

# Observing the Unfolding Transition of $\beta$ -Hairpin Peptides with Nonlinear Infrared Spectroscopy

by

Adam W. Smith

B.S. Chemistry

University of Utah, 2002

Submitted to the Department of Chemistry  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy

at the

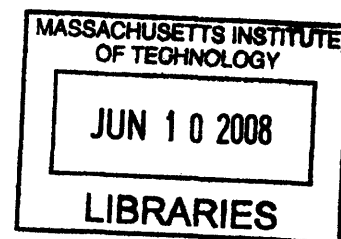
Massachusetts Institute of Technology

February 2008

[June 2008]

© 2008 Massachusetts Institute of Technology.

All rights reserved.



**ARCHIVES**

Signature of Author: \_\_\_\_\_  
Adam W. Smith  
February 21, 2008

Certified by: \_\_\_\_\_  
Andrei Tokmakoff  
Professor of Chemistry  
Thesis Supervisor

Accepted by: \_\_\_\_\_  
Robert W. Field  
Professor of Chemistry  
Chair, Departmental Committee on Graduate Studies

This doctoral thesis has been examined by a committee of the Department of Chemistry that included:

Professor Robert G. Griffin \_\_\_\_\_  
Chair

Professor Keith A. Nelson \_\_\_\_\_

Professor Andrei Tokmakoff \_\_\_\_\_  
Thesis Supervisor

# Observing the Unfolding Transition of $\beta$ -Hairpin Peptides with Nonlinear Infrared Spectroscopy

by

Adam W. Smith

Submitted to the Department of Chemistry on  
February 21, 2008  
in partial fulfillment of the requirements for the degree of  
Doctor of Philosophy

## Abstract

The biological function of a protein is in large measure determined by its three-dimensional structure. To date, however, the transition of the protein between the native and non-native conformations is not well-understood. Part of the difficulty is the large conformational space available to a poly-peptide chain, and a general lack of experimental probes that can access local structural information on the time scale of the transition. Single domain peptides are excellent model systems that reduce the size and complexity of the problem, while maintaining the essential physical interactions. In this thesis,  $\beta$ -hairpin peptides are used as model systems for studying  $\beta$ -sheet secondary structure. Hairpin folding has been studied for a number of years, but there is still debate in the literature about the relative importance of the cross-strand hydrogen bonds, tertiary side chain contacts, and  $\beta$ -turn in the folding pathway. In addition, the denatured state is very poorly understood, which complicates any attempt to describe the folding pathway.

In this work, amide I vibrational spectroscopy is used to resolve the secondary structure of  $\beta$ -hairpin peptides during thermal denaturation. Spectroscopic modeling is presented to describe the amide I band of  $\beta$ -hairpins and relate it to structural features. Three spectroscopic methods are used to probe the amide I band: Fourier transform infrared (FTIR) spectroscopy, two-dimensional infrared (2D IR) spectroscopy, and dispersed vibrational echo (DVE) spectroscopy. 2D IR and DVE spectroscopy are 3<sup>rd</sup> order-nonlinear methods that interrogate the system with a series of ultrafast (100 fs) laser pulses. 2D IR spectra reveal vibrational couplings and measure spectral dynamics on a picosecond time scale. The 2D IR spectra of TZ2 and PG12 are used to identify  $\beta$ -sheet structure during thermal denaturation and to measure the amide I homogeneous line width changes with temperature. The transient folding of TZ2 and PG12 is also probed with 2D IR and DVE spectroscopy following a 10 to 20 °C temperature jump. In order to increase the structural sensitivity of amide I spectroscopy, <sup>13</sup>C and <sup>18</sup>O isotope labels are

incorporated into specific peptide amide groups. The isotope labels red-shift vibrational frequencies and help resolve local structure at the turn and mid-strand regions of the peptides. The transient folding at each labeled site is also measured following a temperature jump.

Together, the results of this work identify folding rates for the thermal disordering transition. For PG12, the unfolding time at the mid-strand region of the peptide is 130 ns, and the turn is found to be stable throughout the transition. For TZ2, the kinetic folding rates at each of the labeled sites are found to be very similar to the global unfolding time ( $\sim 1 \mu\text{s}$ ). Temperature jump 2D IR spectroscopy of TZ2 reveals that the disordering mechanism is unique for different regions of the peptide. The band corresponding to the turn region decouples from the other vibrations, but does not show signs of disorder. In the mid-strand region of the peptide, the isotope-shifted band decouples from the main amide I band and also broadens significantly. Local disordering and decoupling both occur on a  $1 \mu\text{s}$  time scale. The observations in this work combined with previous measurements are used to describe the folding as a hybrid zipper. TZ2 folding is initiated with the formation of the  $\beta$ -turn, following which the tryptophan side chains form a compact, but non-native hydrophobic core. Next, the backbone native contacts are formed and finally the tryptophan side chain packing reaches the native configuration.

Thesis Supervisor: Andrei Tokmakoff  
Title: Associate Professor of Chemistry

*This work is dedicated to Lesli,  
Who has heroically stood by my side from it's beginning to end.*

## Acknowledgements

Being a member of the Tokmakoff group over the last five years has been a tremendous pleasure. I have had the opportunity to work with brilliant scientists who have each been an integral part of this thesis. My advisor, Andrei, has been a valuable source of guidance and criticism. He always encouraged me to take chances and go after big problems, and I have learned a lot from his approach to research questions. He is a man of integrity and an asset to the community.

I could write pages of acknowledgements for the help and guidance I've received from my companions in the Tokmakoff group. Together we have braved the basement trenches of MIT. My first project was to build a regenerative amplifier during which I continually relied on the advice of the original set of graduate students. Chris, Joel, Munira, and Nuri were examples of hard work and set an exceedingly high standard for those of us who followed them. Most of the work that went into this thesis was done in the Big Lab, as 6-048 is affectionately known. Hoi Sung was my chief instructor there, and I am grateful for the time he put into showing me how to run the equipment. He was very patient as I worked my way through each piece of hardware, and the advances presented in the following chapters would have been impossible without his trail blazing studies of protein folding. As the group shifted to a new generation, I have been grateful for Joe, Lauren, Sean, Ziad, Becky, and Krupa. They are extremely gifted scientists, and have each in their own way contributed to my development. Some of my recent work has been done in close collaboration with Kevin and Josh. Kevin has mastered the art of the Big Lab, and Josh's chemistry skills quickly outpaced my initial work in peptide synthesis. Their efforts contributed to Chapters 7 and 8, and it was a pleasure to work with them. Finally, I also am indebted to a fantastic set of postdocs who have passed through our group: Chris Cheatum, Matt Decamp, Poul Petersen, and Benjamin Dietzek.

During my time here I have received unimaginable support from my friends and family. The friends Lesli and I made while living in Cambridge helped us feel rich and happy. I have so many good memories here that I can't imagine leaving it behind. It has been a gift to associate with so many good people.

I want to thank my mom and dad for the years that they spent setting my trajectory toward this point. With children of my own I am beginning to understand the time and energy that it takes, and I admire their efforts. They have believed in me when there was very little evidence to justify it, and I will forever be proud to be their son.

I also want to thank my children, Addison and Eleanor, who have both been an integral part of my graduate research. Each of them has a wit and clarity of mind that is unmatched. Addison approaches and solves problems with textbook skill. He has never backed down from complex questions, and he continually teaches me that there are no upper bounds to curiosity. Eleanor has a razor sharp gift of discernment that she uses to understand the world around her. She has an impressive capacity for abstract thought, and I admire her strength as she defends her points of view. My time here with Addison and Eleanor has been a lot of fun. I'll never forget our bike rides along the Charles River, our trips to the museums, or our adventures in the canoe. I've benefited tremendously from these activities, and they have been a substantive part of my success here.

Finally, I want to thank Lesli, who has put an amazing amount of effort into helping me reach this point. She suffered through the inevitable lows of graduate school without much reward. In the face of this, she worked hard to build a life that has been both meaningful and fun. Lesli truly is a person of strong character. She is a skilled teacher and has a gift for connecting with people. This has served her both as a mother and as a teacher at SCALE. I admire her as a person, and perhaps more selfishly, I am glad that she chooses to spend her life with me.

# Contents

<b>List of Figures.....</b>	<b>13</b>
<b>List of Tables .....</b>	<b>30</b>
<b>1. Introduction &amp; Background .....</b>	<b>31</b>
1.1. Current Issues in Protein Folding .....	31
1.2. Hairpin Folding Simulations.....	38
1.3. Experimental Folding Kinetics .....	44
1.3.1. GB1 <sub>41-56</sub> folding kinetics.....	46
1.3.2. Trpzip folding kinetics.....	47
1.3.3. Folding kinetics of other peptides.....	52
1.4. Experimental Aims .....	53
1.5. References.....	57
<b>2. Peptide Synthesis.....</b>	<b>62</b>
2.1. An Introduction to Solid Phase Peptide Synthesis.....	62
2.2. Resins and Linkers .....	65
2.3. N <sub>α</sub> Deprotection .....	67
2.4. Activation & Coupling.....	68
2.5. Side Chain Protecting Groups.....	69
2.6. Cleavage and Isolation.....	70
2.7. Synthesis Procedures .....	71



2.7.1.	Synthesis of TZ2 .....	71
2.7.2.	Synthesis of PG12 .....	75
2.8.	References .....	78
2.9.	Appendix .....	80
<b>3.</b>	<b>Experimental Infrared Probes .....</b>	<b>87</b>
3.1.	FTIR Spectroscopy .....	87
3.2.	2D IR Spectroscopy .....	88
3.2.1.	Pulse generation and interferometer. ....	88
3.2.2.	2D IR signal detection .....	91
3.3.	DVE Spectroscopy .....	94
3.4.	Temperature-jump .....	96
3.5.	Temperature-Jump Probes .....	98
3.6.	Temperature-Jump Calibration from FTIR .....	100
3.7.	Sample Cell .....	103
3.8.	References .....	106
3.9.	Appendix .....	107
3.9.1.	Cell Housing (brass) .....	107
3.9.2.	Mounting Plate (steel) .....	108
3.9.3.	Insulator Plate (acrylic) .....	109
3.9.4.	Cell Mount – Front (brass) .....	110
3.9.5.	Cell Mount – Rear (brass) .....	112

<b>4. Amide I 2D IR Spectroscopy of <math>\beta</math>-Hairpin Peptides.....</b>	<b>113</b>
4.1. Abstract.....	113
4.2. Introduction.....	114
4.3. Experiment.....	118
4.3.1. Samples & preparation.....	118
4.3.2. Spectroscopy.....	119
4.4. Calculations.....	120
4.4.1. Calculation of 1D and 2D IR spectra.....	120
4.4.2. Visualization of vibrational modes .....	123
4.5. Results.....	124
4.5.1. TZ2.....	124
4.5.2. TZ2-T3*T10* .....	128
4.5.3. PG12 .....	131
4.5.4. Gramicidin S.....	135
4.6. Discussion.....	137
4.6.1. Previous IR studies of $\beta$ -hairpins.....	137
4.6.2. Testing coupling models.....	138
4.6.3. Assignment of site energies .....	140
4.6.4. Comparison to previous work.....	144
4.6.5. Effect of twist on hairpin spectra.....	145
4.7. Conclusion .....	146
4.8. Acknowledgements.....	148
4.9. References.....	149

<b>5. Residual Native Structure in a Thermally Denatured <math>\beta</math>-Hairpin .....</b>	<b>156</b>
5.1. Abstract.....	156
5.2. Introduction.....	157
5.3. Results & Discussion .....	159
5.4. Conclusion .....	165
5.5. Acknowledgements.....	165
5.6. References.....	166
<b>6. Probing Local Structural Events in <math>\beta</math>-Hairpin Unfolding With Transient</b>	
<b>Nonlinear Infrared Spectroscopy .....</b>	<b>168</b>
6.1. Abstract.....	168
6.2. Introduction.....	169
6.3. Experimental Details.....	171
6.4. Amide I Spectroscopy of Folded PG12 .....	172
6.5. Thermal Denaturation .....	178
6.6. Transient Unfolding.....	184
6.7. Conclusions.....	188
6.8. References.....	190
<b>7. Equilibrium Thermal Denaturation of Isotope-Labeled TZ2 .....</b>	<b>193</b>
7.1. Introduction.....	193
7.2. Materials & Methods .....	195
7.2.1. Peptide samples.....	195

7.2.2.	Spectroscopic methods.....	196
7.3.	Amide I Spectroscopy of TZ2 at 25 °C .....	196
7.4.	Simulated Spectra .....	202
7.5.	Thermal Denaturation .....	210
7.5.1.	FTIR.....	210
7.5.2.	2D IR.....	215
7.5.3.	DVE .....	221
7.6.	Thermal Response of Isotope-shifted Bands. ....	227
7.6.1.	TZ2-K8 .....	227
7.6.2.	TZ2-S1 .....	230
7.6.3.	TZ2-TT .....	231
7.7.	Conclusions.....	233
7.8.	References.....	235
7.9.	Appendix.....	237
7.9.1.	Simulated FTIR spectra .....	237
7.9.2.	DVE from 2D IR.....	238
<b>8.</b>	<b>Transient Unfolding of TZ2 .....</b>	<b>240</b>
8.1.	Introduction.....	240
8.2.	Methods.....	243
8.3.	Transient Folding Results .....	244
8.3.1.	T-jump DVE at $T_0 = 45\text{ °C}$ .....	244
8.3.2.	T-jump DVE at variable $T_0$ .....	249

8.3.3.	T-jump 2D IR.....	254
8.4.	Discussion.....	262
8.4.1.	Transient DVE time scales.....	262
8.4.2.	Signatures of $\beta$ -sheet disordering .....	263
8.4.3.	Rise time of the homogeneous line width.....	265
8.4.4.	Isotope-shifted peaks in T-jump 2D IR .....	267
8.5.	Conclusion .....	270
8.6.	Acknowledgements.....	276
8.7.	References.....	277
8.8.	Appendix.....	279

## List of Figures

<b>Figure 1.1</b>	Stick diagram of GB1 <sub>41-56</sub> . Dashed lines represent H-bonds observed in simulations described in the text.....	41
<b>Figure 1.2</b>	Simulated amide I spectra for two peptides in different configurations. The cross-strand interactions of the $\beta$ -hairpin (left) lead to two modes split by approximately 30 cm <sup>-1</sup> . The IR spectrum of a random coil peptide (right) displays a broad peak centered at 1655 cm <sup>-1</sup> .....	46
<b>Figure 1.3</b>	Stick diagram of TZ2C. Circles represent amide groups that were targeted for <sup>13</sup> C isotope labeling .....	51

<b>Figure 2.1</b>	Flow chart for solid phase peptide synthesis. The red hexagon symbolizes the N $_{\alpha}$ -Fmoc protecting group. The PG-labeled triangle represents side chain (SC) protecting groups where appropriate (see Section 2.5). The solid support resin is symbolized by the blue-filled circle. ....	63
<b>Figure 2.2</b>	Diagrams of common chemicals in SPPS. Dimethylformamide (DMF) is the primary reaction solvent. Piperidine and N-methylmorpholine are organic bases used in the deprotection and activation steps respectively. An example Fmoc amino acid is shown for alanine, and the activator HBTU (O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate). ....	65
<b>Figure 2.3</b>	(Left) Wang resin with the first amino acid attached. The circle P represents the polystyrene matrix. (Right) Rink Amide MBHA Resin linker with the C-terminal amide group protected with an Fmoc protecting group. ....	66
<b>Figure 2.4</b>	Reaction scheme for the removal of an Fmoc group from the peptide N $_{\alpha}$ using piperidine in DMF. ....	67
<b>Figure 2.5</b>	Reaction scheme for OBt activation. ....	68
<b>Figure 2.6</b>	Stick diagram of TZ2 is pictured with the three-letter amino acid code above. Colored circles represent the position of the three amide group isotope labels: TZ2-S1, with $^{18}\text{O}$ -labeled Ser1 (green); TZ2-TT, with $^{13}\text{C}$ -labeled Thr3 and Thr10 (blue); and TZ2-K8, with $^{13}\text{C}$ -labeled Lys8 (red). ....	73
<b>Figure 2.7</b>	Stick diagram of PG12 is pictured with the three-letter amino acid code above. Colored circles represent the position of the two amide group isotope labels: PG12-V3 with $^{13}\text{C}$ -labeled Val3 (blue); and PG12-V5 with $^{13}\text{C}$ -labeled Val5 (red). ....	76

<b>Figure 2.a.1</b>	TZ2-UL mass spectrum recorded on an Applied Biosystems Voyager MALDI-TOF mass spectrometer on February 3, 2004. The expected mass of TZ2-UL is 1608 g/mol, and the recorded mass is 1607.6 g/mol .....	80
<b>Figure 2.a.2</b>	TZ2-K8 mass spectrum recorded on an Applied Biosystems Voyager MALDI-TOF mass spectrometer on November 21, 2006. The expected mass of TZ2-K8 is 1609 g/mol. The recorded mass is 1608.3 g/mol, which is approximately 1 a.m.u. higher than the recorded mass of TZ2-UL (Fig. 2.a.1)...	81
<b>Figure 2.a.3</b>	TZ2-S1 mass spectrum recorded on an Applied Biosystems Voyager MALDI-TOF mass spectrometer on February 22, 2007. The expected mass of TZ2-S1 is 1610 g/mol. The recorded mass is 1609.6 g/mol, which is approximately 2 a.m.u. higher than the recorded mass of TZ2-UL (Fig. 2.a.1)...	82
<b>Figure 2.a.4</b>	TZ2-TT mass spectrum provided by Anaspec Inc. The expected mass of TZ2-TT is 1610 g/mol, which agrees with the recorded mass of 1610.7.....	83
<b>Figure 2.a.5</b>	Mass spectrum of <sup>18</sup> O labeled serine. Because there are two labile oxygen atoms in the carboxylic acid group, three peaks appear in the mass spectrum. The peaks correspond to the <sup>16</sup> O, <sup>16</sup> O isotope (326.1 a.mu.), the <sup>16</sup> O, <sup>18</sup> O isotope (328.0 a.m.u), and the <sup>18</sup> O, <sup>18</sup> O isotope (330.0 a.m.u).....	85
<b>Figure 3.1</b>	Diagram of interferometer. Collimated 6.0 μm light enters the setup from the left. Figure adapted from Ref. 3.....	90
<b>Figure 3.2</b>	Pulse arrangement and relative timings as they are directed through the sample. ....	91
<b>Figure 3.3</b>	2D IR and DVE spectra are shown for PG12-UL. The first two panels are the real (absorptive) and imaginary (dispersive) 2D IR spectra of PG12-UL at 25	

°C, and the third panel shows the absolute value of the sum of the real and imaginary 2D IR spectra. Panel four plots the directly-detected DVE spectrum at 25 °C and the DVE spectrum reproduced from the 2D IR. .... 96

**Figure 3.4** Diagram of T-jump laser pulse beam path. The magenta line traces the path of the 1.98 μm T-jump laser. Labels refer to: (L1) 2", UV fused silica lens, 100 mm fl, AR2010; (L2) 2", UV fused silica lens, 300 mm fl, no AR; (L3) 2", UV fused silica lens, 75 mm fl, AR2010; (MR) 1", BK7 / protected gold mirror, 75 mm radius; (S) sample cell. .... 97

**Figure 3.5** Enlarged illustration of the sample cell region. The spot size of the T-jump laser is 850 μm in diameter, which is larger than the spot size of the probe pulses (100 μm in diam.). The third order signal is generated by the interaction of the three femtosecond pulses a, b, and c, and emitted to the wave vector matched direction  $\mathbf{k}_s = -\mathbf{k}_a + \mathbf{k}_b + \mathbf{k}_c$ . Three incoming pulses are blocked by the mask to reduce the scattering. The local oscillator pulse passes through the sample at the same spot as the other three pulses. (Figure adapted from Ref. 4) ..... 99

**Figure 3.6** FTIR spectra of D<sub>2</sub>O solvent as a function of temperature from 5 °C (blue) to 95 °C (red) at 5 °C increments in a 50 μm pathlength cell. An air background was used as the reference spectrum to calculate absorbance. Solvent was prepared with a pH of 2.5 as described in Chapter 7. The green bar shows the spectral region of primary interest in this thesis. .... 101

**Figure 3.7** Change in absorbance from FTIR of D<sub>2</sub>O solvent as a function of temperature from 25 °C (blue) to 95 °C (red) at 5 °C increments. Each spectrum shown on left is obtained by subtracting off the spectrum at T<sub>0</sub> = 25 °C as a



reference. An air background was used as the reference spectrum to calculate absorbance. Solvent was prepared with a pH of 2.5 as described in Chapter 7. 102

**Figure 3.8** Calibrated changes to the solvent temperature during a transient, T-jump event. Left plots show the transient absorbance (top) and temperature (bottom) determined from the calibration procedure described above. The fitted parameters for each starting temperature are:  $B_1 = 1.36e^{-5}$ ,  $B_2 = -2.27e^{-3}$ , and  $B_3 = -1.28e^{-4}$  at 25 °C;  $B_1 = 1.39e^{-5}$ ,  $B_2 = -2.02e^{-3}$ , and  $B_3 = 1.55e^{-3}$  at 35 °C; and  $B_1 = 1.59e^{-5}$ ,  $B_2 = -1.85e^{-3}$ , and  $B_3 = 7.32e^{-5}$  at 45 °C. The right plot shows the calibrated temperature on a log scale from 1 ns to 50 ms..... 103

**Figure 3.9** Three-dimensional rendering of sample cell used in the work shown in Chapters 4 & 5. Four screws held the faceplate to the housing, and three O-ring grooves sealed the sample cell from the surrounding environment..... 104

**Figure 3.10** Three-dimensional rendering of sample cell used in the work shown in Chapters 6, 7, and 8..... 105

**Figure 4.1** (a) Experimental amide I spectra of TZ2 (left). 2D IR contour lines are at  $\pm 80\%$  of the band maximum for this and all other 2D IR spectra in this paper. Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates prior to incorporating Gaussian disorder are plotted as lines with the intensities weighted by the relative strength of the transition moment,  $\mu^2$ . (b) Stick and ribbon diagrams. Ribbon diagrams made with MolMol software. Ellipses overlay the peptide stick diagram to highlight how the local peptide amide groups map on to the box plots below. The color scheme of

the ellipses is that same as the  $\nu_{\perp} = 1633 \text{ cm}^{-1}$  eigenstate. (c) Selected eigenstates from the spectral simulation..... 126

**Figure 4.2** (a) Experimental amide I spectra of TZ2-T3\*T10\* (left). Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates calculated and plotted as in Fig. 1. (b) Selected eigenstates from the spectral simulation. The boxes around the  $^{13}\text{C}'$  labeled Thr3 and Thr10 amide groups are in bold. Particularly notable is the splitting of the two  $\nu_{\perp}$  bands by  $23 \text{ cm}^{-1}$ . The isotope shift deduced for the labeled oscillators based on the site energy model is  $34 \text{ cm}^{-1}$ ..... 130

**Figure 4.3** (a) Experimental amide I spectra of PG12 (left). Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates calculated and plotted as in Fig. 1. (b) Stick and ribbon diagrams. (c) Selected eigenstates from the spectral simulation. .... 133

**Figure 4.4** (a) Experimental amide I spectra of Gramicidin S (left). Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates calculated and plotted as in Fig. 1. (b) Stick and ribbon diagrams. (c) Selected eigenstates from the spectral simulation. .... 136

**Figure 4.5** Site energy plot for TZ2 using several different models. Solid red trace shows the site energies as calculated from the heuristic model used to simulate the spectra in this report. The blue trace shows the site energies calculated by projecting the electric field onto each amide unit.<sup>28</sup> Green squares are estimates based on isotope labels in this report and cyan diamonds are from the earlier study.<sup>24</sup> Black circles are estimates based on structural similarities with isotope

labeled sites. For the remaining dashed black trace, site energies were assigned from the heuristic model. The bar on the left of the graph shows the width ( $2\sigma$ ) of the Gaussian disorder used in the spectral simulations..... 142

**Figure 5.1** FTIR, DVE, and 2D IR spectra of trpzip2 are shown at three representative temperatures: 25, 63, and 82 °C. FTIR and DVE spectra are offset for clarity. For 2D IR spectra, 20 equally spaced contour levels are drawn to  $\pm 60\%$  of peak intensity. .... 161

**Figure 5.2** Second component spectra,  $c_N^{(2)}(\omega)$ , calculated from the SVD analysis of the FTIR (A) and 2D IR (B) data. The dashed guidelines intersect the axes at 1629 and 1680  $\text{cm}^{-1}$  (C) Melting curves from each experiment are plotted as the amplitude of the second SVD component at each temperature. The solid lines are guides for the eye. For clarity, the 2D IR and DVE melting curves are offset by -0.6 and -1.2 respectively. .... 163

**Figure 6.1** (a) The FTIR absorbance spectrum is shown for each PG12 isotopologue. Red and blue highlights show the position of the  $\nu_P$  and  $\nu_{V3}$  peaks respectively. Each spectrum was recorded in a pH 3.8 deuterated acetate buffer. (b) The sequence of PG12 is highlighted with colored circles to show the position of the labelled amide groups. (c) Absorptive 2D IR spectrum is shown for PG12-V3V5 with contours at  $\pm 3, 6, 9, 12, 16, 24, 32, 45, 65,$  and  $85\%$ ). .... 174

**Figure 6.2** FTIR and absorptive 2D IR spectra of PG12-UL (left), PG12-V5 (left-middle), PG12-V3 (right-middle), and PG12-V3V5 (right). Contour levels are the same as in Fig. 6.1..... 175

**Figure 6.3** 2DIR and DVE spectra are shown for PG12-UL. The first two panels are the real and imaginary 2D IR spectra for PG12-UL at 25 °C, and the third panel shows the absolute value of the sum of the real and imaginary 2D IR spectra. Panel four plots the directly detected DVE spectrum at 25 °C and the DVE spectra reproduced from the 2D IR spectra at 25 and 35 °C. .... 177

**Figure 6.4** (a) Amide I FTIR spectra of PG12-V3V5 from 5 to 90 °C. The spectra are colored blue from 5 to 40 °C, black at 45 °C, and red from 50 to 95 °C. The frequency axis is the same for plots a and b. (b) DVE spectra of PG12-V3V5 from 5 to 85 °C. The coloring scheme is the same as the FTIR spectra except the red spectra which range from 50 to 85 °C. (c) Melting curves calculated using the FTIR and DVE spectra from 1540 to 1710  $\text{cm}^{-1}$ . The melting curves are plotted as the value of the 2nd component at each temperature. The DVE curve is offset from the FTIR data by -0.25 for visual clarity..... 178

**Figure 6.5** (Left panels) Spectra are shown for PG12 from 5 to 90 °C and PG12-V3V5 from 5 to 85 °C. Spectra are recorded as detailed in the section 6.3. (Right panels) Equilibrium melting curve constructed as the 2<sup>nd</sup> SVD component of the spectra shown to the right. .... 179

**Figure 6.6** Intensity of DVE peaks as a function of temperature. The traces plotted are the intensity of the signal at the following frequencies:  $\nu_P = 1558 \text{ cm}^{-1}$ ,  $\nu_{V3} = 1598 \text{ cm}^{-1}$ ,  $\nu_{\perp} = 1630 \text{ cm}^{-1}$ , and  $\nu_{\parallel} = 1674 \text{ cm}^{-1}$ . Each trace is normalized to the intensity of the peak at 15 °C..... 181

**Figure 6.7** Absorptive (real) 2D IR spectra are shown for PG12-V3V5 from 15 to 85 °C. Spectra shown are collected in the zzzz polarization to make a more direct

comparison to the DVE spectra which are also collected in the zzzz polarization configuration. Contours are placed at  $\pm 3, 6, 9, 12, 16, 24, 32, 45, 65,$  and  $85\%$  of the amide I peak maximum. .... 181

**Figure 6.8** Thermal change in line broadening for several peaks in the 2D IR spectra of PG12-V3V5 from 15 to 85 °C. Antidiagonal ( $\Gamma$ ) and diagonal ( $\sigma$ ) half width at half maximum of the peaks indicated by the legend are shown with linear fits with the following slopes:  $\Gamma\text{-}\nu_{\perp} = 0.040 \text{ cm}^{-1}/^{\circ}\text{C}$ ,  $\Gamma\text{-}\nu_{V3} = 0.033 \text{ cm}^{-1}/^{\circ}\text{C}$ , and  $\sigma\text{-}\nu_{V3} = 0.009 \text{ cm}^{-1}/^{\circ}\text{C}$ . .... 183

**Figure 6.9** Equilibrium DVE spectra of PG12 (a) and PG12-V3V5 (b) from 15 °C (blue) to 45 °C (red). The solid black plots are the 2<sup>nd</sup> SVD component spectra calculated over the same temperature range. As the temperature is raised, the intensity at  $\nu_{\perp}$  and  $\nu_{\parallel}$  increases, while the  $\nu_P$ ,  $\nu_{SC}$ , and  $\nu_{V3}$  intensity decreases.. 184

**Figure 6.10** T-jump DVE difference data for PG12-UL (left panel) and PG12-V3V5 (right panel). .... 185

**Figure 6.11** T-jump data for PG12-UL averaged over three frequency ranges and  $\nu_P$  from 1598 to 1607  $\text{cm}^{-1}$ . The fitted transient absorbance is scaled and plotted to compare the DVE data to the instantaneous solvent temperature. .... 186

**Figure 6.12** T-jump data for PG12-V3V5 averaged over four frequency ranges:  $\nu_{\perp}$  from 1635 to 1654  $\text{cm}^{-1}$ ,  $\nu_{\parallel}$  from 1663 to 1683  $\text{cm}^{-1}$ ,  $\nu_P$  from 1558 to 1571  $\text{cm}^{-1}$ , and  $\nu_{V3}$  from 1598 to 1607  $\text{cm}^{-1}$ . The fitted transient absorbance is scaled and plotted to compare the DVE data to the instantaneous solvent temperature. .... 186

**Figure 6.13** Transient T-jump DVE signal of PG12-UL and PG12-V3V5 at four frequency channels corresponding to the frequency ranges as in Figs. 6.11 and

6.12. Each data point is the average of over the following frequency ranges: 1558-1567  $\text{cm}^{-1}$  (red), 1598-1607  $\text{cm}^{-1}$  (green), 1639-1648  $\text{cm}^{-1}$  (violet) 1668-1678  $\text{cm}^{-1}$  (blue). The fitted transient absorbance of the solvent is scaled by 150 and plotted in black. Open circles correspond to measured data, and the lines are single exponential fits to the data. The  $t_1$  rise time of the fits to the  $\nu_{\perp}$ ,  $\nu_{\parallel}$ , and  $\nu_P$  data is less than 10 ns, and the fitted  $t_1$  decay time of the PG12-V3V5  $\nu_{V3}$  peak is  $135 \pm 50$  ns. The  $\nu_{\perp}$  and  $\nu_{\parallel}$  of PG12-V3V5 each have subsequent decay times of 610 and 680 ns respectively. The  $\nu_{\perp}$  peak of PG12-UL has a long decay time of 360 ns. 187

**Figure 7.1** Stick diagram of TZ2 is pictured with the three-letter amino acid code above. Colored circles represent the position of the three amide group isotope labels: TZ2-S1, with  $^{18}\text{O}$ -labeled Ser1 (green); TZ2-TT, with  $^{13}\text{C}$ -labeled Thr3 and Thr10 (blue); and TZ2-K8, with  $^{13}\text{C}$ -labeled Lys8 (red). ..... 195

**Figure 7.2** Equilibrium FTIR (a) and DVE (c) spectra are shown for each TZ2 isotopologue at 25 °C and pH 2.5. Each spectrum is baseline corrected with a linear subtraction, and area normalized for comparison. Difference data (b & d) is obtained by subtracting the unlabeled spectrum (TZ2-UL) from each of the spectra. .... 197

**Figure 7.3** Experimental equilibrium 2D IR spectra for each TZ2 isotopologue at 25°C as described in section 7.2. Each data set shown is taken in the ZZZZ polarization geometry, and contours are plotted at  $\pm 3, 6, 9, 12, 16, 24, 32, 45, 65,$  and 85% of the band maximum. .... 200

<b>Figure 7.4</b>	Structures of Markov states that were used for spectral simulations. Structures were provided by William Swope and have been published in Ref. 6.	205
<b>Figure 7.5</b>	Simulated 2D IR spectra of TZ2 isotopologues for the NMR structure and four Markov states. Methods and notation are described in text. For each spectrum, 27 contours are plotted from $\pm 75\%$ of the band maximum.....	206
<b>Figure 7.6</b>	Temperature dependent FTIR spectra of each TZ2 isotopologue taken from 5 to 95 °C in 5 °C increments (pH = 2.5). Spectrum line color shows progression from 5 °C (blue) to 95 °C (red). .....	211
<b>Figure 7.7</b>	SVD analysis of the temperature-dependent FTIR spectra shown in Fig. 7.6. The SVD component spectra are not scaled, and the melting curves are normalized to the 5 °C value.....	213
<b>Figure 7.8</b>	(Left) The frequency of the FTIR amide I maximum is plotted for each TZ2 system. The resolution in the vertical frequency dimension is $1 \text{ cm}^{-1}$ , which is defined by the FTIR instrument resolution during data collection. (Right) TZ2 FTIR peak intensities are plotted vs. temperature. The $^{12}\text{C}$ peak intensities are obtained by averaging the absorbance around the peak positions found in the left panel ( $\nu \pm \sim 5 \text{ cm}^{-1}$ ). The $\nu_{\text{K8-1}}$ trace corresponds to frequencies from 1588 to 1598 $\text{cm}^{-1}$ , the $\nu_{\text{K8-2}}$ trace is averaged from 1612 to 1619 $\text{cm}^{-1}$ , and the $\nu_{\text{TT}}$ trace corresponds to frequencies from 1603-1611 $\text{cm}^{-1}$ . Each trace is normalized to the peak intensity at 5°C.....	214

- Figure 7.9** 2D IR spectra of each TZ2 isotopologue are shown at 15 °C (top row) and 75 °C (bottom row). Black lines show the anti-diagonal slices analyzed in Fig. 7.10. 217
- Figure 7.10** For each TZ2 isotopologue, the width of the anti-diagonal slice through the positive diagonal peaks is plotted as a function of temperature. For TZ2-TT (bottom right), the diagonal width of the  $\nu_{TT}$  peak is also shown. Each data point represents the  $\Gamma/2$  value of a Lorentzian function (Equation 7.1) fit to the (anti-) diagonal slices. Solid lines represent linear fits to the HWHM vs. temperature, with the slopes reported in Table 7.1. .... 218
- Figure 7.11** Equilibrium thermal denaturation DVE spectra for each TZ2 isotopologue from 10 (blue) to 85 °C (red) in 5 °C increments. .... 224
- Figure 7.12** SVD analysis of the temperature-dependent DVE spectra shown in Fig. 7.11. The SVD component spectra are not scaled, while the melting curves are normalized to the 10 °C value. .... 226
- Figure 7.13** DVE peak intensities are plotted versus temperature. The data points are obtained by averaging the intensity around each peak position ( $\nu \pm \sim 5 \text{ cm}^{-1}$ ). The  $\nu_{K8-1}$  trace corresponds to the low frequency peak (1588-1598  $\text{cm}^{-1}$ ) and the  $\nu_{K8-2}$  peak is the high frequency  $^{13}\text{C}$ -shifted peak (1612-1619  $\text{cm}^{-1}$ ). The  $\nu_{TT}$  trace is obtained by averaging the frequency range from 1603-1611  $\text{cm}^{-1}$  in the TZT-TT DVE spectra. Each trace is normalized to the peak intensity at 5 °C. .... 227
- Figure 7.14** Zoom in of temperature-dependent 2D IR spectra of TZ2-K8 shown for the region highlighted in the left plot. Contours lines in the right three spectra are



plotted from 20% to -20% of the amide I maximum at the respective temperature.

229

**Figure 7.15** Zoom in of temperature-dependent 2D IR spectra of TZ2-TT shown for the region highlighted in the left plot. Contours lines in the right three spectra are plotted from 40% to -40% of the amide I maximum at the respective temperature.

232

**Figure 7.a.1** Simulated FTIR spectra of unlabeled TZ2-UL for the NMR structure and four Markov states. Methods and notation are described in Section 7.4. ... 237

**Figure 7.a.2** (Top) Simulated FTIR spectra of TZ2-S1 for the NMR structure and E-D. (Bottom) Difference spectra obtained by subtracting the simulated TZ2-UL spectra (Fig. 7.a.1.) from the TZ2-S1 spectra in the top panel. .... 238

**Figure 7.a.3** Equilibrium thermal denaturation of DVE spectra reconstructed from 2D IR data taken from 15 to 85 °C in 10 °C increments. .... 239

**Figure 7.a.4** SVD analysis of the temperature dependent (reconstructed) DVE spectra shown in Figure 7.a.3. The SVD component spectra and are not scaled, while the  $n^{\text{th}}$  component melting curves are normalized to the 15 °C value. .... 239

**Figure 8.1** (Previous page) (Left column) Transient difference DVE spectra of each TZ2 isotopologue is shown from 1 ns (black) to 3  $\mu$ s (green). The calibrated T-jump,  $\Delta T$ , for TZ2-UL and TZ2-TT is 18.5 °C, and for TZ2-K8 and TZ2-S1 is 18.2 °C. The equilibrium temperature,  $T_0$ , for each data set is 45 °C. Difference spectra are obtained as described in Section 8.2, with the  $\Delta$ Signal defined as the difference DVE spectrum normalized by the equilibrium band maximum. The top trace is the equilibrated DVE spectrum obtained by averaging the 48-50<sup>th</sup> pulses

following the T-jump. (Right column) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles, and solid lines represent single-exponential fits to the data from 12 ns to 12  $\mu$ s. Fitted parameters are displayed in Table 8.1. The black circles represent a sample transient absorbance trace and the solid black line is an exponential fit to those data, which are scaled to show the temporal profile of the solvent temperature..... 247

**Figure 8.2** (Left) Transient difference DVE spectra of TZ2-UL shown from 1 ns to 3  $\mu$ s at 25, 35, and 45  $^{\circ}$ C. Difference spectra are obtained as described in Section 8.2, with the  $\Delta$ Signal defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu$ s. Time constants from the transient fits are shown in Table 8.2. The intensity fluctuations at 1700  $\text{cm}^{-1}$  are from anomalous electronic noise. .... 251

**Figure 8.3** Summary of time constants for TZ2 T-jump data at  $T_0 = 25, 35,$  and 45  $^{\circ}$ C (40  $^{\circ}$ C for TZ2-S1). Each set of data points is the average of several channels similar to Fig 8.1. The solid lines represent the fit of the kinetics to the Arrhenius equation. The activation energies are shown in Table 8.3..... 253

**Figure 8.4** T-jump 2D IR data for TZ2-UL at  $T_0 = 45$   $^{\circ}$ C, and  $\Delta T = 20.8$   $^{\circ}$ C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference

signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra. .... 256

**Figure 8.5** T-jump 2D IR data for TZ2-UL at  $T_0 = 10$  °C, and  $\Delta T = 17.3$  °C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra. .... 257

**Figure 8.6** (Top) Line width of the  $\nu_{\perp}$  and  $\nu_{\parallel}$  bands of TZ2-UL as a function of T-jump delay time. Each data point represents the  $\Gamma/2$  value of a Lorentzian function (Equation 7.1) fit to the anti-diagonal slices. Error bars are the standard deviation of the fitted widths for six consecutive anti-diagonal slices. .... 258

**Figure 8.7** T-jump 2D IR data for TZ2-K8 at  $T_0 = 45$  °C, and  $\Delta T = 20.8$  °C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra. .... 260

**Figure 8.8** T-jump 2D IR data for TZ2-S1 at  $T_0 = 45$  °C, and  $\Delta T = 24.7$  °C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra. .... 261

**Figure 8.9** T-jump 2D IR data for TZ2-TT at  $T_0 = 45$  °C, and  $\Delta T = 19.8$  °C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and

50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra. .... 262

**Figure 8.10** (Top panel) The  $\omega_1$  shift of peak A maximum is shown as a function of delay time from 10 ns to 10  $\mu$ s. The  $\Delta\omega_1$  value is obtained by subtracting the peak max position at 10 ns from each of the following time points. (Bottom panel) Intensity of the peak A maximum relative to the equilibrated 2D IR amide I band maximum as a function of T-jump delay time. .... 264

**Figure 8.11** Intensity of the peak C maximum relative to the equilibrated 2D IR amide I band maximum as a function of T-jump delay time from 10 ns to 10  $\mu$ s. .... 265

**Figure 8.12** (Top) Line width of the  $\nu_{\perp}$  band of each TZ2 isotopologue as a function of T-jump delay time. Each data point represents the  $\Gamma/2$  value of a Lorentzian function (Equation 7.1) fit to the anti-diagonal slices. The  $\Delta$ width is obtained by subtracting the width of the equilibrated spectrum at each time point. Error bars are the standard deviation of the fitted widths for six consecutive anti-diagonal slices. 267

**Figure 8.13** (Left) Intensity of peak D as a function of delay time (relative to equilibrated 2D IR maximum), which corresponds to the intensity change of the  $\nu_{K8-2}$  band. (Right) The  $\nu_{K8}$  band region of the T-jump difference 2D IR spectrum at  $t = 10 \mu$ s is plotted with 27 contours from  $\pm 65\%$  of the spectral region maximum. .... 269

**Figure 8.14** (Left) Intensity of peak D as a function of delay time (relative to equilibrated 2D IR maximum), which corresponds to the intensity change of the

$\nu_{TT}$  band. (Right) The  $\nu_{TT}$  band region of the T-jump difference 2D IR spectrum at  $t = 10 \mu\text{s}$  is plotted with 27 contours from  $\pm 65\%$  of the spectral region maximum. .... 270

**Figure 8.a.1** (Left) Transient difference DVE spectra of TZ2-K8 shown from 1 ns to 3  $\mu\text{s}$  at 25, 35, and 45 °C. Difference spectra are obtained as described in Section 8.2, with the  $\Delta\text{Signal}$  defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu\text{s}$ . Time constants from the transient fits are shown in Table 8.2. .... 279

**Figure 8.a.2** (Left) Transient difference DVE spectra of TZ2-UL shown from 1 ns to 3  $\mu\text{s}$  at 25, 35, and 45 °C. Difference spectra are obtained as described in Section 8.2, with the  $\Delta\text{Signal}$  defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu\text{s}$ . Time constants from the transient fits are shown in Table 8.2. .... 280

**Figure 8.a.3** (Left) Transient difference DVE spectra of TZ2-UL shown from 1 ns to 3  $\mu\text{s}$  at 25, 35, and 45 °C. Difference spectra are obtained as described in Section 8.2, with the  $\Delta\text{Signal}$  defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line

are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu$ s. Time constants from the transient fits are shown in Table 8.2..... 281

## List of Tables

<b>Table 2.1</b>	Side-chain protecting groups for amino acids used in peptide synthesis. The table includes side-chain protecting groups used to synthesize peptides for the current work.....	70
<b>Table 7.1</b>	The rate of line broadening with temperature is shown for diagonal peaks in the TZ2 2D IR spectra. Values shown represent the slope of the line fit to the 2D IR peak width versus temperature shown in Fig. 7.10.....	221
<b>Table 8.1</b>	Results of single exponential fits to data shown in Figs. 8.1. Error values for $\tau_1$ represent the error in fit to Equation 8.2. *All values are reported in $\mu$ s. ....	249
<b>Table 8.2</b>	Summary of single exponential fits to the $T_0$ -dependant T-jump DVE data. $\Delta T$ is the calibrated rise in temperature. *All time scales are shown in $\mu$ s. ....	252
<b>Table 8.3</b>	Summary of activation energies obtained by fitting the temperature-dependant rates to the linear form of the Arrhenius equation: $\ln(k) = -E_a/(RT) + \ln(A)$ . *Values are in units of kJ/mol.....	254

# Chapter 1

## Introduction & Background

### 1.1. Current Issues in Protein Folding

A protein's structure and dynamics define the way that it participates in biological function. Predicting these properties is an enormous challenge, and is generally referred to as the protein folding problem. To date, no global solution has been developed to describe and predict the folding transition, which is in part limited by the large number of degrees of freedom in biological polymers. For an average sized protein of 300 residues and approximately 17 atoms per amino acid, there are over 5000 molecular degrees of freedom. Describing those degrees of freedom during a folding event requires high temporal resolution over long observation times. To solve this problem, biologists, chemists, and computer scientists have pushed the limits of synthesis, spectroscopy, and computational power to work toward an accurate model for this complex structural transition.

