5 rings in A and B.

Once again, spectra in the amide I/II regions confirm the structural assignment (Fig. 3f). The transition at 1796 cm $^{-1}$ is due to the C5 C=O stretch of the COOH group. The transitions at 1737 and 1705 cm $^{-1}$ are interference from the two major transitions of conformer A, leaving the pair of transitions at 1750 and 1733 cm $^{-1}$ yet to assign. The unique geometry of the *amide-stacked* planes produces an interesting intensity pattern in the other two carbonyl stretches, both of which are involved in the stack. These two C=O groups are nearly anti-parallel to one another, and are in similar environments, so that coupling between them produces two delocalized C=O stretch modes that are best described as in- and out-of-phase linear combinations of the two C=O stretch motions. The dipole derivatives nearly cancel in the mode in which the two C=O groups stretch in-phase, but constructively interfere in the out-of-phase motion, leading to a large asymmetry in calculated intensity. On this basis, we assign the strong and weak transitions at 1750 and 1733 cm $^{-1}$ to the out-of-phase and in-phase doublet of amide stacked C=O stretch modes. The amide II region has two transitions at 1609 and 1532 cm $^{-1}$, which are assigned to the NH_2 and C5 amide NH bending fundamentals, respectively.

3.1.3 Z-glutamine-NHMe: IR spectroscopy and conformational assignment. Fig. 5a and b show the RIDIR spectra recorded for the sole conformer found experimentally for Z-Gln-NHMe (conformer A). The NH stretch spectrum has four clear transitions at 3560, 3442, 3402, and 3351 cm $^{-1}$ (Table S1, ESI,† Fig. 5a). As was the case in the Z-Gln-OH conformers, the

transitions at 3560 and 3442 cm $^{-1}$ are readily assigned to the free antisymmetric and symmetric stretches of the NH_2 group, pointing towards a structure with a free NH_2 group. The remaining transitions arise from the backbone amide NH's which are involved in one weak and one strong hydrogen bond, strongly reminiscent of the C5//*C8* motif of Z-Gln-OH conformer A.

Comparing experiment with the predictions of calculations (Fig. 5) enables the assignment of conformer A of Z-Gln-NHMe to an extended backbone C5//*C8* structure, where the sidechain forms a hydrogen bond with the backbone towards the C-terminus. The assigned structure, shown in Fig. 5c, is 0.1 kJ mol $^{-1}$ above the calculated global minimum. The final two transitions can now be assigned to the backbone C5 (3402 cm $^{-1}$) and the *C8* (3351 cm $^{-1}$) backbone-to-sidechain amide NH stretch fundamentals.

The amide I region comprises three well-resolved transitions at 1740, 1721, and 1709 cm $^{-1}$, assignable based on the calculations to the (nominally) free, C5, and *C8* C=O stretch fundamentals, respectively (Fig. 5b). The amide II region is similarly made up of three transitions at 1605, 1584, and 1510 cm $^{-1}$ due to the free NH_2, *C8* amide NH, and C5 amide NH bend fundamentals, respectively. The large intensity of the C5 NH bend and the intensity asymmetry with its *C8* counterpart arise from two effects. First, the C5 NH bend oscillates toward and away from the C=O group involved in the C5 H-bond. This interaction enhances the dipole derivative, increasing the intensity of the C5 NH bend.[61] Second, the anti-parallel arrangement of successive amide groups linked by a C5 H-bond enhances coupling between the two NH groups (despite their different H-bonding environments), producing partially delocalized NH bend modes, with the in-phase oscillation gaining in intensity, much as it does in β-strand and β-sheet structures.[33,61]

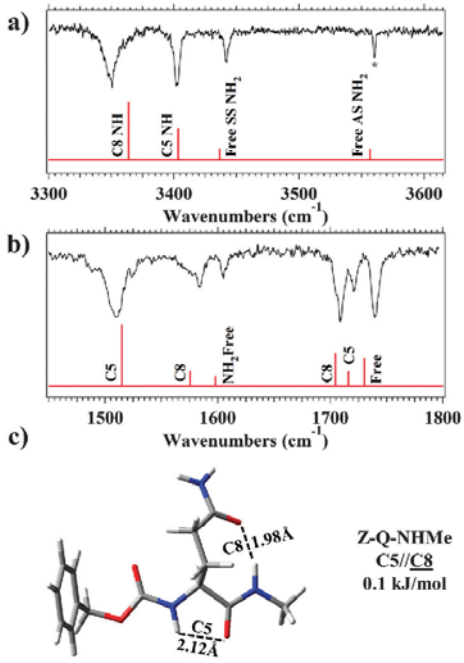


Fig. 5 RIDIR spectra of the single observed conformation of Z-Gln-NHMe in the NH-stretch region (a, seeded) and amide I/II region (b, seeded). Stick spectra below each experimental trace are scaled harmonic vibrational frequency calculations for the assigned C5//*C8* (c) conformation. Starred (*) transition was used to record the IR-UV holeburning spectrum shown in Fig. 2.

It is worth noting that, although the structural families of conformer A of Z-Gln-OH and Z-Gln-NHMe are both C5//*C8*, the former molecule forms a *C8* using its C-terminal OH rather than an amide NH. The OH···O=C *C8* is much shorter than NH···O=C (1.72 Å *vs.* 1.98 Å), producing a much larger shift to lower frequency in Z-Gln-OH (−430 cm $^{-1}$) than in Z-Gln-NHMe (−120 cm $^{-1}$, Table S1, ESI† and Table 1).

While only a single conformer was characterized experimentally, several other low energy structures were calculated for Z-Gln-NHMe including *amide-stacked* (0.00 kJ mol $^{-1}$), C5//*C9/C8* with an extended backbone (1.17 kJ mol $^{-1}$), C7//*C7/C8* with a γ-turn backbone (8.38 kJ mol $^{-1}$), and *C7* with a turned backbone (8.19 kJ mol $^{-1}$). Structures for these unobserved structures are included in ESI,† (Fig. S3). While it is possible that minor conformers could be present and were obscured by overlap with bands from conformer A (Fig. 2a), the single dominant conformer is the C5//*C8* conformer just described.

3.2. Ac-glutamine-NHBn

3.2.1. UV spectroscopy. Fig. 6a shows the R2PI spectrum recorded for the Ac-Gln-NHBn molecule. The spectrum was recorded by monitoring ion yield in the parent mass channel (m/z = 272) as the laser was tuned through the -NHBn S_0-S_1 origin region (~37 300–37 740 cm $^{-1}$). The resulting spectrum shows a very dense region with sharp UV transitions riding on a broad background likely caused by incomplete cooling of a portion of the desorbed sample. UV transitions due to three unique conformations were observed in the molecular beam as identified by RIDIR spectroscopy and IR-UV holeburning, with their origin transitions at 37 548 (A), 37 619 (B), and 37 554 (C) cm $^{-1}$, labeled in Fig. 6a. IR-UV holeburning (Fig. 6b) was performed on conformer A using a unique IR transition at 3402 cm $^{-1}$. Much of the dense vibronic activity in the R2PI spectrum is associated with conformer A, as is the broad background. The IR-UV holeburning scan is missing the transitions ascribed to conformers B and C, identifying them for RIDIR scans which proved that they were due to different conformational isomers. Holeburning scans were not performed on these conformers due to their minor population and our limited sample size. The transition at 37 425 cm $^{-1}$ was investigated minimally, but its carrier could not be identified. It is likely a water cluster built off of conformer A based on cursory investigation in the IR.

3.2.2. IR spectroscopy and conformational assignments. Fig. 7 presents the NH stretch (upper) and amide I/II (lower)

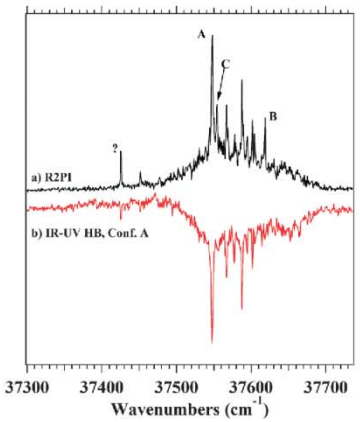


Fig. 6 R2PI (a) of Ac-Gln-NHBn. IR-UV holeburning spectrum (b) for conformer A. Transitions used to record RIDIR spectra for conformers A, B, and C are labeled with A, B, and C, respectively.

RIDIR spectra recorded for the three conformers of Ac-Gln-NHBn. Fig. 7a shows the experimental spectrum recorded in the NH stretch region, with four transitions at 3562, 3443, 3402, and 3310 cm $^{-1}$ (Table S1, ESI†). This spectrum is nearly identical to the spectrum recorded for conformer A of Z-Gln-NHMe {C5//*C8*}, with the three higher-frequency IR transitions appearing within 1 cm $^{-1}$ of each other. Calculations confirm that conformer A of Ac-Gln-NHBn is also a C5//*C8* structure. This structure, shown in Fig. 8a, is calculated to be just 1.54 kJ mol $^{-1}$ above the global minimum. The primary difference between the C5//*C8* conformers of Z-Gln-NHMe and Ac-Gln-NHBn is the length of the *C8* hydrogen bond formed between the sidechain and the backbone, which is 1.98 Å in Z-Gln-NHMe, but shrinks to 1.96 Å in Ac-Gln-NHBn (Table 1). The shorter H-bond in Ac-Gln-NHBn leads to a larger shift in frequency of the *C8* NH stretch from its free position down to 3310 cm $^{-1}$, compared to 3364 cm $^{-1}$ in Z-Gln-NHMe.