Protein folding research has answered some fundamental questions. For example, the Levinthal Paradox once asked how a protein could fold so quickly despite the fact

that a random walk through configuration space would take an astronomically long time.<sup>1</sup> To resolve this, simple models show that an energy bias is sufficient to allow folding on a biologically relevant time scale.<sup>2,3</sup> Statistical mechanical models further demonstrate that the protein is not moving on a flat energy surface, but is instead exploring a multi-dimension energy landscape with a bias toward the native state. This picture is often described as a funnel landscape, where enthalpic and entropic forces combine to guide the protein efficiently to the native state.<sup>3-6</sup>

Researchers have also made considerable progress toward solving the structural prediction problem, especially for single domain proteins of average size. One approach has been to use the expanding library of known protein structures to statistically relate protein structure to amino acid sequence. In this way, the structure of a protein can be predicted by a rigorous comparison to other protein structures. These prediction tools are fast and have been effective in both accurately predicting unknown structures and designing *de novo* proteins with pre-determined folds.<sup>7-9</sup> Physics based models have lagged behind bioinformatics tools because of the computational challenge alluded to earlier. However, state-of-the-art simulation methods are now effective at predicting protein structure.<sup>7,10-13</sup> Solving the structural prediction problem is intertwined with the description of the folding transition, but it does not necessarily lead to direct mechanistic insight. Structure prediction tools must be linked with dynamical observations in order to elucidate protein folding mechanisms.

Progress in describing protein structure and dynamics has led to statements implying that the protein folding problem is no longer limited by fundamental barriers.<sup>14</sup> The appropriateness of such optimism is debatable, but it is clear that there are a number



of unanswered questions. For example, it is widely accepted that the folding pathway is heterogeneous in the sense that within an ensemble of disordered proteins, each molecule will fold via a distinct set of conformational changes to reach the native state.<sup>3-6</sup> However, it is not clear, especially from experiments, how the protein navigates barriers and traps within the energy landscape. What are the precise structural changes that favor folding and how do we assess their relative contributions to the energy of the system?

One way to describe the free energy of the folding transition is as a sum of contributions assigned to specific structural features.<sup>15,16</sup> Three major contributions are the hydrophobic effect, hydrogen bonding, and configurational entropy.<sup>17</sup> The hydrophobic effect is the tendency non-polar side chains to aggregate within the interior of a protein to reduce the surface area exposed to water. It is estimated that for globular proteins, the hydrophobic effect can typically stabilize the folded state by 500 kJ/mol (at 25 °C).<sup>17</sup> Hydrogen bonding to or from the amide oxygen and amide proton is another stabilizing force that plays an important role in the formation of  $\alpha$ -helices and  $\beta$ -sheets. It is difficult to assess the energetic contribution of individual hydrogen bonds (H-bonds) to the stability and folding of a protein because they each depend on the charge and orientation of the bonded units, and because hydrogen bonding to water competes with the intra-protein H-bonds. Configurational entropy describes the role of the many degrees of freedom available to a poly-peptide chain. It generally contributes -800 kcal/mol (at 25 °C) to destabilize the protein.<sup>17</sup> There are several other important interactions, including disulfide bridges, electrostatic forces between charged side chains, and the role of side chains in secondary structure propensities. Understanding how each of these effects contributes to the folding free energy is essential to describe folding

pathway. However, such questions are still a matter of debate and continuing research is needed to build a satisfying and complete picture of protein folding.<sup>14,18</sup>

A number of folding paradigms have been developed, which highlight specific events that drive the folding transition. Two early views are the hydrophobic collapse model and the framework model. The framework model, also called the kinetic zipper model, posits that the rate-limiting step is the nucleation of local secondary structural elements such as  $\alpha$ -helices and  $\beta$ -hairpins.<sup>19-21</sup> After the nucleation of local structure, larger secondary structural domains form. At the final stage they associate together through tertiary contacts into helical bundles,  $\beta$ -sheets, and more complicated mixed-structure domains. This mechanism is typically described as a two-state transition from the unfolded to the native state. In the hydrophobic collapse model, polar side chain groups first collapse to form a molten globule-like state, after which backbone rearrangements guide the protein to its native conformation.<sup>22,23</sup> Hydrophobic collapse is often modeled using three-state kinetics, with the molten globule state representing an on-pathway intermediate. Both the kinetic zipper and hydrophobic collapse models are often considered to be overly simplistic. Local secondary structure can be unstable without significant contribution from tertiary contacts. Hydrophobic contacts are non-specific and can either initiate non-native secondary structure, or else prevent reorganization of the protein. While it is now more common to describe local secondary structure and tertiary contacts working in concert,<sup>18,22,24</sup> early models continue to serve as paradigms around which modern views are built.

One description of the folding pathway that attempts to improve on the two early paradigms is the nucleation-condensation model.<sup>18,22,24</sup> Nucleation-condensation unites

the two early views in the sense that both local secondary structure and tertiary contacts are present in the transition state. The model was developed to deal with combined observations of high  $\phi$ -values for secondary and tertiary contacts as well as the absence of observable intermediates in some proteins.<sup>18</sup> Similar to the nucleation-condensation model is the zipping and assembly model, which posits that local structural coordinates find their native configurations independently of each other.<sup>25-27</sup> As native structure accumulates, tertiary contacts form and the protein structure zips together. The zipping and assembly mechanism relies on tertiary contacts, which distinguishes it from the framework and kinetic zipper models while maintaining a focus on local secondary structure. The paradigm of zipping and assembly has been used to speed up folding in molecular dynamics simulations.<sup>18</sup>

The key to sorting out these mechanisms and discriminating between the various folding models is to observe the microscopic variables that are most relevant to the structural transition. This includes the formation of local secondary structure such as helices, turns, and loops, as well as tertiary contacts such as intra-strand hydrogen bonding, hydrophobic side chain contacts, disulfide bridges, and salt bridges. Observing these events is limited by the fact that most spectroscopic techniques are performed on large ensembles of states, and thus individual folding trajectories cannot be distinguished. Single molecule studies promise to change this, but such methods are still limited by a lack of probes that can elucidate local structure.<sup>28,29</sup> The second major barrier is that traditional spectroscopic tools lack the ability to resolve sub-molecular structural elements over the broad range of relevant time scales.

One approach used to simplify the problem and make it more amenable to present methods is to study small peptides that form single units of secondary structure.<sup>30</sup> Two common classes of peptides for this approach are  $\alpha$ -helices and  $\beta$ -hairpins, which are reduced models of  $\alpha$ -proteins and  $\beta$ -sheet structures. Early work established a simple and robust model for  $\alpha$ -helices.<sup>31-33</sup> The folding pathway is best described within the framework model as a kinetic zipper in which the folding is initiated by the nucleation of several residues to form a turn after which the  $\alpha$ -structure propagates rapidly along the chain. In this model, the driving force is the formation of the backbone dihedral angles followed by the formation of H-bonds between residues  $i$  and  $i+4$ . For single domain helices this model fits the experimental observations with robust accuracy, and molecular dynamics simulations have further validated this mechanistic description.<sup>34</sup>

Hairpin-forming peptides were historically more difficult to study because of a propensity to aggregate at concentrations necessary for experiments. To this end several robust hairpin systems have been developed and subsequently investigated using a battery of experimental and computational methods.<sup>35-40</sup> The GB1<sub>41-56</sub> is a hairpin fragment clipped from the C-terminus of the B1 domain of protein G. It was one of the first hairpin peptides that was stable in solution in the absence of its parent protein. It has been subject to a large number of theoretical studies that will be reviewed below, but experimental studies have been limited to NMR and fluorescence because of its propensity to aggregate.<sup>36,41</sup> When the first GB1<sub>42-56</sub> experiments were reported in 1997, there were very few hairpin peptides that were stable and monomeric in water solutions. Since then a number hairpin scaffolds have been developed including a series of peptides with a D-proline enhanced type II'  $\beta$ -turn.<sup>38,42</sup> These peptides were designed by Gellman

and coworkers, and since have been studied with optical and NMR spectroscopy to verify the anti-parallel  $\beta$ -sheet structure and quantify their stability.<sup>43,44</sup> In 2001 Andrea Cochran et al, developed the trpzip peptides, which are now among the most studied de novo hairpins.<sup>39</sup> Trpzip peptides are engineered with a novel structural motif of four tryptophan residues arranged with two side chains on each of the opposing  $\beta$ -strands. The intra-strand tryptophan side chains are separated by a threonine residue so that all four tryptophan side chains are on the same face of the  $\beta$ -sheet plane. A favorable hydrophobic attraction between the side chains is enhanced by perpendicular stacking of the indole rings, giving rise to a remarkably stable system. The trpzip scaffold has allowed for numerous studies of effects such as  $\beta$ -strand length and turn sequences on the stability and folding mechanism of  $\beta$ -hairpins.

A surprising result of the research on these  $\beta$ -hairpin systems is that it has been very difficult to establish a consensus view of the folding mechanism. The zipper model used to successfully describe  $\alpha$ -helices was first applied to hairpin folding in 1997,<sup>36</sup> but it was called in to question by subsequent molecular dynamics simulations that observed a hydrophobic collapse mechanism.<sup>45</sup> This complication in describing the folding pathway is precisely what makes hairpin peptides a perfect scaffold for sorting out issues in protein folding. To what degree does the formation of local secondary structure such as  $\beta$ -turns and  $\beta$ -sheet dihedral angles dictate the folding pathway? How strong a driving force is the formation of cross strand tertiary contacts such as hydrophobic interactions and salt bridges? Research on hairpin peptides over the last ten years has lead to a much more detailed understanding of the role of the  $\beta$ -turn in  $\beta$ -sheet formation, and in the role of side chains in forming a stabilizing hydrophobic core. For the remainder of this

chapter theoretical and experimental studies of  $\beta$ -hairpin peptides will be reviewed in order to highlight progress and define the goals of the research presented in the rest of the thesis.

## 1.2. Hairpin Folding Simulations

The ideal description of a folding transition should track the spatial coordinates of each atom in the molecule at time intervals that are small relative to the folding event. To date, however, this has only been possible in computer simulations. Atomistic molecular dynamics simulations are currently the most direct way of accessing atomic-level information at femtosecond time steps. In this section a number of theoretical  $\beta$ -hairpin folding studies will be reviewed and discussed.

The first hairpin folding study was reported by Eaton and co-workers on the transient thermal unfolding of GB1<sub>42-56</sub>.<sup>36</sup> The work combined theory, in the form of statistical mechanical modeling, with experimental kinetics measurements.<sup>46</sup> Time-resolved fluorescence was used to measure tryptophan fluorescence decay as the peptide unfolds following a nanosecond temperature-jump (T-jump). The observed kinetics fit well to a two-state model with a 6  $\mu$ s folding time, which is about ten times slower than observed  $\alpha$ -helix folding kinetics.<sup>30</sup> Statistical mechanical modeling described the hairpin folding mechanism as a kinetic zipper, in which the native state is reached through the initial formation of the  $\beta$ -turn followed by a zipping together of the backbone contacts. As in the protein folding framework model, the primary driving force is provided by intra-strand H-bonds and local secondary structural propensities. Hydrophobic side chain contacts were not explicitly accounted for this model.

The zipper mechanism was initially well-received as a description of hairpin folding because it agreed with the kinetic data and produced a plausible folding mechanism. For example, it explained the long time-scale compared to  $\alpha$ -helix folding because if the  $\beta$ -turn can only form in one part of the peptide, it will take longer to find that one state through conformational searching. For  $\alpha$ -helices, the turn nucleation site can occur at any part of the peptide and requires less large-scale structural rearrangements. The weakness of the model is that it does not account for the contribution of the side chains to the folding pathway.

Molecular dynamics simulations of GB1<sub>41-56</sub> by Dinner et al. produced the first evidence that the collapse of the hydrophobic side chains might in fact be the correct barrier crossing event.<sup>45</sup> This leads to a rather different folding mechanism in which the hydrophobic collapse brings the peptide into a configuration from which the native contacts are made quickly through random fluctuations. This model sparked a debate that has continued in the hairpin folding literature, where the question is how to accurately determine the relative importance and time ordering of the native backbone contacts, H-bonds, and side chain packing. This mirrors the debate that continues in the literature regarding the folding mechanism of large proteins. For example, the zipper model discussed above is very similar to the framework model where parts of the protein spontaneously form regions of stable secondary structure, which propagates along the chain until the native state is reached. The hydrophobic collapse model is also mirrored in the protein folding literature, wherein the first event along the protein folding pathway is proposed to be the collapse into a molten globule state that buries the hydrophobic side chains. By sorting through these competing mechanisms, it is clear that a careful and

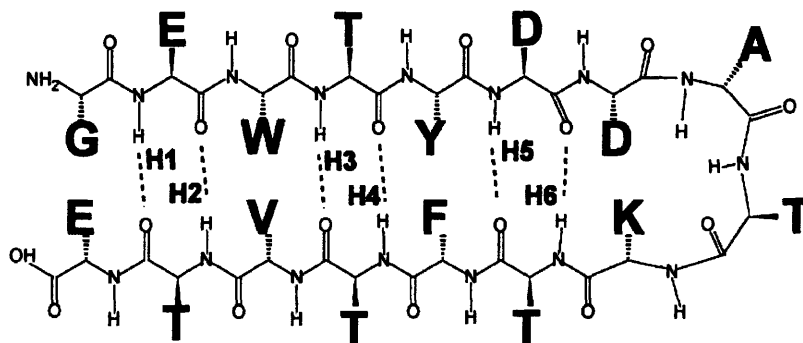
accurate description of hairpin folding can provide valuable information to help solve the protein folding problem.

In a subsequent study, Tsai and Levitt used explicit solvent molecular dynamics simulations of GB1<sub>41-56</sub> to evaluate the role of the salt bridge, by simulating trajectories with and without the final G41 residue.<sup>40</sup> By doing this they built a more complete picture of the forces influencing the folding pathway. Their results suggest that the salt bridges in the GB1<sub>41-56</sub> hairpin contribute approximately 1.3 kcal/mol to the final structure, but that the bridging does not seem to guide the folding trajectory. In their folding model, the turn forms first, which is stabilized by the side chain salt bridge. Then the non-polar side chains collapse to form the hydrophobic core, after which the cross-strand H-bonds form. Levitt calls the model a hybrid zipper, where the hydrophobic side chains are part of the zipper. The ‘turn first’ mechanism is rationalized because without it the unguided hydrophobic collapse could easily lead to a non-native turn. A similar hybrid model was seen in simulations by Zhou et al. in which there was a high degree of cooperativity between the formation of the  $\beta$ -turn and the hydrophobic core.<sup>47</sup>

Work by Wei et al employed an activation-relaxation technique to produce a large number of simulated folding trajectories for the GB1<sub>41-56</sub> peptide.<sup>48</sup> Within the set of pathways, three general mechanisms were found. Mechanism I begins with the partial collapse of the hydrophobic core with the turn forming in the correct place in the chain. The radius of gyration ( $R_g$ ) is near native and several non-native H-bonds stabilize the structure. Further in the trajectory, two native H-bonds near the turn (H5 and H4, Fig. 1.1) form. In the next step the core reorganizes to a well packed state very near the native configuration. The second and third H-bonds from the termini (H2 and H3, Fig.



1.1) form as well as the H-bond nearest the turn (H6). H6 forms and breaks until the end of the run, and never stabilizes completely. The last step is the formation of the H-bond between the two terminal groups (H1).



**Figure 1.1** Stick diagram of GB1<sub>41-56</sub>. Dashed lines represent H-bonds observed in simulations described in the text.

Mechanism II involves an initial step of the formation of a partial hydrophobic core, and the joining of the N and C termini to form a large loop. The side chain contacts break and reform to reach a more compact state. H1 forms, followed by H2, and then H3. The hydrophobic core then reaches its native packing, and H-bonds H4-H6 form. In several of these pathways, helical secondary structure was observed, which is consistent with a previous study.<sup>49</sup> Helical structure in this set of pathways was short lived, and was only seen about half of the trajectories. Mechanism I and II are not necessarily incompatible. The order of H-bond formation can actually start in the middle (H4), and propagate outwards. A loop formation mechanism is also observed in work done by Zagrovic et al.<sup>50</sup>

The third mechanism is new and distinct to this paper. It involves significant participation of non-native structure. Again, the first step is a rapid collapse of the

hydrophobic side chains, but in this case the turn forms in the wrong place along the chain. Strong non-native H-bonds form across the strand, and then slowly creep up the chain through repetitive making and breaking of non-native H-bonds. This reptation mechanism slowly corrects the asymmetry of the hairpin. At some point the reptation slows to allow the hydrophobic core to rearrange and then four non-native H-bonds break and H1-H4 are formed simultaneously. Once this is done, the native state is reached quickly as H5 and H6 form. This reptation mechanism was challenged recently in a simulation study by Pitera et al, in which they analyzed 300, 10 ns simulations of the trpzip2 (TZ2) peptide.<sup>51</sup> In observing numerous folding and unfolding events, they did not see any signs of reptation-like events. Furthermore they argued that reptation is a mechanism that is only observed in densely packed polymeric systems, and should not play a role in monomeric peptide folding transitions.<sup>51</sup>

In a pair of recent papers, Bolhuis has applied transition path sampling to describe the pathways and transition state ensembles of four meta-stable configurations identified in the GB1<sub>41-56</sub> folding trajectory.<sup>52,53</sup> The four states, in order along the unfolding pathway are: the native (N) state, which matches the NMR structure; the frayed (F) state in which the hydrophobic core and the backbone remain in their native configuration, but the residues near the turn and N and C termini are disordered and their cross-strand H-bonds are broken; the hydrophobic (H) state, where the H-bonds are broken, the turn region is disordered, and the hydrophobic core is intact; and the unfolded (U) state which is a completely disordered (random coil) peptide. The transition from the F to the H state occurs as the H-bonds between the threonine residues break in unison and are immediately (10 fs) solvated, concurrently the hydrophobic core undergoes minor

rearrangement. The transition from the H to the U state is characterized by the complete dissolution of the hydrophobic core. These simulations rely intrinsically on the hydrophobic collapse model. Although the subtlety is in the existence of a quasi stable intermediate state (F), which does require the formation of H-bonds to completely pack the hydrophobic core into its native configuration.<sup>52,53</sup>

The mechanisms observed in each of the computer simulations involve a varying degree of cooperation between the  $\beta$ -turn, intra-strand H-bonds, and tertiary side chain contacts. Many more simulations have been done on the GB1<sub>41-56</sub> peptide and TZ2, but generally they can be grouped into the three following scenarios. The first is a hybrid zipper model, where the  $\beta$ -turn forms first, stabilized (in some systems) by a side chain salt bridge. Then the non-polar side chains collapse to form the hydrophobic core, after which the cross-strand H-bonds form. The second model involves the partial collapse of the hydrophobic core with the turn forming in the correct place in the chain. Next, the two cross strand H-bonds near the turn form, after which the core reorganizes to a well packed state very near the native configuration. The third mechanism begins with the formation of a partial hydrophobic core and the joining of the N and C termini to form a loop. The side chain contacts break and reform to reach a more compact state. The terminal H-bonds form next, after which the hydrophobic core reaches its native packing and the H-bonds form near the turn region. One point that further complicates the picture is that within each reported simulation study, several mechanisms were often observed. This suggests that the free energy surface is very rough, and that the folding pathway is highly heterogeneous. Recently, this was explored explicitly for the TZ2 peptide using

both simulations and experiments.<sup>54,55</sup> The experiments will be reviewed below, but the simulations have found that the free-energy surface is in fact rough compared to  $kT$ .<sup>54,55</sup>

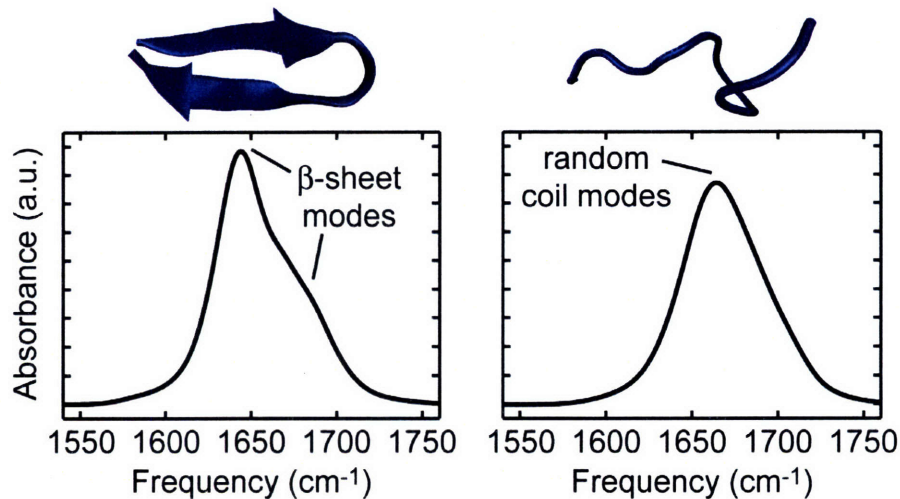
### 1.3. Experimental Folding Kinetics

The folding simulations reviewed above have been an essential part of understanding  $\beta$ -hairpin folding. However, they have not built a consensus view of the general features of the folding mechanism. For example, from simulations alone it is not clear if there is a preferential ordering to the formation of native contacts. It is also unclear to what degree each of those native contacts contributes to the downhill slope of the energy surface, and to what degree they modulate the size of the barriers and bottlenecks. This section will review the experimental approaches that have been taken to address these issues.

For the purpose of answering questions about the microscopic hairpin folding mechanism it is essential to understand precisely what structural information can be obtained from the current experiments. Three types of experiments have been used to study transient  $\beta$ -hairpin folding. Two experiments rely on fast initiation of conformational change followed by a probe of some structural coordinate. The third is an NMR line-broadening study in which time scales as fast as 5  $\mu$ s were extracted from differential exchange broadening.<sup>41</sup> This extraction requires that the chemical shift of the peak in the native state is well separated from that of the denatured state, which inherently requires a substantial population of both the native and denatured states at each temperature.<sup>41,56</sup> For GB1<sub>41-56</sub> and mutants, this condition was met and folding rates were obtained from the Tyr-5-H $\delta$  and Val-14-H $\gamma_{up}$  peaks. The kinetics for each peak were

similar enough for the authors to assume two-state kinetics, allowing them to interpret the Try-5H $\delta$  peak as a probe of native vs. denatured structure. The analysis, therefore, relies on mutation studies to investigate the role of local structure, which is discussed below.

The two other sets of kinetics studies are T-jump initiated kinetics experiments combined with (1) infrared (IR) absorbance and (2) UV fluorescence probes. To date, T-jump fluorescence studies of hairpins have exclusively probed the indole group fluorescence of the tryptophan side chain.<sup>36,54</sup> Tryptophan fluorescence is particularly good at distinguishing side chains exposed to solvent from those packed together in a solvent excluded environment. In this way, the kinetic rates observed in the experiments above are effectively measuring structural rearrangements of the hydrophobic core. The second T-jump probe is amide I IR absorbance. Historically much emphasis has been put on the amide I band because of its long-established relationship to secondary structure.<sup>57,58</sup> In poly-peptides, it is modeled as a set of backbone carbonyl stretches that are weakly coupled via kinetic (through-bond) and electrostatic (through-space) mechanisms. The coupling leads to patterns in the amide I band that report on the presence of  $\beta$ -sheets and  $\alpha$ -helices. For  $\beta$ -hairpins, there are two peaks in the amide I band split by approximately 30  $\text{cm}^{-1}$ . The two-peak structure is generally interpreted as a marker of cross-strand coupling, and the profile of the peaks is sensitive to peptide unfolding (Fig. 1.2).



**Figure 1.2** Simulated amide I spectra for two peptides in different configurations. The cross-strand interactions of the  $\beta$ -hairpin (left) lead to two modes split by approximately  $30\text{ cm}^{-1}$ . The IR spectrum of a random coil peptide (right) displays a broad peak centered at  $1655\text{ cm}^{-1}$ .

### 1.3.1. GB1<sub>41-56</sub> folding kinetics

Transient unfolding of GB1<sub>41-56</sub> was first investigated with T-jump fluorescence of Trp<sub>43</sub>.<sup>36</sup> The fluorescence decay fits well to a single exponential with a time constant of  $6\ \mu\text{s}$  at a starting temperature of  $297\text{ K}$ . Because the GB1<sub>41-56</sub> peptide is susceptible to aggregation at high concentrations, it has not been studied with T-jump IR techniques. Recently, however, Olsen et al reported folding data on GB1<sub>41-56</sub> and mutants using exchange broadening NMR experiments.<sup>41</sup> By observing the broadening of the Tyr-H $\delta$  peak of the folded and unfolded peptides as a function of temperature, a folding rate could be reconstructed. Using this analysis, the folding rate of the native GB1<sub>41-56</sub> hairpin was  $18\ \mu\text{s}$ , which is different from the  $6\ \mu\text{s}$  time scale found earlier. The differences can be explained by the fact that the rate extraction process relies on two-state kinetics and a well-established baseline. In absence of rigorous values, the folding time is expected to

be a slight over-estimate of the actual folding time. However, within a set of mutated peptides, the rates can be compared reliably to one another. The author's find that a proline inserted into the turn region stabilizes the loop and increases the folding rate. Mutations that create a salt bridge between the N and C termini increase the stability of the folded peptide and also increase the folding rate.<sup>41</sup> Because the salt bridge speeds up folding, the author's conclude that native structure forms first at the N and C-terminus rather than at the  $\beta$ -turn.

### **1.3.2. Trpzip folding kinetics**

When the trpzip peptides were first developed, Cochran et al presented UV-CD melting curves to measure the folding thermodynamics of TZ1-TZ6, and solution NMR structures for TZ1, TZ2, and TZ4.<sup>39</sup> The melting curves implied a two-state transition between the folded and unfolded configurations. However, in 2004, Yang et al demonstrated that the folding transition of TZ2 was heterogeneous using a variety of the equilibrium spectroscopies and all-atom molecular dynamics simulations.<sup>55</sup> The conclusion was drawn from the observation that the melting temperature was probe dependant. For example, the UV-CD fluorescence melting curve in the absence of a chemical denaturant has a melting temperature,  $T_m$ , equal to 72 °C, while the FTIR melting curve  $T_m$  is greater than 60 °C. The fluorescence melting curve displays a much lower  $T_m$ , which was found to be less than 15 °C in the absence of chemical denaturants. Molecular dynamics simulations conducted at three temperatures were consistent with the experimental data, and showed that there were a number of local minima separated from native state by relatively large barriers at low temperature. Following this report, there

has been a series of publications reporting the folding kinetics of other trpzip peptides following a small ( $\sim 10$  °C) T-jump. This section will focus on the transient folding of TZ2.<sup>54,59-62</sup>

The observation of heterogeneous folding in the equilibrium experiments was followed up by Yang et al by probing tryptophan fluorescence following an 11 °C T-jump.<sup>54</sup> The fluorescence spectrum of tryptophan is sensitive to the packing and solvent exposure of the indole side chains and is used to describe the formation or breakup of the hydrophobic core. An additional feature of the fluorescence spectrum is that there is a relationship between frequency and solvent exposure: solvent exposed side chains contribute to the red side of the band and tightly packed side chains fluoresce on the blue side of the band. In their study, the kinetic rate was found to be biphasic. The fast timescale ( $\sim 100$  ns) is assigned to side chain flexibility and its amplitude decreases relative to the slow phase with temperature. The slow time scale ( $\sim 1$   $\mu$ s) is assigned to structural rearrangements within the hydrophobic core, and its amplitude increases with temperature relative to the slow phase. An additional feature of the microsecond phase is a temperature and frequency dependence that ranges from  $k_{\text{red}}^{-1} \approx k_{\text{blue}}^{-1} = 670$  ns at 27.7 °C, to  $k_{\text{red}}^{-1} = 2.1$   $\mu$ s and  $k_{\text{blue}}^{-1} = 2.7$   $\mu$ s at 57.5 °C. The red side of the band decaying faster than the blue side of the band indicates that the more solvent exposed side chains will break out of the core faster than the less exposed side chains. The difference between the frequency-dependent time scales grows with temperature and is used to put a lower bound on the folding free energy surface roughness ( $\delta^2G$ ) of  $0.6 k^2T^2$ . TZ2 is thus proposed as a system in which free energy surface roughness is large enough for experiments to elucidate the details of free energy microstructure.



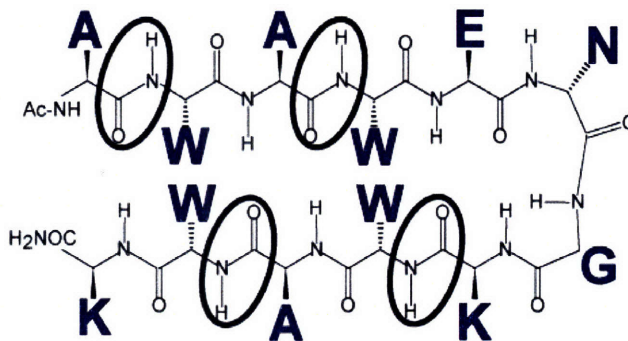
Shortly before the report of Yang et al, a report by Snow et al presented T-jump, IR and fluorescence probe data for TZ1, TZ2, and TZ3 in combination with folding simulations.<sup>59</sup> The folding rates at 21-23 °C for the three systems observed with fluorescence are  $k_f^{-1} = 6.3 \mu\text{s}$  for TZ1,  $k_f^{-1} = 1.8 \mu\text{s}$  for TZ2, and  $k_f^{-1} = 830 \text{ ns}$  for TZ3. Each peptide differs only in turn type, each having the same strand sequence and four-tryptophan hydrophobic core. The fastest time scale is the 830 ns folding time of TZ3, which has a *D*-proline enhanced type II'  $\beta$ -turn that is known to speed up folding in other peptides.<sup>63-65</sup> TZ2 (Type I' turn) is faster than TZ1 (Type II' turn), which reflects the higher stability of TZ2 relative to TZ1 and also suggests that the EN GK turn sequence is a better turn promoter than the EG NK sequence. In their report T-jump IR data was only obtained for TZ2, which displayed a folding rate  $k_f^{-1} = 2.5 \mu\text{s}$  at 23°C. The slower  $k_f^{-1}$  of the IR probe relative to the fluorescence probe is evidence for hydrophobic core-mediated folding mechanism. Folding simulations in the paper generally support a zipper-like mechanism. Although the time-ordering of events in each folding trajectory is heterogeneous, evidence for the zipper mechanism is observed as contacts form near the turn on average faster than those near the N and C-termini. Also evident in the simulations is that the Trp2 side chain is much more flexible than Trp4 and Trp9, and disorders first in the trajectories. The authors also report that the simulations do not show significant free energy traps. Comparing the simulations of TZ1, TZ2, and TZ3 reveals that while the turn type does alter the folding speed in the same way as observed in the experiments, there is no significant deviation in the folding mechanism.

After the report above, Gai and coworkers continued the same type of analysis by reporting T-jump, IR probe data for TZ1, TZ2, TZ3, TZ4, and a new variant, TZ3I in

order to investigate the effects of turn type and hydrophobic cluster position on hairpin folding.<sup>61</sup> The TZ3I peptide was designed with the same strand sequence as TZ3, but the turn sequence was replaced with DATK, which is the same turn sequence as in the GB1<sub>42-56</sub> peptide and TZ4. In comparing the rates of the TZ4 and TZ3I peptides within a two-state kinetics approximation, it was found that while the folding rate of TZ3I was 3 times *slower* than TZ4, the unfolding rate of TZ3I was 6 times *faster* than TZ4. In addition, the folding rate of GB1<sub>42-56</sub> is only ~2 times faster than TZ4, but the unfolding rate is 40 times slower than TZ4. These observations lead to the conclusion that stabilizing role of the four-tryptophan hydrophobic core is to prevent unfolding of the peptide. In comparing the kinetics of the four peptides that only differ in the sequence of the four-residue turn region (TZ1, TZ2, TZ3, and TZ3I) the data shows that the unfolding rate is very similar from peptide to peptide. In contrast, the folding rates show a large amount of variance, especially between TZ3 and TZ3I. This suggests that the turn region is the rate limiting step in peptide folding and is consistent with a zipper mechanism. Such a conclusion is further supported by dividing the entropy into a contribution from the turn and strand region ( $\Delta S_f = \Delta S_{\text{turn}} + \Delta S_{\text{strand}}$ ), and assuming that the  $\Delta S_{\text{strand}}$  term is constant between the peptides. The results show a linear correlation between the turn entropy and the folding rate, supporting a folding mechanism that is driving by turn nucleation.

In 2006, Gai and coworkers did an exhaustive mutation study of TZ4 in which they confirmed that the hydrophobic core was not the key event in peptide folding, but that its role was to limit peptide unfolding.<sup>60</sup> This was evidenced by the very low  $\phi$ -value for mutations that removed a single tryptophan from the hydrophobic core. Furthermore, a

high  $\phi$ -value (0.77) was measured for an aspartic acid residue near the turn region, which supports the picture in which  $\beta$ -turn formation is an integral part of the folding transition.



**Figure 1.3** Stick diagram of TZ2C. Circles represent amide groups that were targeted for  $^{13}\text{C}$  isotope labeling.

The most recent trpzip folding experiment is a T-jump IR study of  $^{13}\text{C}$  labeled TZ2C, a triple alanine variant of TZ2 developed by Keiderling and coworkers (Fig. 1.3).<sup>62</sup> For the unlabeled TZ2C, folding rates were measured for two parts of the amide I band: the low frequency region (termed  $\nu_{\perp}$  in this thesis) around  $1630\text{ cm}^{-1}$  and the high frequency region (termed  $\nu_{\text{dis}}$  in this thesis) around  $1660\text{ cm}^{-1}$ . The  $\nu_{\perp}$  band is interpreted as directly reporting on  $\beta$ -sheet secondary structure within the hairpin, and the  $\nu_{\text{dis}}$  band reports on unstructured amide groups within the peptide. The major finding of this work is that the kinetics of the two groups are distinct and show a quantifiable temperature dependence. For the unlabeled peptide,  $k^{-1}(\nu_{\perp})$  is  $3.0\ \mu\text{s}$  at  $25\text{ }^{\circ}\text{C}$  and  $1.4\ \mu\text{s}$  at  $42\text{ }^{\circ}\text{C}$ , while  $k^{-1}(\nu_{\text{dis}})$  is  $1.9\ \mu\text{s}$  at  $25\text{ }^{\circ}\text{C}$  and  $0.9\ \mu\text{s}$  at  $42\text{ }^{\circ}\text{C}$ . The fact that the rates converge with temperature is attributed to a kinetic intermediate present at low temperatures and becoming less significant as the temperature is increased. Isotope labels were placed at locations that report on the turn (A3K8), mid-strand (A3A10), and termini (A1A10). The

kinetic rates of the A1A10 peak ( $1616\text{ cm}^{-1}$ ) as a function of temperature followed those of the  $\nu_{\text{dis}}$  peak. The rates of the A3A10 peak ( $1616\text{ cm}^{-1}$ ) match the  $\nu_{\perp}$  peak kinetics, which is interpreted as reporting on the hydrophobic core. Finally, the A3K8 peak rates are found to be intermediate between the  $\nu_{\perp}$  and  $\nu_{\text{dis}}$  rates. A comparison of the folding rates for the various isotope peaks suggests that unfolding progresses from the turn and termini toward the center. This appears to be consistent with the findings of Wang et al.<sup>61</sup>

### 1.3.3. Folding kinetics of other peptides

Studies of Trpzip and GB1<sub>41-56</sub> peptides dominate the folding literature, but there have been a number of folding studies on other peptides. Dyer et al measured the folding of cyclic  $\beta$ -hairpin peptides with T-jump amide IR spectroscopy.<sup>65</sup> The folding rates of the cyclic peptides were found to be on the order of 100 ns, which is much faster than for GB1<sub>41-56</sub> and TZ2. This fast time scale can be rationalized by realizing that unfolding of a cyclic peptide is a sterically-hindered event, and will probably involve small structural rearrangements. Kinetic studies have also been performed on peptide1, a 15 residue peptide with a proline enhanced  $\beta$ -bulge turn region. For this peptide, Xu et al measured a folding time of 800 ns.<sup>66</sup> The comparably fast folding times of peptide1 and the cyclic hairpins is further evidence for the role of proline in significantly speeding up the peptide folding rate, which suggests that  $\beta$ -turn formation is the rate limiting step in hairpin folding. These issues will be discussed further in Chapters 6, 7, and 8.

## 1.4. Experimental Aims

The work in this thesis is directed at accessing the microscopic details of hairpin folding. Vibrational spectroscopy is an ideal probe because it is explicitly dependent on nuclear coordinates and it can be designed to resolve very fast (sub-picosecond) structural changes. By expanding the information content of amide I spectroscopy, we will be able to probe the structural changes within the peptide backbone and put together a folding mechanism that takes into account both secondary and tertiary contacts. In Chapter 4 of this thesis the amide I spectroscopy of several  $\beta$ -hairpins is simulated in order to investigate the nature of the vibrational eigenstates. A key aspect of that work is to simulate disorder in the peptide by including Gaussian fluctuations of the amide I site energies, and to investigate the local mode structure of the eigenstates. This work is used to interpret the changes to the hairpin amide I band during thermal denaturation, which is discussed in Chapter 5. One essential result of the spectral simulations is that the vibrations are delocalized across the entire hairpin backbone. As a result, the amide I band cannot directly distinguish backbone conformational change in the turn from that near the terminal groups.

There are two limiting factors to gain local structural information from an amide I spectrum. The first is that amide I vibrations are delocalized as stated above. One way to interrupt the delocalized states is to selectively shift the frequencies of individual amide groups. This can be achieved with the use of  $^{13}\text{C}$  or  $^{18}\text{O}$  isotope labels, which shift the amide I frequency by  $\sim 42$  and  $\sim 39$   $\text{cm}^{-1}$  respectively.<sup>67</sup> This approach has been used a great deal in the past, but had never been applied to  $\beta$ -hairpins until 2006, when two single isotope labels were placed in the TZ2 peptide.<sup>68</sup> Following this study, the work in

Chapter 4 was published in which TZ2 was labeled with two  $^{13}\text{C}$  isotopes at two residues that are strongly coupled across the  $\beta$ -strands.<sup>69</sup> Chapter 5 reports the first thermal unfolding study of an isotope labeled hairpin peptide with T-jump kinetics.<sup>64</sup> The goal of that research was to demonstrate the local probes within the peptide backbone could reveal residue-level structural dynamics.

Isotope-labeled peptides have allowed us to develop a greater understanding of the amide I spectroscopy of  $\beta$ -hairpins. By comparing spectral simulations of the labeled and unlabeled peptides to the experimental spectra we were able to refine the site energy and distinguish between two electrostatic coupling models. These improvements to the spectral modeling will be discussed in Chapter 4. The next step is to use the isotopically-shifted amide I peaks to build up a complete description of the T-jump induced unfolding kinetics in order to sort through the residue-level events that are used to describe the folding pathway. This was the aim of the work by Hauser et al described above, where the folding kinetics of TZ2C were investigated with  $^{13}\text{C}$ -alanine labels.<sup>62</sup> In order to make more direct comparisons with the numerous simulations of TZ2 we have directly labeled three sites in the peptide without alanine mutations. The final chapter describes the spectral information of these isotope-labeled species and uses it to analyze the T-jump kinetics.

The second limiting factor in interpreting the amide I spectrum is that the infrared absorbance band is complicated not only by the large number of vibrations within the band, but also because of the large line widths inherent to these transitions. The line width is temperature dependant, which is in part due to thermally-induced structural changes. Structural contributions to the homogeneous line width are fast conformational

fluctuations ( $\sim 1$  ps) and the interaction of the vibrational modes with the surrounding solvent. The structural contribution to the inhomogeneous line width is the quasi-static distribution of structures within the ensemble of molecules. These two types of disorder lead to spectral broadening, but cannot be distinguished directly from a linear absorption spectrum. An accurate description of these effects would help describe the nature of the thermal folding transition. For example, if the peptide folding transition does in fact go from a large distribution of disordered structures to a small distribution of folded states, then there would be a large decrease in the inhomogeneous line width upon peptide folding. If solvent fluctuations play an important role in the thermal melting transition than the homogeneous line width should show a dependence on the temperature of the system.

Throughout this thesis these questions are directly addressed through the use of two-dimensional infrared (2D IR) spectroscopy. 2D IR spectroscopy is a third-order nonlinear technique that correlates the vibrations in one time period to those in a third time period.<sup>70</sup> The time periods are marked by the interaction of the system with four light fields that coherently drive vibrational transitions. A Fourier transform of the first and third time periods produces a two-dimensional frequency axis that correlates frequencies between the two periods. The vibrational transitions are observed along the diagonal axis of the spectrum, and the width of the peaks in two dimensions distinguishes inhomogeneous from homogeneous contributions to the line widths. Additional advantages of the technique are added structural resolution and the direct visualization of coupling between vibrational modes which is seen in cross peaks in the off-diagonal part

of the spectrum connecting diagonal peaks. 2D IR spectroscopy will be more completely introduced in Chapter 3.

The goal of this research is to combine isotope labeling techniques with 2D IR spectroscopy to build a more complete description of the hairpin folding pathway. Isotope labels allow us to observe residue-level backbone structure throughout the thermal melting transition. 2D IR spectroscopy then contributes to a dynamic description of the peptide during the transition. In this way we can elucidate the microscopic folding routes explored by the peptides and quantify the heterogeneity of these pathways. Such insight will contribute immediately to verifying and improving molecular dynamics simulations. It will also refine our understanding of the larger questions regarding protein folding mechanisms.



## 1.5. References

- (1) Anfinsen, C. B. *Science* **1973**, *181*, 223-230.
- (2) Zwanzig, R.; Szabo, A.; Bagchi, B. *Proceedings of the National Academy of Sciences* **1992**, *89*, 20-22.
- (3) Karplus, M. *Folding and Design* **1997**, *2*, S69-S75.
- (4) Dill, K. A.; Chan, H. S. *Nat. Struct. Biol.* **1997**, *4*, 10-19.
- (5) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct. Funct. Genet.* **1995**, *21*, 167.
- (6) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. *Proteins* **1998**, *32*, 136-158.
- (7) Bradley, P.; Misura, K. M. S.; Baker, D. *Science* **2005**, *309*, 1868-1871.
- (8) Kuhlman, B.; Dantas, G.; Ireton, G. C.; Varani, G.; Stoddard, B. L.; Baker, D. *Science* **2003**, *302*, 1364-1368.
- (9) Yang Zhang, A. K. A., Jeffrey Skolnick, *Proteins: Structure, Function, and Bioinformatics* **2005**, *61*, 91-98.
- (10) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740-744.
- (11) Snow, C. D.; Zagrovic, B.; Pande, V. S. *J. Am. Chem. Soc.* **2002**, *124*, 14548-14549.
- (12) Pitera, J. W.; Swope, W. *Proceedings of the National Academy of Sciences* **2003**, *100*, 7587-7592.
- (13) Lee, M. R.; Tsai, J.; Baker, D.; Kollman, P. A. *J. Mol. Biol.* **2001**, *313*, 417-430.
- (14) Dill, K. A.; Ozkan, S. B.; Weikl, T. R.; Chodera, J. D.; Voelz, V. A. *Curr. Opin. Struct. Biol.* **2007**, *17*, 342-346.

- (15) Dill, K. A. *Biochemistry* **1990**, *29*, 7133.
- (16) Kauzmann, W. In *Adv. Protein Chem.*; Anfinsen, C. B., Anson, M. L., Bailey, A. K., Edsall, J. T., Eds.; Academic Press, 1959; Vol. 14, pp 1-63.
- (17) Murphy, K. P. In *Protein Structure, Stability, and Folding*; Murphy, K. P., Ed.; Humana Press: Totowa, NJ, 2001; Vol. 168, pp 1-16.
- (18) Daggett, V.; Fersht, A. R. *Trends Biochem. Sci.* **2003**, *28*, 18-25.
- (19) Karplus, M.; Weaver, D. L. *Protein Sci.* **1994**, *3*, 650-668.
- (20) Kim, P. S.; Baldwin, R. L. *Annu. Rev. Biochem.* **1982**, *51*, 459-489.
- (21) Ptitsyn, O. B. *J. Protein Chem.* **1987**, *6*, 273-293.
- (22) Dobson, C. M.; Sali, A.; Karplus, M. *Angew. Chem., Int. Ed.* **1998**, *37*, 868-893.
- (23) Ptitsyn, O. B. *Adv. Protein Chem.* **1995**, *47*, 83-229.
- (24) Fersht, A. R. *Curr. Opin. Struct. Biol.* **1997**, *7*, 3-9.
- (25) Ozkan, S. B.; Wu, G. A.; Chodera, J. D.; Dill, K. A. *Proceedings of the National Academy of Sciences* **2007**, *104*, 11987-11992.
- (26) Vincent A. Voelz, K. A. D. *Proteins: Structure, Function, and Bioinformatics* **2007**, *66*, 877-888.
- (27) Weikl, T. R.; Dill, K. A. *J. Mol. Biol.* **2003**, *332*, 953-963.
- (28) Day, R.; Daggett, V. *J. Mol. Biol.* **2007**, *366*, 677-686.
- (29) Leite, V. B. P.; Onuchic, J. N.; Stell, G.; Wang, J. *Biophys. J.* **2004**, *87*, 3633-3641.
- (30) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76-88.
- (31) Scholtz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95.

- (32) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H. *Biochemistry* **1996**, *35*, 691-697.
- (33) Thompson, P. A.; Eaton, W. A.; Hofrichter, J. *Biochemistry* **1997**, *36*, 9200-9210.
- (34) Monticelli, L.; Tieleman, D. P.; Colombo, G. *J. Phys. Chem. B* **2005**, *109*, 20064-20067.
- (35) Searle, M. S.; Williams, D. H.; Packman, L. C. **1995**, *2*, 999-1006.
- (36) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196-199.
- (37) Eaton, W. A.; Munoz, V.; Thompson, P. A.; Henry, E. R.; Hofrichter, J. *Acc. Chem. Res.* **1998**, *31*, 745-753.
- (38) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236-4237.
- (39) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **2001**, *99*, 9081-9081.
- (40) Tsai, J.; Levitt, M. *Biophysical Chemistry* **2002**, *101*, 187-201.
- (41) Olsen, K. A.; Fesinmeyer, R. M.; Stewart, J. M.; Andersen, N. H. *Proceedings of the National Academy of Sciences* **2005**, *102*, 15483-15487.
- (42) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* **2002**, *11*, 1492-1505.
- (43) Hilario, J.; Kubelka, J.; Syud, F. A.; Gellman, S. H.; Keiderling, T. A. *Biopolymers* **2002**, *61*, 233-236.
- (44) Hilario, J.; Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2003**, *125*, 7562-7574.
- (45) Dinner, A. R.; Lazaridis, T.; Karplus, M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9068-9073.

- (46) Munoz, V.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5872-5879.
- (47) Zhou, R.; Berne, B. J.; Germain, R. *Proceedings of the National Academy of Sciences* **2001**, *98*, 14931-14936.
- (48) Wei, G. H.; Mousseau, N.; Derreumaux, P. *Proteins* **2004**, *56*, 464-474.
- (49) Garcia, A. E.; Sanbonmatsu, K. Y. *Proteins* **2001**, *42*, 345-354.
- (50) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153-164.
- (51) Pitera, J. W.; Haque, I.; Swope, W. C. *J. Chem. Phys.* **2006**, *124*, 141102.
- (52) Bolhuis, P. G. *Biophys. J.* **2005**, *88*, 50-61.
- (53) Bolhuis, P. G. *Proc. Nat. Acad. Sci., USA* **2003**, *100*, 12129-12134.
- (54) Yang, W. Y.; Gruebele, M. *J. Am. Chem. Soc.* **2004**, *126*, 7758-7759.
- (55) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* **2004**, *336*, 241-251.
- (56) Millet, O.; Loria, J. P.; Kroenke, C. D.; Pons, M.; Palmer, A. G. *J. Am. Chem. Soc.* **2000**, *122*, 2867-2877.
- (57) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181-364.
- (58) Miyazawa, T.; Blout, E. R. *J. Am. Chem. Soc.* **1961**, *83*, 712-719.
- (59) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4077-4082.
- (60) Du, D.; Tucker, M. J.; Gai, F. *Biochemistry* **2006**, *45*, 2668-2678.
- (61) Du, D.; Zhu, Y.; Huang, C.-Y.; Gai, F. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15915-15920.

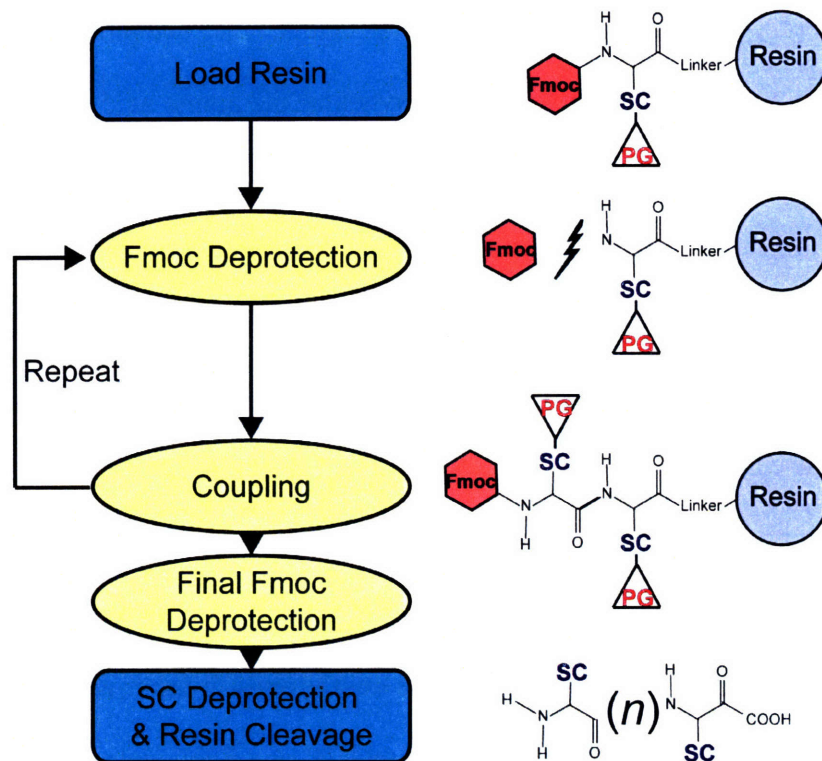
- (62) Hauser, K.; Krejtschi, C.; Huang, R.; Wu, L.; Keiderling, T. A. *J. Am. Chem. Soc.* **2007**, *In Review*.
- (63) Xu, Y.; Purkayastha, P.; Gai, F. *J. Am. Chem. Soc.* **2006**, *128*, 15836-15842.
- (64) Smith, A. W.; Tokmakoff, A. *Angewandte Chemie International Edition* **2007**, *46*, 7984-7987.
- (65) Maness, S. J.; Franzen, S.; Gibbs, A. C.; Causgrove, T. P.; Dyer, R. B. *Biophys. J.* **2003**, *84*, 3874-3882.
- (66) Xu, Y.; Oyola, R.; Gai, F. *J. Am. Chem. Soc.* **2003**, *125*, 15388-15394.
- (67) Decatur, S. M. *Acc. Chem. Res.* **2006**, *39*, 169-175.
- (68) Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. *J. Phys. Chem. B* **2006**, *110*, 7545-7555.
- (69) Smith, A. W.; Tokmakoff, A. *J. Chem. Phys.* **2007**, *126*, 045109.
- (70) Khalil, M.; Demirdoven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258-5279.

# Chapter 2

## Peptide Synthesis

### 2.1. An Introduction to Solid Phase Peptide Synthesis

The peptide samples investigated in this work were each obtained through a synthesis procedure. This chapter outlines general synthetic principles and details specific protocols for each peptide for which spectroscopic data was taken. Solid phase peptide synthesis (SPPS) was pioneered by R. B. Merrifield in the early 1960's, when it was shown that an amino acid chain could be synthesized by anchoring the first residue to a polystyrene resin and adding each subsequent residue one by one.<sup>1</sup> Since its early development, SPPS has undergone dramatic improvements that have made it a standard tool in many chemistry labs. The development of 9-fluorenylmethoxycarbonyl (Fmoc) chemistry has made the technique available to an even broader set of users by reducing cost and eliminating many of the hazardous chemicals needed in the Boc chemistry first developed by Merrifield.<sup>2</sup>



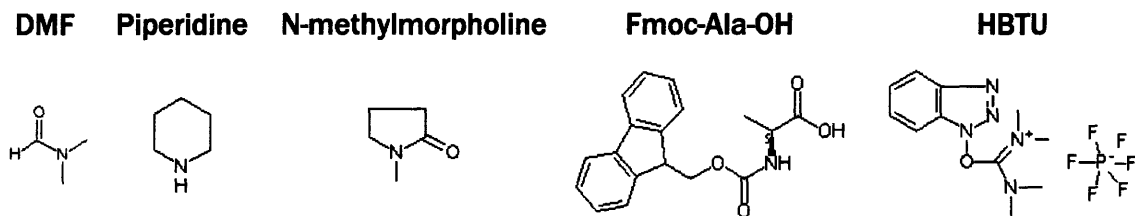
**Figure 2.1** Flow chart for solid phase peptide synthesis. The red hexagon symbolizes the  $N_{\alpha}$ -Fmoc protecting group. The PG-labeled triangle represents side chain (SC) protecting groups where appropriate (see Section 2.5). The solid support resin is symbolized by the blue-filled circle.

The general Fmoc-SPPS method is diagrammed in Fig. 2.1, and some common chemicals are shown in Fig. 2.2. This section will overview the general procedure, and then details of each step will be given in Sections 2.2-2.6. First, an acid labile linker unit is used to bind the first amino acid to the resin. At this stage, the  $\alpha$ -nitrogen ( $N_{\alpha}$ ) is bound to a base-labile Fmoc protecting group to prevent unwanted coupling at the terminal amine. An organic base is then used to remove the Fmoc group in preparation for the coupling of the next amino acid. After Fmoc removal, the next amino acid in the peptide sequence and an activating agent are combined in N,N-dimethylformamide

(DMF). For synthesis procedures shown below, the activating agent creates a benzotriazole ester (OBt) on the carboxyl oxygen of the amino acid as shown in Fig. 2.5 to increase the efficiency of the reaction. The DMF/N-methylmorpholine/activator/amino acid solution is added to the reaction vessel and peptide bond formation occurs via a condensation reaction between the amine group of the resin bound peptide and the carbonyl carbon of the amino acid in solution. These steps are repeated for each residue in the peptide chain.

One important consideration is that many amino acids have side chains that are reactive in the coupling step. Examples of such side chains are the secondary amine in tryptophan, the primary amine in lysine, the acid groups in aspartic and glutamic acids, and others. To prevent unwanted reaction pathways, an orthogonal synthesis procedure is used in which the side chains of these amino acids are protected with acid-labile groups that remain bound throughout the peptide coupling reactions. After removing the final Fmoc group, a concentrated trifluoroacetic acid (TFA) solution is used to cleave the peptide from the resin and also removes the side-chain protecting groups. The TFA/peptide solution is treated with cold ether to crystallize the peptide so that it can be dried, washed and purified for further use. Many reviews and laboratory manuals have been written on this subject,<sup>3</sup> so below I will detail only the procedures used in our laboratory.

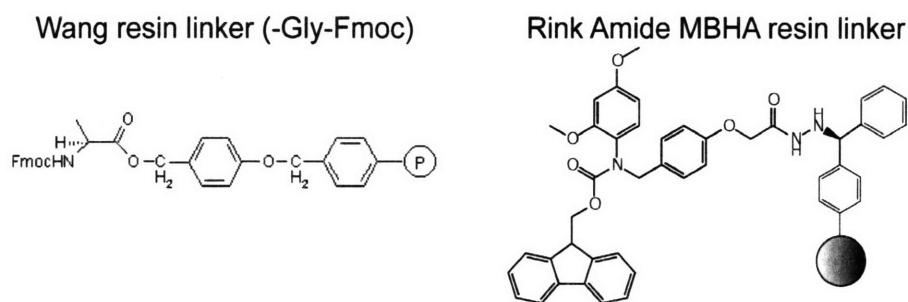




**Figure 2.2** Diagrams of common chemicals in SPPS. Dimethylformamide (DMF) is the primary reaction solvent. Piperidine and N-methylmorpholine are organic bases used in the deprotection and activation steps respectively. An example Fmoc amino acid is shown for alanine, and the activator HBTU (O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate).

## 2.2. Resins and Linkers

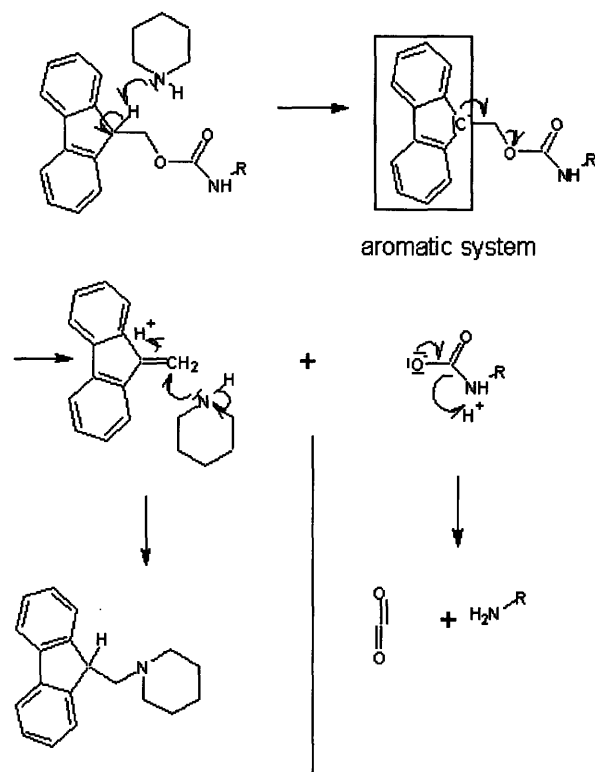
Support resins for SPPS have been researched extensively over that past 40 years, and several reviews are available.<sup>3,4</sup> Two key considerations for choosing a resin are the way the resin surface will interact with the peptide side chains, and the linker chemistry used to attach the first amino acid to the resin. One of the most commonly used resins for Fmoc-based synthesis is the Wang resin, a cross-linked polystyrene polymer functionalized with p-benzyloxybenzyl alcohol.<sup>5</sup> Linkage to the resin occurs through an esterification reaction between the p-benzyloxybenzyl alcohol and the hydroxyl oxygen of the amino acid carboxyl group. This linkage is acid labile and can be easily cleaved with concentrated TFA to produce a carboxyl-terminated peptide. Typically, the resin is purchased with the C-terminal amino acid already attached because the esterification reaction conditions are relatively harsh and incorrect coupling can lead to enantiomerization, low substitution rates, and dipeptide formation.



**Figure 2.3** (Left) Wang resin with the first amino acid attached. The circle P represents the polystyrene matrix. (Right) Rink Amide MBHA Resin linker with the C-terminal amide group protected with an Fmoc protecting group.

For the peptides investigated in this thesis, it is often desirable to terminate the peptide with an amide group in order to eliminate strong electrostatic interactions between the zwitterionic N and C termini. An ideal resin for amidating the C-terminus is the Rink Amide MBHA resin, which is based on a copoly(styrene) matrix.<sup>6,7</sup> The linker is a modified Rink amide attached to the resin with norleucine. Resin is purchased with an Fmoc group protecting what will become the C-terminal amide group. The first step in the synthesis in this case is to remove the amide Fmoc group before coupling the first amino acid. Both of these resins provide a solid support with minimal aggregation at the resin surface, and linkers that can be easily cleaved using TFA.

## 2.3 N<sub>α</sub> Deprotection



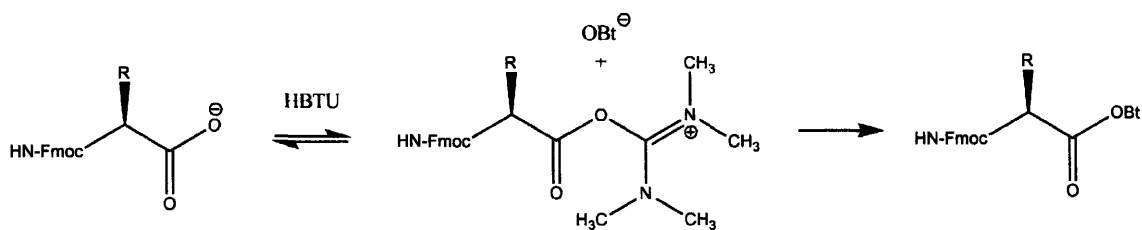
**Figure 2.4** Reaction scheme for the removal of an Fmoc group from the peptide N<sub>α</sub> using piperidine in DMF.

To prepare the amide nitrogen (N<sub>α</sub>) for peptide coupling, the Fmoc group must be removed. This is done in basic conditions, typically with ~20% piperidine in DMF. A reaction scheme is presented in Fig. 2.4 for Fmoc removal with piperidine. In the first step, piperidine deprotonates the 9-carbon of the fluorene group to allow for the formation of a carbamate bond. Nucleophilic attack by another piperidine molecule to the carbamate carbon stabilizes the fluorene byproduct that is removed when the solvent is flushed. Electrons from the ester bond move to the peptide carboxyl group and initiate

the removal of CO<sub>2</sub> as a by-product. This leaves the N<sub>α</sub> exposed and ready for the peptide coupling reaction.

## 2.4. Activation & Coupling

The coupling reaction is the formation of a new peptide bond through a condensation reaction between the carboxyl group of the solution phase amino acid and the exposed N<sub>α</sub> of the solid-supported N-terminal amino acid of the peptide chain. In practice, the carboxyl group needs to be chemically activated in drive the acylation reaction to completion at an efficient rate. A number of activation methods have been developed, but here we focus on OBt ester activation as it represents an extremely efficient and cost effective solution.<sup>3</sup> The OBt ester formation occurs before the activated ester is exposed to the supported peptide system. A common activating agent used in our lab is O-Benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU).<sup>8</sup> After the peptide is deprotonated with N-methylmorpholine, HBTU reacts quickly with the amino acid derivative to create an OBt ester.



**Figure 2.5** Reaction scheme for OBt activation.

Once the OBt ester is exposed to the amide nitrogen of the peptide, the coupling reaction proceeds to completion within 40-60 minutes. This step is handicapped if the

growing peptide has aggregated to the resin surface because the  $N_{\alpha}$  nitrogen is sterically blocked from the solution-phase, activated amino acid. If this occurs, alternative resins or activating agents need to be considered.

## **2.5. Side Chain Protecting Groups**

Many of the naturally-occurring amino acid side chains have chemical groups that are reactive under the deprotection, activation, and coupling steps outlined above. Unwanted side reactions can lead to peptide chains branching out from the side chains, which has a dramatic effect on the final peptide yield. Therefore, chemical modifications have been developed to protect these groups during the synthesis. Side chain protecting groups are designed with reaction chemistry that is orthogonal to the peptide coupling and  $N_{\alpha}$  deprotection steps. In the case of base-catalyzed Fmoc-mediated synthesis, acid labile groups are used to protect the side chains. In both reaction schemes, the side chain protecting groups are cleaved in the same solution that cleaves the peptide from the resin. Table 2.1 summarizes the protecting groups used for each amino acid in the synthesis done by our lab.

Protecting Group	Abbr.	Side Chains
2,2,5,7,8-pentafluorophenyl	Pbf	Arg
Trityl	Trt	Asn, Cys, Gln, His
O-tert-butyl	OtBu	Glu, Asp
tert-butoxycarbonyl	Boc	Lys, Trp
tert-butyl	tBu	Ser, Thr, Tyr

**Table 2.1** Side-chain protecting groups for amino acids used in peptide synthesis. The table includes side-chain protecting groups used to synthesize peptides for the current work.

## 2.6. Cleavage and Isolation

For the resin linkers discussed here, the final cleavage step is performed with concentrated ( $\geq 95\%$ ) TFA. As stated in section 2.1.5, this reaction step also removes the acid-labile side chain protecting groups. The cleavage step produces a number of highly reactive cationic species that can chemically modify the exposed side chain groups. Examples of such species are cations released upon cleavage of *t*-butyl groups, the Rink amid linker, trityl groups, and pbf groups. These cleaved side chains and remnant products of the synthesis procedure are highly reactive, and so the cleavage solution is prepared with additional chemicals that act as scavengers to neutralize these by-products. Various reagents have been employed, but in the current work they were limited to triisopropylsilane (TIPS), 1,2-ethanedithiol (EDT), anisol, phenol, and water. A common solution is 95% TFA and 5% TIPS, which provides an effective alternative to the mephitic EDT.

After the cleavage/deprotection reaction comes to completion, the peptide/TFA solution is filtered from the resin, and the filtrate is cooled in an ice bath. Cold ether is

added to the solution, encouraging the peptide to precipitate out of solution, and the precipitate is washed and collected using centrifugation. The sample is then dissolved in 20% acetic acid in water, and lyophilized to remove the solvent and TFA. Further purification was outsourced to the MIT biopolymers laboratory, where the peptide was purified using reverse phase HPLC in a C18 column. Final purity is approximately 95-98%.

After HPLC purification the peptide was further lyophilized to remove residual TFA, which has an intense vibration at  $1672\text{ cm}^{-1}$  that interferes with the amide I' spectrum. To remove the TFA, HPLC-purified peptide was dissolved in 10-50 mM DCl in D<sub>2</sub>O, and lyophilized. This was repeated several (3-4) times. Another result of this repeated lyophilization is to exchange the labile peptide protons for deuterons, which is important for obtaining high quality amide I' spectra.

## **2.7. Synthesis Procedures**

### **2.7.1. Synthesis of TZ2**

Trpzip2 (TZ2) is a twelve-residue peptide designed and first reported by Cochran et al. It has become one of the most studied hairpin peptides, receiving significant attention in the literature, with multiple experimental studies and numerous theoretical treatments.<sup>9-20</sup> Within the last several years, TZ2 has become a benchmark system which is now used in the development of new simulation techniques.<sup>15,21</sup> The trpzip peptides were engineered with a novel structural motif of four tryptophan residues arranged with two side chains on each of the opposing  $\beta$ -strands.<sup>9</sup> The same-strand tryptophan side

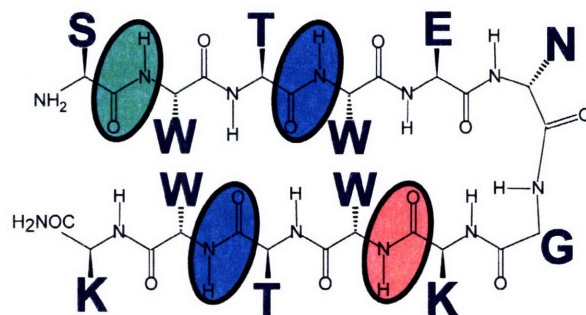
chains are separated by a threonine residue so that all four side chains are on the same side of the  $\beta$ -sheet plane. A favorable hydrophobic attraction between the side chains is enhanced by perpendicular stacking of the indole rings, and gives rise to a remarkably stable system. The trpzip scaffold has allowed for numerous studies of effects such as  $\beta$ -strand length and turn sequences on the stability and folding mechanism of  $\beta$ -hairpins.<sup>17-</sup>  
<sup>19</sup> The sequence of TZ2 from the N to C-terminus is: Ser1-Trp2-Thr3-Trp4-Glu5-Asn6-Gly7-Lys8-Trp9-Thr10-Trp11-Lys12-NH<sub>2</sub> (SWTWENGKWTWK). The turn region, formed by residues ENGK, folds into a type I'  $\beta$ -turn. In this report, TZ2 was prepared as a C-terminal amide, which is consistent with previous experimental studies.

TZ2 was prepared in house using Fmoc-based SPPS on a PS3 peptide synthesizer (Protein Technologies Incorporated). The solid support was a Rink Amide MBHA resin (01-64-0037, Novabiochem), which has a 100-200 mesh and a resin load of 0.40 - 0.80 mmole/g. Fmoc-protected amino acids were obtained from PTI Instruments in pre-packaged cartridges at four molar excess HBTU activator and amino acid. The primary reaction solvent was DMF (EM-DX1732; EMD Chemicals; 99.9% purity). Removal of the Fmoc groups at each step was done with 20% piperidine in DMF (PTI Instruments) with a 7 minute reaction time, and the coupling reaction was carried out in 4M N-methylmorpholine/DMF activator solution (PTI Instruments) with a 45min reaction time. The coupling reaction is done under nitrogen atmosphere, and the reaction is mixed every five minutes by bubbling nitrogen through the reaction vessel. Amino acid and activator were added at a four-fold molar excess relative to the resin load. After the coupling reactions were completed and the final Fmoc group removed, the loaded resin is washed with dichloromethane (DCM), DMF, and cold ether several times. A cleavage cocktail of



95% TFA and 5% TIPS was used to remove the peptide from the resin and the acid-labile protecting groups from the side chains. Purification was performed by the MIT biopolymers laboratory with a C18 HPLC column and two-phase buffer gradient: (Buffer A) 0.1% TFA in H<sub>2</sub>O and (Buffer B) 80% acetonitrile, 0.085% TFA in H<sub>2</sub>O. The mass spectrum of the final product is provided in the appendix (Fig. 2.a.1). For infrared studies the TFA was removed by lyophilizing against 20mM DCl in D<sub>2</sub>O repeated at least three times. This also served to exchange the labile protons for deuterons. Final peptide concentrations for IR studies were 10-20 mg/mL, or ~6-12 mM.

### Ser-Trp-Thr-Trp-Glu-Asn-Gly-Lys-Trp-Thr-Trp-Lys



**Figure 2.6** Stick diagram of TZ2 is pictured with the three-letter amino acid code above. Colored circles represent the position of the three amide group isotope labels: TZ2-S1, with <sup>18</sup>O-labeled Ser1 (green); TZ2-TT, with <sup>13</sup>C-labeled Thr3 and Thr10 (blue); and TZ2-K8, with <sup>13</sup>C-labeled Lys8 (red).

Three isotopologues of TZ2 were synthesized to probe vibrations localized to the turn, N-terminus, and mid-strand regions of the peptide. The mid-strand isotope-labeled compound (TZ2-TT) was prepared with a <sup>13</sup>C in the C' carbon of the Thr3 and Thr10 residues (SWTWENGKWTWK), which shifts the site energy of the Thr3-Trp4 and Thr10-Trp11 amide groups. TZ2-TT (TZ2-T3\*T10\* in Chapter 5) was synthesized and

purified by Anaspec Inc. (San Jose, CA) using L-Threonine-1-<sup>13</sup>C (Icon Isotopes, Summit, NJ), and the mass spectrum is provided in the appendix (Fig. 2.a.3). The turn-region isotope-labeled compound (TZ2-K8) was prepared with a <sup>13</sup>C atom in the C' position of the Lys8 residue (SWTWENG**K**WTWK). Isotope-labeled lysine (L-Lysine- $\alpha$ -N-FMOC, $\epsilon$ -N-t-BOC) was purchased from Cambridge Isotope Labs (CLM-6194, 1-<sup>13</sup>C, 99%) and used without any modification in synthesis and purification procedures identical to that of unlabeled TZ2 above. The mass spectrum of synthesized TZ2-K8 is provided in the appendix (Fig. 2.a.2).

The N-terminal label (TZ2-S1) was prepared with an <sup>18</sup>O label in the carbonyl oxygen of the Ser1 residue (SWTWENG**K**WTWK). Labeled serine was synthesized in our laboratory directly from Fmoc-Serine-OH (Novabiochem) using acid catalyzed hydrolysis. Anhydrous HCl gas is bubbled through Fmoc-Ser-OH in H<sub>2</sub><sup>18</sup>O:acetonitrile (approximately 1:4 by volume), and then heated at 80°C for four weeks. MALDI mass spectrometry of the amino acid was used to verify that the substitution was successful, with an approximately 75% isotopic enrichment of the carboxyl oxygen atoms (Appendix Fig. 2.a.5).<sup>22</sup> For use in peptide synthesis, the serine hydroxyl group is protected by a silylation reaction using *N*-methyl-*N*-(*tert*-butyldimethylsilyl)-trifluoroacetamide (MTBSTFA) in the procedure of Madson et al.<sup>23</sup> Specifically, 139 mg of <sup>18</sup>O labeled serine is dissolved in 15 mL DMF and 3 mL MTBSTFA, where it silylates both the side chain hydroxyl group and the carboxylic acid. The reaction solution is lyophilized and the remaining peptide is then suspended in hexane (20 mL). Ethanol (24 mL) is added to selectively de-silylate the carboxylic acid. The resulting amino acid was used in an Fmoc-SPPS synthesis identical to that of TZ2-UL. MALDI\_TOF mass spectrometry of

the labeled TZ2-S1 verified the mass of the isotope-labeled peptide indicating that the isotope labeled Ser was incorporated into the peptide (Appendix Fig. 2.a.4). FTIR spectra for each compound are shown in Chapter 7. The amide I spectrum of each isotopologue displays the spectral changes that are expected from the isotopic substitutions.

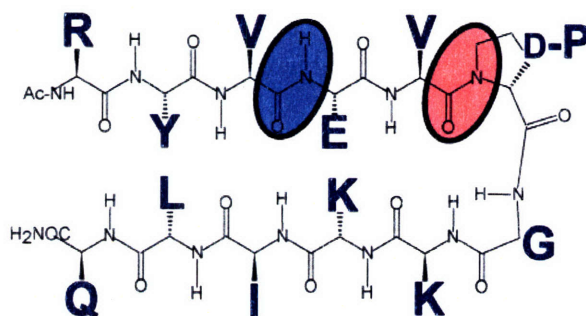
### 2.7.2. Synthesis of PG12

One of the early sets of stable hairpin peptides was developed by S. Gellman and coworkers.<sup>24,25</sup> The most stable of these species uses the non-natural D-isomer of proline to sterically enforce a type II'  $\beta$ -turn. Optical and NMR spectroscopy studies have verified the anti-parallel  $\beta$ -sheet structure of these peptides, and quantified their stability.<sup>26,27</sup> The D-proline turn was used in one of the trpzip peptides (TZ3), although the tryptophan core and D-proline turn stabilities did not turn out to be additive.<sup>9</sup> The sequence of PG12 from the N to C-terminus is: Arg1-Tyr2-Val3-Glu4-Val5-D-Pro6-Gly7-Lys8-Lys9-Ile10-Leu11-Gln12-NH<sub>2</sub> (RYVEVpGKKILQ). As with TZ2, PG12 was synthesized as a C-terminal amide.

PG12 was also prepared in house using Fmoc-based SPPS on a PS3 peptide synthesizer (Protein Technologies Incorporated). The resin used was Rink Amide MBHA (Novabiochem), and the Fmoc-protected amino acids were obtained from PTI Instruments in pre-packaged cartridges at four molar excess HBTU activator and amino acid. Reactions were carried out in DMF (EM-DX1732; EMD Chemicals; 99.9% purity). Removal of the Fmoc groups at each step was done with 20% piperidine in DMF (7 min. reaction time), and the coupling reaction was carried out in 4M N-

methylmorpholine/DMF activator solution (45 min. reaction time). Amino acid and activator were added at a four-fold molar excess relative to the resin load. After the coupling reactions were completed and the final Fmoc group removed, a cleavage cocktail of 95% TFA and 5% TIPS was used to remove the peptide from the resin and the acid-labile protecting groups from the side chains. Purification was performed by the MIT Biopolymers Laboratory with a C18 HPLC column and two-phase buffer gradient: (A) 0.1% TFA in H<sub>2</sub>O and (B) 80% acetonitrile, 0.085% TFA in H<sub>2</sub>O. For infrared studies the TFA was removed by lyophilizing against 20 mM DCl in D<sub>2</sub>O three to four times.

### Arg-Tyr-Val-Glu-Val-D-Pro-Gly-Lys-Lys-Ile-Leu-Gln



**Figure 2.7** Stick diagram of PG12 is pictured with the three-letter amino acid code above. Colored circles represent the position of the two amide group isotope labels: PG12-V3 with <sup>13</sup>C-labeled Val3 (blue); and PG12-V5 with <sup>13</sup>C-labeled Val5 (red).

Local structure is obtained from PG12 by incorporating <sup>13</sup>C labels at the C' position of the valine residues. Three isotopologues were synthesized: PG12-V3 and PG12-V5 with single <sup>13</sup>C' labels at the Val3 and Val5 sites respectively, and PG12-V3V5 with <sup>13</sup>C' labels at both valines. Infrared spectra for each isotopologue are shown in Chapter 5. Isotope-labeled valine was purchased from Cambridge Isotope Labs (CLM-3640-1; Lot

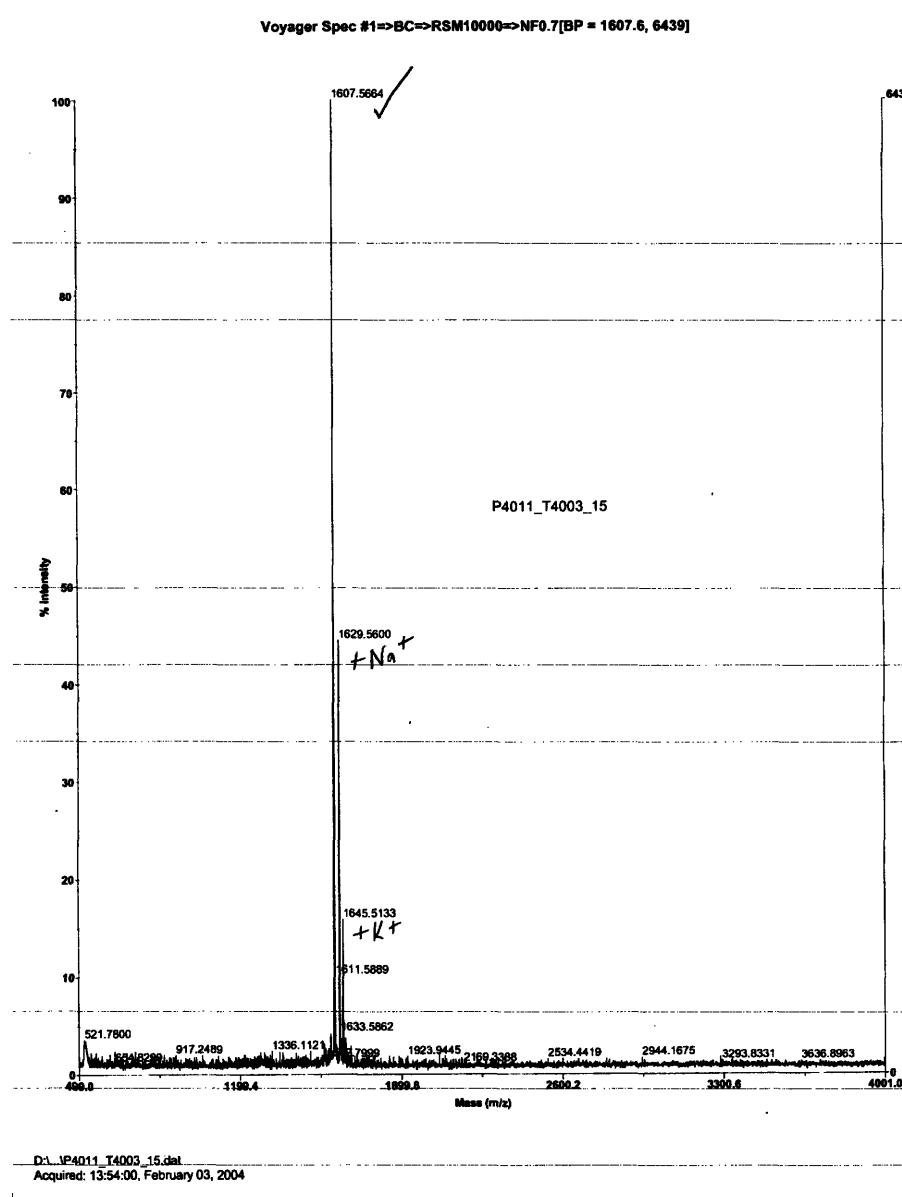
PR-14774) and used without further purification. Each sample was synthesized in our lab using the procedures outlined above. Final purification was performed by the MIT Biopolymers Laboratory as described above.

## 2.8. References

- (1) Merrifield, R. B. *J. Am. Chem. Soc.* 1963, *85*, 2149-2154.
- (2) Fields, G. B.; Noble, R. L. *Int. J. Pept. Protein Res.* 1990, *35*, 161.
- (3) Chan, W. C.; White, P. D., Eds. *Fmoc Solid Phase Peptide Synthesis*; Oxford University Press: Oxford, 2000.
- (4) Sherrington, D. C. *Chem. Commun.* 1998, 2275-2286.
- (5) Wang, S.-S. *J. Org. Chem.* 1976, *41*, 3258-3261.
- (6) Rink, H. *Tetrahedron Letters* 1987, *28*, 3787-3790.
- (7) Bernatowicz, M. S.; Daniels, S. B.; Koster, H. *Tetrahedron Letters* 1989, *30*, 4645-4648.
- (8) Knorr, R.; Trzeciak, A.; Bannwarth, W.; Gillessen, D. *Tetrahedron Letters* 1989, *30*, 1927-1930.
- (9) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* 2001, *99*, 9081-9081.
- (10) Streicher, W. W.; Makhatadze, G. I. *J. Am. Chem. Soc.* 2006, *128*, 30-31.
- (11) Pitera, J. W.; Haque, I.; Swope, W. C. *J. Chem. Phys.* 2006, *124*, 141102.
- (12) Bour, P.; Keiderling, T. A. *J. Phys. Chem. B* 2005, *109*, 23687-23697.
- (13) Yang, W. Y.; Gruebele, M. *J. Am. Chem. Soc.* 2004, *126*, 7758-7759.
- (14) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* 2004, *336*, 241-251.
- (15) Ulmschneider, J. P.; Jorgensen, W. L. *J. Am. Chem. Soc.* 2004, *126*, 1849-1857.