Interestingly, while the C5//*C8* conformers of Ac-Gln-NHBn (Fig. 7b) and Z-Gln-NHMe (Fig. 5b) have amide I spectra that are very similar in appearance, the difference in strength of the *C8* H-bond and the differing positions of the aromatic caps shifts the absolute frequencies, re-shuffles the order, and changes the coupling between the three C=O stretch transitions. The stronger *C8* H-bond involving the side-chain C=O group shifts the *C8* C=O stretch up in frequency from 1709 to 1720 cm $^{-1}$, and localizes its motion on the *C8* C=O oscillator. The nominally "free" and C5 amide C=O groups in the extended backbone are nearly degenerate with, and anti-parallel to, one another. Coupling between them produces in-phase and out-of-phase modes delocalized over the pair of C=O groups, with frequencies of 1695 and 1703 cm $^{-1}$ for the out-of-phase (oop) and in-phase (ip) combinations, respectively, as labeled in the figure. This coupling is reflected in the asymmetric intensity

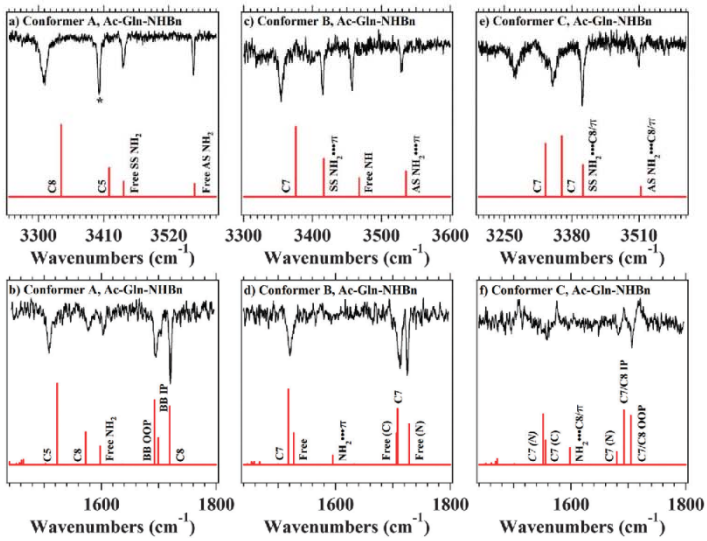


Fig. 7 RIDIR spectra recorded for the three conformers of Ac-Gln-NHBn in the NH-stretch region (top row, unseeded) and amide I/II region (bottom row, unseeded): conformers A (a and b), B (c and d), and C (e and f). Stick-spectra below each experimental trace are scaled harmonic vibrational frequency calculations for the assigned conformations. The starred (*) transition was used to record the IR-UV holeburning spectrum shown in Fig. 6.

distribution between the pair, with the out-of-phase combination, with its constructive interference, gaining intensity at the expense of the in-phase combination.

The amide II region of conformer A is also very similar in appearance to its C5//*C8* counterpart in Z-Q-NHMe, with the C5 NH bend at low frequency (1509 cm $^{-1}$) and possessing its characteristic large intensity.

Conformer B RIDIR spectra are shown in Fig. 7c and d. The NH stretch spectrum bears little resemblance to the spectra recorded for the other molecules, indicating that it arises from a new structural family. The NH stretch fundamentals appear at 3529, 3457, 3415, and 3354 cm $^{-1}$ (Table S1, ESI†). The NH_2 symmetric and antisymmetric stretch transitions both appear shifted down in frequency slightly from their nominally free value, consistent with the presence of a weak H-bond involving the aromatic π cloud. The remaining transitions at 3354 and 3457 cm $^{-1}$ are representative of a strongly hydrogen bonded backbone NH and a nearly free amide NH, respectively.

Comparison of the NH stretch spectra with theory produces an assignment of a *C7/π* structure in which the only H-bonds are those between the Gln side chain and the peptide backbone (*C7*, N-term NH to Gln C=O) or aromatic cap (Gln NH_2 to Bn π), with no backbone-to-backbone hydrogen bonds. The assigned structure is the calculated global minimum (Fig. 8b), with calculated IR transitions that provide an excellent match in both the NH stretch and amide I/II regions. As anticipated the single strong H-bond is the *C7* H-bond between the N-terminal NH and the Gln C=O group.

The amide I and II regions provide corroborating support for the assignment, but due to the small signal size and lower laser power several of the lower intensity transitions are not resolved experimentally. The amide I region shows two clearly resolved transitions at 1725 and 1713 cm $^{-1}$, while the amide II region displays only a single broad transition at 1521 cm $^{-1}$ (Table S1, ESI,† and Fig. 7d). Comparison with calculations makes it possible to assign these transitions and partially explain the reason for the missing transitions. In the amide I region, the transition at 1725 cm $^{-1}$ is the free C=O stretch of the N-terminal carbonyl. The broader peak at 1713 cm $^{-1}$ likely contains two nearly degenerate C=O stretch transitions of the sidechain *C7* and C-terminal free carbonyls. The predicted splitting of 2 cm $^{-1}$ is likely unresolvable in the experimental spectrum. The amide II region shows a similar near degeneracy between the free amide NH and *C7* amide NH, which are predicted to be split by 9 cm $^{-1}$, and are experimentally unresolved, producing a single broadened transition at 1521 cm $^{-1}$.

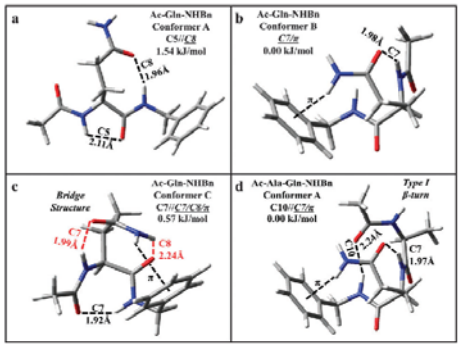


Fig. 8 Three assigned structures for the three conformers of Ac-Gln-NHBn (a–c) and the single conformer of Ac-Ala-Gln-NHBn (d). Also shown for each of the assigned conformers are their relative energies, hydrogen bonding pattern and bond lengths. The hydrogen bonds labeled in red indicate a hydrogen bond formed when glutamine forms a hydrogen bonded "bridge" motif.

Conformer C is the final minor conformer of Ac-Gln-NHBn, with RIDIR spectra shown in Fig. 7e and f. The NH stretch spectrum of this conformer is also unique compared to those so far observed. All four NH stretch transitions (3509, 3400, 3343, and 3270 cm $^{-1}$ (Table S1, ESI†), are shifted down in frequency from their nominally free values (3560/3445 for NH_2 and 3450–3480 for amide NH), suggesting that the NH_2 and two amide NH groups are all involved in H-bonds.

Comparison with theory leads to an assignment for conformer C to the C7//*C7/C8/π* structure shown pictorially in Fig. 8c. This structure is only 0.57 kJ mol $^{-1}$ above the global minimum, and is a highly compact structure in which every group is involved in a hydrogen bond. The backbone C7 ring is locked in place by a set of two sidechain-to-backbone H-bonds, with the side-chain C=O forming a *C7* ring with the N-terminal NH, and the NH_2 group engaged in a *C8* H-bond with the C-terminal C=O. This double-bridge capping arrangement of the Gln side-chain stabilizes the C7 peptide backbone, which forms what is formally an inverse γ-turn, with Ramachandran angles of (−82.4, 56.3)°. Such structures have been previously identified in the literature by several groups.[33,39,62] The two hydrogen bonded transitions at 3343 and 3270 cm $^{-1}$ are now identified as arising from the *C7* sidechain-to-backbone hydrogen bond and the C7 hydrogen bond formed within the backbone, respectively. The predicted NH stretch spectrum shown in Fig. 7e shows an excellent match to the experimental spectrum.

The recorded amide I/II spectrum (Fig. 7f) is complicated by the small signal levels, and the proximity of the S_0–S_1 origin transition (37 554 cm $^{-1}$) to the origin transition of conformer A (37 548 cm $^{-1}$). The depletion signals for conformer C are observed amidst gains in ion intensity caused by the neighboring major conformer. This phenomenon has been observed previously and complicates the experimental spectrum.[63] Still, two identifiable transitions are present in the amide I region at 1706 and 1682 cm $^{-1}$. The lower frequency transition is assigned to the C=O stretch involved in the backbone C7. The amide II transitions are difficult to observe in the midst of the gains from conformer A. The only identifiable transition is a weak, broad absorption at 1558 cm $^{-1}$. This transition likely contains both the amide NH (C-terminal) C7 and the *C7* amide NH (N-terminal) which are predicted to be nearly degenerate at 1556 and 1552 cm $^{-1}$, respectively.

The backbone C7 in conformer C produces the lowest frequency NH stretch observed so far (3270 cm $^{-1}$), providing

evidence that this H-bond is particularly strong. This backbone C7 is also substantially stronger than the C7 hydrogen bonds formed between the sidechain and the backbone, which show shifts of only ~120 cm $^{-1}$ from their nominally free position. This strength is borne out by the calculations, which predict a C7 H-bond length of 1.92 Å, compared to 1.98/1.99 Å for the sidechain-to-backbone C7's in conformers B and C respectively (Table 1).

3.3. Ac-alanine-glutamine-NHBn

3.3.1. UV spectroscopy. Fig. 9a shows the R2PI and IR-UV holeburning spectra for Ac-Ala-Gln-NHBn. The R2PI spectrum (black) shows very little background and is relatively uncongested. IR-UV holeburning (red) proves that all transitions belong to a single conformation with an S_0–S_1 origin transition at 37 621 cm $^{-1}$. The spectrum shows a short Franck–Condon progression involving a 19 cm $^{-1}$ vibration, presumably arising from a low frequency mode involving the backbone and sidechain groups vibrating against the π-cloud of the UV chromophore.