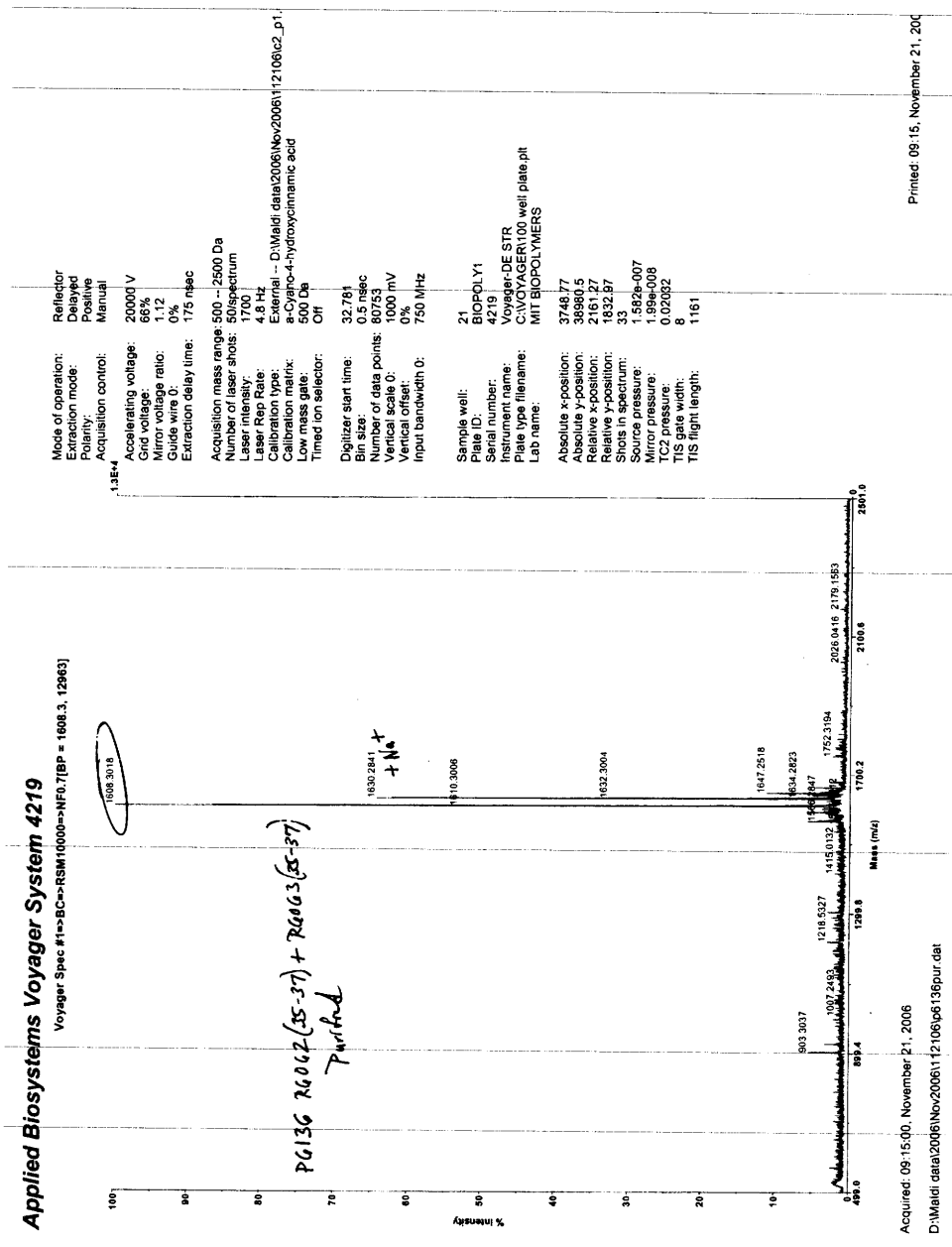
- (16) Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. *J. Am. Chem. Soc.* 2003, *125*, 388-395.
- (17) Du, D.; Tucker, M. J.; Gai, F. *Biochemistry* 2006, *45*, 2668-2678.
- (18) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 4077-4082.
- (19) Du, D.; Zhu, Y.; Huang, C.-Y.; Gai, F. *Proc. Natl. Acad. Sci. USA* 2004, *101*, 15915-15920.
- (20) Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. *J. Phys. Chem. B* 2006, *110*, 7545-7555.
- (21) Okur, A.; Strockbine, B.; Hornak, V.; Simmerling, C. *J. Comput. Chem.* 2002, *24*, 21-31.
- (22) Jaume Torres, A. K., Jonathan M. Goodman, Isaiah T. Arkin, *Biopolymers* 2001, *59*, 396-401.
- (23) Mawhinney; Madson *J Org Chem* 1982, *47*, 3336-3339.
- (24) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* 1998, *120*, 4236-4237.
- (25) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* 2002, *11*, 1492-1505.
- (26) Hilario, J.; Kubelka, J.; Syud, F. A.; Gellman, S. H.; Keiderling, T. A. *Biopolymers* 2002, *61*, 233-236.
- (27) Hilario, J.; Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* 2003, *125*, 7562-7574.

## 2.9. Appendix

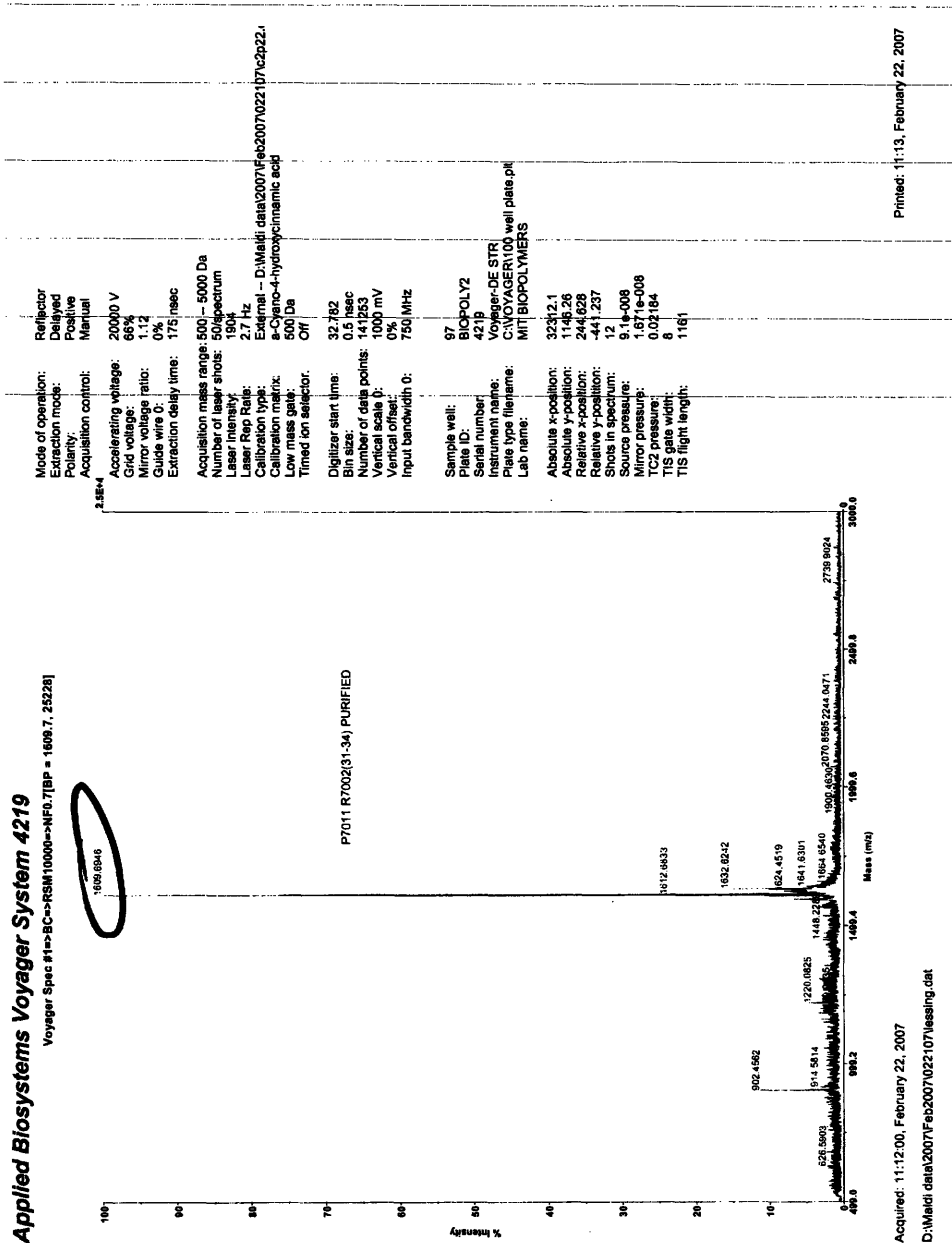


**Figure 2.a.1** TZ2-UL mass spectrum recorded on an Applied Biosystems Voyager MALDI-TOF mass spectrometer on February 3, 2004. The expected mass of TZ2-UL is 1608 g/mol, and the recorded mass is 1607.6 g/mol

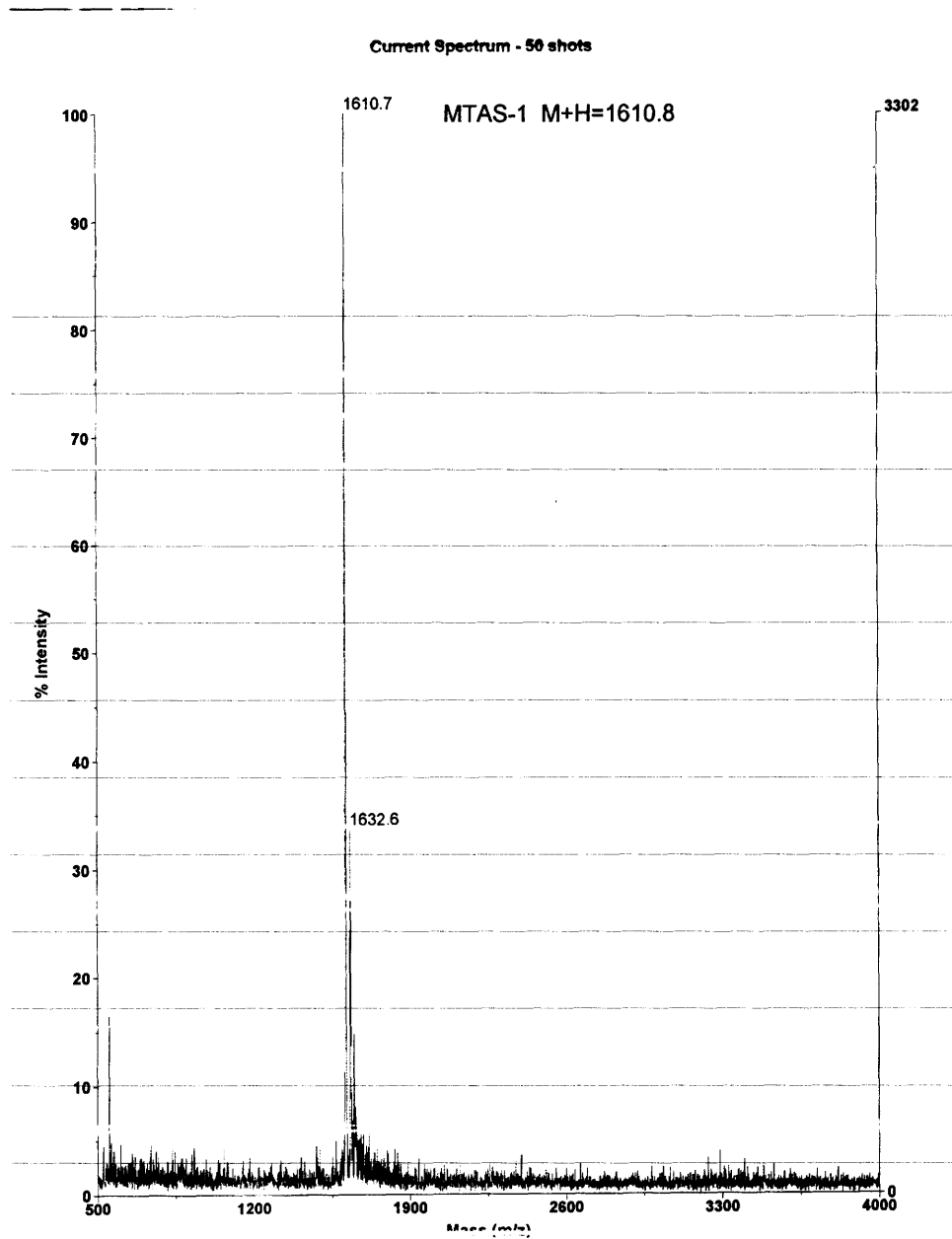




**Figure 2.a.2** TZ2-K8 mass spectrum recorded on an Applied Biosystems Voyager MALDI-TOF mass spectrometer on November 21, 2006. The expected mass of TZ2-K8 is 1609 g/mol. The recorded mass is 1608.3 g/mol, which is approximately 1 a.m.u. higher than the recorded mass of TZ2-UL (Fig. 2.a.1).

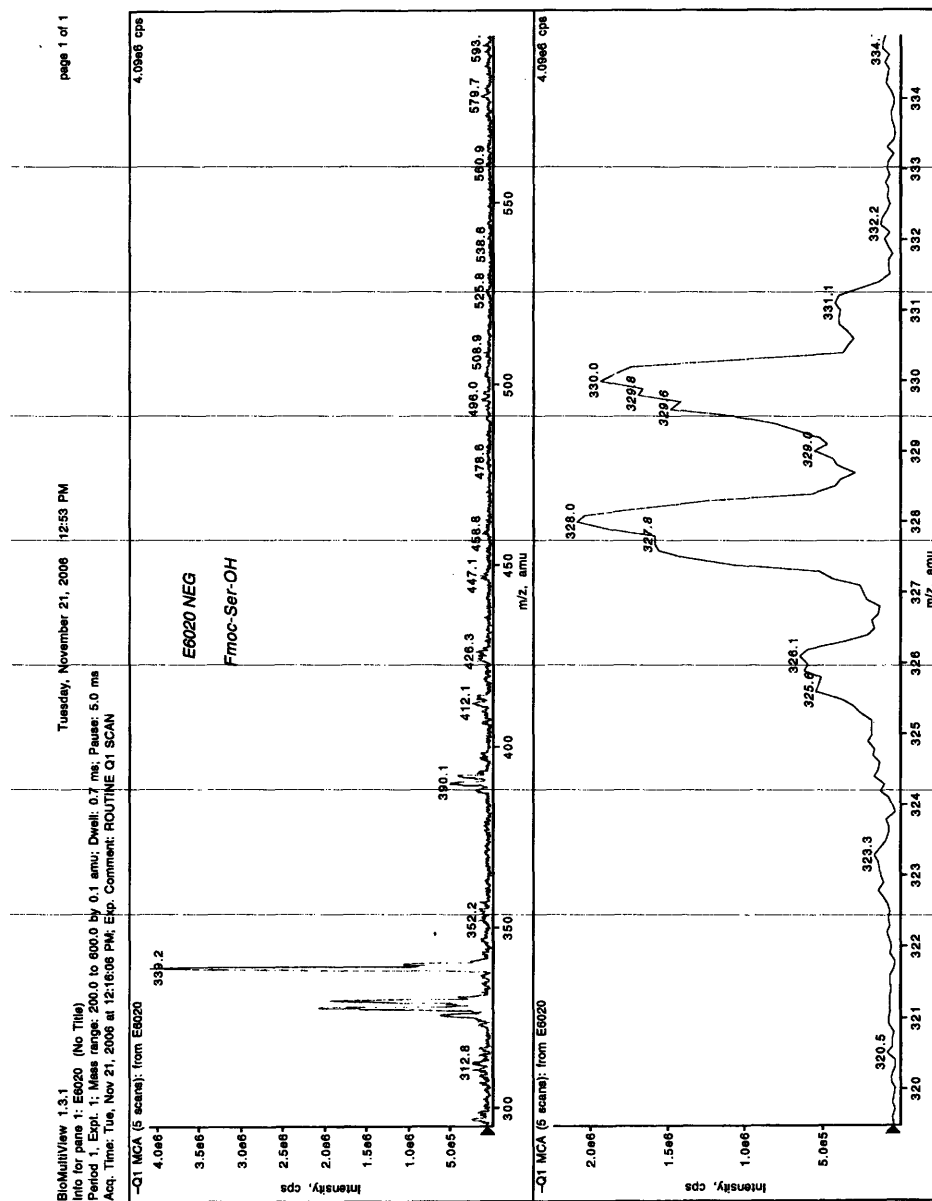


**Figure 2.a.3** TZ2-S1 mass spectrum recorded on an Applied Biosystems Voyager MALDI-TOF mass spectrometer on February 22, 2007. The expected mass of TZ2-S1 is 1610 g/mol. The recorded mass is 1609.6 g/mol, which is approximately 2 a.m.u. higher than the recorded mass of TZ2-UL (Fig. 2.a.1).



**Figure 2.a.4** TZ2-TT mass spectrum provided by Anaspec Inc. The expected mass of TZ2-TT is 1610 g/mol, which agrees with the recorded mass of 1610.7.





**Figure 2.a.5** Mass spectrum of  $^{18}\text{O}$  labeled serine. Because there are two labile oxygen atoms in the carboxylic acid group, three peaks appear in the mass spectrum. The peaks correspond to the  $^{16}\text{O}$ ,  $^{16}\text{O}$  isotope (326.1 a.m.u.), the  $^{16}\text{O}$ ,  $^{18}\text{O}$  isotope (328.0 a.m.u), and the  $^{18}\text{O}$ ,  $^{18}\text{O}$  isotope (330.0 a.m.u).



# Chapter 3

## Experimental Infrared Probes

### 3.1. FTIR Spectroscopy

FTIR spectroscopy is a linear spectroscopy that reports the absorbance of the system as function of frequency. At each spectral position, the absorbance is linearly proportional to the concentration of the peptide, the molar absorptivity of the sample, and the pathlength of the sample cell. For the data presented here, the sample solution was sandwiched between two CaF<sub>2</sub> windows using a 50 μm thick Teflon spacer that defines the cell pathlength. The 50 μm pathlength was chosen to be long enough to maximize peptide absorbance, but also short enough to avoid significant absorbance by the solvent. Sample solutions were prepared in buffered D<sub>2</sub>O solutions to eliminate the H<sub>2</sub>O bend absorbance peaks in the amide I region of the spectrum. FTIR data were collected on two spectrometers. The first is a Mattson Infinity FTIR spectrometer (Model #960M0013, Mattson Instruments, Madison WI), optimized for measurement of mid-IR wavelengths at up to 0.5 cm<sup>-1</sup> resolution. It uses a tungsten halogen lamp as the IR light source, and a

mercury cadmium telluride element at room temperature as the detector. The second instrument is a Nicolet 380 FTIR spectrometer (Thermo-Electron Corporation) with a high resolution add-on giving  $0.5 \text{ cm}^{-1}$  of spectral resolution.

Temperature dependant IR spectra are reported throughout this work. To obtain these data, sample temperatures are controlled by mounting the sample cell in a brass mount and circulating water through the housing. The temperature of the water bath is controlled with a recirculating water chiller (Neslab RTE-10; Thermo Scientific). The design of the sample cell is discussed in detail below (Sec. 3.6). To measure the actual temperature of the sample cell, a thermocouple was attached to the brass insert.

## **3.2. 2D IR Spectroscopy**

Two-dimensional infrared (2D IR) spectroscopy and dispersed vibrational echo (DVE) spectroscopy are both used extensively in this work to characterize the amide I vibrational band of  $\beta$ -hairpin peptides. Both methods are third-order nonlinear experiments in which the sample is interrogated with three input light fields producing a signal field that is emitted in a wave-vector matched direction. The details of the experiment are found in previous publications from our group.<sup>1,2</sup> Below are the details that pertain to the work presented in this thesis.

### **3.2.1. Pulse generation and interferometer.**

Multidimensional IR spectroscopy is a four-wave mixing experiment, in which three laser pulses interrogate a sample and generate a coherent signal with a defined frequency and wavevector.<sup>1</sup> Here, the excitation and signal fields have the same center

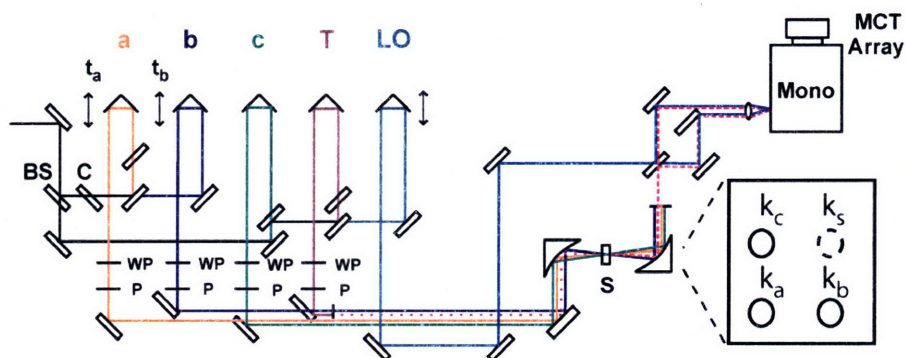


frequency,  $\sim 6 \mu\text{m}$ , and are generated by a two-stage OPA seeded with an amplified, 90 fs,  $\lambda = 790 \text{ nm}$  pump source. The 790 nm light is generated in Ti:Sapphire, mode-locked laser (Tsunami, Spectra-Physics), which is pumped with the output of a continuous wave 532 nm diode laser (Millenia V, Spectra-Physics). The oscillator produces an 82 MHz pulse train with 375 mW of average power. A portion ( $\sim 30\%$ ) of this light is amplified at a kHz repetition rate in a Ti:Sapphire regenerative amplifier (Spitfire, Spectra-Physics). Pulses are inserted into the cavity with a Pockels cell, and pass through the gain medium an average of 16 times before being ejected from the cavity with a second Pockels cell. The amplifier is pumped with an intracavity-doubled Nd:YLF diode pumped laser (Empower, Spectra-Physics). Pump powers of 9.2 W were used to obtain amplified pulse energies of 1.1 mJ, of which 740  $\mu\text{J}$  is used to seed the OPA.

Upon entering the OPA, 1 % of the 790 nm light is focused through a Ti:Sapphire window to generate white light. Fifteen percent of the 790 nm light is then overlapped with the white light in a type II  $\beta$ -barium borate (BBO) crystal ( $\theta = 27^\circ$ , 3 mm-thick), to generate a signal ( $\lambda = 1.4 \mu\text{m}$ ) and idler ( $\lambda = 1.82 \mu\text{m}$ ). The idler is separated from the signal and overlapped with the remaining 790 nm light in a second pass through the BBO crystal. Before mixing the signal and idler, the pulse energies are  $\sim 90$  and  $\sim 40 \mu\text{J}$  respectively. Difference frequency mixing of the signal and idler is realized in a  $\text{AgGaS}_2$  crystal (type II,  $\theta = 41.5^\circ$ ,  $\phi = 0^\circ$ ,  $8 \times 8 \times 1 \text{ mm}$ , Eksma). Typical pulses at  $6.0 \mu\text{m}$  have energies of 3  $\mu\text{J}$ , and bandwidths of  $160 \text{ cm}^{-1}$  (FWHM).

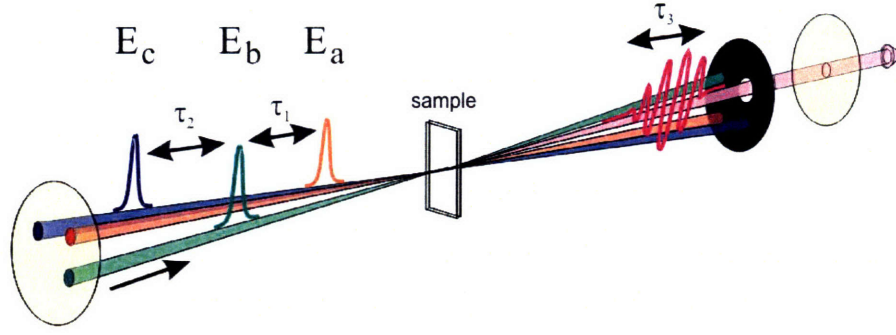
Output from the OPA is overlapped with a helium neon laser for alignment purposes and expanded to a beam diameter of 10-12 mm. With a series of beam splitters and optical delay lines the  $6 \mu\text{m}$  laser pulses are divided into five beams (Fig. 3.1) Three

beams with electric fields  $\mathbf{E}_a$ ,  $\mathbf{E}_b$ , and  $\mathbf{E}_c$ , and wavevectors  $\mathbf{k}_a$ ,  $\mathbf{k}_b$ , and  $\mathbf{k}_c$ , are arranged in a boxcar geometry, with the delay between  $\mathbf{E}_a$  and  $\mathbf{E}_b$  defined as  $\tau_1$  and the timing between  $\mathbf{E}_b$  and  $\mathbf{E}_c$  defined as  $\tau_2$  (Fig. 3.2). Pulse energies for each beam range from 150 to 200 nJ at the sample.



**Figure 3.1** Diagram of interferometer. Collimated 6.0  $\mu\text{m}$  light enters the setup from the left. Figure adapted from Ref. 3.

The three light fields are focussed to a 100  $\mu\text{m}$  diameter beam waist at the sample cell. The light fields interact with the sample and drive transitions between the 1 and 2 quantum vibrational eigenstates of the system. The oscillations after the third pulse generate a coherent field ( $\mathbf{E}_s$ ) that radiates from the sample in the direction of the fourth corner of the box ( $\mathbf{k}_s$ ). The timing between  $\mathbf{E}_c$  and  $\mathbf{E}_s$  is defined as  $\tau_3$ .



**Figure 3.2** Pulse arrangement and relative timings as they are directed through the sample.

### 3.2.2. 2D IR signal detection

For 2D IR spectra, the signal is heterodyne detected in order to resolve the phase information contained in the nonlinear signal. Upon emission from the sample,  $\mathbf{E}_s$  is overlapped with the fourth laser pulse called the local oscillator ( $\mathbf{E}_{LO}$ ,  $\mathbf{k}_{LO}$ ), and is directed through the spectrometer onto a 64 element array detector with a pixel size of 0.1 x 1 mm and 50  $\mu\text{m}$  dead space between pixels. For data presented in this work, two detection schemes are used. In the first, a 70/30 ZnSe beamsplitter is used to overlap  $\mathbf{E}_s$  and  $\mathbf{E}_{LO}$ . The signal component that is 70% reflected and overlapped with the transmitted local oscillator is imaged through a monochromator (Triax 190, Jobin Yvon) and on to a single 64-element mercury cadmium telluride (MCT) IR array (IR-0144, Infrared Systems Development). The grating used in the monochromator has a groove spacing of 40 lines/mm giving  $\sim 5 \text{ cm}^{-1}$  of resolution at 6.0  $\mu\text{m}$ . The signal is chopped at 500 Hz so that the local oscillator intensity can be subtracted from the combined  $\mathbf{E}_{LO}$  and  $\mathbf{E}_s$  fields ( $|E_s|^2 \ll |E_{LO}E_s|$ ):

$$S = |E_{LO} + E_s|^2 = |E_{LO}|^2 + 2 \text{Re}(E_{LO}E_s^*) + |E_s|^2. \quad (3.1)$$

Time delays between pulses are controlled with an independent optical delay line for each beam. To generate a 2D IR spectrum, the heterodyned interference signal is collected with the timing between  $E_s$  and  $E_{LO}$  ( $\tau_3$ ) set to zero. In this data collection scheme, the grating of the spectrometer effectively Fourier transforms the  $\tau_3$  time period to generate the  $\omega_3$  frequency axis. The waiting time,  $\tau_2$ , is held constant for each 2D IR spectrum, and is set to 0 or 100 fs for most of the data in this work. The evolution time,  $\tau_1$ , is step-scanned in 4 fs increments from 0 to 2 ps, and each time point is averaged for 1000 laser shots. This time window allows for approximately  $2.66 \text{ cm}^{-1}$  resolution in  $\omega_1$  ( $\Delta\omega/2\pi = 1/2\pi c\Delta t$ ;  $\Delta t = 2 \text{ ps}$ ).

The geometry and time-ordering between the first two pulses is used to collect two components of the final spectrum. First is the rephasing spectrum in which oscillations during  $\tau_1$  acquire a conjugate phase during  $\tau_3$ , giving rise to a macroscopic rephasing of the oscillations. In the non-rephasing spectrum, phase acquired in  $\tau_1$  is identical that of  $\tau_3$ . The rephasing and non-rephasing spectra are collected separately and added together in post-processing to give a 2D IR correlation spectrum with absorptive peaks.<sup>1</sup>

The second detection method used in this work is hereafter referred to as balanced detection. Balanced detection significantly reduces the signal-to-noise ratio of the heterodyned measurement by eliminating the baseline noise in the DC component of the interferograms. In this mode, a 50/50 ZnSe beamsplitter is used to combine the signal and local oscillator fields (Fig. 3.1). After the beamsplitter, the combined signal/local oscillator of each arm is independently directed to one of two 64 element arrays, which are vertically displaced from each other by 5mm. The signal measured on the upper and lower stripe,  $S^U$  and  $S^L$ , is given by the following equation:

$$\begin{aligned}
S^U &= |E_{LO} + E_s|^2 = |E_{LO}|^2 + 2 \operatorname{Re}(E_{LO} E_s^*) + |E_s|^2 \\
S^L &= |E_{LO} - E_s|^2 = |E_{LO}|^2 - 2 \operatorname{Re}(E_{LO} E_s^*) + |E_s|^2
\end{aligned}
\tag{3.2}$$

As in the regular detection mode, the signal on each stripe is chopped to remove the  $E_{LO}$  contribution. After this subtraction, the heterodyned signal of the lower strip is  $90^\circ$  out-of-phase with the upper strip, and the 2D signal is obtained by subtracting the two stripes:

$$S^{2D} = S^U - S^L \tag{3.3}$$

In practice, the signal to noise ratio using balanced detection is 10-15 times higher than in the single stripe detection.<sup>4</sup> Because of this, data is typically collected by undersampling the interferograms in  $\tau_1$  at 14 fs steps. By reducing the total experimental collection time, undersampling opens up a practical route for obtaining temperature-jump 2D IR data as described below.

After the data has been collected, a Matlab routine is used to find the precise  $\tau_1 = 0$  by fitting the projection of the complex 2D IR spectrum to the pump probe spectrum.<sup>4</sup> The pump-probe spectrum is collected by measuring the intensity of the tracer beam, which follows the  $\mathbf{k}_s$  beam path, with and without the probe beam,  $\mathbf{k}_c$ . The error in  $\tau_1$  is used as a phase factor,  $\exp(i\omega_1 \Delta \tau_1)$ , in the Fourier transform of the rephasing and non-rephasing spectra.

Once the heterodyned signal is collected and processed, the rephasing and non-rephasing are added together to produce a 2D IR spectrum that reveals information hidden or ambiguous in an FTIR spectrum. The bandwidth and multiple photon interactions of the 2D IR experiment allow experimental access to both the one and two-

quantum vibrational eigenstates of the system. Along the diagonal axis of the spectrum, two peaks are seen for each vibrational mode. The positive peak records stimulated emissions processes during the detection period ( $\tau_3$ ), while negative peaks represent bleaching transitions during  $\tau_3$ . Because of this, the splitting between diagonal peaks in a well-resolved spectrum is a direct measure of the vibrational anharmonicity. In the off-diagonal region of the spectrum, cross peaks connect the diagonal peaks of anharmonically coupled eigenstates. Peak widths are a probe of the system dynamics and can be used to separate out heterogeneous line-broadening from homogeneous line-broadening. Heterogeneity, for example, leads to elongation of the peaks in the diagonal dimension of the 2D IR spectrum because the eigenenergy is static on the time-scale of the pulsed experiment (1-2 ps). For vibrations that fluctuate on this picosecond time scale, the peaks are broadened symmetrically in the spectrum, most obviously in the anti-diagonal dimension. The information content of the peptide 2D IR spectra will be discussed in later chapters.

### 3.3. DVE Spectroscopy

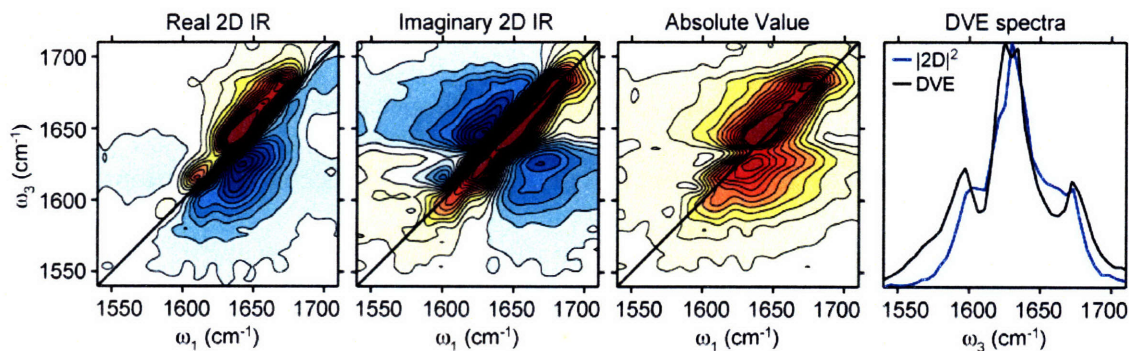
DVE spectroscopy is also a four-wave mixing method, but without the heterodyned detection used in 2D IR. For the data shown here, the experiments were performed identically to the 2D IR methods above, but the signal beam,  $E_s$ , is not overlapped with a local oscillator pulse. Instead the signal is dispersed directly onto the array, often without passing through the beamsplitter optic. Pulse delays are set to  $\tau_1 = 0$  and  $\tau_2 = 100$  fs. The non-zero  $\tau_2$  value is chosen to avoid spectral contributions from non-resonant scattering. For a typical data set, each spectrum is averaged for 200,000 laser shots. By not stepping

the  $\tau_1$  timing, DVE spectra are collected in about 3 minutes compared with the approximately 30-minute collection time for a 2D IR spectrum.

The DVE signal is extractable from a full 2D IR spectrum as it represents the homodyne component of the pre-processed 2D spectrum at  $\tau_1 = 0$  and  $\tau_2 = 100$  fs. The two methods are related to each other through the projection slice theorem of Fourier transforms. Numerically, the DVE spectrum can be reproduced from a projection of the complex 2D IR spectrum  $S_{2D} = \text{Re}[S_{2D}] + i\text{Im}[S_{2D}]$  onto the detection axis  $\omega_3$ :<sup>2</sup>

$$S_{DVE}(\omega_3) = \int d\omega_1 |S_{2D}(\omega_1, \omega_3)|^2. \quad (3.4)$$

With this in mind, features in the DVE spectrum are related not only to the transition dipoles of the eigenstates as in FTIR, but are also representative of the features of the 2D IR spectrum. For example, off-diagonal intensity in the 2D IR spectrum adds to diagonal peak intensity with the final effect of sharpening the peaks of the DVE spectrum relative to FTIR. Interference between the positive and negative peaks in the 2D IR spectrum lead to shifted peak positions in DVE relative to the measured frequencies in 2D IR and FTIR spectra. Also, diagonal and anti-diagonal line widths in 2D IR can directly affect DVE peak intensities, with broader widths generally leading to lower DVE intensity. Fig. 3.3 compares the directly measured DVE spectrum and the DVE spectrum reproduced from the 2D IR data of PG12.



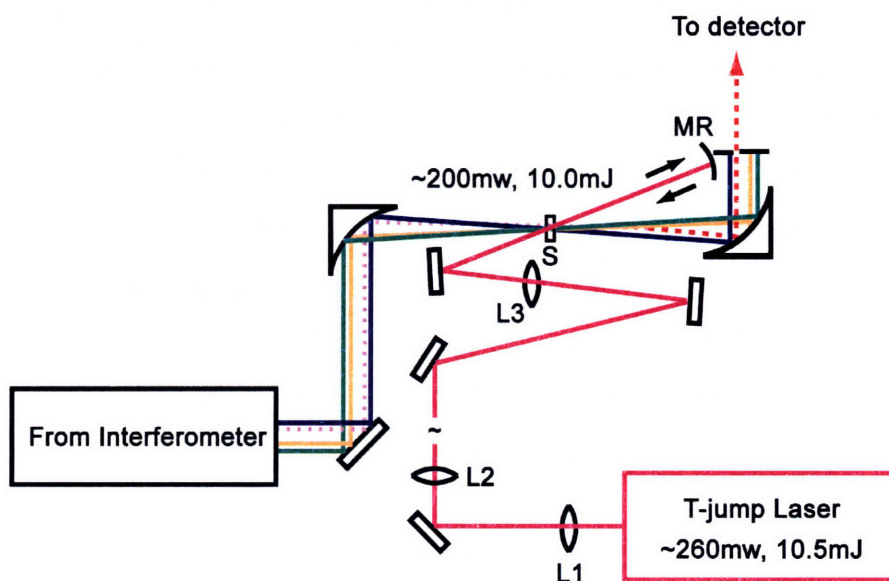
**Figure 3.3** 2D IR and DVE spectra are shown for PG12-UL. The first two panels are the real (absorptive) and imaginary (dispersive) 2D IR spectra of PG12-UL at 25 °C, and the third panel shows the absolute value of the sum of the real and imaginary 2D IR spectra. Panel four plots the directly-detected DVE spectrum at 25 °C and the DVE spectrum reproduced from the 2D IR.

### 3.4. Temperature-jump

To measure the fast kinetics of the thermal folding transition, a temperature-jump (T-jump) perturbation was effected on the sample using a nanosecond laser pulse. The T-jump laser pulse was generated by pumping an OPO with the second harmonic output of a Q-switched Nd:YAG laser (YG981c, OPOTEK) at a repetition rate of 20 Hz. The OPO idler output ( $\lambda = 1.98 \mu\text{m}$ ) was used to pump OD stretch overtones in the  $\text{D}_2\text{O}$  solvent, which thermally equilibrate within the six nanosecond pulse duration. Delivery of the T-jump light to the sample cell is described in previous publications for data shown in Chapter 6.<sup>5,6</sup> For later data, some of the optics were replaced with new substrates and coatings to reduce to absorptive and reflective loss. In this new setup, the light from the OPO was expanded and collimated with a two-lens Keplerian telescope (L1: PLCX-50.8-51.5-UV-2010nm, CVI laser; L2: LA4855, Thorlabs Inc.) to a beam diameter of approximately 30 mm. Light was focused to a  $\sim 850 \mu\text{m}$  beam waist diameter in the



sample cell using a 75 mm fl lens (PLCX-50.8-38.6-UV-2010nm). This lens setup delivered 180-200 mW (9-10 mJ/pulse) of 1.98  $\mu\text{m}$  laser light at the sample cell position. An example of a calibrated temperature rise in this single pass geometry is  $9.8 \pm 0.8$   $^{\circ}\text{C}$  (Chapter 8). Details of the calibration procedure will be given below.



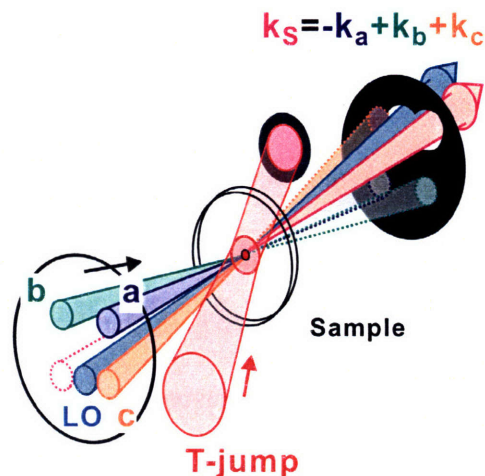
**Figure 3.4** Diagram of T-jump laser pulse beam path. The magenta line traces the path of the 1.98  $\mu\text{m}$  T-jump laser. Labels refer to: (L1) 2", UV fused silica lens, 100 mm fl, AR2010; (L2) 2", UV fused silica lens, 300 mm fl, no AR; (L3) 2", UV fused silica lens, 75 mm fl, AR2010; (MR) 1", BK7 / protected gold mirror, 75 mm radius; (S) sample cell.

To further increase the size of the T-jump, a retro-reflecting mirror (PG-SMCC-1037-0.075-C, CVI Laser Corp) was positioned behind the sample cell to redirect the transmitted light back into the sample (Fig. 3.4). The transmittance of the sample solution is typically 85-90 %, and so the second pass nearly doubles the energy absorbed by the sample. In this new geometry, T-jumps of 18 to 24  $^{\circ}\text{C}$  were recorded.

### 3.5. Temperature-Jump Probes

DVE and 2D IR spectroscopy were both used to measure the amide I spectrum of hairpin peptides following the T-jump perturbation described above. Experimental details have been described extensively in a previous publication.<sup>5</sup> The synchronization of the 1 kHz, 6.0  $\mu\text{m}$  IR probe pulse train with the 20 Hz, 1.98  $\mu\text{m}$  T-jump pulse train was controlled with an electronic delay generator (DG535, Stanford Research Systems). The difference between the two pulse frequencies means that every T-jump laser pulse is followed by 50 IR probe pulses. The delay between the T-jump pulse and the first of the 50 probe pulses is the T-jump delay time ( $\tau$ ).

For DVE experiments,  $\tau$  was stepped in small increments from -10 ns to 0.9 ms typically with 3-4 time points per decade of delay time. At each delay, data was collected in-phase with an optical chopper for 5000 shots and 180° out-of-phase for 2000 shots. Data collected in the two phases was subtracted to produce the final signal as described in previous reports.<sup>5</sup> The total time to step through all timings was about two hours, which was repeated sequentially at least five times for each of the data sets shown in subsequent chapters.



**Figure 3.5** Enlarged illustration of the sample cell region. The spot size of the T-jump laser is 850  $\mu\text{m}$  in diameter, which is larger than the spot size of the probe pulses (100  $\mu\text{m}$  in diam.). The third order signal is generated by the interaction of the three femtosecond pulses a, b, and c, and emitted to the wave vector matched direction  $\mathbf{k}_s = -\mathbf{k}_a + \mathbf{k}_b + \mathbf{k}_c$ . Three incoming pulses are blocked by the mask to reduce the scattering. The local oscillator pulse passes through the sample at the same spot as the other three pulses. (Figure adapted from Ref. 4)

T-jump 2D IR was collected in the same way as reported previously by our group.<sup>5</sup> The primary difference between the T-jump 2D spectra and those taken at equilibrium is that the  $\tau_1$  time axis was undersampled at 14 fs steps rather than the 4 fs steps used to collect data in the equilibrium spectra. This made it experimentally possible to collect T-jump 2D IR spectra on a reasonable time-scale. Typically,  $\tau_1$  was collected out to 1.5 and 1.2 ps for rephasing and non-rephasing geometries respectively. Each  $\tau_1$  time point was averaged for 5000 shots leading to a 2.5 hour collection time for one 2D spectrum at a single  $\tau$  delay time. Each T-jump 2D IR spectrum consists of at least four 2.5-hour averages.

Because 2D IR is a heterodyned measurement, it relies on differencing the local oscillator and the local oscillator plus signal. This is usually performed by optically chopping the signal at half the repetition rate of the IR pulse train. The temperature variation between sequential probe pulses following the T-jump laser pulse excludes the use of sequentially chopped pulses to produce the heterodyne signal. Instead, the difference is obtained by subtracting the measured signal of the first pulse from that of the 49<sup>th</sup> pulse to obtain a 2D difference signal. This is shown in the following expression for the recorded signal,  $S$ , as a function of T-jump delay ( $\tau$ ) and pulse number,  $n$ , following the T-jump.

$$\begin{aligned}\Delta S(\tau, n) &= [S_0(\tau, n) - S_\pi(\tau, n)] - [S_0(\tau, 49) - S_\pi(\tau, 49)], & \text{for odd } n \\ &= [S_\pi(\tau, n) - S_0(\tau, n)] - [S_\pi(\tau, 50) - S_0(\tau, 50)], & \text{for even } n\end{aligned}\quad (3.5)$$

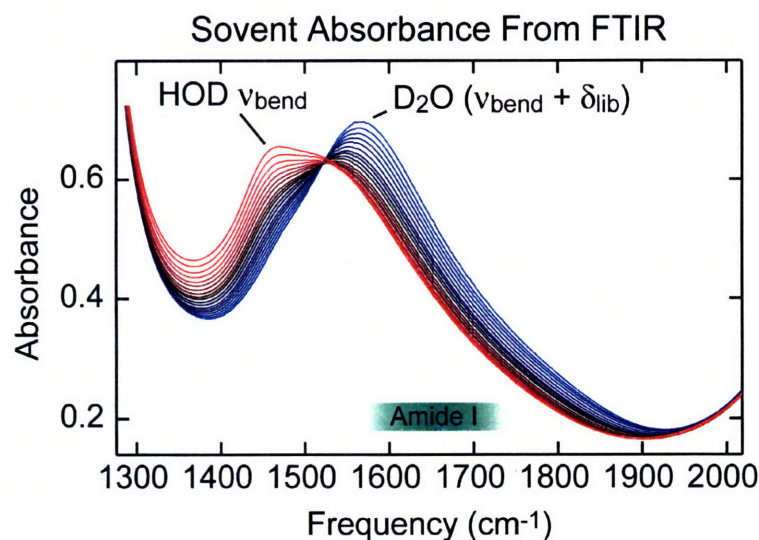
When full 2D spectra are displayed at a delay time,  $\tau$ , they are reconstructed by adding the directly detected difference signal to the ‘equilibrated’ spectrum, which is constructed with the 49<sup>th</sup> and 50<sup>th</sup> pulses.

### 3.6. Temperature-Jump Calibration from FTIR

Calibration of the T-jump relies on the well-characterized temperature-dependant absorbance of the D<sub>2</sub>O solvent. This is shown for the spectral region around amide I in Fig. 3.6. As can be seen, absorbance of the solvent decreases with temperature from 1550 to 1900 cm<sup>-1</sup>. To calibrate the size of the T-jump, the intensity ( $I(\tau, \omega)$ ) of a tracer beam was measured as a function of the delay time,  $\tau$ , following the T-jump laser pulse. The differential absorbance was calculated as:

$$\Delta A(t, \omega) = -\log_{10} \left( \frac{I(\tau, \omega)}{I_0(\omega)} \right), \quad (3.6)$$

where  $I_0(\omega)$  is the intensity of the probe pulses after the solvent temperature has relaxed to equilibrium.  $I_0(\omega)$  is obtained by averaging pulses the last few pulses following the T-jump pulse (i.e.  $n = 48, 49,$  and  $50$ ).



**Figure 3.6** FTIR spectra of D<sub>2</sub>O solvent as a function of temperature from 5 °C (blue) to 95 °C (red) at 5 °C increments in a 50 μm pathlength cell. An air background was used as the reference spectrum to calculate absorbance. Solvent was prepared with a pH of 2.5 as described in Chapter 7. The green bar shows the spectral region of primary interest in this thesis.

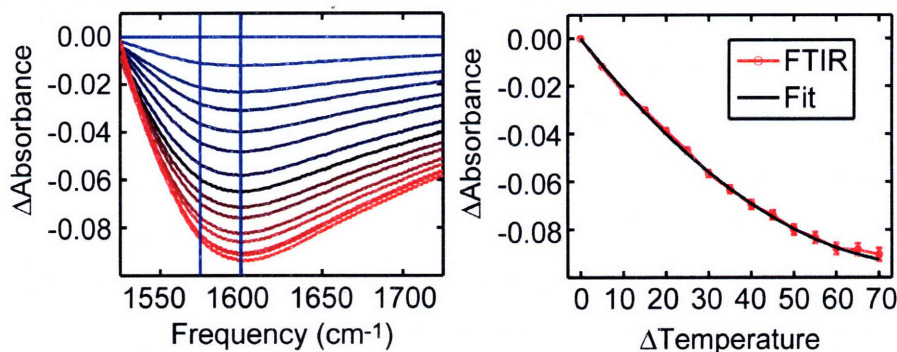
The temperature-dependant absorbance of the buffer in the amide I region of the spectrum tracks thermal changes to a D<sub>2</sub>O combination band ( $\nu_{\text{bend}} + \delta_{\text{lib}}$ ), and an HOD bend.<sup>7</sup> Thermally-induced intensity changes to the absorbance at  $\sim 1600 \text{ cm}^{-1}$  can be fit to a second-order polynomial in temperature.<sup>8</sup> In the equation below, we fit the differential absorbance and differential temperature,

$$\begin{aligned} \Delta A(\Delta T, \omega) &= A(T, \omega) - A(T_0, \omega) \\ \Delta T &= T - T_0 \end{aligned} \quad (3.7)$$

to a second-order polynomial:

$$\Delta A(\Delta T, \omega) = B_1(\omega) \cdot \Delta T^2 + B_2(\omega) \cdot \Delta T + B_3(\omega). \quad (3.8)$$

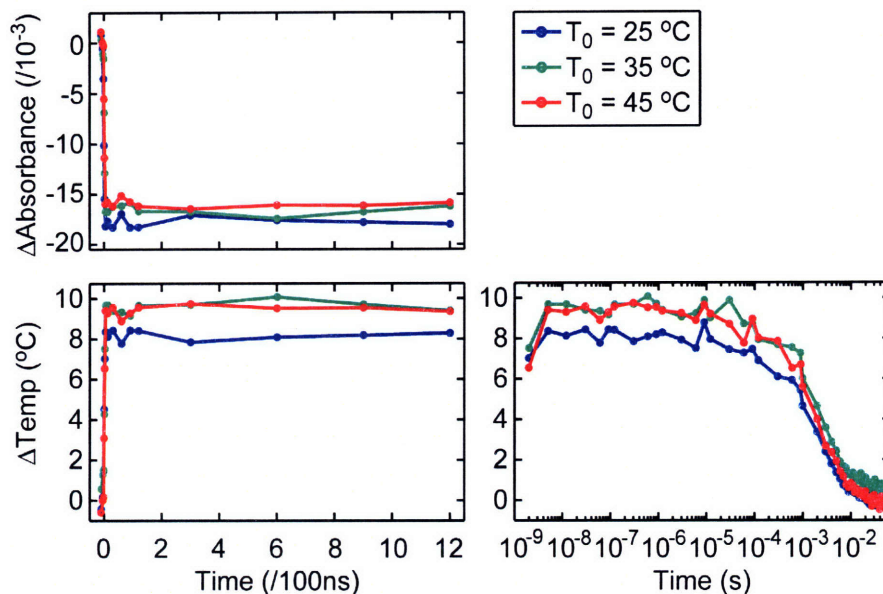
The constants are determined for a given equilibrium temperature,  $T_0$ , and a set of frequencies, after which the equation can be inverted to translate the T-jump transient absorbance to transient temperature.



**Figure 3.7** Change in absorbance from FTIR of  $D_2O$  solvent as a function of temperature from 25 °C (blue) to 95 °C (red) at 5 °C increments. Each spectrum shown on left is obtained by subtracting off the spectrum at  $T_0 = 25$  °C as a reference. An air background was used as the reference spectrum to calculate absorbance. Solvent was prepared with a pH of 2.5 as described in Chapter 7.

Calibration of a T-jump scan is shown below for a pH 2.5 buffer solution at three temperatures. The rise time of the temperature is 6 ns. At a time delay of 75 ns there is a repeatable decrease in the temperature that we interpret as a density wave propagating through the sample. The speed of sound in  $D_2O$  is 1500 m/s, which would lead to a 40-50 ns time to propagate across the diameter of the probe beam.<sup>5</sup> The temperature stays relatively constant to delay times of 10 to 100  $\mu$ s. Re-equilibration of the sample

temperature is best fit to a stretched exponential with a time constant of 2.7 ms and a stretching factor,  $\beta$ , equal to 0.63.<sup>5</sup>

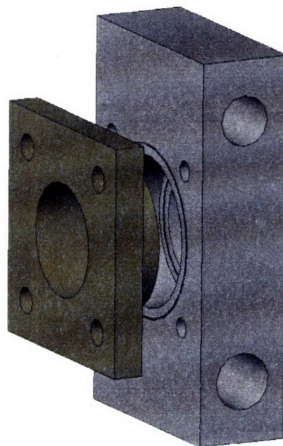


**Figure 3.8** Calibrated changes to the solvent temperature during a transient, T-jump event. Left plots show the transient absorbance (top) and temperature (bottom) determined from the calibration procedure described above. The fitted parameters for each starting temperature are:  $B_1= 1.36e^{-5}$ ,  $B_2= -2.27e^{-3}$ , and  $B_3= -1.28e^{-4}$  at 25 °C;  $B_1= 1.39e^{-5}$ ,  $B_2= -2.02e^{-3}$ , and  $B_3= 1.55e^{-3}$  at 35 °C; and  $B_1= 1.59e^{-5}$ ,  $B_2= -1.85e^{-3}$ , and  $B_3= 7.32e^{-5}$  at 45 °C. The right plot shows the calibrated temperature on a log scale from 1 ns to 50 ms.

### 3.7. Sample Cell

The conformational changes studied within this work are induced by controlling the temperature of the sample. Raising the temperature of the solvent by 20-60 °C above room temperature denatures the hairpin peptide relative to the known folded structure. This section briefly describes a new sample cell design, which is used to control the sample temperature and provide greater flexibility and transferability between

spectrometers. The previous design in our group used a brass housing that was drilled to allow water from a recirculating chiller to control the temperature. There was a cylindrical opening machined into the mount where the CaF<sub>2</sub> windows were secured into place with a faceplate (Fig. 3.9).

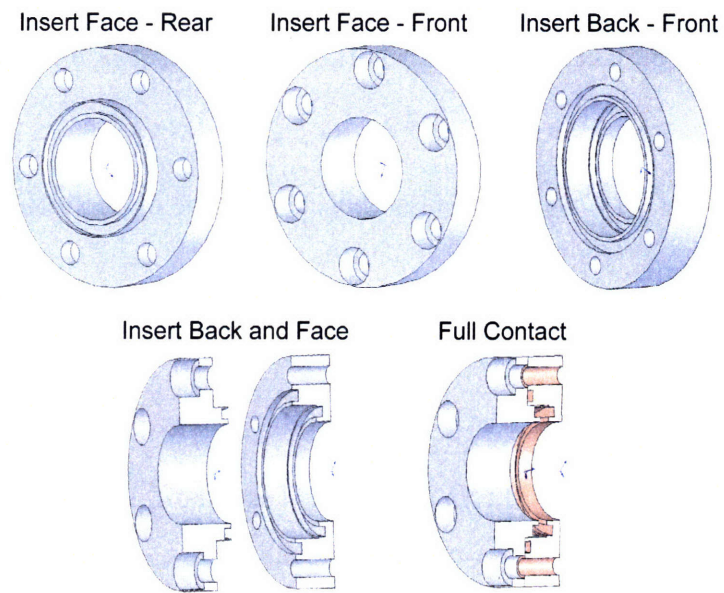


**Figure 3.9** Three-dimensional rendering of sample cell used in the work shown in Chapters 4 & 5. Four screws held the faceplate to the housing, and three O-ring grooves sealed the sample cell from the surrounding environment.

Once the sample was loaded and circulating water lines were in place, it was prohibitive to transfer the entire cell between the FTIR and 2D IR spectrometers because of the attached circulating water lines. To overcome this, a removable cell was designed and placed in a similar housing for temperature control. The cell is made of two parts that can be sealed on a bench top and then placed in the housing dedicated to the spectrometer of interest. In this way, FTIR and 2D IR spectra can be taken on identical sample volumes without wasting peptide sample to fill two cells. Furthermore, sample can be left in the sealed cells, while spectra are taken for other systems. This is important not only for exact spectral comparisons, but also to avoid losing expensive isotope-



labeled sample. Fig. 3.10 shows the three-dimensional rendering of the sample cell holder. The housing (not shown) was a rectangular slab of brass with holes drilled for water circulation and the positioning of the sample cell holder. Drawings for the housing and sample cell are included in the Appendix.



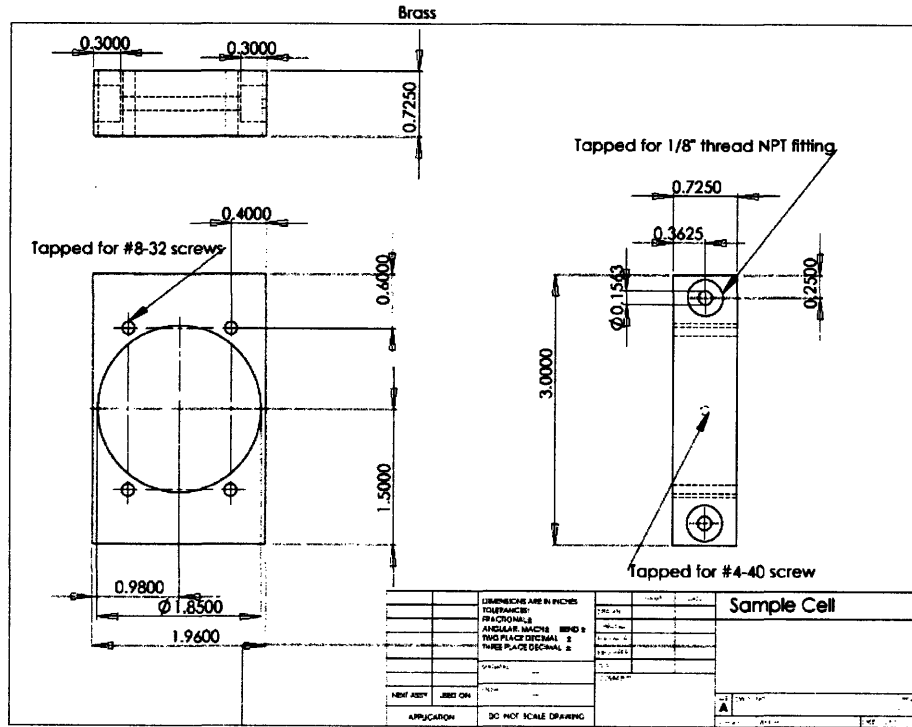
**Figure 3.10** Three-dimensional rendering of sample cell used in the work shown in Chapters 6, 7, and 8.

### 3.8. References

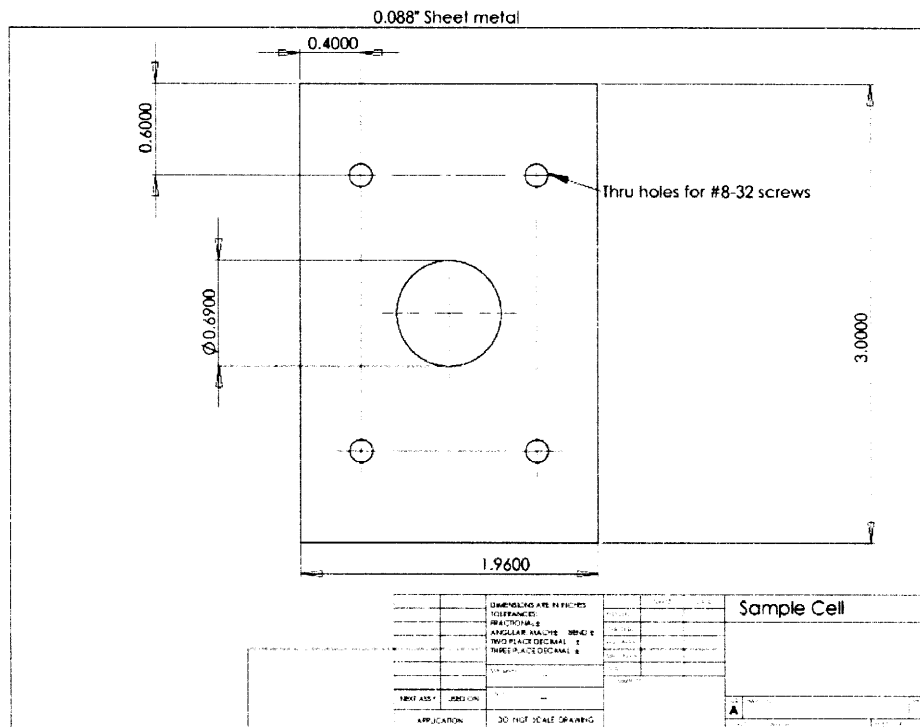
- (1) Khalil, M.; Demirdoven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258-5279.
- (2) Chung, H. S.; Khalil, M.; Tokmakoff, A. *J. Phys. Chem. B* **2004**, *108*, 15332-15343.
- (3) Loparo, J. J. In *Chemistry*; Massachusetts Institute of Technology: Cambridge, MA, 2007, p 211.
- (4) Chung, H. S. In *Chemistry*; Massachusetts Institute of Technology: Cambridge, MA, 2007, p 231.
- (5) Chung, H. S.; Khalil, M.; Smith, A. W.; Tokmakoff, A. *Rev. Sci. Instrum.* **2007**, *78*.
- (6) Smith, A. W.; Tokmakoff, A. *Angewandte Chemie International Edition* **2007**, *46*, 7984-7987.
- (7) Walrafen, G. E. *J. Chem. Phys.* **1964**, *40*, 3249.
- (8) Gai, F.; Du, D.; Xu, Y. In *Protein Folding Protocols*; Bai, Y., Nussnov, R., Eds.; Humana Press: Totowa, NJ, 2007; Vol. 350, pp 1-20.

### 3.9. Appendix

#### 3.9.1. Cell Housing (brass)



### 3.9.2. Mounting Plate (steel)

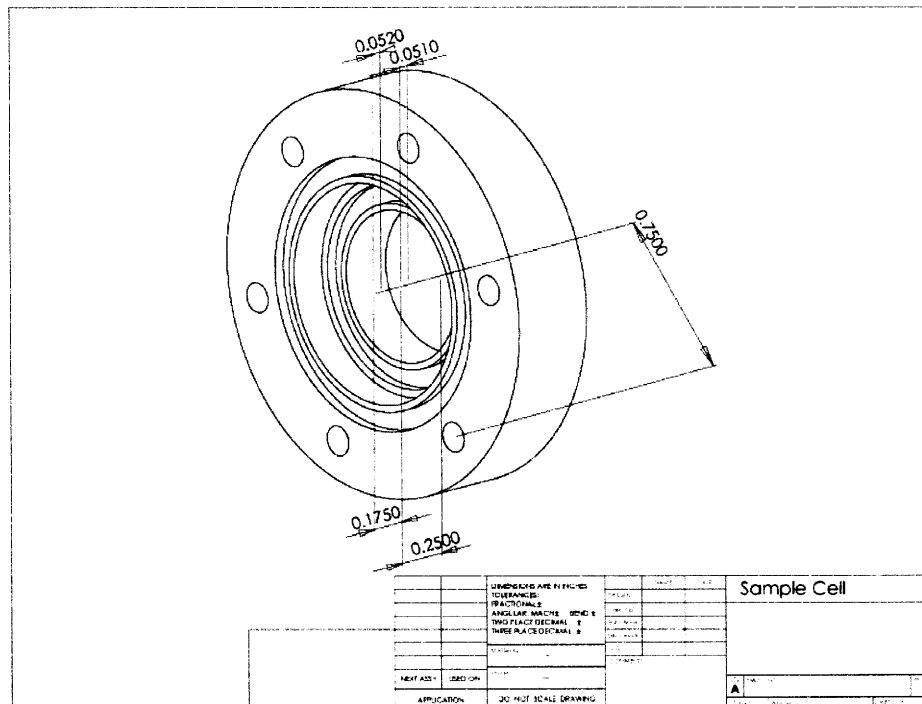
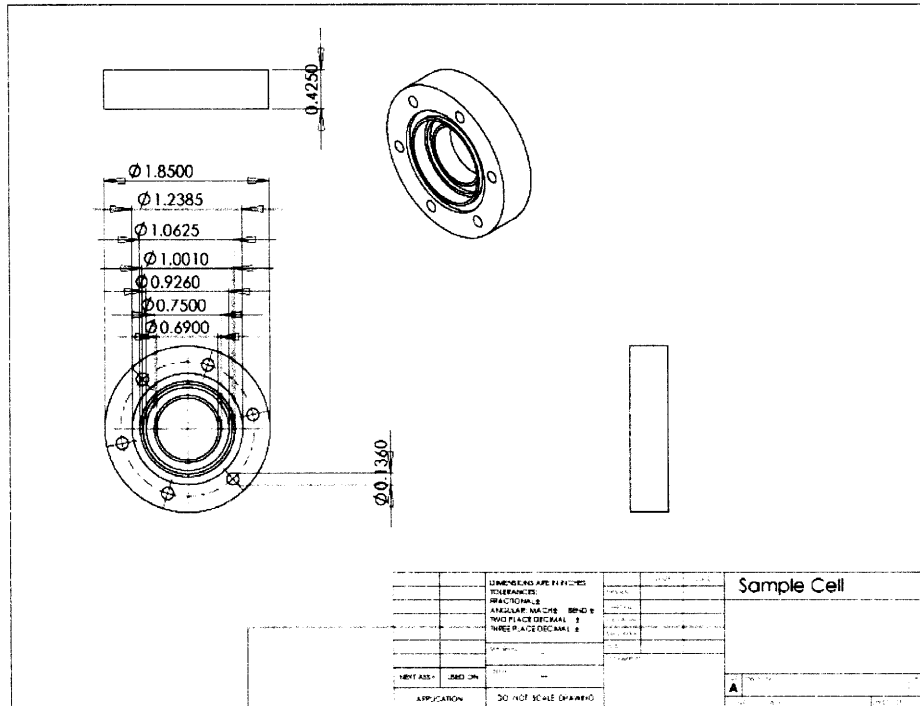








### 3.9.5. Cell Mount – Rear (brass)





# Chapter 4

## Amide I 2D IR Spectroscopy of $\beta$ -Hairpin Peptides

The work presented in this chapter has been published in the following paper:

“Amide I two-dimensional infrared spectroscopy of  $\beta$ -hairpin peptides,” A. W. Smith and A. Tokmakoff, *J. Chem. Phys.* **126**, 45109 (2007).

### 4.1. Abstract

In this report, spectral simulations and isotope labeling are used to describe the 2D IR spectroscopy of  $\beta$ -hairpin peptides in the amide I spectral region. 2D IR spectra of Gramicidin S, PG12, Trpzip2 (TZ2), and TZ2-T3\*T10\* (TZ2-TT in other chapters), a dual  $^{13}\text{C}$  isotope label, are qualitatively described by a model based on the widely used local mode amide I Hamiltonian. Our model includes methods for calculating site energies for individual amide oscillators on the basis of hydrogen bonding, nearest neighbor and long-range coupling between sites, and disorder in the site energy. The dependence of the spectral features on the peptide backbone structure is described using disorder-averaged eigenstates, which are visualized by mapping back onto the local

amide I sites.  $\beta$ -hairpin IR spectra are dominated by delocalized vibrations that vary by the phase of adjacent oscillators parallel and perpendicular to the strands. The dominant  $\nu_{\perp}$  band is sensitive to the length of the hairpin and the amount of twisting in the backbone structure, while the  $\nu_{\parallel}$  band is composed of several low symmetry modes that delocalize along the strands. The spectra of TZ2-T3\*T10\* are used to compare coupling models, from which we conclude that transition charge coupling (TCC) is superior to transition dipole coupling (TDC) for amide groups directly hydrogen bound across the  $\beta$ -strands. The 2D IR spectra of TZ2-T3\*T10\* are used to resolve the red-shifted amide I band and extract the site energy of the labeled groups. This allows us to compare several methods for calculating the site energies used in excitonic treatments of the amide I band. Gramicidin S is studied in dimethyl sulfoxide to test the role of solvent on the spectral simulations.

## 4.2. Introduction

$\beta$ -hairpin secondary structure is common in proteins, and represents a subunit of larger, anti-parallel  $\beta$ -sheet motifs.  $\beta$ -hairpin peptides are small enough to be subject to rigorous simulation methods, while containing a complex network of weak stabilizing interactions known to guide protein folding processes. Recent work has sought to study these systems in isolation in order to describe the factors that stabilize native structure<sup>1-4</sup> and guide structural isomerization during folding.<sup>5-10</sup> These efforts have led to a clearer description of the stabilizing role of mesoscopic structural features such as the four to six residue turn region and the packing together of hydrophobic side chains.<sup>1-10</sup>

In spite of the variety of methods used to characterize structure and stability, it has been difficult to obtain residue-level information about the folding dynamics of these systems. Their fast folding times (100 ns-10  $\mu$ s) make the direct observation of folding events inaccessible to NMR. Optical methods are fast enough but offer more qualitative structural information. For example, UV fluorescence spectroscopy is useful for reporting on the solvent environment of side chain chromophores.<sup>8,11</sup> UV circular dichroism and infrared spectroscopy report on larger scale structure and can be especially informative for peptide systems with a well-defined structural morphology.<sup>7,12</sup> More detailed insight can be gained from comprehensive mutation studies, which can be carried out at several important sites to determine the effect of side chains at specific parts of the structure.<sup>1,13</sup>

The introduction of two-dimensional infrared (2D IR) spectroscopy<sup>14,15</sup> of amide I vibrations<sup>16</sup> has provided a new technique that increases structural information while retaining the high time resolution required for folding studies. 2D IR spectroscopy reveals vibrational couplings within the amide I band, which depend on the spatial configuration and bonding of peptide units within the protein backbone. On this basis, 2D IR has been used to reveal the backbone configurations of di- and tri-peptides.<sup>17,18</sup> However, because of the excitonic nature of the amide I transition,<sup>19,20</sup> the structural interpretation for larger systems is generally focused on large-order secondary structure. For instance, the pair of amide I peaks characteristic of anti-parallel  $\beta$ -sheets has been used to observe the thermal unfolding of proteins and peptides with 2D IR and related spectroscopies.<sup>21-23</sup> When combined with site-specific isotope labeling of the amide I group, 2D IR spectroscopy provides a more local probe of couplings and peptide configurations.<sup>18,24,25</sup> This approach offers a way to probe individual amide groups within

a larger peptide at equilibrium and during a folding event. One unique advantage of 2D IR spectroscopy over traditional FTIR measurements is that it can distinguish homogeneous and inhomogeneous contributions to the spectral line shape.<sup>26</sup> In this way, a 2D IR spectrum can be used to describe the structural heterogeneity of the peptide, which leads to the broad lineshapes observed in infrared experiments.<sup>24,27,28</sup>

The interpretation of protein amide I 2D IR spectra relies on a spectroscopic model, the development and parameterization of which is an active area of research.<sup>17,29-32</sup> Since *ab initio* calculation of the vibrational spectra of proteins is difficult to scale up to biologically relevant proteins, approximate methods are used to reproduce the band positions and line shapes.<sup>19,30,33,34</sup> One approach is to use semi-empirical simulations combined with a Hessian matrix reconstruction to calculate amide I frequencies and intensities.<sup>29</sup> Another method performs full quantum mechanical simulations of peptide fragments from which local interactions are extracted.<sup>35,36</sup> An alternate approach that is commonly used and is easily transferable to very large systems is based on a local mode Hamiltonian that treats couplings between amide I vibrations of the individual peptide units.<sup>17,19,20,30,34,37</sup> In this model, a protein structure is used to parameterize the diagonal and off-diagonal elements (site energies and couplings), and the energy eigenvalues and transition moments obtained from this Hamiltonian are used to compare to the spectroscopy. The site energy of these oscillators has been shown to be dependent on hydrogen bonding to the amide unit.<sup>38,39</sup> More generally, it is dependant on electrostatic interactions with the surrounding solvent, which forms the basis for a number of site energy models.<sup>40-44</sup> Transition dipole coupling was originally used to calculate couplings between oscillators.<sup>45</sup> More accurate models now go beyond the dipole approximation to

include higher-order electrostatic interactions for through space couplings, and assign through-bond interactions between adjacent peptide units from *ab initio* calculations.<sup>17,30,46</sup> We have recently evaluated a number of local amide Hamiltonian models, and concluded that site energy variation and the disorder in site energies that arise from structural variation or fluctuations in the peptide and surrounding solvent are the crucial elements to simulating protein 2D IR spectra.<sup>27,28</sup>

In this report, we model the 2D IR spectra of a number of  $\beta$ -hairpins and describe the local mode symmetry of the amide I eigenstates. The spectroscopic model draws on a straightforward combination of methods for choosing site energies, nearest neighbor coupling and long range coupling.<sup>34</sup> The spectral simulations are used to prepare eigenstate visualizations according to the participation ratio and relative vibrational phase of each local site. The symmetry of the eigenstates allows us to describe the main spectral features present in the 2D IR data. The salient features in a  $\beta$ -hairpin 2D IR spectrum are the  $\nu_{\perp}$  and  $\nu_{\parallel}$  diagonal peaks with the cross-peak ridge connecting the two frequency regions. These modes have been characterized in detail in the context of  $\beta$ -sheets.<sup>47-50</sup> The  $\nu_{\perp}$  peak (1635-1645  $\text{cm}^{-1}$ ) consists of amide I sites oscillating in-phase with the amide group across the strand and out of phase with their in-strand neighbor. The  $\nu_{\parallel}$  peak (1665-1680  $\text{cm}^{-1}$ ) involves in-phase oscillations along the strand. These peaks are sensitive to hairpin structure and stability including twisting of the  $\beta$ -sheet strands.<sup>49-51</sup>

We perform our studies on three types of  $\beta$ -hairpin peptides with well-defined folded structures. The tryptophan zipper trpzip2 (TZ2)<sup>52</sup> and PG12<sup>53</sup> are tightly folded *de novo* designed peptides that differ in the  $\beta$ -turn type, relative stability, and degree of

twisting in the  $\beta$ -strands. Spectra of a TZ2  $^{13}\text{C}'$  isotopomer is used to study the influence of isotope labeling on the amide I eigenstates and 2D IR spectroscopy. Gramicidin S, a cyclic  $\beta$ -hairpin peptide,<sup>54</sup> is chosen to test the transferability of amide I spectroscopic models to non-aqueous solvents. A comparison of the spectroscopy and vibrational modes of these peptides offers insight into the generality of spectral assignments and amide I modeling for  $\beta$ -hairpins.

## 4.3. Experiment

### 4.3.1. Samples & preparation

The TZ2 (SWTWENGKWTWK-NH<sub>2</sub>) and PG12 (RYVEV-D-Pro-GKKILQ-NH<sub>2</sub>) samples were synthesized and purified by the MIT Biopolymers Laboratory. Both peptides were synthesized using standard Fmoc protocols as C-terminal amides, purified with HPLC, and tested for purity using MALDI-TOF mass spectrometry. To prepare for infrared measurements, the samples were repeatedly lyophilized against 50 mM (TZ2) or 20mM (PG12) DCl in D<sub>2</sub>O to remove residual trifluoroacetic acid, which exhibits a strong IR transition at 1673 cm<sup>-1</sup>. The PG12 solution consisted of a 10 mM deuterated acetate buffer at pH 3.85 (pD = 4.25). TZ2 was dissolved in a 50 mM deuterated phosphate buffer at pH 7.0 (pD = 7.4). The peptide concentration for the infrared measurements was approximately 12mg/mL (~7 mM) for TZ2 and 25 mg/mL (~17 mM) for PG12. Gramicidin S (cyclo(-Val-Orn-Leu-D-Phe-Pro-)<sub>2</sub>) was obtained from Prof. Glenn Jones of Holy Cross College, Worcester, MA, and used without further

purification. For infrared measurements, it was dissolved in dimethylsulfoxide-d<sub>6</sub> (Cambridge Isotope Labs) at a concentration of 25 mg/mL (~22 mM). TZ2-T3\*T10\* was synthesized and purified by Anaspec Inc. (San Jose, CA) using L-Threonine-1-<sup>13</sup>C (Icon Isotopes, Summit, NJ) at residues Thr3 and Thr10). Experimental sample conditions were the same as for the unlabeled TZ2. For TZ2, TZ2-T3\*T10\*, and PG12, the sample temperature was held to 25 °C. The Gramicidin S spectrum was taken at room temperature (~22 °C).

### 4.3.2. Spectroscopy

Infrared measurements were done in a temperature controlled sample cell consisting of two 25x1 mm CaF<sub>2</sub> windows separated by a 50 μm Teflon spacer. FTIR spectra were collected on Mattson Infinity FTIR spectrometer. Buffer background and sample spectra were taken at 0.5 cm<sup>-1</sup> resolution and averaged for 64 scans.

The 2D IR instrument and methods used have been described in detail previously.<sup>26</sup> Briefly, a series of three 90 fs, 6.0 μm laser pulses drive and probe transitions that are resonant with the pulse spectrum, including the peptide amide I  $\nu_i = 1 \leftrightarrow 0$  and  $\nu_i = 2 \leftrightarrow 1$  transitions, as well as the two quantum combination bands ( $\nu_i + \nu_j = 2 \leftrightarrow \nu_i = 1$ ). The data shown in this report are absorptive 2D IR spectra which correlate the excitation and detection frequencies.<sup>26</sup> The excitation ( $\omega_1$ ) axis is obtained by a numerical Fourier transform of  $\tau_1$ , the time delay between the first and second probe pulses. The signal is collected for  $\tau_1$  delays up to 2000 fs in the rephasing geometry and 1500 fs in the non-rephasing geometry. For the data shown in this report, the waiting time ( $\tau_2$ ), which is the delay between the second and third pulses, is set to zero. To obtain the detection axis,  $\omega_3$ ,

we overlap the emitted signal field with a local oscillator pulse, with the time delay between them designated as  $\tau_3$ . In practice, we set  $\tau_3=0$  and send the overlapped signal and local oscillator through spectrometer and onto a 64-channel MCT array detector. The grating acts to Fourier transform  $\tau_3$  to produce the  $\omega_3$  axis.<sup>26</sup> Spectra are collected in the ZZYY polarization geometry.

To find the zero times between the probe pulses, we first measure the cross correlation between pulses 1, 2, and 3, which gives a temporal resolution on the order of  $\pm 10$  fs. After the data is collected, we run a least squares fitting routine to match the slices in  $\tau_1$  with the dispersed pump probe. The  $\tau_3$  timing is found by setting the  $\tau_1$  and  $\tau_2$  equal to zero and scanning the local oscillator timing until the signal matches the dispersed pump probe. The numerical Fourier transform of the  $\tau_1$  axis is done using a triangle apodization function with zero padding.

## 4.4. Calculations

### 4.4.1. Calculation of 1D and 2D IR spectra

Spectral calculations were done using a local amide I Hamiltonian (LAH) formalism, which is commonly used to describe protein and peptide amide I spectroscopy.<sup>19,20,34,37,55</sup> The LAH treats the amide I vibrations as a set of  $N$  coupled oscillators in the basis of the amide I oscillators of the peptide residues. The zero-order vibrational frequency of each amide group is included as a diagonal matrix element with the coupling to other groups included in the off-diagonal matrix elements. Amide I site energies are influenced by the hydrogen bonding of the amide group, which can be



described in terms of the electric field of the surroundings acting on the amide I coordinate.<sup>40-44</sup> Long-range couplings are dominated by electrostatic interactions, and neighboring groups along the  $\beta$ -strand are dominated by coupling through the peptide backbone.

The matrix elements of the LAH are assigned on the basis of previously determined hairpin structures. The NMR structure of TZ2 was obtained from the Protein Data Bank (ID 1le1).<sup>52</sup> We used the first of the twenty available structures for all calculations shown here. The structure of PG12 was provided by Samuel Gellman at UW-Madison. It was determined using experimental NMR constraints to guide a MD simulation (DYNA software package). The structure of Gramicidin S has not been measured experimentally because it has not been possible to produce a high quality crystal sample. We used the crystal structure of a Gramicidin S analog, where the ornithine side chains were capped with m-bromobenzoic acid.<sup>56</sup>

To choose site energies we employed a heuristic model that is based on hydrogen bonding geometries.<sup>19</sup> Starting from a zero-order frequency ( $\omega_0 = 1688 \text{ cm}^{-1}$ ), the energy of a particular site ( $\omega_n$ ) is red shifted if the amide oxygen participates in a hydrogen bond to an amide proton or to a solvent hydrogen bond donor ( $\Delta\omega_{CO,n}$ ). An additional red shift is included if the amide proton is hydrogen bound to another amide oxygen or to a solvent hydrogen bond acceptor ( $\Delta\omega_{NH,n}$ ):<sup>38</sup>

$$\omega_n = \omega_0 - \Delta\omega_{NH,n} - \Delta\omega_{CO,n}$$

The frequency shifts are

$$\Delta\omega_{CO,n,b} = 30 \text{ cm}^{-1} \cdot (2.6 - r_{O(n)\cdots H(b)} / \text{\AA})$$

$$\Delta\omega_{NH,n} = 10 \text{ cm}^{-1}$$

Here  $r_{O(n)\dots H(b)}$  is the hydrogen bonding distance in Ångstroms. For groups with hydrogen bonds to solvent we set  $\Delta\omega_{CO,n} = 20 \text{ cm}^{-1}$ .<sup>38</sup> For intra-molecular hydrogen bonds,  $r_{O(n)\dots H(b)}$  is obtained from the experimentally determined structure. If the amide oxygen or proton does not participate in a hydrogen bond, we estimate that the long-range solvent contribution to the site energy is  $\Delta\omega_{CO,n} = 10 \text{ cm}^{-1}$ .<sup>38</sup> For the present systems, this only occurs for Gramicidin S in DMSO, where there is no solvent proton available to H-bond with the amide oxygen.

Vibrational couplings take into account through-bond and through space interactions. Coupling terms between each nearest neighbor pair along the peptide chain,  $\beta_{n, n+1}$ , are obtained from an amide I coupling map parameterized as a function of the dihedral angles using DFT calculations on glycine dipeptide.<sup>46</sup> The *ab initio* parameterization inherently treats all through-bond and through-space interactions. All other couplings are calculated using transition charge coupling (TCC), which was developed as a higher order treatment of the electrostatics compared to the transition dipole coupling (TDC) model.<sup>17</sup> Similar approaches have been used recently by several groups to simulate amide I infrared spectra in a variety of peptide and protein systems.<sup>24,34,55,57,58</sup>

Calculation of 2D IR spectra requires knowledge of the two quantum states in the LAH. The  $2\nu$  site energies are assigned as  $2\omega_n - A$ , where the anharmonicity  $A$  is set to  $16 \text{ cm}^{-1}$ .<sup>19</sup> The couplings between the  $2\nu$  states are the same as the one-quantum coupling terms but harmonically scaled by  $\sqrt{2}$ .<sup>19</sup> The linear and 2D IR spectra were calculated using the transition dipole moments and energy eigenvalues of the vibrational eigenstates

obtained from diagonalizing the LAH. The methods for this calculation involved an eigenstate summation over 1D or 2D Lorentzian lineshapes ( $\Gamma = 9.0 \text{ cm}^{-1}$ ), and have been described in detail elsewhere.<sup>28,49</sup>

As we have demonstrated recently,<sup>28</sup> it is not possible to properly simulate amide I 2D IR spectroscopy without accounting for disorder in the site energies. We incorporated the effect of static disorder in the spectral calculations by adding random Gaussian disorder to the calculated site energies. This approximation is based on the recent work of Ganim et al, where they found that the dominant source of disorder is in the site energies and that the disorder in coupling due to structural fluctuations has a negligible effect on the spectroscopy.<sup>28</sup> A similar observation was also made by Cho.<sup>29</sup> For each spectral calculation (*i*) a Gaussian weighted random number was subtracted from the original site energy:

$$\omega_{n,i} = \omega_n - \delta\omega_i$$

The standard deviation of the random site energy distribution was chosen based on qualitative agreement with the experimental 2D IR data and was set to  $\sigma = 12 \text{ cm}^{-1}$ . For each of the systems, 10,000 disorder realizations were averaged to give the final calculated spectra. The methods used in the disorder calculation of 1D and 2D spectra are described elsewhere.<sup>28</sup>

#### **4.4.2. Visualization of vibrational modes**

To develop an interpretation of the vibrational modes in different spectral regions in terms of peptide backbone structure, we visualize the calculated eigenstates by projecting them back onto the peptide local modes. This visualization scheme adapts a

method used by Dijkstra et al. and Chung et al., in which the magnitude and phase of each amide group is color coded onto a grid.<sup>27,34</sup> Each local amide group corresponds to a box in the grid. The color of the box (red or blue) reflects the phase of the oscillator, while the depth of color encodes the amplitude of that group in the vibrational eigenstate.

Since disorder is crucial to reproducing the amide I spectroscopy, the visualizations are performed on disorder-averaged vibrational eigenstates. The eigenvectors obtained for a calculation without site energy perturbations ( $\delta\omega_i = 0$ ) are used as reference eigenstates. For each disorder realization, the overlap integral between the realization and reference eigenvectors is calculated, and maximum overlap is used to sort the eigenstates. The eigenvectors from each disorder step are binned and averaged to create the visualizations.

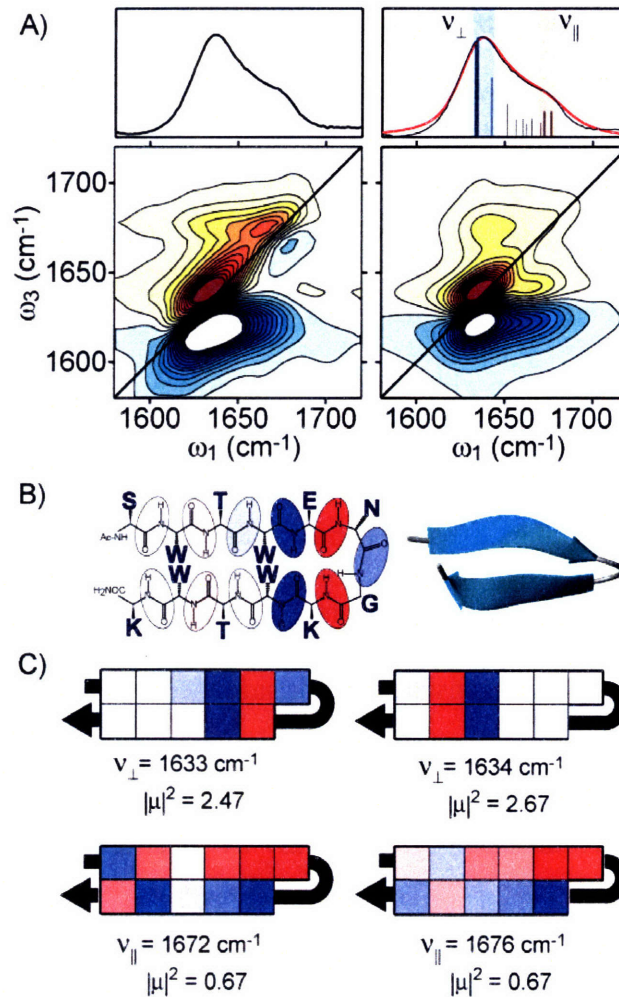
## 4.5. Results

### 4.5.1. TZ2

TZ2 is an excellent candidate for the purposes of investigating structure-spectroscopy relationships in  $\beta$ -hairpins. Previous investigations of TZ2 have demonstrated its exceptional stability and resistance to structural change.<sup>7,8,21,52</sup> Through fluorescence measurements it was shown that both high temperatures and a chemical denaturant were necessary to completely dissociate the side chain contacts.<sup>6</sup> Initial 2D IR work also showed that the hydrogen bond network maintains some degree of native structure at high, typically denaturing temperatures.<sup>21</sup> TZ2 is an important benchmark

because it is well-characterized experimentally and is accessible to a wide range of theoretical methods and computational methodologies<sup>6,8,59-61</sup>

The amide I FTIR spectrum of TZ2, shown in Fig. 1 together with a diagram of the structure, has two distinguishing features: a band maximum at 1636  $\text{cm}^{-1}$  and the prominent shoulder near 1673  $\text{cm}^{-1}$ .<sup>6,8</sup> As described previously<sup>21,48-50,62</sup> and shown in detail below, the eigenstates giving rise to the intense low frequency peak involve in-phase vibration of amide oscillators directly across the strand and in-phase vibration of backbone nearest neighbors. The spectral feature is thus termed the  $\nu_{\perp}$  peak and the contributing vibrations are collectively termed the  $\nu_{\perp}$  modes. Neighboring oscillators within the high frequency eigenstates comprising the shoulder around 1673  $\text{cm}^{-1}$  vibrate in-phase with each other. Consequently, we designate the shoulder as the  $\nu_{\parallel}$  peak, with its constituent  $\nu_{\parallel}$  vibrational modes. The 2D IR spectrum of TZ2 is shown in Fig. 1. Along the diagonal axis there are two well-resolved positive peaks, corresponding to the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks inferred from the FTIR spectrum. The negative diagonal peaks corresponding to the  $\nu_{\perp}$  and  $\nu_{\parallel}$  overtone transitions are also resolvable, although the  $\nu_{\perp}$  overtone peak is quite intense and blurs out the cross peaks around  $\omega_1 = \nu_{\parallel}$  and  $\omega_3 = \nu_{\perp}$ . The ridge extending out to  $\omega_1 = \nu_{\perp}$  and  $\omega_3 = \nu_{\parallel}$  is the result of the cross peak intensity mixing with that of the positive diagonal peaks. As reported previously, the presence of these two coupled vibrational modes in the 2D IR spectrum and the resultant Z-shaped contour profiles are a direct indication of anti-parallel  $\beta$ -sheet secondary structure.<sup>21,50,63</sup>



**Figure 4.1** (a) Experimental amide I spectra of TZ2 (left). 2D IR contour lines are at  $\pm 80\%$  of the band maximum for this and all other 2D IR spectra in this paper. Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates prior to incorporating Gaussian disorder are plotted as lines with the intensities weighted by the relative strength of the transition moment,  $\mu^2$ . (b) Stick and ribbon diagrams. Ribbon diagrams made with MolMol software. Ellipses overlay the peptide stick diagram to highlight how the local peptide amide groups map on to the box plots below. The color scheme of the ellipses is that same as the  $\nu_{\perp} = 1633 \text{ cm}^{-1}$  eigenstate. (c) Selected eigenstates from the spectral simulation.