3.3.2. IR spectroscopy and conformational assignment. Fig. 9b and c present the RIDIR spectra recorded for the single observed conformer of Ac-Ala-Gln-NHBn in the NH stretch and amide I/II regions, respectively. The NH stretch spectrum shows three sharp transitions at 3520, 3470, and 3408 cm $^{-1}$ and a broad band with a barely-resolved splitting at 3361/3357 cm $^{-1}$ (Table S1, ESI†). The overall spectrum is similar in appearance to that of conformer B of Ac-Gln-NHBn. Following that lead, the transitions at 3520 and 3408 cm $^{-1}$ can be tentatively assigned as symmetric and antisymmetric stretch fundamentals of the NH_2 group, slightly shifted from their free positions by virtue of a π H-bond. The structure also possesses a completely free amide NH at 3470 cm $^{-1}$. The broadened transition at 3361/3357 cm $^{-1}$ shows a slight splitting which indicates that two nearly degenerate transitions are present at a frequency typical of both C7 and C10 hydrogen bonds.[33,38,39,62,64]

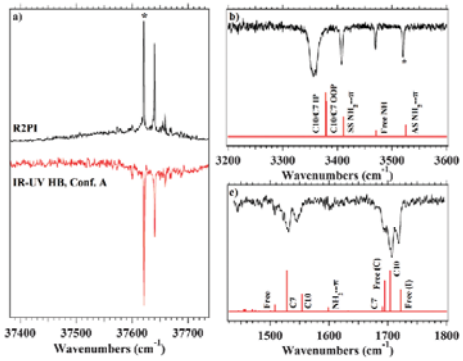


Fig. 9 Experimental spectra recorded for Ac-Ala-Gln-NHBn. R2PI (a) and IR-UV holeburning (a) for the sole observed conformation. Experimental RIDIR spectra recorded in the NH-stretch (b, unseeded) and amide I/II (c, unseeded) regions shown above the labeled red stick spectra from the assigned structure's harmonic frequency calculation. Starred transitions (*) indicated transitions used to record RIDIR and IR-UV holeburning spectra.

The stick spectra shown below the experimental spectra in Fig. 9b present an excellent fit, and confirm the tentative assignments. We assign this conformer to the C10//C7/π structure shown in Fig. 8d, which is calculated to be the global minimum. This structure has a backbone C10 hydrogen bond that links the N-terminal C=O and the C-terminal amide NH, forming a type I β-turn, the most common β-turn in biological settings.[65] The backbone dihedrals for this conformer are {−70, −13.3}° and {−73, −11}° for $i+1$ and $i+2$ Ramachandran angles, respectively (Table 2), close to those for an idealized type I β-turn [{−60, −30}° and {−90, 0}°].[65] The glutamine residue here behaves in a similar fashion to conformer B of Ac-Gln-NHBn, with the sidechain forming a C7 hydrogen bond towards the N-terminus nearest-neighbor, and the NH_2 forming a hydrogen bond with the π-cloud of the NHBn cap. The C10 and C7 NH stretch transitions are nearly degenerate, and highly coupled, leading to the broad, slightly split pair of transitions at 3361/3357 cm $^{-1}$. A detailed discussion of the similarities between the gas-phase and crystal/NMR β-turn structures will be taken up in Section 4.1.3.

The amide I/II region (Fig. 9c) also shows an excellent match to theory in both absolute frequency and intensity pattern. The amide I region has transitions at 1718, 1706 and 1694 cm $^{-1}$ with a weak shoulder at ~1680 cm $^{-1}$. The highest frequency transition at 1718 cm $^{-1}$ is assigned to the free C=O stretch of the interior amide group, while the transition at 1694 cm $^{-1}$ is from the free C-terminal C=O stretch. The remaining two C=O stretch modes at 1706 (high intensity) and 1681 cm $^{-1}$ (very weak) show a large intensity asymmetry resulting from the nearly anti-parallel orientation of the Ac and Gln C=O groups in this structure, which are engaged in C10 and C7 H-bonds, respectively. The normal modes are nearly equal mixes of the local mode oscillations, with the in-phase combination at 1681 cm $^{-1}$ near zero intensity, and the out-of-phase counterpart at 1706 cm $^{-1}$ with unusual intensity. The amide II region has four well resolved transitions at 1601, 1546, 1531, and 1508 cm $^{-1}$ which are assigned to the NH_2 bend, the C10, the C7, and free amide NH bending vibrations, respectively.

4. Discussion

Using single-conformation spectroscopy, we have recorded infrared spectra of a total of eight conformations of four short Gln-containing peptides of different lengths and capping arrangements: three in Z-Gln-OH, one in Z-Gln-NHMe, three in Ac-Q-NHBn, and a single conformer in Ac-AQ-NHBn. The jet-cooled gas-phase data includes representative examples of five unique H-bonding architectures: three C5//C8 conformers, one C5//C9, one amide-stacked structure, one C7//C7/C8/π double-bridge cap, and a C10//C7/π type I β-turn. Representative examples of each of these structural types are shown in Fig. 10. A sixth type, the C7/π conformer of Ac-Gln-NHBn, shares the same stabilizing C7 Gln H-bond and backbone bend

Table 2 Ramachandran angle comparison of type-I β-turns with sequence X(AQ)X

Structure	Turn sequence/residue #	Turn residues $(i+1)$ $(\phi,\psi)^\circ$	$(i+2)$ $(\phi,\psi)^\circ$
Ac-Q-NHBn, conf. B[a]	X(XQ)X		(−87.2, −9.1)
Ac-AQ-NHBn[a]	X(AQ)X	(−70, −13.3)	(−73, −11)
Complex, hydrolase product[b]	F(AQ)L 9–12	(−57.0, −21.7)	(−98.4, 1.67)
Nuclear protein[c]	K(AQ)F 92–95	(−57.1, −34.1)	(−64.0, −23.7)
Signaling protein, cell cycle[d]	E(AQ)G 23–26	(−66.6, −19.1)	(−91.8, −2.8)
Viral protein, anti-viral protein[e]	E(AQ)Q 560–563	(−65.5, −43.1)	(−64.4, −31.3)
Xylanase[f]	A(AQ)N 69–72	(−56.8, −33.3)	(−84.5, −14.1)
Electron transport (heme protein)[g]	C(AQ)C 14–17	(−63.3, −15.0)	(−76.6, −8.8)
Toxoflavin binding protein[h]	G(AQ)H 26–29	(−52.9, −27.8)	(−68.8, −17.8)
Histocompatibility anitgen i-ak[i]	F(AQ)L 48–51	(−65.2, −6.1)	(−93.9, −21.3)
Iron–sulfur-binding protein[j]	I(AQ)G 69–72	(−64.3, −27.5)	(−90.2, 2.0)
Hydrolase, ligase/inhibitor[k]	V(AQ)E 146–149	(−54.5, −39.4)	(−78.4, −7.7)
Anti-oncogene protein[l]	D(AQ)A 352–355	(−67.6, −23.0)	(−83.5, −10.6)
Viral protein, metal binding protein[m]	F(AQ)C 34–37	(−58.9, −30.9)	(−78.8, −40.3)
Structural genomics, signaling protein[n]	L(AQ)A 37–40	(−48.7, −38.4)	(−73.1, −35.8)
Average of crystal structures	X(AQ)X	(−59.9, −27.7)	(−80.5, −16.2)

[a] This work. [b] (PDB ID: 1af2).[66] [c] (PDB ID: 3g36).[67] [d] (PDB ID: 1mb3).[68] [e] (PDB ID: 4i2L).[69] [f] (PDB ID: 1w3h).[70] [g] (PDB ID: 5cyt).[71] [h] (PDB ID: 3ouL).[72] [i] (PDB ID: 1iak).[73] [j] (PDB ID: 3zyy).[74] [k] (PDB ID: 4dm9).[75] [l] (PDB ID: 1oLh).[76] [m] (PDB ID: 2a51).[77] [n] (PDB ID: 2cs4).[78]

as in the second half of the type I β-turn, and therefore is not shown separately in the figure (Fig. S6, ESI†).

The active role played by the Gln side chain in dictating structure, with its amide group capable of forming H-bonds to the peptide backbone amides, is clear from the figure. In each structural motif, the Gln sidechain is involved in hydrogen bond(s) that are either the primary H-bonding interaction stabilizing the structure or that enhance the stability of an otherwise common peptide backbone structure.

Fig. 11 presents a series of energy level diagrams that compare the calculated relative energies of the conformational minima within 15 kJ mol ^{1} of the global minimum for each of the capped amino acids (Z-Gln-OH, Z-Gln-NHMe, Ac-Gln-NHBn) and the sole capped dipeptide studied here (Ac-Ala-Gln-NHBn).

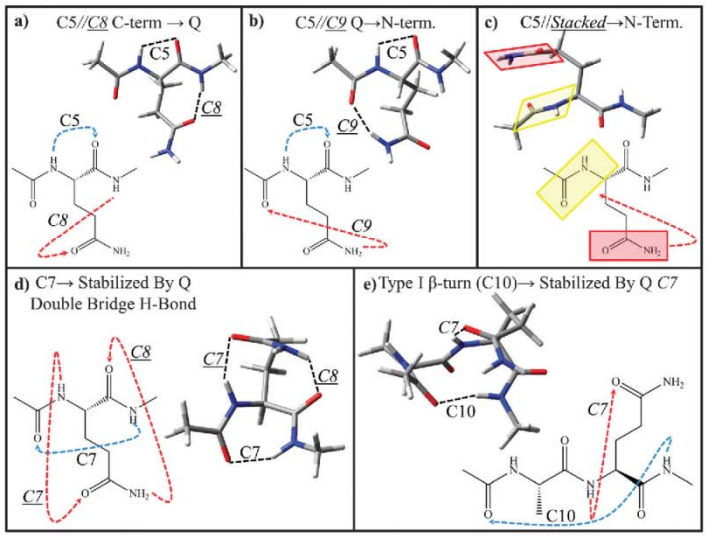


Fig. 10 Hydrogen bonding motifs of the glutamine sidechain: (a) C8 sidechain-to-backbone hydrogen bond towards the N-terminus, (b) C9 sidechain-to-backbone hydrogen bond towards the C-terminus, (c) sidechain amide stacking towards the N-terminus, (d) backbone C7 hydrogen bond stabilized by a pair of Gln side-chain hydrogen bonds (C7/C8) that bridge adjacent amide groups in the peptide backbone, and (e) a type-I β-turn (C10) stabilized through a C7 sidechain-to-backbone hydrogen bond towards the N-terminus.

The figure also includes calculations in which the cap containing the aromatic ring is removed, in order to assess the effects of the aromatic ring on the stabilities of the structures. In Section 4.1 we use Fig. 10 and 11 to assess the ways in which a single glutamine side chain interacts with, and stabilizes various local peptide backbone architectures, including the β-turn structure found in longer XAQY sequences in proteins. In Section 4.2, we will take a wider view of the potential energy surfaces, comparing the observed structures with other competing possibilities, and assessing the role played by the Z- and –NHBn caps.