Fig. 1 shows the calculated 1D and 2D IR spectra of TZ2 along side the experimental data. The lineshape parameters used in the simulation ( $\sigma = 12 \text{ cm}^{-1}$  and  $\Gamma = 9 \text{ cm}^{-1}$ ) were chosen to fit the experimental diagonal and antidiagonal line widths

respectively. The spectral simulation reproduces the experimental  $\nu_{\perp}$  positive peak width and position. The calculated overtone peak intensity is much smaller than the experiment, although the slope of the node between the fundamental and overtone peaks agrees quite well. The intensity in the  $\omega_1 = \nu_{\perp}$  and  $\omega_3 = \nu_{\parallel}$  region of the spectrum is also similar in the experiment and the simulation. The most significant discrepancy between the spectra is the amplitude of the  $\nu_{\parallel}$  diagonal peak relative to the  $\nu_{\perp}$  diagonal feature. The 2D IR simulation underestimates the  $\nu_{\parallel}$  amplitude by almost half. This mismatch could be due to a number of factors, besides the model. One possibility is that the extra intensity is from either the Asn6 side chain or C-terminal amide groups, which are known to absorb near this frequency in neutral pH solutions.<sup>64</sup> Calculations including these additional amide oscillators improves the comparison with experiment, as detailed in the supplementary material.<sup>65</sup>

From the calculations on TZ2, there are two intense eigenstates that carry over 50% of the total amide I oscillator strength and are pictured in Fig. 1. One of the states is localized on oscillators 4, 5, 7, and 8 ( $1633 \text{ cm}^{-1}$ ), while the other is localized on oscillators 2, 3, 9, and 10 ( $1634 \text{ cm}^{-1}$ ). The reason for the separation into two modes on either end of the hairpin is the high degree of twisting, which disrupts the symmetry that in an ideal  $\beta$ -hairpin would delocalize the low energy mode across the peptide. The two modes are not resolved because of their near degenerate frequencies. Both of the modes contribute to the  $\nu_{\perp}$  region – so called because the oscillators are in-phase across the strand and the transition dipole of these modes points perpendicular to the  $\beta$ -strand. The high frequency region generally contains highly mixed vibrational states. The two eigenstates with the highest energy and significant intensity show phase symmetry

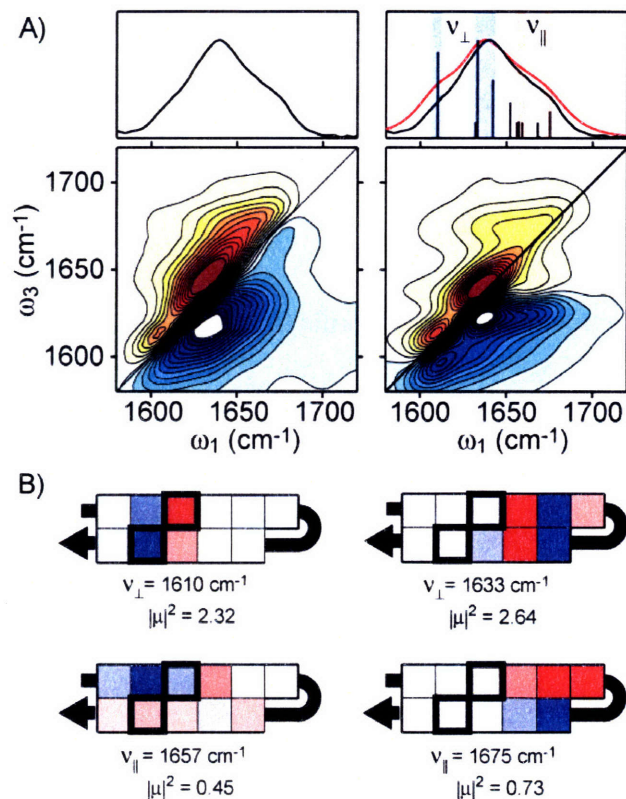
similar to that expected for ideal  $\nu_{||}$  modes, where the neighboring oscillators on the same strand are vibrating in phase and the two strands are out of phase.<sup>34,49</sup> In this case, the modes are delocalized over the entire hairpin. While the agreement between spectral calculations remains largely qualitative, we find that the description of the eigenstates is robust. As demonstrated in the supplementary material,<sup>65</sup> the symmetry and local mode distribution of the dominant eigenstates is relatively insensitive to minor changes in site energy and coupling models.

#### **4.5.2. TZ2-T3\*T10\***

Backbone  $^{13}\text{C}'=^{16}\text{O}$  isotope labeling introduces a  $\sim 40\text{ cm}^{-1}$  red shift to the amide I' vibration. The technique has been used extensively for peptide and protein FTIR spectroscopy,<sup>66-68</sup> although its application to  $\beta$ -hairpin peptides is limited.<sup>24,69</sup> This is due not only to the relatively small number of systems, but also because many of the amino acids in the primary sequence of well-studied hairpin peptides are difficult to obtain commercially. We chose to study a pair-wise  $^{13}\text{C}$  label of TZ2 at the C' atom of Thr3 and Thr10 for two purposes. First, the spectroscopy of the dual label ensures that shift of the isotope labeled peak relative to the amide I maximum is sensitive to coupling between the oscillators. Studying the two closely spaced aligned oscillators provides test of coupling models. For example, TCC predicts the Thr3-Thr10 coupling to be  $6\text{ cm}^{-1}$ , while the TDC model predicts it to be over  $12\text{ cm}^{-1}$ . The other reason to choose this isotope combination is because it will isolate a spectral signature directly related to the Thr3-Thr10 inter-strand hydrogen bond contacts.



The experimental FTIR and 2D IR spectra of TZ2-T3\*T10\* are shown in Fig. 2. The FTIR spectrum of the isotope labeled peptide shows a shoulder near  $1613\text{ cm}^{-1}$  and a well-resolved diagonal peak in the 2D IR spectrum. The amide I maximum shifts from  $1636\text{ cm}^{-1}$  to  $1640\text{ cm}^{-1}$ , while the  $\nu_{\parallel}$  band contracts to the red, becoming less prominent in the FTIR spectrum and unresolved in the 2D IR spectrum. While a distinct cross peak is not observed, it is present as ridges elongated along  $\omega_3 = \nu_{\parallel}$  and  $\nu_{\perp}$ , and can be seen in a  $\omega_1 = \nu_{\parallel}$  slice through the 2D spectrum.



**Figure 4.2** (a) Experimental amide I spectra of TZ2-T3\*T10\* (left). Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates calculated and plotted as in Fig. 1. (b) Selected eigenstates from the spectral simulation. The boxes around the <sup>13</sup>C labeled Thr3 and Thr10 amide groups are in bold. Particularly notable is the splitting of the two  $\nu_{\perp}$  bands by 23 cm<sup>-1</sup>. The isotope shift deduced for the labeled oscillators based on the site energy model is 34 cm<sup>-1</sup>.

The parameters from the spectral calculation of the unlabeled compound were transferred without modification except the site energy of residues 3 and 10, which were red shifted by 33 cm<sup>-1</sup> to obtain agreement with the experiment. This empirical isotope shift is lower than 39 cm<sup>-1</sup> shift estimated by the difference in the reduced mass of a <sup>12</sup>C=<sup>16</sup>O oscillator compared to <sup>13</sup>C=<sup>16</sup>O. It is also lower than the reported 44 cm<sup>-1</sup> shift based on a DFT (B3LYP/6-31+G\*\* level) calculation of N-methylacetamide in vacuum.<sup>24</sup> As discussed in section V.C., we suspect that the unusually small shift is due

to our original site energy model, rather than a fundamental disagreement with the earlier predictions.

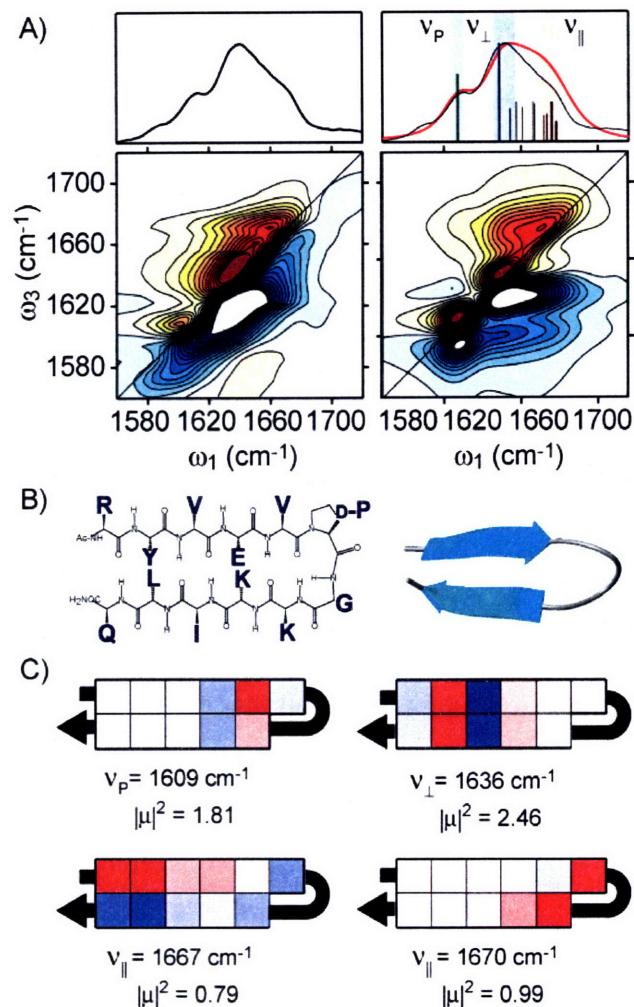
The calculated eigenstates of the T3\*T10\* isotope labeled compound can be used to examine the nature of the red-shifted amide I vibration. From the phase plot in Fig. 2, we see that the 1613  $\text{cm}^{-1}$  eigenstate is localized on the 2, 3, 9, & 10 amide groups, with most of the intensity on the labeled groups 3 & 10. This localization of the  $^{13}\text{C}$  shifted  $\nu_{\perp}$  vibration to the residues near the peptide terminal region causes the eigenstate at 1637  $\text{cm}^{-1}$ , or the unlabeled  $\nu_{\perp}$  vibration, to localize on residues 4, 5, 7, & 8. Thus, while the 1613  $\text{cm}^{-1}$  peak reports on the cross-strand hydrogen bonds near the N and C termini, the amide I spectral maximum directly probes the hydrogen bond pair near the turn region. In other words, the effect of the T3\*T10\* isotope labels is to remove the near degeneracy of the two  $\nu_{\perp}$  modes of the unlabeled peptide. As expected, the 1613  $\text{cm}^{-1}$  mode is an asymmetric stretch of the labeled 3-10 oscillators; the calculations suggest that the weak symmetric vibration of the labeled oscillators at 1613  $\text{cm}^{-1}$  is coincident with the strong  $\nu_{\perp}$  transition.

### 4.5.3. PG12

We also investigate the amide I spectroscopy of a 12 residue *de novo* hairpin peptide designed by Gellman and coworkers.<sup>53</sup> The PG12 hairpin structure is stabilized by a D-Pro enhanced type II'  $\beta$ -turn, and has been optimized for stability. In fact, the structure has proven reliable enough to be used as a template for the design of larger three and four stranded  $\beta$ -sheet peptides.<sup>70</sup> A previous infrared spectroscopy study of this peptide was done to systematically analyze the stability of D-PG12 by varying the length

of the peptide, substituting an L-Pro residue into the turn region, and testing it against cyclic hairpin analogues.<sup>71</sup>

The FTIR spectrum of PG12 is shown in Fig. 3, and is similar to that reported in the previous work.<sup>71</sup> The amide I maximum, or  $\nu_{\perp}$  peak, is centered at  $1640\text{ cm}^{-1}$ , while the shoulder, or  $\nu_{\parallel}$  peak, appears near  $1670\text{ cm}^{-1}$ . The energy of the  $\nu_{\perp}$  mode of D-PG12 is about  $4\text{ cm}^{-1}$  higher than that of TZ2. Other resolvable features include the peak at  $1613\text{ cm}^{-1}$ , which was previously assigned to the D-Pro residue  $\nu_{\text{P}}$ .<sup>71</sup> Two other weak features at  $1588$  and  $1705\text{ cm}^{-1}$  are attributed to the side chain carbonyl vibrations of Gln12 and Glu4 respectively.<sup>64</sup>



**Figure 4.3** (a) Experimental amide I spectra of PG12 (left). Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates calculated and plotted as in Fig. 1. (b) Stick and ribbon diagrams. (c) Selected eigenstates from the spectral simulation.

The 2D IR spectrum displays positive diagonal peaks that correspond to the peaks in the FTIR spectrum. The  $\nu_{\parallel}$  peak is resolved in the 2D IR spectrum, but its spectral position and intensity relative to the  $\nu_{\perp}$  peak differs from that of TZ2. The cross peak ridge is clearly visible, but lacks the prominence of the ridge in the TZ2 2D IR spectrum. These effects are primarily caused by interference with intensity from the broad  $\sim 1660$

$\text{cm}^{-1}$  peak corresponding to disordered peptide structure. We submit this interpretation based on comparison to the high temperature spectrum of TZ2,<sup>21</sup> where disorder has been thermally introduced to the system. Coupling of the proline to the remainder of the backbone amides is observed through interference effects that lead to extended  $\omega_3 = \nu_{\parallel}$  and  $\nu_{\perp}$  ridges. Also, the negative ridge observed for ( $\omega_1 \approx 1600$ ;  $\omega_3 = 1620 \text{ cm}^{-1}$ ) arises from the cross peak between the proline and  $\nu_{\perp}$  resonances.

While significant differences to experiment exist, the spectral simulations of PG12 have many similar features. In the simulated 2D IR spectrum, the frequencies and relative intensities of the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks agree, as do the intensity and profile of the positive diagonal cross peak ridge. On the other hand, the correct relative intensity of the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks is not captured well in the simulated 1D absorption spectrum. For the spectral calculations, the D-Pro peak near  $1613 \text{ cm}^{-1}$  was produced by empirically shifting the site energy of amide group 5 by  $40 \text{ cm}^{-1}$ . In order to reproduce the inhomogeneous linewidth of this peak it was necessary to reduce the width of the disorder to  $\sigma = 6 \text{ cm}^{-1}$ . Other sites were left with  $\sigma = 12 \text{ cm}^{-1}$ , yet the simulated  $1640 \text{ cm}^{-1}$  positive diagonal peak is significantly narrower than the experiment. This suggests is that the disorder varies from the turn to the ends of the hairpin, and is consistent with a tightly folded turn and fraying ends of the hairpin. This was also observed in TZ2 by Wang et al, who showed that the 2D IR measured homogeneous and inhomogeneous linewidths of residues Trp2 and Gly7 are significantly different.<sup>24</sup>

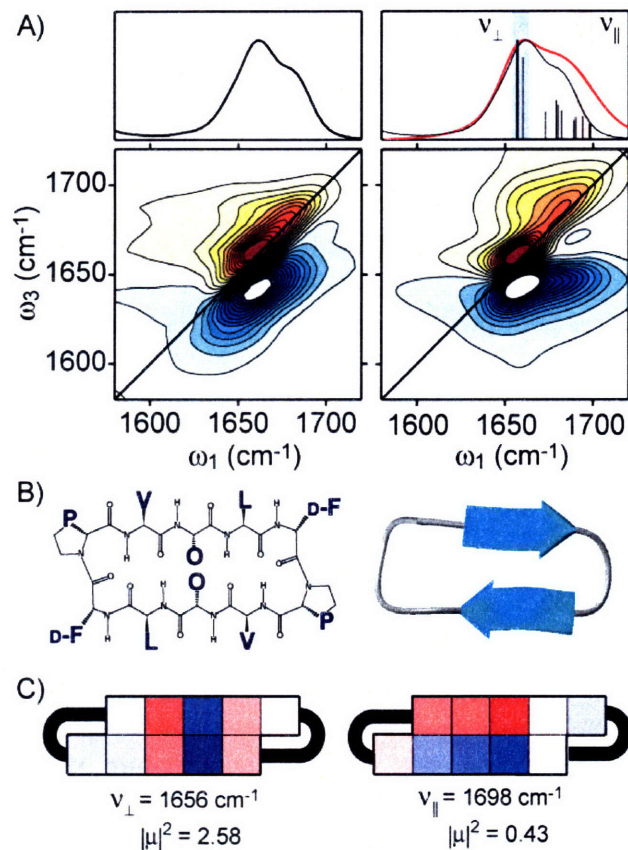
The eigenstates visualizations in Fig. 3 show that the  $1613 \text{ cm}^{-1}$  peak is largely located on the Site 5 (Val5-Pro6) amide group, but that there is visible intensity on neighboring oscillators. Because the off-diagonal intensity in the simulated 2D IR

spectrum does not match the experiment, we suspect that the calculated couplings near the turn region are inaccurate. This means that the vibration may in fact be more localized on Site 5 than the simulations suggest. Looking at the other eigenstates we see that the  $\nu_{\perp}$  vibration is observed to delocalize across the entire hairpin, with most of the intensity on the groups not involved in the  $1613\text{ cm}^{-1}$  vibration. The  $\nu_{\parallel}$  modes are highly mixed, although the  $1667\text{ cm}^{-1}$  vibration does show the expected  $\nu_{\parallel}$  phase pattern on oscillators 1-3 and 8-11.<sup>72</sup>

#### 4.5.4. Gramicidin S

Gramicidin S is an antimicrobial peptide that induces cell death by aggregating in and disrupting the outer membrane.<sup>72</sup> No crystal structure of unmodified Gramicidin S has been measured, but early NMR and spectroscopic work showed that it does form an antiparallel pleated sheet, connected by two type II'  $\beta$ -turns.<sup>54,73,74</sup> Recent crystal structures of trichloroacetyl and m-bromobenzoyl Gramicidin S confirm the earlier structure models.<sup>56</sup> Gramicidin S is not water soluble, but it can be dissolved in DMSO, for which the FTIR spectrum has been reported previously.<sup>75</sup>

The FTIR and 2D IR spectra of Gramicidin S are presented in Fig. 4. The amide I peak in the FTIR spectrum is  $1661\text{ cm}^{-1}$  while the high frequency shoulder appears near  $1680\text{ cm}^{-1}$ , which correspond to the  $\nu_{\perp}$  and  $\nu_{\parallel}$  bands, respectively. The most striking difference with the earlier systems is that the center frequency is blue shifted and width of the amide I band is narrower. These effects can both be attributed to weaker interactions with the DMSO solvent. Similar shifts in frequency and linewidth are observed for the amide I vibration of N-methylacetamide.<sup>76</sup>



**Figure 4.4** (a) Experimental amide I spectra of Gramicidin S (left). Simulated amide I spectra (right). The simulated FTIR spectrum is plotted in red and the underlying eigenstates calculated and plotted as in Fig. 1. (b) Stick and ribbon diagrams. (c) Selected eigenstates from the spectral simulation.

The simulated amide I spectra of Gramicidin S are shown in Fig. 4. The line width parameters ( $\Gamma = 9 \text{ cm}^{-1}$ ,  $\sigma = 12 \text{ cm}^{-1}$ ), and coupling model are the same as were used for PG12 and TZ2. The coupling and lineshape parameters in are chosen to give a consistent comparison to the simulations of the other hairpin spectra in this report. As can be seen, the calculated spectrum overestimates the  $\nu_{\perp} - \nu_{\parallel}$  splitting by about  $10 \text{ cm}^{-1}$ . The dominant eigenstates do have the same oscillator phase pattern as in the previous system. The only assumption made in transferring these calculations to the non-hydrogen bonding



DMSO solvent is that site energies show a weaker solvent red-shift. The qualitative agreement we observe suggests that electrostatic site energy models should be transferable to different peptide environments.

## **4.6. Discussion**

### **4.6.1. Previous IR studies of $\beta$ -hairpins**

As recently as ten years ago, there were very few stable  $\beta$ -hairpin peptides that could be studied in aqueous environments. The model systems that did exist were difficult to study with infrared spectroscopy because the required experimental concentrations would lead to significant aggregation over the timescale of the experiments. Consequently, applications of infrared spectroscopy to study  $\beta$ -hairpins were rare. Only in the last 5-7 years has there been a widespread use of FTIR methods to study hairpin structure, stability, and folding.<sup>6-9,77,78-80</sup> Much of this work has taken advantage of the distinctive amide I line-shape exemplified by the systems presented in this report. FTIR spectroscopy of hairpin peptides has been used most frequently to compare the relative stability of  $\beta$ -hairpin systems. This work has led to a greater understanding of the stabilizing role of the  $\beta$ -turn and the amino acid sequences that enhance this stability. The thermal melting transition has also been investigated extensively with FTIR techniques, including transient temperature jump studies to investigate folding kinetics.<sup>9,81</sup> These studies have helped develop a description of thermal denaturation, including insight into the structural regions that contribute significantly to the folding rate.

In spite of these advances, the detailed interpretation of the amide I spectral features remains somewhat ambiguous. Generally, the structural interpretation of these data has been focused on the stability of the cross-strand hydrogen bond contacts. For example, a number of reports have focused on the low frequency (1615-1645  $\text{cm}^{-1}$ ) region. Often, this consists of using the presence of a low frequency peak simply as a marker of  $\beta$ -hairpin structure.<sup>8,77,80</sup> Other groups have used band deconvolution to assign a low frequency peak to the turn region as well to distinguish the carbonyl groups hydrogen-bound across the strand from those hydrogen bound to solvent.<sup>9,78</sup> Another method of interpretation is based on electrostatic coupling methods used in this report. For example, the high frequency region has used to quantify  $\beta$ -hairpin stability,<sup>7</sup> and was the basis for previous interpretations of the 2D IR spectroscopy of TZ2.<sup>21</sup> The vibrational spectroscopy of TZ1 was recently studied using an *ab initio* fragmentation method.<sup>35</sup>

#### 4.6.2. Testing coupling models

Isotope labeling is often employed in vibrational spectroscopy to clarify spectral assignments and decongest broad vibrational bands. Spectra of the isotope labeled TZ2 peptide can be used to investigate the way that vibrational coupling is treated in spectral models. For TZ2, the TDC calculated coupling between amide groups 3 and 10 is 12.9  $\text{cm}^{-1}$ , while the TCC coupling model predicts a value of 6.3  $\text{cm}^{-1}$ . This large difference in predicted coupling does not produce significant differences between the TCC and TDC simulated amide I line shapes of unlabeled TZ2. While largely a result of spectral congestion, this insensitivity is also because the other matrix elements are not as sensitive to coupling models. On the other hand, when the two models are used to calculate the

spectroscopy of the isotope labeled compound, where interaction between Thr3 and Thr10 can be spectrally isolated from the main amide I band, there is a marked difference. Specifically, the TDC calculation predicts that the frequency of the peak corresponding to the isotope labeled amide groups is  $1607\text{ cm}^{-1}$ , while the calculated frequency in the TCC model is  $1613\text{ cm}^{-1}$  – equal to the experimental value.

In order to help isolate coupling effects from the original choice of site energy, we can also compare the difference between the red-shifted peak and the amide I maximum in the positive diagonal region of the 2D IR spectra. By finding the maxima of the two peaks we can compute the frequency splitting between the two peaks on the 2D IR surface. The value obtained from the experimental 2D IR spectrum  $37.5\text{ cm}^{-1}$ . For the TCC simulated 2D IR spectrum, the value is  $36.8\text{ cm}^{-1}$ , while the TDC simulation gives a difference of  $47.4\text{ cm}^{-1}$ . This suggests that the TDC coupling model overestimates the strength of coupling between hydrogen bonded amide groups, although the test cannot be absolutely definitive because of uncertainties in the original site energy model.

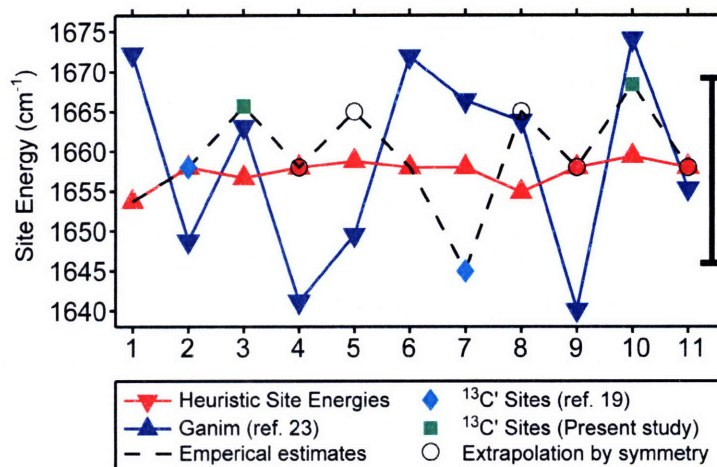
The discrepancy between the two models can be understood in terms of the limits of the dipole approximation, which is expected to break down as the distance between the dipoles approaches their length.<sup>41,82</sup> If we approximate the size of the amide I dipole to be the distance from the amide proton to the amide oxygen ( $\sim 3.15\text{ \AA}$ ), we see that it is very close to the distance between cross-strand amide groups 3&10 ( $\sim 3.73\text{ \AA}$ ). This effect was the primary motivation for the incorporation of the DFT parameterization to simulate nearest neighbor coupling.<sup>17</sup> Here we have shown that for peptides in tight  $\beta$ -strand registry, the TDC approximation breaks down appreciably.

Generally speaking, there are still questions that remain regarding the accuracy of the existing coupling models. In all of the calculated spectra, the relative ratio of diagonal to cross peak amplitudes and strong horizontal cross peak ridges suggest that the relative magnitude of the couplings relative to the site energy variation still is somewhat high. This is an issue that remains to be resolved in the parameterization of amide I coupling and site energy models so that quantitative agreement to experiment can be achieved.

#### **4.6.3. Assignment of site energies**

Recently, two single isotope labeled variations of TZ2 have been used to study the line broadening dynamics of the Trp2 and Gly7 residues using 2D IR spectroscopy.<sup>24</sup> Hochstrasser and co-workers found that the <sup>13</sup>C labeled Trp2 amide I vibration is more inhomogeneous than that of Gly7. Further insights were gained by comparing the separation between the positive and negative diagonal peaks to explore the anharmonicity of the <sup>13</sup>C labeled amide I vibration. The authors also concluded that the amide I frequency of Gly7 is  $\sim 16\text{ cm}^{-1}$  lower than that of Trp2, demonstrating that the zero order site energies are not degenerate. By assuming a  $47\text{ cm}^{-1}$  red shift of the amide I vibration upon <sup>13</sup>C substitution, a site energy is assigned to amide groups 2 and 7 (1658.5 and 1645  $\text{cm}^{-1}$  respectively). Based on a symmetry argument, the authors also posit that the site energy of sites Glu5 and Gly7 should be the same; however this seems unreasonable since the hydrogen bond geometry and solvent exposure of sites 5 and 7 are very different. The symmetry argument does, however, seem applicable to the Trp residues 2, 4, 9, & 11 that have the same side chain environment as well as the carbonyl oxygen hydrogen bound to solvent rather than across the strand.

The TZ2 isotope presented in this work can similarly lead to a reasonable approximation for the site energies of residues 3 & 10. Assuming an amide I  $^{13}\text{C}$  isotope shift of  $42\text{ cm}^{-1}$ ,<sup>66</sup> and adding that to the empirical site energy used to calculate the TZ2-T3\*T10\* eigenstates in Fig. 2 gives a value of  $1666\text{ cm}^{-1}$  for Thr3 and  $1668\text{ cm}^{-1}$  for Thr10. We can draw on these values to obtain a rough approximation for the site energies of Glu5 and Lys8, which have very similar local geometries and hydrogen bonding partners. Our site energy assignments combined with those from the previous study<sup>24</sup> are summarized in Fig. 5. Only three main chain amide group site energies are left unassigned: Ser1, Lys12, and Asn6. These predictions show considerably more site-to-site variation than the heuristic model. The alternating pattern can be understood in terms of an additional red shift to sites with solvent exposed oxygen atoms, and has been recognized by other groups.<sup>29,35,78</sup>



**Figure 4.5** Site energy plot for TZ2 using several different models. Solid red trace shows the site energies as calculated from the heuristic model used to simulate the spectra in this report. The blue trace shows the site energies calculated by projecting the electric field onto each amide unit.<sup>28</sup> Green squares are estimates based on isotope labels in this report and cyan diamonds are from the earlier study.<sup>24</sup> Black circles are estimates based on structural similarities with isotope labeled sites. For the remaining dashed black trace, site energies were assigned from the heuristic model. The bar on the left of the graph shows the width ( $2\sigma$ ) of the Gaussian disorder used in the spectral simulations.

The pattern in the empirically obtained site energies can be compared to other methods for assigning TZ2 site energies. Ganim et al. have recently used electrostatic methods for calculating site energies and introducing disorder by using snapshots from MD trajectories in explicit solvent.<sup>28</sup> The site energy shift is calculated from the electric field or electrostatic potential of the solvent and solute atoms acting on the individual amide I sites. The site energies of Ganim et al. obtained for the electric field site energy model of Skinner et al.<sup>42</sup> are compared to our heuristic and deduced TZ2 site energies in Fig. 5. The MD based method predicts a strong alternation in site energies between adjacent oscillators which arises from additional red shift to solvent exposed oscillators. A similar alternation of site energies, albeit with smaller variation, was observed by Hahn

and Cho in calculations of  $\beta$ -hairpin amide I spectra from MD simulations using a electrostatic potential model.<sup>29</sup>

The 2D IR spectral simulations of TZ2 using the same coupling model but treating site energy disorder from MD simulation<sup>28</sup> leads to a somewhat better agreement with experiment than this heuristic model with Gaussian random disorder.<sup>65</sup> Since off-diagonal disorder was observed to be insignificant, this observation suggests that site specific disorder plays an important role in modeling the experiment. This would be expected in the case that the TZ2 turn is tightly folded, while the ends are frayed.

To better understand the influence of these site energy assignments we have compared the calculated linear and 2D IR spectra of TZ2 and eigenstate visualizations for four different site energy models.<sup>65</sup> The heuristic site energy model used in this report has little variation in site energies, and requires relatively large Gaussian disorder to reproduce the experimental amide I linewidth. The spectrum calculated by using empirical site energies deduced from the TZ2 isotopomers shows little change in the shape of the IR and 2D IR spectra, or in the nature of the underlying eigenstates. The calculated TZ2 2D IR spectra using site energies obtained from MD simulations has a broader than observed site energy distribution with a particularly high  $\nu_{||}$  frequency. Finally, we calculate results by modifying the heuristic model to add an additional  $-20$   $\text{cm}^{-1}$  shift for solvent exposed oscillators. This draws on the fact that on average water will make two hydrogen bonds to the peptide CO, and the observation that the site energy shift is additive in the number of hydrogen bonds to water.<sup>40,83</sup> This reproduces the large alternating shift of site energies; however, there is little difference in the shape of the calculated 2D IR spectrum when compared to only one hydrogen bond allowed to

solvent. We note here that the new eigenstates do not differ greatly from those shown in Fig. 1. The degree of delocalization does depend on the site energy assignments, but the underlying symmetry is retained. This allows us to conclude that the eigenstate analysis used to interpret the structure-spectrum relationships in the results sections remains valid even with differences in the assignment of site energies.

Finally, we note that the calculations above do not take into account amide oscillators of the amino acid side chains. In the supplementary material<sup>65</sup> we show a spectrum calculated by using empirical site energies and including the Asn6 side chain amide group and the terminal (Lys12-NH<sub>2</sub>) amide group. The additional oscillators do improve the qualitative agreement to experiment through more defined  $\nu_{\parallel}$  diagonal and cross peaks.

#### 4.6.4. Comparison to previous work

There have been two previous studies of the simulated 2D IR spectroscopy of  $\beta$ -hairpin peptides.<sup>24,29</sup> In both studies, site energies are reported using semi-empirical calculations<sup>24,29</sup> and electric field projections within MD trajectories.<sup>29</sup> The site energies of the strand regions show remarkably similar patterns to those we posit from the isotope labeled systems. Specifically, there is a consistent 10 cm<sup>-1</sup> red shift for site energies of amide groups with carbonyl groups pointing out toward the solvent relative to groups whose carbonyl group points in toward the opposing  $\beta$ -strand. If a heuristic site energy model is further developed, it must be able to account for this effect.

In the earlier reports, the  $\nu_{\perp}$  band was attributed primarily to oscillators in the  $\beta$ -strands, which is consistent with the present interpretation.<sup>24,29</sup> The phase relationships



we assign to the local sites are also consistent with previous work.<sup>24,84</sup> In the 2D IR simulation of the 16-residue GB1 hairpin, Hahn et al. show that the eigenstates at the high frequency region of the spectrum are localized on the hairpin turn residues.<sup>29</sup> Wang et al. made a similar observation based on the spectral distribution of the local modes.<sup>24</sup> For the simulations shown in Fig. 1, the eigenstates in the  $\nu_{\parallel}$  spectral region are highly delocalized and there is no clear evidence that the residues in the turn region preferentially contribute to the high frequency eigenstates. In the case of the GB1 hairpin, the site energies in the turn region are much higher than those in the  $\beta$ -strands, and thus unsurprisingly contribute more to the high frequency eigenstates. The difference in the spectral interpretation is reasonable as the GB1 hairpin comprises a large type I  $\beta$ -turn stabilized by unique side chain hydrogen bonds, while the TZ2 turn is a tight, type II'  $\beta$ -turn.<sup>52</sup> The TZ2 spectral contribution calculated by Wang et al. is unexpected as they set the site energy of Gly7 and Glu5 to 1645  $\text{cm}^{-1}$  leaving only Asn13 (side chain) and Asn6 to account for the spectral intensity.<sup>24</sup> Further isotope labeling schemes could help resolve these issues; however, this is further evidence that the  $\nu_{\parallel}$  eigenstates are highly mixed states whose structural origins are difficult to model with current methods.

#### **4.6.5. Effect of twist on hairpin spectra**

Spectral simulations and eigenstate analysis resolve structural details by relating gross features in the amide I spectrum to the relative participation of the local amide groups in the underlying eigenstates. This type of analysis can be used to compare PG12 to TZ2, and explain the structural basis for the subtle differences in their respective

spectra. The low frequency amide I maximum, or  $\nu_{\perp}$  peak, for TZ2 is  $1636\text{ cm}^{-1}$ , while that of PG12 is  $1640\text{ cm}^{-1}$ , a difference of about  $4\text{ cm}^{-1}$ . When we compare the eigenstates (Fig. 1c and Fig. 2c) we see that the  $\nu_{\perp}$  modes of TZ2 are largely localized on four oscillators, while that of PG12 extends across the sheet. The difference in the delocalization is largely a result of the relative degree of twist in the  $\beta$ -strands. Less twist leads to stronger coupling between the amide groups near the turn and those near the termini, which causes the vibration to delocalize further along the strands. As the vibration delocalizes along the direction of the  $\beta$ -strands, there is a corresponding blue shift in the amide I  $\nu_{\perp}$  maximum. The frequency shift of the amide I band as a function of hairpin length was demonstrated in earlier spectral modeling of ideal hairpins by Cheatum et al.<sup>49</sup> The effects of twist on hairpin spectroscopy have been more directly investigated in an *ab initio* computational study by Bour and Keiderling.<sup>51</sup>

## 4.7. Conclusion

We have investigated the amide I vibrational spectroscopy of several  $\beta$ -hairpin peptides and performed spectral simulations for each system to connect distinct spectral features to peptide structure. The general features in a  $\beta$ -hairpin 2D IR spectrum are the  $\nu_{\perp}$  and  $\nu_{\parallel}$  diagonal peaks with the cross-peak ridge connecting the two frequency regions. The  $\nu_{\perp}$  consists of amide I sites oscillating out of phase with their in-strand neighbor, and in phase with the amide group across the strand. For 12-16 residue hairpin peptides we expect the  $\nu_{\perp}$  frequency to be in the range of  $1635\text{-}1645\text{ cm}^{-1}$ . The  $\nu_{\perp}$  peak is sensitive to both the strength of the cross strand coupling and the delocalization of the vibration along

the direction of the strands. The  $\nu_{\parallel}$  region, or shoulder in the FTIR spectrum, is a much more complicated set of vibrations. The symmetry of these vibrations is less well-defined, but is generally delocalized over the entire hairpin and consists of amide I oscillations largely in phase along the direction of the strand. The cross peak ridge extending from the positive  $\nu_{\parallel}$  peak can also be used as a marker of coupling strength and can help identify the  $\nu_{\perp}$ ,  $\nu_{\parallel}$  doublet when it is not well-resolved in the 2D IR spectrum. The 2D IR spectrum of TZ2-T3\*T10\* show that local  $\nu_{\perp}$  modes can be spectrally isolated and are a good probe of cross-strand hydrogen bond contact. However, isotope labeling must be interpreted carefully, since the isotope induced shift of a subset of oscillators within the hairpin also alters the nature of the remaining delocalized amide I modes.

The amide I model used in this report is a useful tool to analyze the spectroscopy and pattern of eigenstates to deduce structure level details including stability and the relative degree of twist in the  $\beta$ -strands. The heuristic method of choosing site energies based on the hydrogen bond geometries of the amide group qualitatively reproduces the features in the spectrum. As discussed in the previous section, however, there is evidence that the heuristic model does not properly account for the difference between inward pointing amide carbonyl groups and those pointing out toward the solvent. Studies of site energy shifts and heterogeneity drawing on MD simulations will continue to be an important tool for building empirical site energy models. Calculations on the basis of MD simulation also offer an approach to revealing as yet unanswered questions about the importance of accounting for motional narrowing in the spectroscopy.

The 2D IR spectrum of TZ2-T3\*T10\* compared to simulations suggests that TCC is superior to TDC when accounting for the electrostatic coupling of the amide I vibrations. The coupling model used in this report overall gives qualitative agreement with the data. Experimental data generally have diagonal peaks that are more inhomogeneous and cross peak ridges that are less pronounced, which indicates that further refinements are needed to obtain quantitative agreement. It is possible that such improvements may require explicitly treating the anharmonic couplings to other amide vibrations of the peptide.<sup>85</sup> The spectral simulation of Gramicidin S also suggests that the model is transferable to other solvents, and perhaps more significantly, to the unique environments of amide groups buried within folded proteins.

We have shown that a combination of 2D IR spectroscopy, isotope labeling, and spectral modeling has revealed information about site energy distributions, coupling models, and hairpin structure and stability. Such insight is essential as vibrational spectroscopy is increasingly being applied to solve complex problems in protein folding and other questions in biopolymer physics.

## **4.8. Acknowledgements**

The authors would like to thank Prof. Sam Gellman for providing structures of PG12 and providing samples early in the investigations. The authors also would like to thank Prof. Glenn Jones for providing samples of Gramicidin S as well as guidance early in the studies. This work was supported by the NSF (CHE-0316736 and CHE-0616575) and the David and Lucile Packard Foundation.

## 4.9. References

- (1) Searle, M. S.; Ciani, B. *Curr. Opin. Struct. Biol.* **2004**, *14*, 458-464.
- (2) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* **2002**, *11*, 1492-1505.
- (3) Andersen, N. H.; Olsen, K. A.; Fesinmeyer, R. M.; Tan, X.; Hudson, F. M.; Eidenschink, L. A.; Farazi, S. R. *J. Am. Chem. Soc.* **2006**, *128*, 6101-6110.
- (4) Santiveri, C. M.; Santoro, J.; Rico, M.; Jimenez, M. A. *J. Am. Chem. Soc.* **2002**, *124*, 14903-14909.
- (5) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196-199.
- (6) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4077-4082.
- (7) Wang, T.; Xu, Y.; Du, D.; Gai, F. *Biopolymers* **2004**, *75*, 163-172.
- (8) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* **2004**, *336*, 241-251.
- (9) Dyer, R. B.; Maness, S. J.; Peterson, E. S.; Franzen, S.; Fesinmeyer, R. M.; Andersen, N. H. *Biochemistry* **2004**, *43*, 11560-11566.
- (10) Olsen, K. A.; Fesinmeyer, R. M.; Stewart, J. M.; Andersen, N. H. *Proceedings of the National Academy of Sciences* **2005**, *102*, 15483-15487.
- (11) Yang, W. Y.; Gruebele, M. *J. Am. Chem. Soc.* **2004**, *126*, 7758-7759.
- (12) Radhakrishnan Mahalakshmi, G. S., Prasad L. Polavarapu, Padmanabhan Balaram, *ChemBioChem* **2005**, *6*, 2152-2158.
- (13) Du, D.; Tucker, M. J.; Gai, F. *Biochemistry* **2006**, *45*, 2668-2678.