4.1. Sidechain-to-backbone hydrogen bonds and local peptide backbone structure

4.1.1. Extended backbone structures. Five of seven assigned conformations of the capped Gln-containing amino acids (Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn) contain an extended peptide backbone, with its characteristic nearest-neighbor C5 hydrogen bond. When the nearest-neighbor amide NH on the C-terminal side of the Gln residue forms an H-bond with the Gln C=O group, a *C8* H-bond is formed, as shown in Fig. 10a. The flexibility and length of the Gln side chain is such that this *C8* H-bond is formed with good strength while maintaining an extended peptide backbone. C5//*C8* structures are formed in Z-Gln-NHMe and Ac-Gln-NHBn, while its OH···O=C analog is formed in Z-Gln-OH. In each case, these structures are among the calculated lowest-energy structures (Fig. 11a).

Alternatively, when the Gln NH_2 group acts as H-bond donor to the amide C=O group on the N-terminal side, a *C9* H-bond is formed (Fig. 10b). This side-chain to backbone H-bond is also best formed with the peptide backbone in an extended C5 conformation. Interestingly, the *C9* H-bonded ring is similar to that formed in γ-peptides studied previously by our group (Ac-γ^2-hPhe-NHMe and its cyclically-constrained counterpart),[47–51] which share the same three-carbon alkyl group spacing between amide groups with the same relative directions of the two amide groups necessary for C9 H-bond formation (–CONH–C–C–C–$CONH_2$). Indeed, the calculated H-bond lengths (Table 1) and dihedral angles are very similar to those in the γ-peptides, as is the frequency of the C9 NH stretch fundamental.

The C5//*stacked* structure of Z-Gln-OH conformer C is the first amide-stacked structure experimentally observed in a natural amino acid. This structure is shown in generic form in Fig. 10c, where the two amide stacking planes are identified. Like the *C9* H-bond just considered, amide stacking was first discovered and characterized in a synthetic γ-peptide foldamer.[47–50] It is an intriguing structure because it contains no amide-amide

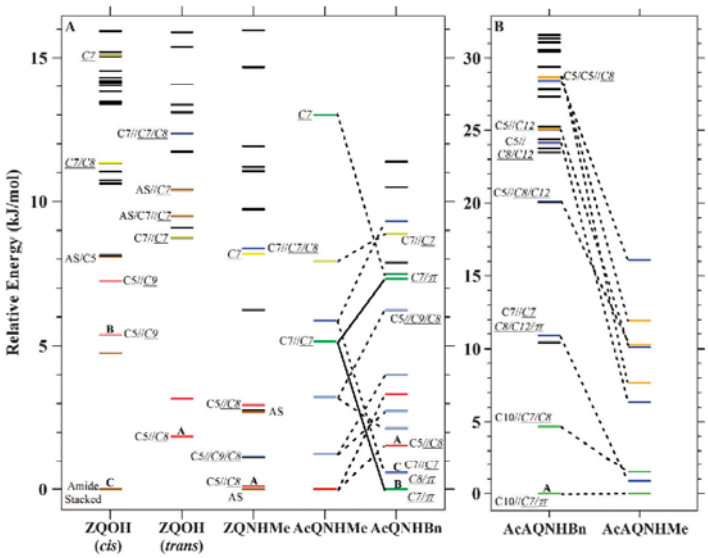


Fig. 11 Energy level diagram for (A) the capped amino acids Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn and (B) capped dipeptide Ac-Ala-Gln-NHBn. The relative zero-point corrected energies of conformational minima are calculated at the DFT M05-2X/6-31+G(d) level of theory. The molecules observed experimentally are compared with Ac-Gln-NHMe and Ac-Ala-Gln-NHMe in the figure in order to provide a point of comparison of the effects of the aromatic cap at the N- and C-terminus. Levels connected with a dotted line retain the same geometry in the –NHMe capped structure as the –NHBn structures, solid lines indicate a change in geometry. Selected levels are labeled with their hydrogen bonding motif. The energy levels are color-coded by structural family: nominally amide-stacked (brown), C5/*C9* (pink), C5/*C8* (red), C5/C5 containing and C5/*C12* (orange), general glutamine bridging geometries (navy blue), C7/*C7/C8* containing (blue), C5/*C9/C8* (light blue), general turned structures (yellow), *C7/π* (green), C7/*C7* (gold), β-turns (light green), general extended structures (maroon), and unclassified (black). Energy levels of assigned conformers are indicated by their letter labels.

H-bonds, but is held together by electrostatic and dispersive interactions between the anti-parallel planes of the two amide groups. These amide planes are 0.07–0.10 Å further separated than those in the γ-peptide structure of Ac-γ²-hPhe-NHMe (Table 1), with dihedral angles along the 3-carbon spacer that differ by no more than a few degrees.[47–50] Although this is the first identification of an amide-stacked geometry in a natural amino-acid, the structure receives additional stabilization beyond the interactions present in the amide stack from an $NH_2\cdots\pi$ interaction with the Z-cap aromatic ring, which would not be present when this cap is removed. As we shall see, this energetic stabilization from the cap makes it less likely that single Gln residues would take up this configuration in natural environments without additional stabilizing interactions (Fig. S4, ESI†). As Fig. 11a shows, an amide-stacked structure is also calculated to be among the lowest energy structures of Z-Gln-NHMe, but is not observed experimentally. Instead, a single C5//*C8* structure is found, suggesting that the C5//*C8* structure is the global minimum, and that interconversion of the amide stacked conformer to the C5//*C8* minimum may occur during the cooling process.

The C5//*C8*, C5//*C9* and C5//*stacked* structures all utilize sidechain to backbone interactions to stabilize the extended (C5) peptide backbone geometry characteristic of a β-sheet structure. It is possible that these local interactions of the Gln side chain with the peptide backbone could play a role in enabling β-sheet formation by stabilizing the extended backbone structure locally so that H-bonding sites along the extended backbone on the opposite side from the Gln side chain(s) are free to engage in H-bonds that lead to β-sheet formation. If this stabilization of extended structures is so prevalent in Gln residues, it may be one contributing factor to glutamine's prevalence in fibril forming regions of peptides.

4.1.2. Turned backbone structures. While the majority of observed structures in our X-Gln-Y series (X = Z, Ac; Y = OH, NHMe, NHBn) incorporate the extended C5 peptide backbone, the single Gln side chain can also interact with the peptide amide groups to support other local peptide backbone architectures. Indeed, two of the three structures identified in Ac-Gln-NHBn involve a turned peptide backbone that is stabilized by strong sidechain-to-backbone hydrogen bonds: the inverse γ-turn C7//*C7/C8/π* geometry taken up by conformer C (Fig. 8c) and the *C7/π* geometry of conformer B, (Fig. 8b), both of which are calculated to be slightly lower in energy than the C5//*C8* conformer A (Fig. 8a and 11a).

The C7//*C7/C8* structure, shown in schematic form in Fig. 10d, involves the peptide backbone in an inverse γ-turn with its characteristic C7 H-bond. This C7 hydrogen bond is stabilized and strengthened by the sidechain forming a doubly hydrogen bonded bridge in which the Gln side-chain C=O and NH groups form *C7* and *C8* H-bonds, respectively, to the NH and C=O groups on adjacent amide groups in the peptide backbone. In order to accommodate this pair of additional H-bonds, the peptide backbone must tip these groups out-of-plane towards the bridging Gln side chain, thereby enforcing formation of a peptide C7 H-bond. In so doing, the Gln side chain completes a hydrogen-bonded cycle involving three H-bonds that have the potential for cooperative strengthening, as each amide group is involved in H-bonds both as donor and acceptor.

In Ac-Gln-NHBn, the structure gains additional stabilization from the NHBn cap, which forms a π H-bond with the remaining NH group of the Gln NH_2 (not shown in Fig. 10d). When the –NHBn group is removed to form Ac-Gln-NHMe (Fig. 11a), the corresponding structure is about 5 kJ mol^{-1} higher in energy than the C5//*C8* structure (Table S2, ESI†), which becomes the global minimum.

With or without the NHBn cap, the inverse γ-turn found here has an unusually short H-bond, with an NH···O=C H-bond length (1.92 Å) about 0.1 Å smaller than C7 H-bonds in other natural peptides or synthetic foldamers studied to date, as summarized in Table 1. This is mirrored in its large frequency shift of the NH stretch fundamental, down to 3270 cm^{-1}, almost 200 cm^{-1} below its free value. The glutamine sidechain of this inverse γ-turn structure also forms a *C7* H-bond which is quite strong, with an NH···O=C H-bond length of 1.99 Å, while the *C8* H-bond is quite weak (2.24 Å), in part because the π H-bond is also present to the same NH_2 group.

The *C7/π* conformation has a backbone bend which contains no standard hydrogen bonds and Ramachandran angles of (−87.2, −9.1)° (Table 2). This unusual backbone geometry is stabilized by a sidechain-to-backbone *C7* with a hydrogen bond length of 1.98 Å. This hydrogen bond is very similar to the one formed in the inverse γ-turn structure, with an NH stretch frequency (3354 cm^{-1} *vs.* 3343 cm^{-1}) that reflects that similarity. Once again, the π H-bond between the other NH bond in the NH_2 group and the phenyl ring of the NHBn cap provides additional stabilization. When the phenyl ring is replaced by a methyl group to form Ac-Gln-NHMe, calculations predict a destabilization of about 5 kJ mol^{-1}, as shown in the energy level diagrams of Fig. 11a (Table S2, ESI†). Interestingly, the unusual backbone turn reconfigures to form a C7 H-bond in the process. As we shall see in the next section, the peptide backbone makes use of an unusual turn in forming the *C7/π* structure that reappears as an element of the energetically preferred global minimum in Ac-Ala-Gln-NHBn and its simpler parent Ac-Ala-Gln-NHMe, marking it as an important structural element in Gln's interaction with longer peptide sequences. To this we now turn.

4.1.3. Type I β-turn. The results on Ac-Ala-Gln-NHBn are especially important, because they offer a first glimpse at the conformational preferences imposed by a single Gln residue in a longer peptide sequence, here extended by the inclusion of an Ala residue on the N-terminal side. As we have seen, the entire population of Ac-Ala-Gln-NHBn is funneled into a single structure, the C10//*C7/π* conformer shown in schematic form (minus the Bn cap) in Fig. 10e.

As we have already noted, the C10 H-bond constitutes a type I β-turn, an important secondary structural element that is made all the more relevant by its role in linking successive anti-parallel β-sheet segments in the fibrils formed in neurodegenerative diseases related to poly-glutamine repeat disorders including Huntington's disease.[26,27,79] In many peptide

sequences, formation of a β-turn with its C10 H-bond leaves the central amide group of the turn without an amide-amide H-bond. However, the C=O group of the Gln side chain forms a strong *C7* H-bond to this central amide NH, the same *C7* H-bond formed in the double-bridge C7//*C7*/*C8*/π and *C7*/π structures of Ac-Gln-NHBn. This additional H-bond stabilizes the C10//*C7*/π β-turn structure, making it the global minimum in Ac-Ala-Gln-NHBn with only one other structure, also a β-turn, within 11 kJ mol^{-1}, as shown in Fig. 11b. The observed structure remains the global minimum when the phenyl cap is replaced by methyl (Fig. 11b), confirming that the Gln side chain adds to the stability of the type I β-turn principally through the Gln side chain H-bond to the interior amide NH rather than the π-hydrogen bond (Table S3, ESI†).

The assigned C10//*C7*/π type-I β-turn structure of Ac-Ala-Gln-NHBn can be compared against similar β-turns found in protein crystal and solution NMR structures with Ala and Gln as the $i + 1$ and $i + 2$ residues, respectively. One such comparison of this type is shown in Fig. 12a and b. In these overlays the backbone geometry of Ac-Ala-Gln-NHBn is compared with the backbone of a type-I β-turn from the myoglobin and cytochrome *C* crystal structure (PDB ID: 5cyt, C(AQ)C sequence).[71] The type-I β-turn from the crystal structure also contains a similar *C7* hydrogen bond between the interior amide NH and the Gln sidechain. These overlays demonstrate a striking agreement of the gas phase type-I β-turn with both the idealized type-I β-turn motif of the backbone (Fig. 12a) but also representative type-I β-turn crystal structures of the same sequence (Fig. 12b). This excellent agreement extends to several XAQY type-I β-turns found in the X-ray crystal and solution NMR structures of a number of proteins, as summarized in Table 2. The average $i + 1$ and $i + 2$ Ramachandran angles of the crystal and solution NMR structures are $\{-59.9, -27.7\}^\circ$ and $\{-80.5, -16.2\}^\circ$, in close correspondence with the gas-phase values of $\{-70, -13.3\}^\circ$ and $\{-73, -11\}^\circ$ (literature ranges for type-I β-turns are $\{-60, -30\}^\circ$ and $\{-90, 0\}^\circ$).[65] At the same time, the majority of the PDB structures in Table 2 show the Gln amide group forming hydrogen bond(s) with other nearby parts of the protein or with its solvent environment. Thus, while the Gln sidechain H-bond to the center of the β-turn is stabilizing, it is not normative in solution or crystalline environments.

The similarities in backbone structures between gas-phase and β-turns in proteins indicate that the inherent conformational preferences uncovered by the present gas-phase studies are robust to at least some of the environmental perturbations that arise from the protein environment and crystal formation. The present results provide IR spectra of this Gln-containing β-turn, which may facilitate its identification in solution, and contribute to a refinement of the theoretical models of the amide I/II regions on which certain condensed phase structural deductions are built (for examples see ref. 33 and references found within).

4.2. Exploring the potential energy surface

The energy level diagrams in Fig. 11 have provided firm evidence that the observed conformations are in every case among the lowest energy structures of each molecule calculated at the DFT M05-2X/6-31+G(d) level of theory. At the same time, in cases where more than one conformation is observed (Z-Gln-OH and Ac-Gln-NHBn), the observed conformations represent lowest-energy examples of different H-bonding architectures, consistent with a rugged potential energy landscape in which population is trapped, at least in part, behind barriers that prevent isomerization between certain families during the collisional cooling process. Assured by the reasonable correspondence between observation and calculation, in this section we take a more global view of the molecules' potential energy landscapes afforded by the near-exhaustive set of conformational minima calculated at this level of theory.

While the observed conformations are always among the lowest-energy computed structures, there are other structures that are also low in energy, yet are not observed. As others have noted, under thermalized conditions, it is the free energy landscape that is of primary importance for determining population abundances.[80,81] In the initial desorption plume, tightly-folded minima with stiffer vibrational frequencies will be discriminated against on entropic grounds relative to more extended structures. In addition, it is possible that population initially in one conformational minimum will be transferred into a lower-energy minimum during the collisional cooling process.[82]

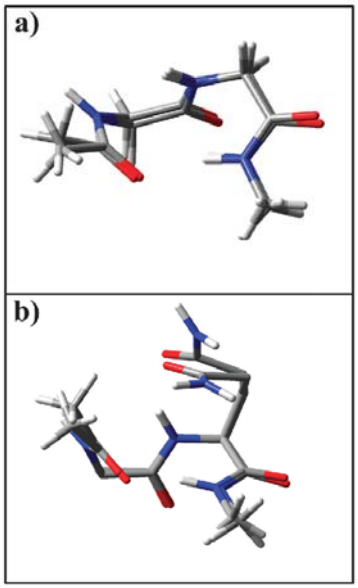


Fig. 12 Two views (a and b) of the assigned gas-phase type-I β-turn of Ac-Ala-Gln-NHBn overlaid on top of the type-1 β-turn (residues 14–17) of myoglobin and cytochrome *C* (PDB ID: 5cyt):[71] (a) view of the peptide backbones without the Gln sidechain, and (b) including the Gln side chains, showing their shared C10//*C7* motif.

Despite not being experimentally observed, it is worth commenting briefly on the other types of low-energy hydrogen-bonded families identified by the calculations, since they add to the rich variety of possibilities afforded by a single glutamine residue. Figures showing representative structures in these families are included in ESI† (Fig. S2–S5).

In the triad of capped amino acids (Z-Gln-NHMe, Ac-Gln-NHMe, and Ac-Gln-NHBn), C5//*C9/C8* conformations exist (labeled in light blue) with calculated energies within 2–3 kJ mol^{-1} of the global minimum. Conformers in this family engage the Gln amide group as a doubly H-bonded bridge to the C5 peptide backbone, with the C=O on the N-terminus and NH on the C-terminus forming *C9* and *C8* H-bonds, respectively, offering yet a fourth means by which a single Gln residue can stabilize an extended peptide backbone.

In Ac-Gln-NHBn, the *C7/π* structure is calculated to be the global minimum. As we have noted, this structure is highly unusual in that it has a turned backbone that has no backbone H-bond (*i.e.*, neither C5 nor C7), and is stabilized by the bridging *C7* and π H-bonds formed by the Gln side chain. When the benzyl cap is removed to form Ac-Gln-NHMe, the backbone reconfigures to form a C7//*C7* structure, which less stable than the C5//*C8* global minimum by 5 kJ mol^{-1} (Fig. 11a, Table S2, ESI†). This structure uses a different Gln side chain configuration in forming the C7 backbone-to-sidechain H-bond (Fig. S4, ESI†), than its C7//*C7/C8/π* counterpart, and the two possibilities are close in energy in the absence of an aromatic cap.

Finally, in Ac-Ala-Gln-NHBn, a second tightly folded structure, labeled as C7//*C7/C8/C12/π* drops in energy to within 2 kJ mol^{-1} of the C10//*C7/π* global minimum when the NHBn cap is removed in this larger peptide (Fig. 11b and Table S3, ESI†). This structure, shown in Fig. S5b (ESI†), engages the Gln side chain in three H-bonds involving both hydrogens of the NH_2 group and the C=O group. This illustrates the expanding possibilities that exist when the peptide backbone is lengthened.

5. Conclusion

The central issue we have sought to address in this work is the way in which the sidechain of a glutamine residue interacts with nearby amide groups, and the impact of such sidechain-to-backbone interactions on local conformational preferences. The structural families observed demonstrate the unique ability of the glutamine sidechain to form a variety of hydrogen bonds with the peptide backbone. The molecules that contain only a single Gln residue probe the preferences of the backbone immediately adjacent to the Gln sidechain, while Ac-Ala-Gln-NHBn reveals possibilities for longer range interactions.

A total of seven conformers of Z-Gln-OH, Z-Gln-NHMe, and Ac-Gln-NHBn have been experimentally assigned, including (i) C5//*C9*, C5//*C8*, and C5//*amide-stacked* structures that incorporate an extended peptide backbone, (ii) the *C7/π* turn in which the H-bond formed between the Gln side chain and the peptide backbone is the only one present, and (iii) the C7//*C7/C8/π* inverse γ-turn with its Gln doubly H-bonded bridge. In the larger Ac-Ala-Gln-NHBn molecule, all the population resides in a single type-I β-turn (C10//*C7/π*) conformation in which the Gln residue stabilizes the structure by forming a *C7* H-bond with the central amide group in the β-turn, which otherwise would be without a hydrogen-bonding partner. With the single exception of the *C7/π* conformer, in every case the sidechain-to-backbone hydrogen bonds are either the primary interaction stabilizing the molecule, or provide additional stabilization to a structure already prevalent in the absence of glutamine.