(14) 2D IR spectroscopy is a coherent Fourier transform infrared spectroscopy analogous to COSY methods used in NMR. This differs from the 2D correlation spectroscopy originally proposed by Noda,(ref 15) which is a correlation analysis of spectral changes to a time series of FTIR spectra in the presence of a perturbation. The axis labeling scheme used in our work uses  $\omega_1$  and  $\omega_3$  to refer to the Fourier transform variables of the first (excitation) and third (detection) time periods respectively. This differs from other reports (for instance Ref. 24) where the  $\omega_1$  axis is termed  $\omega_t$  and is plotted on the ordinate while the  $\omega_3$  axis is labeled as  $\omega_t$  and plotted on the abscissa. 2D contour plots surface plots can be compared by reflecting about the diagonal axis.

(15) Noda, I. *Appl. Spectrosc.* **1993**, *47*, 1329.

(16) Hamm, P. K.; Lim, M.; Hochstrasser, R. M. *Abstr. Pap. Am. Chem. Soc.* **1999**, *218*, 104-PHYS.

(17) Hamm, P.; Woutersen, S. *Bull. Chem. Soc. Jpn.* **2002**, *75*, 985-988.

(18) Woutersen, S.; Hamm, P. *J. Chem. Phys.* **2001**, *114*, 2727-2737.

(19) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123-6138.

(20) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181-364.

(21) Smith, A. W.; Chung, H. S.; Ganim, Z.; Tokmakoff, A. *J. Phys. Chem. B* **2005**, *109*, 17025-17027.

(22) Chung, H. S.; Khalil, M.; Smith, A. W.; Ganim, Z.; Tokmakoff, A. In *Ultrafast Phenomena XIV*; Kobayashi, T., Okada, T., Kobayashi, T., Nelson, K. A., DeSilvestri, S., Eds.; Springer-Verlag, Berlin: Niigata, Japan, 2005, pp 551-553.

(23) Chung, H. S.; Khalil, M.; Smith, A. W.; Ganim, Z.; Tokmakoff, A. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 612-617.

- (24) Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. *J. Phys. Chem. B* **2006**, *110*, 7545-7555.
- (25) Paul, C.; Wang, J.; Wimley, W. C.; Hochstrasser, R. M.; Axelsen, P. H. *J. Am. Chem. Soc.* **2004**, *126*, 5843-5850.
- (26) Khalil, M.; Demirdoven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258-5279.
- (27) Dijkstra, A. G.; Knoester, J. *J. Phys. Chem. B* **2005**, *109*, 9787-9798.
- (28) Ganim, Z.; Tokmakoff, A. *Biophys. J.* **2006**, *91*, 2636-2646.
- (29) Hahn, S.; Ham, S.; Cho, M. *J. Phys. Chem. B* **2005**, *109*, 11789-11801.
- (30) Moran, A.; Mukamel, S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 506-510.
- (31) Gorbunov, R. D.; Kosov, D. S.; Stock, G. *J. Chem. Phys.* **2005**, *122*, 224904-224912.
- (32) Jansen, T. I. C.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. *J. Chem. Phys.* **2006**, *125*, 044312-044319.
- (33) Lee, C.; Cho, M. *J. Phys. Chem. B* **2004**, *108*, 20397-20407.
- (34) Chung, H. S.; Tokmakoff, A. *J. Phys. Chem. B* **2006**, *110*, 2888-2898.
- (35) Bour, P.; Keiderling, T. A. *J. Phys. Chem. B* **2005**, *109*, 23687-23697.
- (36) Bour, P.; Keiderling, T. A. *J. Phys. Chem. B* **2005**, *109*, 5348-5357.
- (37) Torii, H.; Tasumi, M. *J. Chem. Phys.* **1992**, *96*, 3379-3387.
- (38) Torii, H.; Tatsumi, T.; Tasumi, M. *Mikrochim. Acta [Suppl.]* **1997**, *14*, 531-533.
- (39) Torii, H.; Tatsumi, T.; Tasumi, M. *J. Raman Spectros.* **1998**, *29*, 537-546.
- (40) Ham, S.; Kim, J.-H.; Lee, H.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 3491-3498.
- (41) Ham, S.; Cho, M. *J. Chem. Phys.* **2003**, *118*, 6915-6922.

- (42) Schmidt, J. R.; Corcelli, S. A.; Skinner, J. L. *J. Chem. Phys.* **2004**, *121*, 8887-8896.
- (43) Jansen, T. I. C.; Knoester, J. *J. Chem. Phys.* **2006**, in press.
- (44) Bour, P.; Keiderling, T. A. *J. Am. Chem. Soc.* **1993**, *115*, 9602-9607.
- (45) Krimm, S.; Abe, Y. *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 2788-2792.
- (46) Ham, S.; Cha, S.; Choi, J.-H.; Cho, M. *J. Chem. Phys.* **2003**, *119*, 1451-1461.
- (47) Chirgadze, Y. N.; Nevskaya, N. A. *Biopolymers* **1976**, *15*, 607-625.
- (48) Miyazawa, T.; Blout, E. R. *J. Am. Chem. Soc.* **1961**, *83*, 712-719.
- (49) Cheatum, C. M.; Tokmakoff, A.; Knoester, J. *J. Chem. Phys.* **2004**, *120*, 8201-8215.
- (50) Demirdöven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. *J. Am. Chem. Soc.* **2004**, *126*, 7981-7990.
- (51) Bour, P.; Keiderling, T. A. *Theochem* **2004**, *675*, 95-105.
- (52) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **2001**, *99*, 9081-9081.
- (53) Syud, F. A.; Espinoza, J. F.; Gellman, S. H. *J. Am. Chem. Soc.* **1999**, *121*, 11577-11578.
- (54) Waki, M.; Izumiya, N. In *Biochemistry of peptide antibiotics: Recent advances in the biotechnology of b-lactams and bioactive peptides*; Kleinkauf, H., von Dohren, H., Eds., 1990, pp 205-244.
- (55) Woutersen, S.; Hamm, P. *J. Phys.: Condens. Mat.* **2002**, *14*, 1035-1062.
- (56) Doi, M.; Fujita, S.; Katsuya, Y.; Sasaki, M.; Taniguchi, T.; Hasegawa, H. *Arch. Biochem. Biophys.* **2001**, *395*, 85-93.



- (57) Wang, J.; Hochstrasser, R. M. *Chem. Phys.* **2004**, *297*, 195-219.
- (58) Hamm, P.; Lim, M.; DeGrado, W. F.; Hochstrasser, R. M. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 2036-2041.
- (59) Bolhuis, P. G. *Proc. Nat. Acad. Sci., USA* **2003**, *100*, 12129-12134.
- (60) Pitera, J. W.; Haque, I.; Swope, W. C. *J. Chem. Phys.* **2006**, *124*, 141102.
- (61) Ulmschneider, J. P.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2004**, *126*, 1849-1857.
- (62) Barth, A.; Zscherp, C. *Q. Rev. Biophys* **2002**, *35*, 369-430.
- (63) Chung, H. S.; Khalil, M.; Tokmakoff, A. *J. Phys. Chem. B* **2004**, *108*, 15332-15343.
- (64) Barth, A. *Progress in Biophysics and Molecular Biology* **2000**, *74*, 141-173.
- (65) See EPAPS Document No. \_\_\_\_\_ for additional spectral simulations, site energies, and eigenstates. This document can be reached through a direct link in the online article's HTML reference section or via the EPAPS homepage (<http://www.aip.org/pubservs/epaps.html>).
- (66) Decatur, S. M. *Acc. Chem. Res.* **2006**, *39*, 169-175.
- (67) Jaume Torres, A. K., Jonathan M. Goodman, Isaiah T. Arkin, *Biopolymers* **2001**, *59*, 396-401.
- (68) Tiansheng Li, J. T., Lisa Zeni, Robert Rosenfeld, George Stearns, Tsutomu Arakawa, *Biopolymers* **2002**, *67*, 10-19.
- (69) Setnicka, V.; Huang, R.; Thomas, C. L.; Etienne, M. A.; Kubelka, J.; Hammer, R. P.; Keiderling, T. A. *J. Am. Chem. Soc.* **2005**, *127*, 4992-4993.
- (70) Syud, F. A.; Stanger, H. E.; Mortell, H. S.; Espinosa, J. F.; Fisk, J. D.; Fry, C. G.; Gellman, S. H. *J. Mol. Biol.* **2003**, *328*, 303-303.

- (71) Hilario, J.; Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2003**, *125*, 7562-7574.
- (72) Gause, G. F.; Brazhnikova, M. G. *Nature* **1944**, *154*, 703-703.
- (73) Krauss, E. M.; Chan, S. I. *J. Am. Chem. Soc.* **1982**, *104*, 1824-1830.
- (74) Live, D. H.; Davis, D. G.; Agosta, W. C.; Cowburn, D. *J. Am. Chem. Soc.* **1984**, *106*, 1939-1941.
- (75) Lewis, R. N. A. H.; Prenner, E. J.; Kondejewski, L. H.; Flach, C. R.; Mendelsohn, R.; Hodges, R. S.; McElhaney, R. N. *Biochemistry* **1999**, *38*, 15193-15203.
- (76) DeCamp, M. F.; DeFlores, L. P.; McCracken, J. M.; Tokmakoff, A.; Kwac, K.; Cho, M. *J. Phys. Chem. B* **2005**, *109*, 11016-11026.
- (77) Zhao, C.; Polavarapu, P. L.; Das, C.; Balaram, P. *J. Am. Chem. Soc.* **2000**, *122*, 8228-8231.
- (78) Maness, S. J.; Franzen, S.; Gibbs, A. C.; Causgrove, T. P.; Dyer, R. B. *Biophys. J.* **2003**, *84*, 3874-3882.
- (79) Gangani, R. A.; Silva, D.; Sherman, S. A.; Perini, F.; Bedows, E.; Keiderling, T. A. *J. Am. Chem. Soc.* **2000**, *122*, 8623-8630.
- (80) Colley, C. S.; Griffiths-Jones, S. R.; George, M. W.; Searle, M. S. *Chemical Communications (Cambridge)* **2000**, 593-594.
- (81) Du, D.; Zhu, Y.; Huang, C.-Y.; Gai, F. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15915-15920.
- (82) Jackson, J. D. *Classical Electrodynamics*; 2nd ed.; John Wiley and Sons: New York, 1975.
- (83) Torii, H.; Tasumi, M. *J. Raman. Spectrosc.* **1998**, *29*, 537-546.

- (84) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2001**, *123*, 6142-6150.
- (85) Hayashi, T.; Zhuang, W.; Mukamel, S. *J. Phys. Chem. A* **2006**, *109*, 9747-9759.

# Chapter 5

## Residual Native Structure in a Thermally Denatured $\beta$ -Hairpin

The work presented in this chapter has been published in the following paper:

“Residual Native Structure in a Thermally Denatured  $\beta$ -Hairpin,” A. W. Smith, H. S. Chung, Z. Ganim, A. Tokmakoff, *J. Phys. Chem. B* **109**, 17025 (2006).

### 5.1. Abstract

We investigate the thermal denaturation of trpzip2 between 15 and 82 °C using two-dimensional infrared (2D IR) vibrational spectroscopy, dispersed vibrational echo (DVE) spectroscopy, and Fourier transform infrared (FTIR) spectroscopy. The FTIR and DVE spectra of trpzip2 show two resonances in the amide I region of the spectrum, which arise primarily from the interstrand coupling between local amide I oscillators along the peptide backbone. The coupling is directly seen in the 2D IR spectra as the formation of cross peak ridges. Although small shifts of these frequencies occur on heating the sample, the existence of cross-peak ridges at all temperatures indicates that

stable hydrogen bond interactions persist between the two  $\beta$ -strands. These observations indicate a significant amount of native structure in the thermally denatured state of trpzip2.

## 5.2. Introduction

Small peptides that adopt a single secondary structure are valuable model systems to study factors influencing protein folding and stability.<sup>1,2</sup> Of these,  $\beta$ -hairpins represent a fundamental subunit of extended  $\beta$ -structures that are common in naturally occurring proteins. The small size of hairpin-forming peptides make them well-suited for folding studies<sup>3-6</sup> because they can be directly modeled with molecular dynamics simulations and experimental spectra are more easily related to the peptide morphology. Additionally, debate about the folding pathway for  $\beta$ -hairpins mirrors the questions posed for folding in globular proteins, which centers on the time ordering, relative importance, and cooperativity between physical events. For example, recent investigations of hairpin peptides have proposed several different initiating events in the folding pathway such as the collapse of the hydrophobic core or the formation of native contacts in the turn region.<sup>7-9</sup> Further restructuring can occur through the formation of native hydrogen bonds or through more efficient packing of the side chains. These events may happen in concert or in a specific sequence. Experimentally it has been difficult to obtain detailed structural information on the time scale of these events that can discriminate between the various pathways. It is especially difficult to identify and characterize non-native ensembles, which are necessary to describe the pathway from the unfolded conformational ensemble to the more compact native state.<sup>6</sup> In this study we have used two femtosecond

vibrational spectroscopies, two-dimensional infrared (2D IR) spectroscopy and dispersed vibrational echo (DVE) spectroscopy,<sup>10,11</sup> to investigate changes in the amide I band during the equilibrium thermal denaturation of the trpzip2 hairpin. These methods are particularly sensitive to hydrogen bonding contacts between peptides in antiparallel registry. We find that the interstrand hydrogen bond network of trpzip2 is thermally stable and retains a considerable amount of native contacts in the thermally denatured state.

The tryptophan zipper peptides are a family of de novo  $\beta$ -hairpins designed to be especially stable. Much of the stability is attributed to the tryptophan residues that form a tightly folded hydrophobic core.<sup>12</sup> Trpzip2 is a twelve residue peptide in which two  $\beta$ -strands are connected with a type I' turn using the ENGK sequence. It has been studied using molecular dynamics simulations and a variety of experimental methods.<sup>6,13-17</sup> Equilibrium studies of thermal denaturation using circular dichroism, fluorescence, and IR spectroscopy have shown that the energy landscape of trpzip2 is heterogeneous, with a variety of configurational minima.<sup>14</sup> Melting temperatures from each measurement differed, and were interpreted as the transition temperatures for various structural changes. The FTIR data indicated a melting temperature of  $>60$  °C, which was interpreted as the temperature at which backbone hydrogen bonds are broken.<sup>14</sup> Subsequent work characterized IR absorbance changes of the amide I band around  $1676$   $\text{cm}^{-1}$  and gave a melting curve with a much narrower melting transition centered at  $60$  °C.<sup>15</sup> The 2D IR spectroscopy used in this study builds on the information obtained from the earlier FTIR spectroscopy by adding an extra dimension of resolution and increased sensitivity to vibrational coupling. The results of this work show directly that the two

amide I vibrational modes resulting from the native interstrand hydrogen bonds persist up to 82°C.

### 5.3. Results & Discussion

The IR spectra of  $\beta$ -strands in antiparallel registry show two amide I vibrational resonances. Strong electrostatic couplings between local amide I oscillators leads to vibrations that are delocalized across the sheet.<sup>18,19</sup> The IR oscillator strength is carried by two modes whose transition moment is oriented roughly perpendicular or parallel to the strands ( $\nu_{\perp}$  and  $\nu_{\parallel}$ ), and whose splitting is primarily determined by the interstrand coupling.<sup>18</sup> Such features are difficult to discern in traditional IR spectra, which are often broadened by disorder; however, delocalized amide I states can be revealed with 2D IR spectroscopy.<sup>20</sup> 2D IR spectroscopy is a vibrational analogue of multidimensional NMR.<sup>21,22</sup> It uses a sequence of femtosecond mid-infrared pulses to interrogate couplings between molecular vibrations. These are observed as cross-peaks in a Fourier transform 2D IR spectrum. Our prior investigations show that cross peaks are observed between the  $\nu_{\perp}$  and  $\nu_{\parallel}$  modes of antiparallel  $\beta$ -sheets, even when the FTIR spectrum is congested.<sup>10,11</sup> For proteins, the cross peaks elongate and interfere with the diagonal peaks to become ridges, yielding characteristic “Z”-shaped contours in the 2D IR lineshape.<sup>10,11,23</sup> By using 2D IR spectroscopy to investigate trpzip2, it is possible to characterize the changes in interstrand amide I couplings that would arise from complete unfolding of the hairpin.

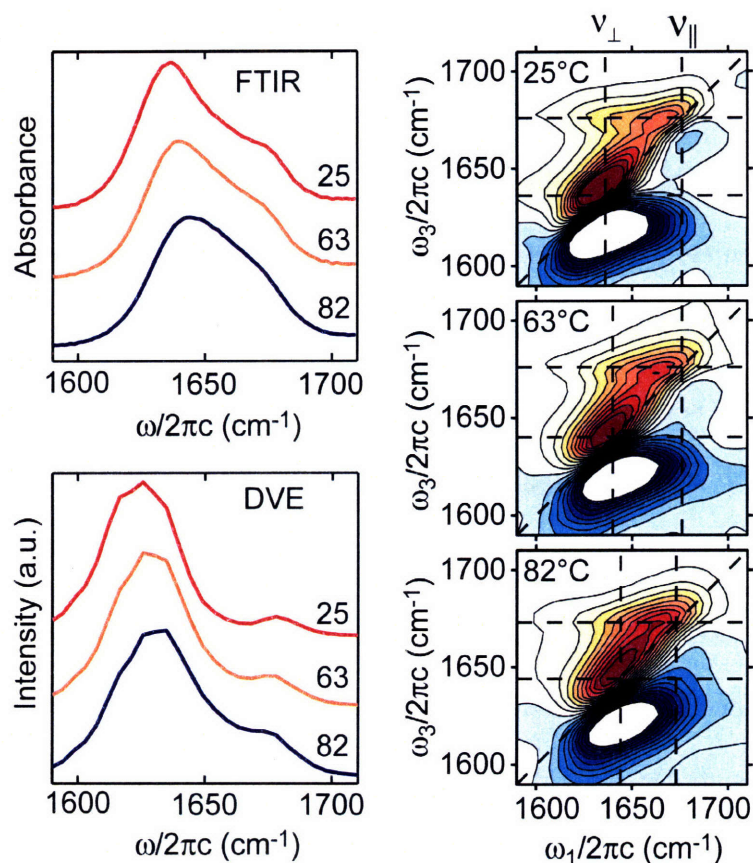
Trpzip2 was synthesized as a C-terminal amide peptide using established Fmoc solid phase synthesis. The peptide was HPLC purified and then further treated by

lyophilization against 50 mM HCl to remove residual trifluoroacetic acid. The sample was dissolved in D<sub>2</sub>O, heated to 60°C for 30 minutes for hydrogen-deuterium exchange, and lyophilized before being used in the experiments. For all experiments, the sample was dissolved in a 50 mM, pH 7.0, potassium deuterium phosphate buffer to a final concentration of about 4 mM. For spectral measurements, the sample was placed between CaF<sub>2</sub> windows separated with a 50 μm Teflon spacer. The filled sample cell was heated to 80°C and then cooled to room temperature before any measurements were taken. The experimental methods used to acquire 2D IR and DVE spectra are described elsewhere.<sup>10</sup> Spectra were collected over temperatures ranging from 15 to 86 °C. Higher temperatures were not accessed because of trpzip2's propensity to aggregate. Under the described conditions, there was no spectral evidence for the formation of trpzip2 aggregates. The sample conditions in this study closely matched those outlined in previous investigations where there also was no evidence of aggregation.<sup>14,15</sup>

Representative FTIR spectra are shown in Fig. 1. In the room temperature spectrum, the amide I absorption maximum is at 1636 cm<sup>-1</sup>, with a shoulder at higher frequencies indicating a second peak at 1676 cm<sup>-1</sup>. The overall changes in the FTIR spectrum with temperature were analyzed by singular value decomposition (SVD). The second component spectrum, representing the spectral changes over the temperature range, is displayed in Fig. 2a, with the corresponding melting curve shown in Fig. 2c. The loss of the two mode structure is seen in the second component spectrum as two positive-going peaks at 1629 and 1680 cm<sup>-1</sup> representing the  $\nu_{\perp}$  and  $\nu_{\parallel}$  modes at 1636 and 1676 cm<sup>-1</sup>. The broad negative peak at 1655 cm<sup>-1</sup> is attributed to increasing random coil-like geometry. Although there is no indication of a sharp transition temperature, these



results suggest that interstrand couplings are being lost during thermal denaturation. Indeed the loss of amplitude in the shoulder was previously observed and interpreted as loss of the high frequency component common to antiparallel  $\beta$ -sheets.<sup>15</sup>

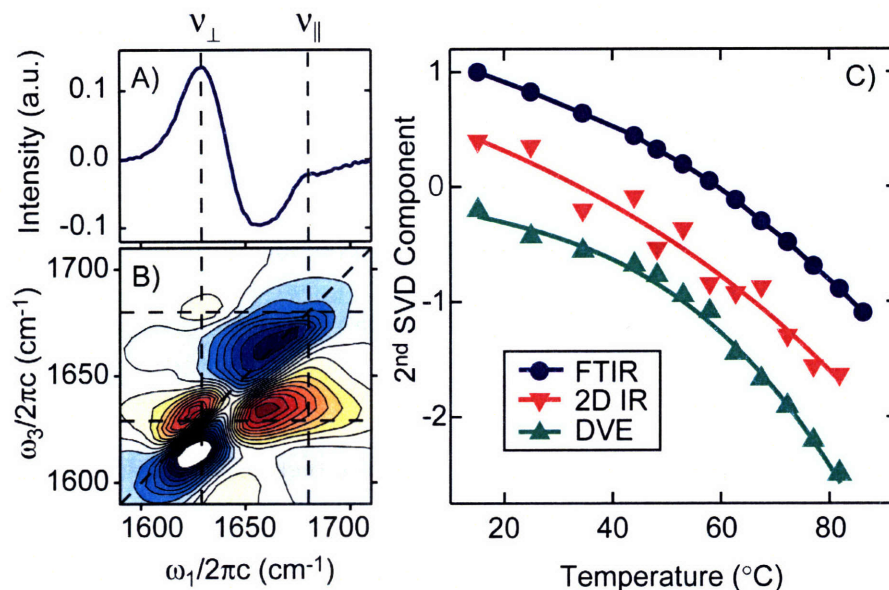


**Figure 5.1** FTIR, DVE, and 2D IR spectra of trpzip2 are shown at three representative temperatures: 25, 63, and 82 °C. FTIR and DVE spectra are offset for clarity. For 2D IR spectra, 20 equally spaced contour levels are drawn to  $\pm 60\%$  of peak intensity.

Representative temperature-dependent 2D IR spectra of trpzip2 are shown in Fig. 1, and results of SVD analysis are shown in Fig. 2 (B&C). At 25 °C, two peaks are observed along the diagonal axis ( $\omega_1 = \omega_3$ ) corresponding to the  $v_{\perp}$  and  $v_{\parallel}$  resonances. (In 2D IR spectra, each resonance is composed of a vertically displaced, oppositely signed doublet). Cross peaks between the two resonances appear in the off-diagonal region of

the spectrum, and are shaped here due to overlap with the main diagonal peaks. In the off-diagonal region centered at  $\omega_1 = \nu_{\perp}$  and  $\omega_3 = \nu_{\parallel}$ , the cross peaks appear as a ridge stretching in the  $\omega_1$  direction at constant  $\omega_3$ . The ridges indicate a distinct two-mode structure characteristic of the splitting induced by vibrational couplings and consistent with the cross-strand interactions of a hairpin with antiparallel symmetry.<sup>11,18</sup> With increasing temperature, the diagonal peaks shift to band center, and some of the ridge structure is lost (Fig. 2b). However, at 82 °C there are still two distinct maxima in the positive diagonal peaks. The splitting decreases with increasing temperature, but the cross peak ridge is observed at all temperatures. This indicates that antiparallel structure and considerable native hydrogen bonding contacts remain in trpzip2 at high temperature.

Similar observations are also apparent in the DVE spectra (Figs. 1 and 2c). DVE measurements are related to the absolute value of the 2D IR surface projected onto the  $\omega_3$  axis.<sup>10</sup> For  $\beta$ -sheets, the projection along cross peak ridges that originate in the interstrand couplings leads to distinct features in the spectrum. The DVE spectra of trpzip2 display two resolvable peaks in the amide I' region of the spectrum (Fig 1). At higher temperatures, the peaks shift towards band center, but retain a spacing of about 30  $\text{cm}^{-1}$  at 82 °C. This result provides further evidence of residual hydrogen bond interactions between the two strands in the denatured state.



**Figure 5.2** Second component spectra,  $c_N^{(2)}(\omega)$ , calculated from the SVD analysis of the FTIR (A) and 2D IR (B) data. The dashed guidelines intersect the axes at 1629 and 1680 cm<sup>-1</sup> (C) Melting curves from each experiment are plotted as the amplitude of the second SVD component at each temperature. The solid lines are guides for the eye. For clarity, the 2D IR and DVE melting curves are offset by -0.6 and -1.2 respectively.

The melting curves for each of the infrared spectroscopy techniques in this study follow the same trends, including a broad melting transition and an approximate melting temperature of 60 °C. Similar trends were observed in previously reported FTIR data.<sup>14,15</sup> Wang et. al. specifically analyzed the high frequency ( $\sim 1676\text{cm}^{-1}$ ) amide I transition in the FTIR spectrum, assigning it to the  $\nu_{\parallel}$  component of the  $\beta$ -sheet doublet and interpreting the loss of the peak as the loss of  $\beta$ -sheet structure. This high frequency mode is directly observable in the 2D IR and DVE spectra without the need for band decomposition or subtraction methods. The added resolution, along with the direct observation of coupling, confirms the peak assignment, while contributing much more detail to the structural interpretation. Specifically, the cross peak ridge observed in the

82°C 2D IR spectrum demonstrates that the symmetry of the antiparallel contacts between strands is preserved and that the proximity of the strands for these contacts remains similar to that at low temperatures. Because both the 2D IR and the DVE spectra are able to explicitly resolve the high frequency component at 82 °C, we conclude that there is a more significant population of molecules with intact interstrand hydrogen bonds than was previously recognized.

The results of our IR measurements point to a picture in which the peptide remains compact, with interstrand hydrogen-bond contacts persisting even at high temperatures. Our high temperature spectra are consistent with fraying of the ends of the hairpin and an intact turn region much like the frayed or F state recently reported in the Bolhuis study of the GB1 hairpin.<sup>7</sup> We can also not rule out increased amplitude of fluctuations around the native geometry, a view consistent with the mean structure hypothesis proposed by Zagrovic et al.<sup>6</sup> More specific conclusions will be possible following careful simulations of the amide I vibrational couplings and spectroscopy, drawing on the results of previous molecular dynamics simulations.<sup>6,14,16,17</sup> This work has shown that the details obtained from the 2D IR and DVE spectra add another experimental observable that can help researchers construct a more accurate model of  $\beta$ -hairpin folding. By incorporating these methods into a transient experiment<sup>24</sup> it will be possible to resolve events such as the formation of the turn region and the formation of interstrand hydrogen bonds. The results may be combined with those obtained from time-resolved fluorescence experiments that probe the collapse and restructuring of the hydrophobic core to more accurately test recent simulations of  $\beta$ -hairpin folding.<sup>7-9</sup>

## 5.4. Conclusion

The gradual changes in the trpzip2 melting curves over the temperature range measured here indicate a thermally stable hydrogen bond network. This observation, combined with the earlier observation of a compact hydrophobic core at high temperatures, leads to the conclusion that temperature dependent studies on  $\beta$ -hairpins, either at equilibrium or as temperature-jump experiments may not be probing large changes in peptide conformation – at least without the presence of additional chemical denaturant. Such considerations must be taken into account when the purpose is to obtain detailed information on folding from an extended state. Ongoing investigations are needed to understand how this observation holds for other systems and how it varies depending on the stability of the system. De novo systems designed for stability may inherently fold differently when compared to naturally occurring systems where the dynamics are shaped through evolutionary mechanisms.

## 5.5. Acknowledgements

This work was supported by the NSF (CHE-0316736), the ACS Petroleum Research Fund, and the David and Lucile Packard Foundation.

## 5.6. References

- (1) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76-88.
- (2) Searle, M. S.; Ciani, B. *Curr. Opin. Struct. Biol.* **2004**, *14*, 458-464.
- (3) Dyer, R. B.; Maness, S. J.; Peterson, E. S.; Franzen, S.; Fesinmeyer, R. M.; Andersen, N. H. *Biochemistry* **2004**, *43*, 11560-11566.
- (4) Munoz, V.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 5872-5879.
- (5) Dinner, A. R.; Lazaridis, T.; Karplus, M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9068-9073.
- (6) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153-164.
- (7) Bolhuis, P. G. *Biophys. J.* **2005**, *88*, 50-61.
- (8) Tsai, J.; Levitt, M. *Biophysical Chemistry* **2002**, *101*, 187-201.
- (9) Wei, G. H.; Mousseau, N.; Derreumaux, P. *Proteins* **2004**, *56*, 464-474.
- (10) Chung, H. S.; Khalil, M.; Tokmakoff, A. *J. Phys. Chem. B* **2004**, *108*, 15332-15343.
- (11) Demirdöven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. *J. Am. Chem. Soc.* **2004**, *126*, 7981-7990.
- (12) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **2001**, *99*, 9081-9081.
- (13) Yang, W. Y.; Gruebele, M. *J. Am. Chem. Soc.* **2004**, *126*, 7758-7759.

- (14) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* **2004**, *336*, 241-251.
- (15) Wang, T.; Xu, Y.; Du, D.; Gai, F. *Biopolymers* **2004**, *75*, 163-172.
- (16) Ulmschneider, J. P.; Jorgensen, W. L. *J. Am. Chem. Soc.* **2004**, *126*, 1849-1857.
- (17) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4077-4082.
- (18) Cheatum, C. M.; Tokmakoff, A.; Knoester, J. *J. Chem. Phys.* **2004**, *120*, 8201-8215.
- (19) Bour, P.; Keiderling, T. A. *J. Chem. Phys.* **2003**, *119*, 11253-11262.
- (20) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123-6138.
- (21) Zanni, M. T.; Hochstrasser, R. M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 516-522.
- (22) Khalil, M.; Demirdoven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258-5279.
- (23) Dijkstra, A. G.; Knoester, J. *J. Phys. Chem. B* **2005**, *109*, 9787-9798.
- (24) Chung, H. S.; Khalil, M.; Smith, A. W.; Ganim, Z.; Tokmakoff, A. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 612-617.

# Chapter 6

## Probing Local Structural Events in $\beta$ -Hairpin Unfolding With Transient Nonlinear Infrared Spectroscopy

The work presented in this chapter has been published in the following paper:

- “Probing Local Structural Events in  $\beta$ -Hairpin Unfolding With Transient Nonlinear Infrared Spectroscopy,” A. W. Smith and A. Tokmakoff, *Angew. Chem. Int. Ed.* **46**, 7984.

### 6.1. Abstract

PG12 is a de-novo designed  $\beta$ -hairpin peptide that forms a  $\beta$ -hairpin stabilized by a D-Pro enhanced type II'  $\beta$ -turn. Using a combination of  $^{13}\text{C}$  isotope labels and temperature-jump nonlinear infrared spectroscopy we observe local folding events at a sub-microsecond time scale. The observations support an unfolding paradigm in which the turn region remains structured during the temperature jump, and mid-strand region opens up on a 130 ns time scale.



## 6.2. Introduction

Contemporary descriptions of protein folding pathways have been greatly enhanced by the study of single domain peptides.<sup>1,2</sup> Hairpins in particular have been used to investigate the role of  $\beta$ -turns in secondary structure stability and have also been used to explore the stabilizing effect of tertiary side chain contacts.<sup>2-5</sup> Transient folding studies are particularly valuable because  $\beta$ -hairpins are frequently used to compare and test theoretical protein folding models.<sup>2,6-8</sup> To date, however, fast folding experiments on the nanosecond to microsecond timescale are limited in their level of structural detail. Previous hairpin folding studies have probed either the transient fluorescence of side chain chromophores<sup>9,10</sup> or the transient infrared absorbance of the amide I band,<sup>11,12</sup> a congested spectral feature consisting of delocalized CO vibrations of the peptide backbone.<sup>13,14</sup> Here we report on an experimental method that increases the structural resolution of amide I' spectroscopy to probe the unfolding kinetics of individual amino acid residues on a sub-microsecond time scale following a temperature jump (T-jump).

A combination of two methods is used to elucidate detailed structural information from T-jump amide I vibrational spectroscopy. First, <sup>13</sup>C' isotope labelling is used to spectrally isolate individual amino acid residues within the hairpin sequence. Amide I' spectral features are separated and assigned with two-dimensional infrared (2D IR) spectroscopy. Second, dispersed vibrational echo (DVE) spectroscopy is used to probe the sample following a nanosecond  $\sim 10$  °C T-jump. The primary advantage of the nonlinear IR techniques over traditional absorbance experiments is that they are sensitive to vibrational couplings and different line broadening mechanisms. In this report, these methods are used to study the turn region and cross-strand contacts of the *de novo*

peptide, PG12.<sup>3</sup> The V3-E4 amide group, which is involved in cross-strand hydrogen bond, is found to unfold on a time scale of 130 ns. Furthermore, it is found that the turn region of PG12 does *not* undergo significant structural change during the T-jump. These experimental observations are consistent with unfolding of the peptide by fraying at the N and C-termini. The results represent an important step toward building an experimentally verified hairpin folding mechanism.

The hairpin folding mechanisms that have been proposed are guided by two principle paradigms. The first is the kinetic zipper model, in which formation of the  $\beta$ -turn is the important nucleation event that triggers the zipping together of cross strand contacts beginning at the turn and continuing to the termini.<sup>15</sup> In the second model, folding begins with the collapse of the hydrophobic side chain groups, following which the backbone contacts adjust to their native state.<sup>16</sup> Deviations from these two models have been observed in simulation studies, where the folding mechanisms are distinguished by more subtle structural events. Some of these events include partial hydrophobic collapse, fraying of the N and C terminal residues, the ordering of the cross strand hydrogen bonds, salt bridging between cross-strand side chains, and rearrangements of the hydrophobic core.<sup>6,8,17-19</sup> The time scale and relative importance of these events has been discussed in a number of publications, but experimental data has not been able to directly distinguish between the proposed models.

While no decisive method has been developed, kinetic experiments have provided considerable insight. Side chain fluorescence studies can resolve structural changes within the hydrophobic core. Such measurements assign folding times ranging from 5-10  $\mu$ s,<sup>9,11,20</sup> although in some systems, this rate has been seen to vary depending on starting

conditions.<sup>10</sup> Amide I transient absorbance methods have been carried out for a number of hairpin peptides. Typical time scales are on the order of 10  $\mu$ s,<sup>12,21,22</sup> although faster kinetics have been observed in cyclic peptides<sup>23</sup> or in peptides with a proline-enhanced  $\beta$ -turn.<sup>22,24,25</sup> In addition, mutation studies have been combined with spectroscopic probes to investigate the role of specific residues on the folding rate.<sup>21</sup> The key information that is unavailable in any of these studies is the folding rate of individual backbone contacts – particularly the hydrogen bond contacts that help stabilize the  $\beta$ -sheet secondary structure.

Isotope-edited amide I spectroscopy is one route towards obtaining site-specific information of polypeptides.<sup>26</sup> By substituting a  $^{13}\text{C}$  atom at the C' position, the amide I vibrational frequency is reduced by approximately 40  $\text{cm}^{-1}$ , which isolates it from the main amide I' band. The frequency shift effectively decouples the oscillator from the other vibrations so that it is approximately 90 % localized on the labelled amide group.<sup>14</sup> This effect has been used to investigate the amide I band of  $\beta$ -hairpin peptides,<sup>14,27,28</sup> and has also been used to study various other peptides and proteins.<sup>26,29,30</sup> Here the technique is used to spectrally isolate two peptide groups on PG12, a 12 residue hairpin with a D-Proline enhanced type II'  $\beta$ -turn.<sup>3</sup>

### 6.3. Experimental Details

PG12 (RYVEV-D-Pro-GKKILQ) samples were synthesized as C-terminal amides using solid phase peptide synthesis, and were HPLC purified by the MIT Biopolymers Laboratory.  $^{13}\text{C}'$  labelled Valine was obtained from Cambridge Isotope labs (99 % isotopic purity). For infrared measurements, the peptides were lyophilized repeatedly in

10 mM DCl to remove trifluoroacetic acid impurities, and to complete the H-D exchange. Experiments were performed at a peptide concentration of 17.5 mM in a deuterated, 10 mM acetate buffer at pH 3.85 (uncorrected for D<sub>2</sub>O).

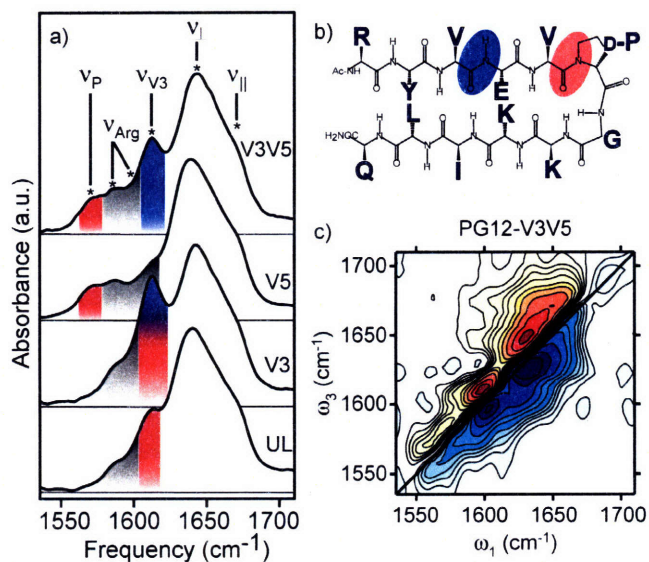
The 2D IR and DVE methods have been published in detail elsewhere.<sup>31,32</sup> The sample cell for all infrared measurements consists of two 1mm thick CaF<sub>2</sub> windows separated with a 50 μm Teflon spacer. The temperature was controlled with a circulating water bath and measured with a thermocouple in contact with the brass housing surrounding the cell. Equilibrium DVE spectra were collected with each beam in the parallel (zzzz) polarization with the waiting time,  $\tau_2$ , set to 100 fs. 2D IR spectra were collected in the perpendicular polarization geometry (zzyy) with a 100 fs waiting time. The evolution time,  $\tau_1$ , was collected for delay times from  $\tau_1=0$  to 2ps and Fourier transformed with a rectangular windowing function.

The T-jump DVE apparatus has been described previously.<sup>31-33</sup> An 8 ns, 6 mJ laser pulse with a frequency of 1.98 μm was used to raise the solvent temperature from a bath-controlled starting temperature,  $T_0$ , of 25.0±0.1 °C to a final temperature of 35±1 °C. The time-dependent temperature profile of the buffer as a function of  $\tau$  was calibrated using transient solvent absorbance of a 6μm laser pulse. The DVE spectrum at each  $\tau$  point was collected for 1000 laser shots, and the entire set of time delays was measured 15 times.

## 6.4. Amide I Spectroscopy of Folded PG12

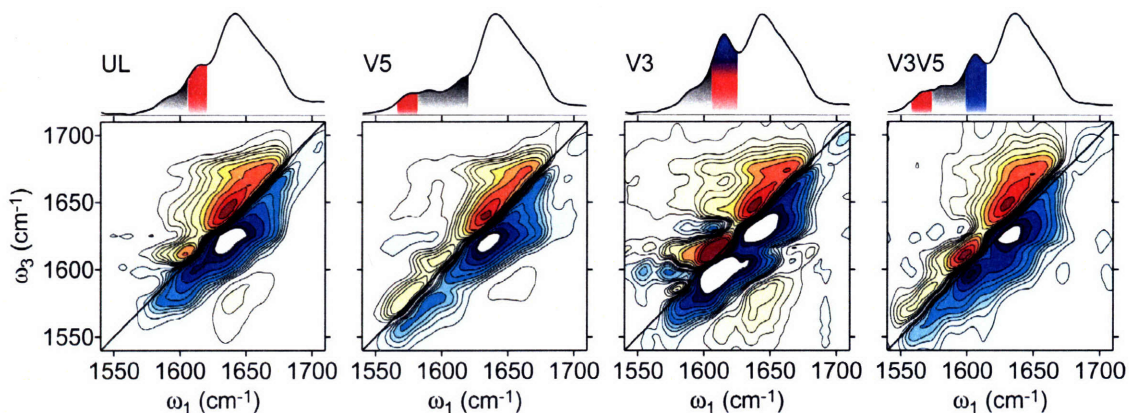
The FTIR spectrum of unlabelled PG12 (PG12-UL) has three amide I peaks (Fig. 6.1). The most intense peak ( $\nu_{\perp}$ ) corresponds to a delocalized vibration in which the

amide groups oscillate in-phase across the strand and out-of-phase with their nearest neighbor.<sup>14</sup> The shoulder region near  $1674\text{ cm}^{-1}$  ( $\nu_{\parallel}$ ) is a more complicated set of vibrations that often involve the amide groups oscillating in phase with their nearest neighbor and out of phase across the strand.<sup>14</sup> The peak at  $1613\text{ cm}^{-1}$  ( $\nu_{\text{P}}$ ) is localized on the Val-Pro amide group.<sup>14</sup> It is red shifted from the main amide I band because the proline  $\text{N}_{\alpha}$  forms a tertiary amide, as compared to the other peptide groups that are secondary amides.<sup>34</sup>



**Figure 6.1** (a) The FTIR absorbance spectrum is shown for each PG12 isotopologue. Red and blue highlights show the position of the  $\nu_P$  and  $\nu_{V3}$  peaks respectively. Each spectrum was recorded in a pH 3.8 deuterated acetate buffer. (b) The sequence of PG12 is highlighted with colored circles to show the position of the labelled amide groups. (c) Absorptive 2D IR spectrum is shown for PG12-V3V5 with contours at  $\pm 3, 6, 9, 12, 16, 24, 32, 45, 65,$  and  $85\%$ .

As can be seen in Fig. 6.1, a single  $^{13}\text{C}'$  substitution at the Val5 position (PG12-V5) shifts the  $\nu_P$  peak to  $1561\text{ cm}^{-1}$  ( $42\text{ cm}^{-1}$ ). The same isotopic substitution at the Val3 position (PG12-V3) brings one of the main amide transitions ( $\nu_{V3}$ ) out to  $1612\text{ cm}^{-1}$ . In PG12-V3,  $\nu_{V3}$  lies on top of the  $\nu_P$  peak, but in the dual labelled species there is approximately  $40\text{ cm}^{-1}$  separation between the  $\nu_P$ ,  $\nu_{V3}$ , and  $\nu_{\perp}$  vibrations respectively. These distinct spectral markers allow the transient folding experiment to simultaneously and independently probe the kinetics of the turn and mid-strand region of the peptide. Finally, we note that a comparison of all FTIR spectra indicates that each also contains two additional weak peaks at  $1590\text{ cm}^{-1}$  and  $1610\text{ cm}^{-1}$ . These are assigned to the symmetric and asymmetric CN stretches of the protonated arginine side chain.<sup>35</sup>



**Figure 6.2** FTIR and absorptive 2D IR spectra of PG12-UL (left), PG12-V5 (left-middle), PG12-V3 (right-middle), and PG12-V3V5 (right). Contour levels are the same as in Fig. 6.1.