In this bottom-up approach to studying peptide structure, we also begin to address whether the structures preferred in the smaller peptides carry over to larger ones, and if so, in what way. Comparison of the *C7/π* structure of Ac-Gln-NHBn (Fig. 8b) with the C10//*C7/π* conformer of Ac-Ala-Gln-NHBn (Fig. 8d) displays the structural 'inheritance' between the two structures (Fig. S6, ESI†). The glutamine side-chain in two structures adopts the same *C7/π* geometry, with backbone Ramachandran angles about the Gln residues [(−87.2, −9.1)° *cf.* (−73, −11°)] and H-bond lengths [1.98 *cf.* 197 Å] that are strikingly similar. Furthermore, comparison of the backbone of the β-turn from Ac-Ala-Gln-NHBn with the backbones of a series of X-AQ-Y β-turns found in the protein data bank reveal them to be highly conserved, proving that the β-turn formed in the isolated molecule is robustly retained in the crystalline protein environment, despite the presence of additional intermolecular interactions in the latter case.

Although this initial work on Gln-containing peptides has uncovered certain aspects of the primary conformational influences of a single Gln residue, several important interactions that are significant in fibril formation have not been explored. Sidechain-to-sidechain interactions between glutamine residues will undoubtedly play an important structure-determining role not only in monomers of glutamine rich peptides, but also in the aggregation of these peptides into fibrils. Studies of the present type on model Gln-rich peptides and their aggregates are needed to more directly probe the pivotal role played by Gln–Gln interactions, including those that stabilize the fibril structures *via* polar zippers.[8–10] Gln–Gln H-bonds are likely also to play a critical role in the progenitor species thought to be the key toxic species neurodegenerative diseases, motivating future work in this direction.

Acknowledgements

PSW, JCD, and TSZ gratefully acknowledge the National Science Foundation (CHE-1213289) for support of this research. CM and SHG gratefully acknowledge NSF grant CHE-1307365. HK was supported by the Basic Science Program (NRF-2013R1A1A2008708) through the National Research Foundation of Korea (NRF) funded by the Ministry of Education.

References

1 E. D. Roberson and L. Mucke, *Science*, 2006, **314**, 781–784.
2 M. Goedert and M. G. Spillantini, *Science*, 2006, **314**, 777–781.

3 M. Goedert, *Science*, 2015, **349**, DOI: 10.1126/science.1255555.
4 C. Soto, *Nat. Rev. Neurosci.*, 2003, **4**, 49.
5 C. A. Ross, *Neuron*, 2002, **35**, 819–822.
6 C. A. Ross and M. A. Poirier, *Nat. Med.*, 2004, **10**, S10–S17.
7 L. Mangiarini, K. Sathasivam, M. Seller, B. Cozens, A. Harper, C. Hetherington, M. Lawton, Y. Trottier, H. Lehrach, S. W. Davies and G. P. Bates, *Cell*, 1996, **87**, 493–506.
8 M. Perutz, *Protein Sci.*, 1994, **3**, 1629–1637.
9 R. Nelson, M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe and D. Eisenberg, *Nature*, 2005, **435**, 773–778.
10 M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, H. T. McFarlane, A. Ø. Madsen, C. Riekel and D. Eisenberg, *Nature*, 2007, **447**, 453–457.
11 C. Bleiholder, N. F. Dupuis, T. Wyttenbach and M. T. Bowers, *Nat. Chem.*, 2011, **3**, 172–177.
12 S.-H. Shim, R. Gupta, Y. L. Ling, D. B. Strasfeld, D. P. Raleigh and M. T. Zanni, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 6614–6619.
13 L. Wang, C. T. Middleton, S. Singh, A. S. Reddy, A. M. Woys, D. B. Strasfeld, P. Marek, D. P. Raleigh, J. J. de Pablo, M. T. Zanni and J. L. Skinner, *J. Am. Chem. Soc.*, 2011, **133**, 16062–16071.
14 C. T. Middleton, P. Marek, P. Cao, C.-c. Chiu, S. Singh, A. M. Woys, J. J. de Pablo, D. P. Raleigh and M. T. Zanni, *Nat. Chem.*, 2012, **4**, 355–360.
15 S. D. Moran, A. M. Woys, L. E. Buchanan, E. Bixby, S. M. Decatur and M. T. Zanni, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 3329–3334.
16 C. A. Ross, M. A. Poirier, E. E. Wanker and M. Amzel, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 1–3.
17 H. Y. Zoghbi and H. T. Orr, *Annu. Rev. Neurosci.*, 2000, **23**, 217.
18 M. F. Perutz, T. Johnson, M. Suzuki and J. T. Finch, *Proc. Natl. Acad. Sci. U. S. A.*, 1994, **91**, 5355–5358.
19 M. T. Bauer, K. A. Gilmore and S. A. Petty, *Biochem. Biophys. Res. Commun.*, 2011, **406**, 348–352.
20 S. Kecel, A. E. Ozel, S. Akyuz and S. Celik, *Spectrosc. Int. J.*, 2010, **24**, 219–232.
21 T. D. Vaden, S. A. N. Gowers, T. S. J. A. de Boer, J. D. Steill, J. Oomens and L. C. Snoek, *J. Am. Chem. Soc.*, 2008, **130**, 14640–14650.
22 A. M. Tarditi, M. W. Klipfel, A. M. Rodriguez, F. D. Suvire, G. A. Chasse, O. Farkas, A. Perczel and R. D. Enriz, *THEOCHEM*, 2001, **545**, 29–47.
23 K. Xiong, D. Punihaole and S. A. Asher, *Biochemistry*, 2012, **51**, 5822–5830.
24 J. Gsponer, U. Haberthür and A. Caflisch, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 5154–5159.
25 A. J. Marchut and C. K. Hall, *Biophys. J.*, 2006, **90**, 4574–4584.
26 L. E. Buchanan, J. K. Carr, A. M. Fluitt, A. J. Hoganson, S. D. Moran, J. J. de Pablo, J. L. Skinner and M. T. Zanni, *Proc. Natl. Acad. Sci. U. S. A.*, 2014, **111**, 5796–5801.
27 K. Kar, C. L. Hoop, K. W. Drombosky, M. A. Baker, R. Kodali, I. Arduini, P. C. A. van der Wel, W. S. Horne and R. Wetzel, *J. Mol. Biol.*, 2013, **425**, 1183–1197.
28 A. K. Thakur and R. Wetzel, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 17014–17019.
29 R. Schneider, M. C. Schumacher, H. Mueller, D. Nand, V. Klaukien, H. Heise, D. Riedel, G. Wolf, E. Behrmann, S. Raunser, R. Seidel, M. Engelhard and M. Baldus, *J. Mol. Biol.*, 2011, **412**, 121–136.
30 Y. Loquais, E. Gloaguen, S. Habka, V. Vaquero-Vara, V. Brenner, B. Tardivel and M. Mons, *J. Phys. Chem. A*, 2015, **119**, 5932–5941.
31 E. Gloaguen, Y. Loquais, J. A. Thomas, D. W. Pratt and M. Mons, *J. Phys. Chem. B*, 2013, **117**, 4945–4955.
32 S. Jaeqx, W. Du, E. J. Meijer, J. Oomens and A. M. Rijs, *J. Phys. Chem. A*, 2013, **117**, 1216–1227.
33 E. G. Buchanan, W. H. James III, S. H. Choi, L. Guo, S. H. Gellman, C. W. Muller and T. S. Zwier, *J. Chem. Phys.*, 2012, **137**, 094301.
34 T. S. Zwier, *Annu. Rev. Phys. Chem.*, 1996, **47**, 205–241.
35 T. S. Zwier, *J. Phys. Chem. A*, 2001, **105**, 8827–8839.
36 N. L. Burke, J. G. Redwine, J. C. Dean, S. A. McLuckey and T. S. Zwier, *Int. J. Mass Spectrom.*, 2015, **378**, 196–205.
37 J. C. Dean, E. G. Buchanan and T. S. Zwier, *J. Am. Chem. Soc.*, 2012, **134**, 17186–17201.
38 E. Gloaguen, R. Pollet, F. Piuzzi, B. Tardivel and M. Mons, *Phys. Chem. Chem. Phys.*, 2009, **11**, 11385–11388.
39 W. Chin, F. Piuzzi, J. P. Dognon, I. Dimicoli and M. Mons, *J. Chem. Phys.*, 2005, **123**, 084301.
40 J. A. Stearns, O. V. Boyarkin and T. R. Rizzo, *J. Am. Chem. Soc.*, 2007, **129**, 13820–13821.
41 V. Brenner, F. Piuzzi, I. Dimicoli, B. Tardivel and M. Mons, *Angew. Chem., Int. Ed.*, 2007, **46**, 2463–2466.
42 J. R. Gord, P. S. Walsh, B. F. Fisher, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. B*, 2014, **118**, 8246–8256.
43 R. Kusaka, D. Zhang, P. S. Walsh, J. R. Gord, B. F. Fisher, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2013, **117**, 10847–10862.
44 M. Alauddin, E. Gloaguen, V. Brenner, B. Tardivel, M. Mons, A. Zehnacker-Rentien, V. Declerck and D. J. Aitken, *Chem. – Eur. J.*, 2015, **21**, 16479–16493.
45 W. H. James III, E. E. Baquero, V. A. Shubert, S. H. Choi, S. H. Gellman and T. S. Zwier, *J. Am. Chem. Soc.*, 2009, **131**, 6574–6590.
46 W. H. James III, E. E. Baquero, S. H. Choi, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2010, **114**, 1581–1591.
47 W. H. James III, C. W. Müller, E. G. Buchanan, M. G. D. Nix, L. Guo, L. Roskop, M. S. Gordon, L. V. Slipchenko, S. H. Gellman and T. S. Zwier, *J. Am. Chem. Soc.*, 2009, **131**, 14243–14245.
48 W. H. James III, E. G. Buchanan, L. Guo, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2011, **115**, 11960–11970.
49 W. H. James III, E. G. Buchanan, C. W. Müller, J. C. Dean, D. Kosenkov, L. V. Slipchenko, L. Guo, A. G. Reidenbach, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2011, **115**, 13783–13798.
50 E. G. Buchanan, W. H. James III, A. Gutberlet, J. C. Dean, L. Guo, S. H. Gellman and T. S. Zwier, *Faraday Discuss.*, 2011, **150**, 209–226.

51 P. S. Walsh, R. Kusaka, E. G. Buchanan, W. H. James, B. F. Fisher, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2013, **117**, 12350–12362.
52 J. C. Dean, R. Kusaka, P. S. Walsh, F. Allais and T. S. Zwier, *J. Am. Chem. Soc.*, 2014, **136**, 14780–14795.
53 D. M. Lubman, *Mass Spectrom. Rev.*, 1988, **7**, 559–592.
54 R. H. Page, Y. R. Shen and Y. T. Lee, *J. Chem. Phys.*, 1988, **88**, 4621–4636.
55 F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440–467.
56 P. K. Weiner and P. A. Kollman, *J. Comput. Chem.*, 1981, **2**, 287–303.
57 Y. Zhao and D. G. Truhlar, *J. Chem. Theory Comput.*, 2007, **3**, 289–300.
58 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, *et al.*, 2009.
59 H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov and P. E. Bourne, *Nucleic Acids Res.*, 2000, **28**, 235–242.
60 European Bioinformatics Institute, PDBeMotif Search Tool (http://www.ebi.ac.uk/pdbe-site/pdbemotif).
61 H. Torii, *J. Phys. Chem. Lett.*, 2012, **3**, 112–116.
62 E. Gloaguen, F. Pagliarulo, V. Brenner, W. Chin, F. Piuzzi, B. Tardivel and M. Mons, *Phys. Chem. Chem. Phys.*, 2007, **9**, 4491–4497.
63 J. C. Dean, E. G. Buchanan, W. H. James III, A. Gutberlet, B. Biswas, P. V. Ramachandran and T. S. Zwier, *J. Phys. Chem. A*, 2011, **115**, 8464–8478.
64 W. Chin, I. Compagnon, J. P. Dognon, C. Canuel, F. Piuzzi, I. Dimicoli, G. von Helden, G. Meijer and M. Mons, *J. Am. Chem. Soc.*, 2005, **127**, 1388–1389.
65 P. N. Lewis, F. A. Momany and H. A. Scheraga, *Biochim. Biophys. Acta, Protein Struct.*, 1973, **303**, 211–229.
66 S. Xiang, S. A. Short, R. Wolfenden and C. W. Carter, *Biochemistry*, 1997, **36**, 4768–4774.
67 X. Wang, Z. Lou, X. Dong, W. Yang, Y. Peng, B. Yin, Y. Gong, J. Yuan, W. Zhou, M. Bartlam, X. Peng and Z. Rao, *J. Mol. Biol.*, 2009, **390**, 530–537.
68 V. Guillet, N. Ohta, S. Cabantous, A. Newton and J.-P. Samama, *J. Biol. Chem.*, 2002, **277**, 42003–42010.
69 X. Yao, H. H. Chong, S. Waltersperger, M. T. Wang, Y. X. He and S. Cui, *New HIV entry inhibitor MTSFT/T23 complex*, 2014, DOI: 10.2210/pdb4i2l/pdb.
70 S. R. Andrews, E. J. Taylor, G. Pell, F. Vincent, V. M.-A. Ducros, G. J. Davies, J. H. Lakey and H. J. Gilbert, *J. Biol. Chem.*, 2004, **279**, 54369–54379.
71 T. Takano, *Refinement of Myoglobin and Cytochrome C*, 1989, DOI: 10.2210/pdb5cyt/pdb.
72 W.-S. Jung, J. Lee, M.-I. Kim, J. Ma, T. Nagamatsu, E. Goo, H. Kim, I. Hwang, J. Han and S. Rhee, *PLoS One*, 2011, **6**, e22443.
73 D. H. Fremont, D. Monnaie, C. A. Nelson, W. A. Hendrickson and E. R. Unanue, *Immunity*, 1998, **8**, 305–317.
74 S. E. Hennig, J.-H. Jeoung, S. Goetzl and H. Dobbek, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, 5235–5240.
75 C. W. Davies, J. Chaney, G. Korbel, D. Ringe, G. A. Petsko, H. Ploegh and C. Das, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 3900–3904.
76 G. M. Clore, J. G. Omichinski, K. Sakaguchi, N. Zambrano, H. Sakamoto, E. Appella and A. M. Gronenborn, *Science*, 1994, **265**, 386–391.
77 N. Morellet, H. Meudal, S. Bouaziz and B. P. Roques, *Biochem. J.*, 2006, **393**, 725–732.
78 T. Nagashima, F. Hayashi and S. Yokoyama, *Solution structure of N-terminal domain of chromosome 12 open reading frame 2*, 2005, DOI: 10.2210/pdb2cs4/pdb.
79 C. L. Hoop, H.-K. Lin, K. Kar, G. Magyarfalvi, J. M. Lamley, J. C. Boatz, A. Mandal, J. R. Lewandowski, R. Wetzel and P. C. A. van der Wel, *Proc. Natl. Acad. Sci. U. S. A.*, 2016, **113**, 1546–1551.
80 V. A. Shubert, E. E. Baquero, J. R. Clarkson, W. H. James, J. A. Turk, A. A. Hare, K. Worrel, M. A. Lipton, D. P. Schofield, K. D. Jordan and T. S. Zwier, *J. Chem. Phys.*, 2007, **127**, 234315.
81 E. Gloaguen, B. de Courcy, J. P. Piquemal, J. Pilme, O. Parisel, R. Pollet, H. S. Biswal, F. Piuzzi, B. Tardivel, M. Broquier and M. Mons, *J. Am. Chem. Soc.*, 2010, **132**, 11860–11863.
82 R. S. Ruoff, T. D. Klots, T. Emilsson and H. S. Gutowsky, *J. Chem. Phys.*, 1990, **93**, 3142.

THE JOURNAL OF PHYSICAL CHEMISTRY A

Article

pubs.acs.org/JPCA

Cyclic Constraints on Conformational Flexibility in γ-Peptides: Conformation Specific IR and UV Spectroscopy

Patrick S. Walsh,[†] Ryoji Kusaka,[†,§] Evan G. Buchanan,[†] William H. James III,[†,#] Brian F. Fisher,[‡] Samuel H. Gellman,[‡] and Timothy S. Zwier*[,†]

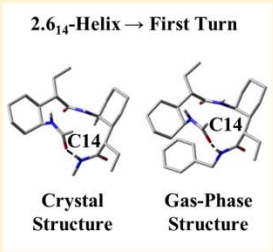
[†]Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, Indiana 47907 United States

[‡]Department of Chemistry, University of Wisconsin-Madison, Madison, Wisconsin 53706 United States

[§]Department of Chemistry, Graduate School of Science, Hiroshima University, Higashi-Hiroshima, 739-8526, Japan.

Supporting Information

ABSTRACT: Single-conformation spectroscopy has been used to study two cyclically constrained and capped γ-peptides: Ac-γ_{ACHC}-NHBn (hereafter γ_{ACHC}, Figure 1a), and Ac-γ_{ACHC}-γ_{ACHC}-NHBn ($\gamma\gamma_{ACHC}$, Figure 1b), under jet-cooled conditions in the gas phase. The γ-peptide backbone in both molecules contains a cyclohexane ring incorporated across each Cβ-Cγ bond and an ethyl group at each Cα. This substitution pattern was designed to stabilize a (g+, g+) torsion angle sequence across the Cα−Cβ−Cγ segment of each γ-amino acid residue. Resonant two-photon ionization (R2PI), infrared−ultraviolet hole-burning (IR−UV HB), and resonant ion-dip infrared (RIDIR) spectroscopy have been used to probe the single-conformation spectroscopy of these molecules. In both γ_{ACHC} and $\gamma\gamma_{ACHC}$, all population is funneled into a single conformation. With RIDIR spectra in the NH stretch (3200−3500 cm^{-1}) and amide I/II regions (1400−1800 cm^{-1}), in conjunction with theoretical predictions, assignments have been made for the conformations observed in the molecular beam. γ_{ACHC} forms a single nearest-neighbor C9 hydrogen-bonded ring whereas $\gamma\gamma_{ACHC}$ takes up a next-nearest-neighbor C14 hydrogen-bonded structure. The gas-phase C14 conformation represents the beginning of a 2.6_{14}-helix, suggesting that the constraints imposed on the γ-peptide backbone by the ACHC and ethyl groups already impose this preference in the gas-phase di-γ-peptide, in which only a single C14 H-bond is possible, constituting one full turn of the helix. A similar conformational preference was previously documented in crystal structures and NMR analysis of longer γ-peptide oligomers containing the γ_{ACHC} subunit [Guo, L., et al. *Angew. Chem. Int. Ed.* **2011**, *50*, 5843−5846]. In the gas phase, the γ_{ACHC}-H_2O complex was also observed and spectroscopically interrogated in the molecular beam. Here, the monosolvated γ_{ACHC} retains the C9 hydrogen bond observed in the bare molecule, with the water acting as a bridge between the C-terminal carbonyl and the π-cloud of the UV chromophore. This is in contrast to the unconstrained γ-peptide-H_2O complex, which incorporates H_2O into both C9 and amide-stacked conformations.



I. INTRODUCTION

Proteins are composed of α-amino acid residues, in which one carbon atom (Cα) separates adjacent amide groups. There has been increasing interest in extending the set of protein-like oligoamides to include other building blocks, including β-amino acid residues (two carbons between amide groups) and γ-amino acid residues (three carbons between amide groups). These protein-inspired oligomers can also contain heterogeneous backbones, in which α, β, and/or γ residues are combined. These synthetic foldamers are designed to adopt secondary structures comparable to those found in biopolymers (e.g., helices, sheets, reverse turns), and to display properties beyond those available to conventional peptides and proteins, including resistance to enzymatic degradation and greater conformational stability.[1−6] Among the structural elements that have benefited most from this development are helices, which can be formed with a variety of helical pitches, radii, and side-chain architecture.

As Figure 1d illustrates, the conformation of a γ residue is determined by four dihedral angles. Two are analogous to the Ramachandran angles used to determine α-amino acid residue conformation (φ, ψ), and the other two arise from the extended γ residue backbone (θ, ζ). This increase in backbone complexity in a γ residue relative to an α residue might be anticipated to result in a wider range of competing low-energy structures, which could lead to a greater diversity in folding preferences for γ-peptides relative to α-peptides. The presence of three positions for substitution within each γ residue also offers a richer palette of side chain-based strategies to control γ residue conformational propensity relative to the options available for α residues. Several previous studies focused attention on γ-peptides with side chains similar to those in α-

Received: August 31, 2013
Revised: October 17, 2013
Published: October 22, 2013

dx.doi.org/10.1021/jp408736t | J. Phys. Chem. A 2013, 117, 12350−12362

THE JOURNAL OF CHEMICAL PHYSICS **142**, 154303 (2015)

Binding water to a PEG-linked flexible bichromophore: IR spectra of diphenoxyethane-$(H_2O)_n$ clusters, n = 2-4

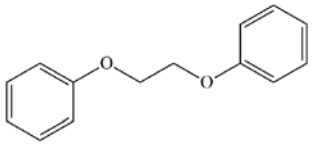
Patrick S. Walsh, Evan G. Buchanan,[a] Joseph R. Gord, and Timothy S. Zwier[b]
Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, Indiana 47907-2084, USA

(Received 9 February 2015; accepted 26 March 2015; published online 17 April 2015)

The single-conformation infrared (IR) and ultraviolet (UV) spectroscopies of neutral 1,2-diphenoxyethane-$(H_2O)_n$ clusters with n = 2-4 (labeled henceforth as 1:n) have been studied in a molecular beam using a combination of resonant two-photon ionization, IR-UV holeburning, and resonant ion-dip infrared (RIDIR) spectroscopies. Ground state RIDIR spectra in the OH and CH stretch regions were used to provide firm assignments for the structures of the clusters by comparing the experimental spectra with the predictions of calculations carried out at the density functional M05-2X/6-31+G(d) level of theory. At all sizes in this range, the water molecules form water clusters in which all water molecules engage in a single H-bonded network. Selective binding to the *tgt* monomer conformer of 1,2-diphenoxyethane (C_6H_5-O-CH_2-CH_2-O-C_6H_5, DPOE) occurs, since this conformer provides a binding pocket in which the two ether oxygens and two phenyl ring π clouds can be involved in stabilizing the water cluster. The 1:2 cluster incorporates a water dimer "chain" bound to DPOE much as it is in the 1:1 complex [E. G. Buchanan *et al.*, J. Phys. Chem. Lett. **4**, 1644 (2013)], with primary attachment via a double-donor water that bridges the ether oxygen of one phenoxy group and the π cloud of the other. Two conformers of the 1:3 cluster are observed and characterized, one that extends the water chain to a third molecule (1:3 chain) and the other incorporating a water trimer cycle (1:3 cycle). A cyclic water structure is also observed for the 1:4 cluster. These structural characterizations provide a necessary foundation for studies of the perturbations imposed on the two close-lying S_1/S_2 excited states of DPOE considered in the adjoining paper [P. S. Walsh *et al.*, J. Chem. Phys. **142**, 154304 (2015)]. © *2015 AIP Publishing LLC.* [http://dx.doi.org/10.1063/1.4917305]

I. INTRODUCTION

1,2-diphenoxyethane (DPOE) is a prototypical flexible bichromophore in which the two phenyl rings responsible for the molecule's ultraviolet (UV) absorptions are joined by a flexible -O-CH_2-CH_2-O- linker characteristic of a polyethylene glycol (PEG) or polyethylene oxide (PEO) chain.[1–8]



The fact that this linkage is flexible imbues DPOE with a pair of close-lying multi-dimensional excited state surfaces that change the distance and orientation of the two aromatic chromophores, thereby shifting their relative energies and modulating the strength of their excitonic coupling. By bringing the molecule into the gas phase and cooling it in a supersonic expansion, the population is funneled into the zero-point levels of low-lying conformational minima where they can be studied via conformation-specific spectroscopy.

While much of the focus of past work[9–25] has been on the vibronic coupling between the close-lying S_1 and S_2 states, a necessary prerequisite for such understanding is establishing the inherent ground-state conformational preferences of the molecule, which serve as the starting geometry for excitation onto the excited state surfaces. To that end, the observed vibronic and infrared (IR) transitions of DPOE have been previously assigned to the *ttt* and *tgt* conformers shown in Figure 1(a).[22,23] These conformers differ primarily in the dihedral angle τ about the central C–C bond, either *trans* (180°) or *gauche* (60°), with the former of C_{2h} symmetry and the latter C_2 symmetry. The two conformers have roughly equal populations in the supersonic expansion, as proven by infrared population transfer spectroscopy.

Armed with this understanding about the isolated molecule, one can then address how the conformational preferences are modified in the presence of solvent, connecting more directly with the solution phase. As nature's solvent, water is an important solvent to consider in this regard. Indeed, DPOE is a fascinating solute for water because it offers several binding sites, most notably the two ether oxygens and the two π clouds of the two phenyl rings. As a first step along this path, we previously studied the infrared and ultraviolet spectroscopies of the 1:1 DPOE-H_2O complex containing a single H_2O molecule.[24] This single water molecule binds selectively to the *tgt* conformer of DPOE (Figure 1(b)), donating both its OH groups to H-bonds to the ether oxygen of one phenoxy group and the π cloud of the other. This asymmetric solvent binding is reflected

[a]Present address: Quantum Electronics and Photonics Division, Physical Measurement Laboratory, National Institute of Standards and Technology, Boulder, Colorado 80305, USA.
[b]Author to whom correspondence should be addressed. Electronic mail: zwier@purdue.edu

THE JOURNAL OF CHEMICAL PHYSICS **142**, 154304 (2015)

Solvent-mediated internal conversion in diphenoxyethane-$(H_2O)_n$ clusters, n = 2-4

Patrick S. Walsh, Evan G. Buchanan,[a] Joseph R. Gord, and Timothy S. Zwier[b]
Department of Chemistry, Purdue University, 560 Oval Drive, West Lafayette, Indiana 47907-2084, USA

(Received 9 February 2015; accepted 26 March 2015; published online 17 April 2015)

1,2-diphenoxyethane (DPOE) is a flexible bichromophore whose excited states come in close-lying pairs whose splitting and vibronic coupling can be modulated by solvent. Building on the ground state infrared spectroscopy of DPOE-$(H_2O)_n$ clusters with n = 2-4 from the adjoining paper [Walsh *et al.*, J. Chem. Phys. **142**, 154303 (2015)], the present work focuses on the vibronic and excited state infrared spectroscopies of the clusters. The type and degree of asymmetry of the water cluster binding to DPOE is reflected in the variation in the magnitude of the S_1/S_2 splitting with cluster size. Excited state resonant ion-dip infrared spectroscopy was performed at the electronic origins of the first two excited states in order to explore how the water clusters' OH stretch spectra report on the nature of the two excited states, and the interaction of the S_2 state with nearby S_1 vibronic levels mediated by the water clusters. The data set, when taken as a whole, provides a state-to-state view of internal conversion and the role of solvent in mediating conversion of electronic excitation between two chromophores, providing a molecular-scale view of Kasha's rule. © *2015 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4917307]

I. INTRODUCTION

Kasha's rule stands as one of the cornerstones of solution-phase photochemistry,[1–4] summarizing the common observation that, for many molecules, independent of the excitation wavelength used, emission occurs primarily from the first excited electronic state. The rule arises because, in many circumstances, the time scale for internal conversion from high lying electronic states $\{S_n\}$ to the lowest excited state of the same multiplicity, S_1, is much shorter than the timescale for fluorescence itself. In solution, vibrational relaxation within a given electronic state is also very fast (ps timescale), leading to a pile-up of thermalized population in the lowest excited state, from which emission occurs. More than 40 years ago, Bixon and Jortner provided a theoretical description of such radiationless transitions.[5,6] Their expression for the rate of internal conversion showed clearly that the small energy difference between the manifold of excited singlet electronic states provides a mechanism for fast internal conversion down to S_1, where fluorescence can compete with the slower internal conversion to S_0 that results from the large energy gap between S_1 and S_0.

In the simplest case in which the molecule contains a single electronic chromophore, internal conversion occurs between states of the same multiplicity localized on this chromophore, where the geometry changes and electronic energy gaps are those of that single chromophore. When more than one chromophore is present, either as separate parts of the same molecule or as separate molecules within the sample, electronic energy transfer between chromophores is possible, leading to fascinating and important new possibilities. Electronic energy transfer between two chromophores, often referred to as Förster resonant energy transfer (FRET), is widely used by the biophysics community to probe the time-dependent changes in distance between donor and acceptor chromophores, with the rate of electronic energy transfer sensitively dependent on the distance between them, their electronic energy gap, and relative orientation.[7–9] In the context of light harvesting and photosynthesis, resonant energy transfer occurs within an array of chromophores in close spatial proximity, funneling electronic excitation from its excitation point to a distant reaction center, where that energy can be used to generate charge separation. Recent advances in 2D electronic spectroscopy have provided a means of probing these energy transfer mechanisms on the femtosecond timescales, leading to qualitative new insights into the pathways for electronic excitation and the possible role of coherences in electronic energy transport.[10–15] Several recent studies have sought to determine whether these coherences play an integral role in the energy transfer process and to what extent particular vibrational modes will facilitate the energy transfer.[16,17]

Most biologically relevant internal conversion or electronic energy transfer processes occur in aqueous solution at room temperature, where competing intramolecular and inter-chromophore photophysical pathways are modified or possibly even re-directed by interactions with aqueous solvent. This has led several groups to study the laser spectroscopy of coupled chromophores in the form of both molecular dimers and as covalently linked bichromophores, prepared under jet-cooled, isolated conditions.[18–27] These experiments enable the in-depth study of the close lying excited states and the underlying vibronic coupling on a state-specific basis devoid of solvent effects. The results serve as a testing ground for emerging theories of excitonic and vibronic couplings.[28–34]

[a]Current address: Quantum Electronics and Photonics Division, Physical Measurement Laboratory, National Institute of Standards and Technology, Boulder, Colorado 80305, USA.
[b]Author to whom correspondence should be addressed. Electronic mail: zwier@purdue.edu