These amide I spectral features are more clearly observed with 2D IR spectroscopy. We use a femtosecond Fourier transform method to obtain an absorptive spectrum that correlates the frequency of vibrational excitation ( $\omega_1$ ) with a detection frequency ( $\omega_3$ ).<sup>36</sup> The spectrum for PG12-V3V5 is shown in Fig. 6.1, and spectra of the other isotopologues is shown in Fig. 6.2. There are several key pieces of information available in a 2D IR spectrum. First, cross peaks in a 2D IR spectrum directly report on the coupling between amide vibrations, and help to diminish problems with spectral congestion. This has been particularly valuable in detecting and understanding the  $v_{\perp}$  and  $v_{\parallel}$  modes of  $\beta$ -sheet proteins and peptides.<sup>37</sup> The cross peak between the two modes (observed as a ridge along  $\omega_3 = \omega_1$ ) has been observed in a number of proteins<sup>38</sup> and was used to identify cross-strand contacts in a  $\beta$ -hairpin at high temperatures.<sup>39</sup> In addition to these features, the 2D IR spectrum of PG12-V3V5 shows two distinct features on the diagonal corresponding to the V3 and V5 labelled peptide groups in the linear spectrum. The cross peak between  $v_p$  and  $v_{\perp}/v_{\parallel}$  is weak indicating that the label effectively

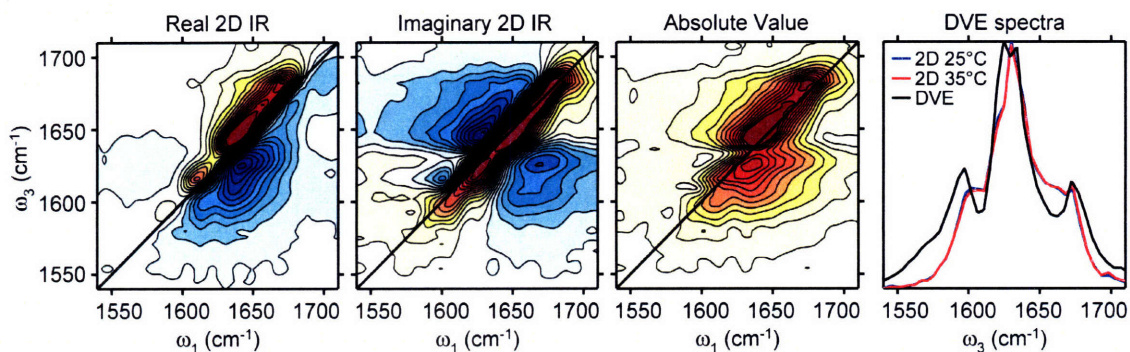
decouples the peptide unit from the other amide vibrations. Cross peaks are visible in the spectral regions connecting the  $\nu_{V3}$  and  $\nu_{\perp}/\nu_{\parallel}$  peaks, indicating coupling between those modes. Additionally, the 2D line shape independently reports on the inhomogeneous (diagonal) and homogeneous (anti-diagonal) contributions to the line width.<sup>36</sup> This information can be used to characterize structural heterogeneity, which is always ambiguous in a one-dimensional absorbance spectrum.

DVE spectroscopy is a one-dimensional variant of 2D IR spectroscopy that is measured by probing the sample with three time-coincident femtosecond infrared pulses and spectrally dispersing the emitted nonlinear signal field.<sup>31,32</sup> For DVE spectroscopy, the pulses are time-coincident, and the detected signal is background free. After being dispersed by a spectrometer, a 64 channel array detects the signal directly, giving full amide I spectral detection at each laser shot. 2D IR spectra are collected by interfering the signal field with a local oscillator field and spectrally dispersing the overlapped beams. The time period between the first and second pulses is scanned and Fourier transformed to give a complex 2D IR surface. The DVE spectrum is related to the 2D IR spectrum through a projection of the complex 2D IR spectrum  $S_{2D} = \text{Re}[S_{2D}] + i \text{Im}[S_{2D}]$  onto the detection axis  $\omega_3$ :

$$S_{DVE}(\omega_3) = \int d\omega_1 |S_{2D}(\omega_1, \omega_3)|^2.$$

Detailed descriptions of these projection relationships are presented in an earlier publication by our group,<sup>32</sup> and can be seen for PG12 in Fig. 6.3.

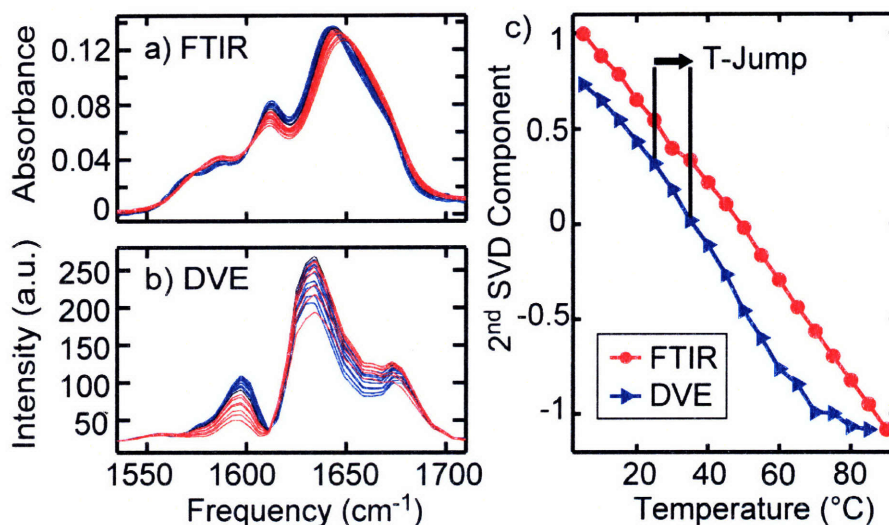




**Figure 6.3** 2DIR and DVE spectra are shown for PG12-UL. The first two panels are the real and imaginary 2D IR spectra for PG12-UL at 25 °C, and the third panel shows the absolute value of the sum of the real and imaginary 2D IR spectra. Panel four plots the directly detected DVE spectrum at 25 °C and the DVE spectra reproduced from the 2D IR spectra at 25 and 35 °C.

DVE spectroscopy has the advantage of being much faster to collect than 2D IR and is background free compared to FTIR. Because the DVE spectrum is equivalent to the projection of the absolute value of the complex 2D IR spectrum onto the  $\omega_3$  axis, as described above, it is possible to use the 2D IR spectrum to interpret the features of a DVE spectrum taken under identical conditions (Fig. 6.3). For example, off-diagonal intensity in the 2D IR spectrum contributes to the diagonal peak intensity to sharpen the features of the DVE spectrum relative to the FTIR spectrum. Interference between the positive and negative peaks of the 2D IR spectrum lead to shifted peak positions in the DVE spectrum relative to the measured frequencies in the 2D IR and FTIR spectra. This is most striking in the  $\nu_P$ ,  $\nu_{V3}$ , and  $\nu_{\perp}$  peaks of PG12-V3V5 (Fig. 6.4) which are red-shifted nearly  $10 \text{ cm}^{-1}$  relative to the FTIR spectrum.

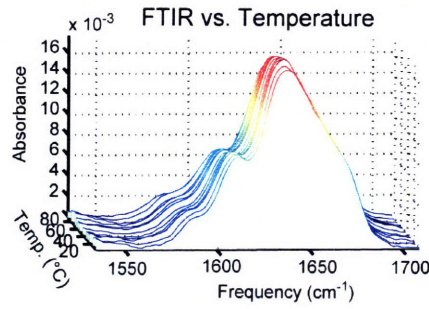
## 6.5. Thermal Denaturation



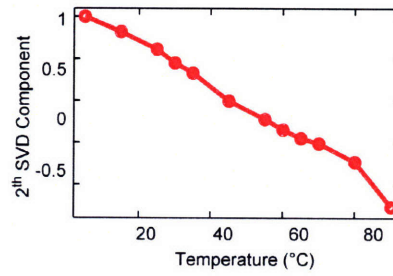
**Figure 6.4** (a) Amide I FTIR spectra of PG12-V3V5 from 5 to 90 °C. The spectra are colored blue from 5 to 40 °C, black at 45 °C, and red from 50 to 95 °C. The frequency axis is the same for plots a and b. (b) DVE spectra of PG12-V3V5 from 5 to 85 °C. The coloring scheme is the same as the FTIR spectra except the red spectra which range from 50 to 85 °C. (c) Melting curves calculated using the FTIR and DVE spectra from 1540 to 1710  $\text{cm}^{-1}$ . The melting curves are plotted as the value of the 2nd component at each temperature. The DVE curve is offset from the FTIR data by -0.25 for visual clarity.

Temperature dependant equilibrium data are shown below for PG12-UL and PG12-V3V5. Each series of spectra are taken under the same experimental conditions as described in the paper: Deuterated acetate buffer, pH 3.85, with a sample cell thickness of 50  $\mu\text{m}$ . The PG12-UL melting curves obtained from singular value decomposition are broad and relatively flat, which is similar to those measured for identical and similar hairpin peptides.<sup>27,34</sup>

**PG12-UL**

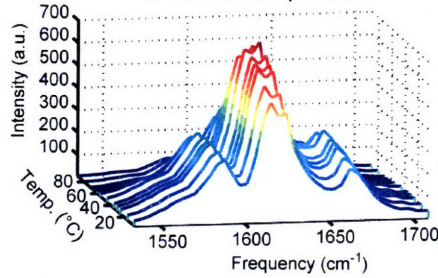


**Melting Curve**

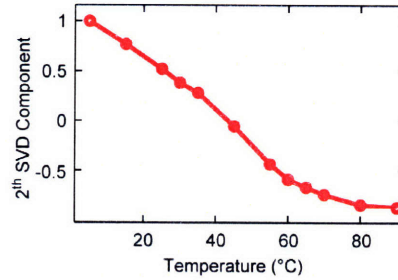


**PG12-UL**

**DVE vs. Temperature**

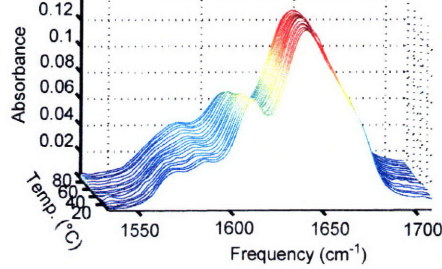


**Melting Curve**

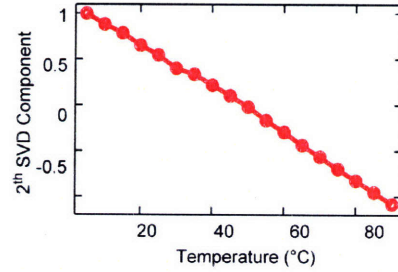


**PG12-V3V5**

**FTIR vs. Temperature**

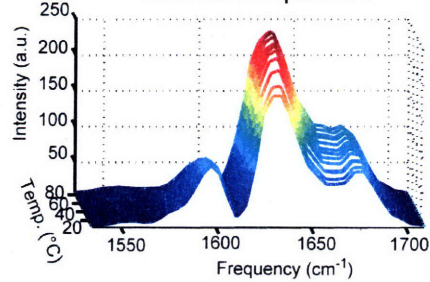


**Melting Curve**

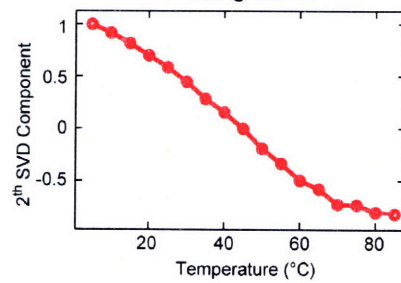


**PG12-V3V5**

**DVE vs. Temperature**

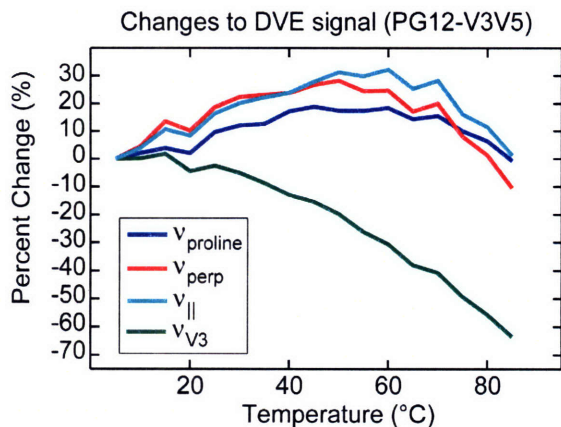


**Melting Curve**

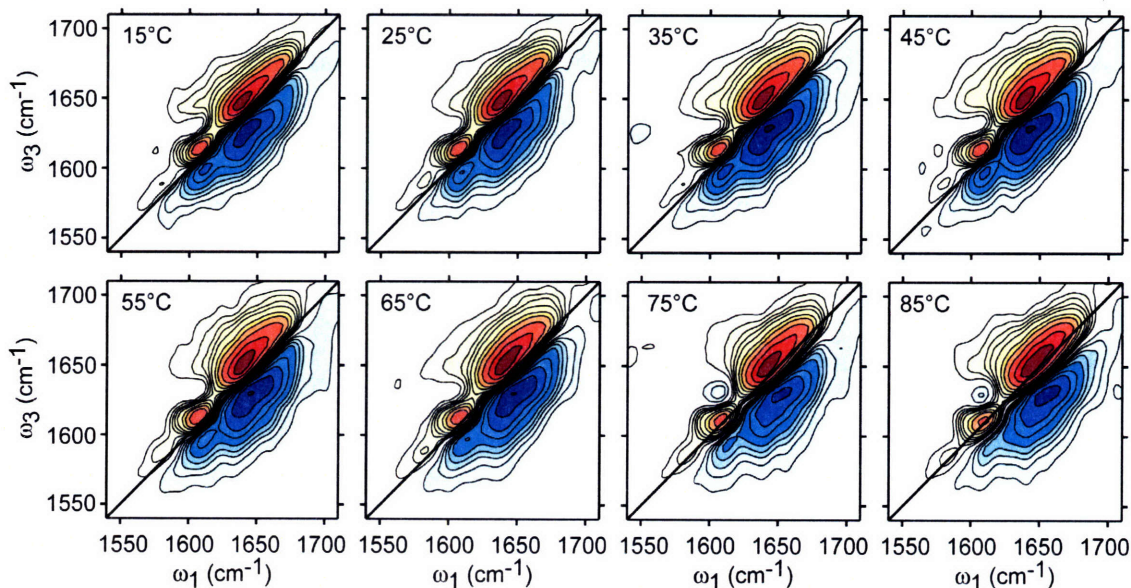


**Figure 6.5** (Left panels) Spectra are shown for PG12 from 5 to 90 °C and PG12-V3V5 from 5 to 85 °C. Spectra are recorded as detailed in the section 6.3. (Right panels) Equilibrium melting curve constructed as the 2<sup>nd</sup> SVD component of the spectra shown to the right.

As shown above (Fig. 6.5), the DVE and FTIR spectra show similar equilibrium melting curves. At the same time, the peak intensities show different trends with temperature. In the PG12-V3V5 FTIR spectra the  $\nu_{\perp}$  and  $\nu_{\parallel}$  bands appear to shift toward band center, and the intensity at frequencies corresponding to disordered peptide increases. The  $\nu_P$  and  $\nu_{V3}$  peaks slowly decrease over the temperature range, while the  $\nu_{Arg}$  peak does not change substantially. There are a number of factors that contribute to the changes in the amide I absorption spectrum including homogeneous and inhomogeneous line widths, vibrational coupling, and vibrational frequency. Peaks in the DVE spectra are influenced by the same factors, but the nonlinear signal displays these differences in a unique way. The first thermal effect seen in the nonlinear spectroscopy is that the solvent transmission increases with temperature. Signal increases from the solvent transmission effects combine with other factors that decrease the signal to create a concave temperature profile at several frequency channels (shown in Fig. 6.6).



**Figure 6.6** Intensity of DVE peaks as a function of temperature. The traces plotted are the intensity of the signal at the following frequencies:  $\nu_p = 1558 \text{ cm}^{-1}$ ,  $\nu_{V3} = 1598 \text{ cm}^{-1}$ ,  $\nu_{\perp} = 1630 \text{ cm}^{-1}$ , and  $\nu_{\parallel} = 1674 \text{ cm}^{-1}$ . Each trace is normalized to the intensity of the peak at 15 °C.

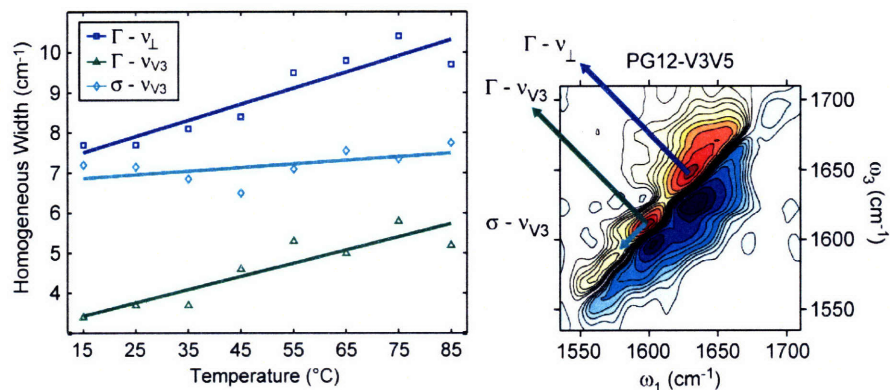


**Figure 6.7** Absorptive (real) 2D IR spectra are shown for PG12-V3V5 from 15 to 85 °C. Spectra shown are collected in the zzzz polarization to make a more direct comparison to the DVE spectra which are also collected in the zzzz polarization configuration. Contours are placed at  $\pm 3, 6, 9, 12, 16, 24, 32, 45, 65,$  and  $85\%$  of the amide I peak maximum.

Signal loss in the DVE spectrum with temperature can best be explained in terms of the corresponding changes to the 2D IR spectra, shown in Fig. 6.7. One cause of loss in

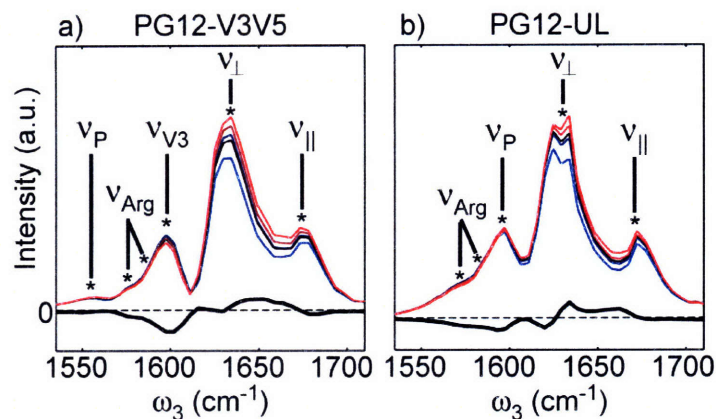
the DVE signal is weaker coupling seen as decreased cross peak intensity in the 2DIR spectrum. A second contribution to lower DVE peaks is a decrease in diagonal peak intensity, which is caused by temperature-dependant changes to the vibrational eigenstates that report on the disordering of the peptide. Finally, any increase in the homogeneous or inhomogeneous line width will lead to a lower DVE signal because anti-diagonal and diagonal line widths project equally onto the  $\omega_3$  axis.

To better understand 2D IR line width contributions, the anti-diagonal line widths from the spectra are plotted in Fig. 6.8. The anti-diagonal width reports on the homogeneous line width, while the diagonal width reports on the inhomogeneous width. Fig. 6.8 illustrates antidiagonal slices taken through the PG12-V3V5 2DIR spectrum at each temperature. The full width at half maximum of the positive peak is plotted as a function of temperature. As can be seen, the homogeneous width increases linearly with temperature with a slope of  $0.019 \text{ cm}^{-1}/^\circ\text{C}$  for the  $\nu_{\text{V}_3}$  peak, and  $0.062 \text{ cm}^{-1}/^\circ\text{C}$  for the  $\nu_{\perp}$  peak. Based on the equilibrium data we do not expect any nonlinear dependence of the DVE signal from changes to the homogeneous line widths.



**Figure 6.8** Thermal change in line broadening for several peaks in the 2D IR spectra of PG12-V3V5 from 15 to 85 °C. Antidiagonal ( $\Gamma$ ) and diagonal ( $\sigma$ ) half width at half maximum of the peaks indicated by the legend are shown with linear fits with the following slopes:  $\Gamma - \nu_{\perp} = 0.040 \text{ cm}^{-1}/^{\circ}\text{C}$ ,  $\Gamma - \nu_{V3} = 0.033 \text{ cm}^{-1}/^{\circ}\text{C}$ , and  $\sigma - \nu_{V3} = 0.009 \text{ cm}^{-1}/^{\circ}\text{C}$ .

The equilibrium DVE spectra of PG12 and PG12-V3V5 are shown in Fig 6.9 for temperatures from 15 to 45 °C. Spectral changes are shown with the 2<sup>nd</sup> component spectrum of a singular value decomposition. As the temperature is raised, the intensity of the  $\nu_{\parallel}$  and  $\nu_{\perp}$  peaks increases. Over this temperature range we assign the change to increasing amide I intensity corresponding to disordered peptide ( $\sim 1650 \text{ cm}^{-1}$ ) and the increased transmission of the excitation and signal pulses due to the temperature dependent solvent absorbance. At 50 °C (Fig. 6.5, 6.6, and 6.9) the  $\nu_{\parallel}$  and  $\nu_{\perp}$  peaks begin to decrease with increasing temperature indicating that the contributions from the increasing disorder and line widths are overcoming the solvent transmittance effects. The  $\nu_p$  peak in both PG12 and PG12-V3V5 exhibits very little change, which we interpret as negligible structural changes in that part of the peptide. The  $\nu_{\text{Arg}}$  and  $\nu_{V3}$  peaks decrease as the temperature is raised, which is assigned to an increase in structural disorder.

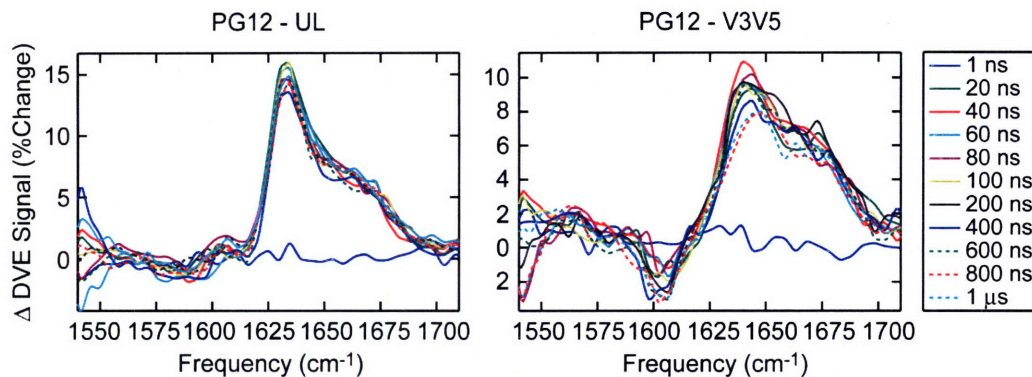


**Figure 6.9** Equilibrium DVE spectra of PG12 (a) and PG12-V3V5 (b) from 15 °C (blue) to 45 °C (red). The solid black plots are the 2<sup>nd</sup> SVD component spectra calculated over the same temperature range. As the temperature is raised, the intensity at  $\nu_{\perp}$  and  $\nu_{\parallel}$  increases, while the  $\nu_p$ ,  $\nu_{SC}$ , and  $\nu_{V3}$  intensity decreases.

## 6.6. Transient Unfolding

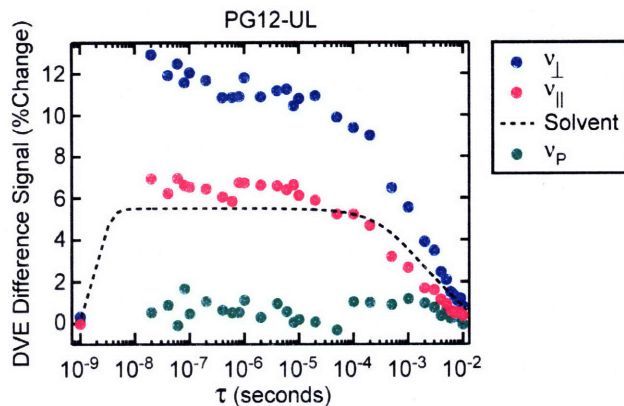
Transient thermal unfolding of PG12 and PG12-V3V5 was probed by recording the entire amide I DVE spectrum at time delays ( $\tau$ ) following a 10 °C T-jump. The starting temperature,  $T_0$ , was 25.0 °C, and the T-jump laser raises the temperature to 35 °C. A 3 % pulse-to-pulse variation in the T-jump laser power leads to an approximate error in the final temperature of  $\pm 0.3$  °C. In Fig. 6.10 the DVE difference spectra are plotted at T-jump delays ranging from 1ns to 1 $\mu$ s.



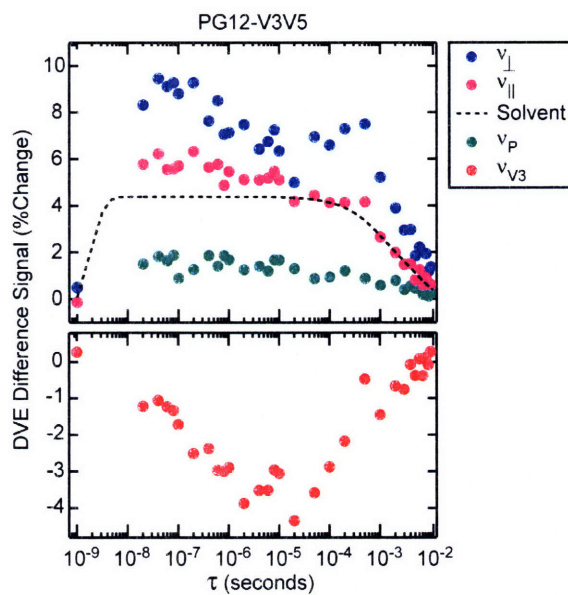


**Figure 6.10** T-jump DVE difference data for PG12-UL (left panel) and PG12-V3V5 (right panel).

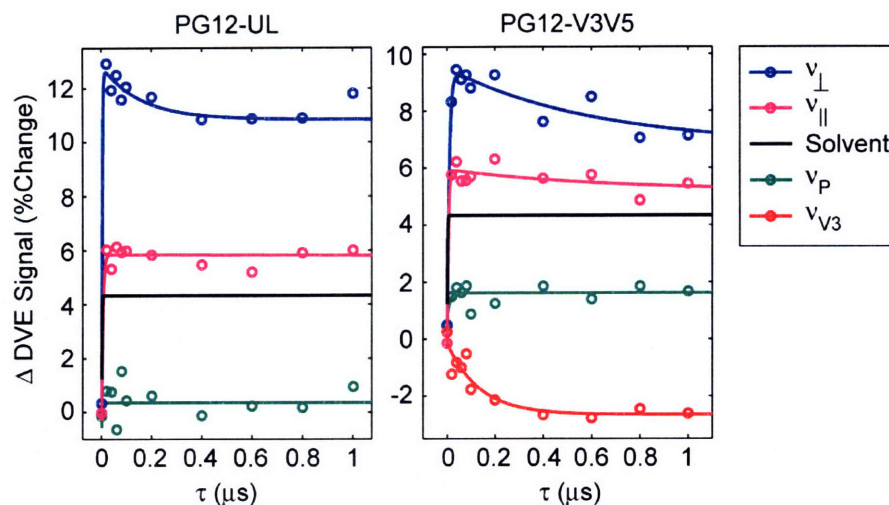
The transient DVE signal is plotted in Figs. 6.11 and 6.12 as a function of T-jump delay time on a semi-log plot for a series of frequency channels. The frequency ranges correspond to the  $\nu_{\perp}$ ,  $\nu_{\parallel}$ ,  $\nu_P$ , and  $\nu_{V3}$  peaks. The signal is averaged over frequency ranges detailed in the figure captions. The transient temperature profile of the sample is plotted as a dashed black line. The signal at each frequency after 100 $\mu$ s was observed to relax with the 2 ms temperature re-equilibration of the sample.



**Figure 6.11** T-jump data for PG12-UL averaged over three frequency ranges and  $\nu_P$  from 1598 to 1607  $\text{cm}^{-1}$ . The fitted transient absorbance is scaled and plotted to compare the DVE data to the instantaneous solvent temperature.



**Figure 6.12** T-jump data for PG12-V3V5 averaged over four frequency ranges:  $\nu_{\perp}$  from 1635 to 1654  $\text{cm}^{-1}$ ,  $\nu_{\parallel}$  from 1663 to 1683  $\text{cm}^{-1}$ ,  $\nu_P$  from 1558 to 1571  $\text{cm}^{-1}$ , and  $\nu_{V3}$  from 1598 to 1607  $\text{cm}^{-1}$ . The fitted transient absorbance is scaled and plotted to compare the DVE data to the instantaneous solvent temperature.



**Figure 6.13** Transient T-jump DVE signal of PG12-UL and PG12-V3V5 at four frequency channels corresponding to the frequency ranges as in Figs. 6.11 and 6.12. Each data point is the average of over the following frequency ranges: 1558-1567  $\text{cm}^{-1}$  (red), 1598-1607  $\text{cm}^{-1}$  (green), 1639-1648  $\text{cm}^{-1}$  (violet) 1668-1678  $\text{cm}^{-1}$  (blue). The fitted transient absorbance of the solvent is scaled by 150 and plotted in black. Open circles correspond to measured data, and the lines are single exponential fits to the data. The  $t_1$  rise time of the fits to the  $\nu_{\perp}$ ,  $\nu_{\parallel}$ , and  $\nu_P$  data is less than 10 ns, and the fitted  $t_1$  decay time of the PG12-V3V5  $\nu_{V3}$  peak is  $135 \pm 50$  ns. The  $\nu_{\perp}$  and  $\nu_{\parallel}$  of PG12-V3V5 each have subsequent decay times of 610 and 680 ns respectively. The  $\nu_{\perp}$  peak of PG12-UL has a long decay time of 360 ns.

Changes to the DVE signal of PG12-UL and PG12-V3V5 from 10 ns to 1  $\mu\text{s}$  are shown in Fig 6.13 for four frequency windows. The transient solvent response is plotted in black as a reference for the temperature profile. The  $\nu_{\perp}$ ,  $\nu_{\parallel}$ , and  $\nu_P$  frequency regions each have rise times  $\leq 10$  ns, which is on the same order as the 8 ns instrument response time. This time scale reflects spectral changes that track the increased solvent temperature. In addition, the  $\nu_{\perp}$  peak displays a small (2 %) decrease that has a fitted decay time of 610 ns. This decay time corresponds generally to thermally induced disorder within the peptide. Changes to the  $\nu_P$  peak are small in magnitude which leads us to conclude that the Val5-Pro6 amide group does not undergo significant structural

changes during the 10 °C T-jump. This conclusion is corroborated by the  $\nu_P$  transient of PG12-UL and the temperature dependant intensity profile of the peak in the equilibrium DVE data plotted in Fig. 6.6.

Changes to the  $\nu_{V3}$  peak in PG12-V3V5 are much slower than the 10ns rise time of the other spectral features. The exponential fit shown in Fig. 6.13 has a time constant of  $135 \pm 50$  ns, and the magnitude and sign of the transient difference signal match the equilibrium spectra. Structurally, the change is indicative of increased disorder at the Val3-Glu4 amide group. Because the amide frequency is closely correlated with hydrogen bonding strength, this time-scale probes disordering of the cross-strand hydrogen bond contacts.

## 6.7. Conclusions

The observed transient DVE signal of the  $\nu_{V3}$  and  $\nu_P$  peaks puts detailed structural constraints on proposed unfolding models for this system. Specifically, the results support an unfolding model for this peptide in which the turn region does not undergo significant structural variation and the mid-strand region is disrupted on a 130 ns time scale. This time scale is about 5 times faster than the slow decay of the  $\nu_{\perp}$  peak (610 ns), and suggests that the mid-strand region denatures before the rest of the peptide. The unfolding time of the mid-strand region is one to two orders of magnitude faster than the observed folding time of many other  $\beta$ -hairpins. However, it is similar to the unfolding times of cyclic hairpins<sup>23</sup> as well as that of a three-stranded  $\beta$ -sheet peptide containing D-Pro enhanced  $\beta$ -turns.<sup>24</sup> It is therefore possible that the T-jump transients describe the

sampling of an increased configurational space about the folded structure at the elevated temperature rather than a full unfolding trajectory.<sup>40</sup> This would also be consistent with the finding that tightly folded hairpins such as PG12 and TrpZip2 have a highly heterogeneous unfolding free energy surface.<sup>10</sup>

The isotope labeling approach used in this study is generally applicable to other peptides and small proteins. By incorporating isotope labels into various structural regions of a given polypeptide it is possible to resolve a residue-level folding mechanism. Nonlinear IR probes like 2D IR and DVE spectroscopy are uniquely sensitive to vibrational coupling and line shapes. The spectral signatures, which report on structural disorder and tertiary contacts such as hydrogen bonding across  $\beta$ -strands, explore essential physical events during thermally induced peptide unfolding. Such studies will significantly accelerate our understanding of peptide and protein folding.

## 6.8. References

- (1) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76-88.
- (2) Searle, M. S.; Ciani, B. *Curr. Opin. Struct. Biol.* **2004**, *14*, 458-464.
- (3) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236-4237.
- (4) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. *Protein Sci.* **2002**, *11*, 1492-1505.
- (5) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* **2001**, *99*, 9081-9081.
- (6) Bolhuis, P. G. *Biophys. J.* **2005**, *88*, 50-61.
- (7) Klimov, D. K.; Thirumalai, D. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 2544-2549.
- (8) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. *J. Mol. Biol.* **2004**, *336*, 241-251.
- (9) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. *Nature* **1997**, *390*, 196-199.
- (10) Yang, W. Y.; Gruebele, M. *J. Am. Chem. Soc.* **2004**, *126*, 7758-7759.
- (11) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 4077-4082.
- (12) Dyer, R. B.; Maness, S. J.; Peterson, E. S.; Franzen, S.; Fesinmeyer, R. M.; Andersen, N. H. *Biochemistry* **2004**, *43*, 11560-11566.
- (13) Hamm, P.; Lim, M.; Hochstrasser, R. M. *J. Phys. Chem. B* **1998**, *102*, 6123-6138.
- (14) Smith, A. W.; Tokmakoff, A. *J. Chem. Phys.* **2007**, *126*, 045109.

- (15) Munoz, V.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5872-5879.
- (16) Dinner, A. R.; Lazaridis, T.; Karplus, M. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9068-9073.
- (17) Tsai, J.; Levitt, M. *Biophysical Chemistry* **2002**, *101*, 187-201.
- (18) Wei, G. H.; Mousseau, N.; Derreumaux, P. *Proteins* **2004**, *56*, 464-474.
- (19) Pitera, J. W.; Haque, I.; Swope, W. C. *J. Chem. Phys.* **2006**, *124*, 141102.
- (20) Munoz, V.; Ghirlando, R.; Blanco, F. J.; Jas, G. S.; Hofrichter, J.; Eaton, W. A. *Biochemistry* **2006**, *45*, 7023-7035.
- (21) Du, D.; Tucker, M. J.; Gai, F. *Biochemistry* **2006**, *45*, 2668-2678.
- (22) Du, D.; Zhu, Y.; Huang, C.-Y.; Gai, F. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 15915-15920.
- (23) Maness, S. J.; Franzen, S.; Gibbs, A. C.; Causgrove, T. P.; Dyer, R. B. *Biophys. J.* **2003**, *84*, 3874-3882.
- (24) Xu, Y.; Purkayastha, P.; Gai, F. *J. Am. Chem. Soc.* **2006**, *128*, 15836-15842.
- (25) Xu, Y.; Oyola, R.; Gai, F. *J. Am. Chem. Soc.* **2003**, *125*, 15388-15394.
- (26) Decatur, S. M. *Acc. Chem. Res.* **2006**, *39*, 169-175.
- (27) Setnicka, V.; Huang, R.; Thomas, C. L.; Etienne, M. A.; Kubelka, J.; Hammer, R. P.; Keiderling, T. A. *J. Am. Chem. Soc.* **2005**, *127*, 4992-4993.
- (28) Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. *J. Phys. Chem. B* **2006**, *110*, 7545-7555.
- (29) Brewer, S. H.; Song, B.; Raleigh, D. P.; Dyer, R. B. *Biochemistry* **2007**, *46*, 3279-3285.

- (30) Werner, J. H.; Dyer, R. B.; Fesinmeyer, R. M.; Andersen, N. H. *J. Phys. Chem. B* **2002**, *106*, 487-494.
- (31) Chung, H. S.; Khalil, M.; Smith, A. W.; Ganim, Z.; Tokmakoff, A. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 612-617.
- (32) Chung, H. S.; Khalil, M.; Tokmakoff, A. *J. Phys. Chem. B* **2004**, *108*, 15332-15343.
- (33) Chung, H. S.; Khalil, M.; Smith, A. W.; Tokmakoff, A. *Rev. Sci. Instrum.* **2007**, 78.
- (34) Hilario, J.; Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* **2003**, *125*, 7562-7574.
- (35) Chirgadze, Y. N.; Fedorov, O. V.; Trushina, N. P. *Biopolymers* **1975**, *14*, 679.
- (36) Khalil, M.; Demirdöven, N.; Tokmakoff, A. *J. Phys. Chem. A* **2003**, *107*, 5258-5279.
- (37) Cheatum, C. M.; Tokmakoff, A.; Knoester, J. *J. Chem. Phys.* **2004**, *120*, 8201-8215.
- (38) Demirdöven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A. *J. Am. Chem. Soc.* **2004**, *126*, 7981-7990.
- (39) Smith, A. W.; Chung, H. S.; Ganim, Z.; Tokmakoff, A. *J. Phys. Chem. B* **2005**, *109*, 17025-17027.
- (40) Zagrovic, B.; Snow, C. D.; Khaliq, S.; Shirts, M. R.; Pande, V. S. *J. Mol. Biol.* **2002**, *323*, 153-164.



# Chapter 7

## Equilibrium Thermal Denaturation of Isotope-Labeled TZ2

### 7.1. Introduction

The experiments used in Chapter 6 to study the transient folding of PG12 provide an approach to measure local structural information with nanosecond temporal resolution. Local structure is accessed with isotope labels that shift the amide I band of specific peptide groups. The amide band is then probed with 2D IR and DVE spectroscopy, which are directly sensitive to vibrational frequencies, couplings and line-broadening mechanisms. In this chapter and in Chapter 8, the same approach will be used to study the thermal unfolding of TZ2. This chapter will focus on equilibrium thermal denaturation, and Chapter 8 will probe the transient thermal unfolding.

The methods in this chapter have been applied to un-labeled TZ2 (TZ2-UL) in Chapter 5. FTIR, 2D IR, and DVE spectroscopy will be used to probe the amide I band of a series of isotopologues designed to probe separate regions of the hairpin structure

(Fig. 7.1). In addition to the dual threonine label introduced in Chapter 4 (TZ2-TT), two other labels will be used to access local structure. The  $\beta$ -turn region will be accessed with TZ2-K8, which was prepared with a  $^{13}\text{C}$  atom in the C' position of the K8 residue (SWTWENG**K**WTWK). The N-terminus is accessed with TZ2-S1, which was prepared with an  $^{18}\text{O}$  label in the carbonyl oxygen of the Ser1 residue (SWTWENGKWTWK).

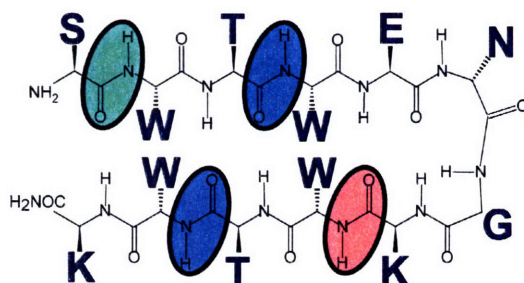
Nonlinear infrared spectroscopy of these three TZ2 isotopologues resolves band structure hidden in the FTIR spectra. The 2D IR cross peaks and line widths show temperature trends that describe local disordering in the thermal denaturation transition. Data presented in this chapter are consistent with the previous observations of heterogeneous folding.<sup>1,2</sup> For example, the amide I spectra of TZ2-K8 resolves two peaks that correspond to unique turn geometries. The thermal changes to these peaks are evidence for at least two unique structures present in the sample over the experimental range of temperatures. Amide I spectra of TZ2-S1 are consistent with an N-terminus that explores a range of configurations. In some of these configurations, the S1 amide group is hydrogen bound to the opposing  $\beta$ -strand, while in others it is solvent exposed. Data from the dual threonine labeled TZ2-TT show that the mid-strand backbone contacts are statically disordered at low temperatures. At high temperatures, the labeled amides partially decouple from the rest of the peptide and explore a larger configurational space. The observations are consistent with a disordering transition described as hairpin fraying, which has been proposed as a mechanism for hairpin unfolding. In Chapter 8, these observations will be combined with temperature-jump experiments to probe the transient kinetics of the TZ2 peptide.

## 7.2. Materials & Methods

### 7.2.1. Peptide samples

TZ2 was prepared using Fmoc-based SPPS as described in Chapter 2, and each of the isotopologues described in Section 2.7.1 will be investigated in this chapter. For FTIR, DVE, and 2D IR spectra, each peptide was dissolved in pure D<sub>2</sub>O, and the pH was adjusted to 2.5 by adding 0.5 - 1.5  $\mu$ L of 25 mM DCl. The low pH was chosen to avoid aggregation during the transient folding experiments. Final peptide concentrations were between 10 and 12 mg/mL. The equilibrium 2D IR spectra of TZ2-UL, TZ2-S1, and TZ2-K8 were taken in a phosphate-buffered pH 7.0 solution at similar concentrations.

#### Ser-Trp-Thr-Trp-Glu-Asn-Gly-Lys-Trp-Thr-Trp-Lys



**Figure 7.1** Stick diagram of TZ2 is pictured with the three-letter amino acid code above. Colored circles represent the position of the three amide group isotope labels: TZ2-S1, with <sup>18</sup>O-labeled Ser1 (green); TZ2-TT, with <sup>13</sup>C-labeled Thr3 and Thr10 (blue); and TZ2-K8, with <sup>13</sup>C-labeled Lys8 (red).

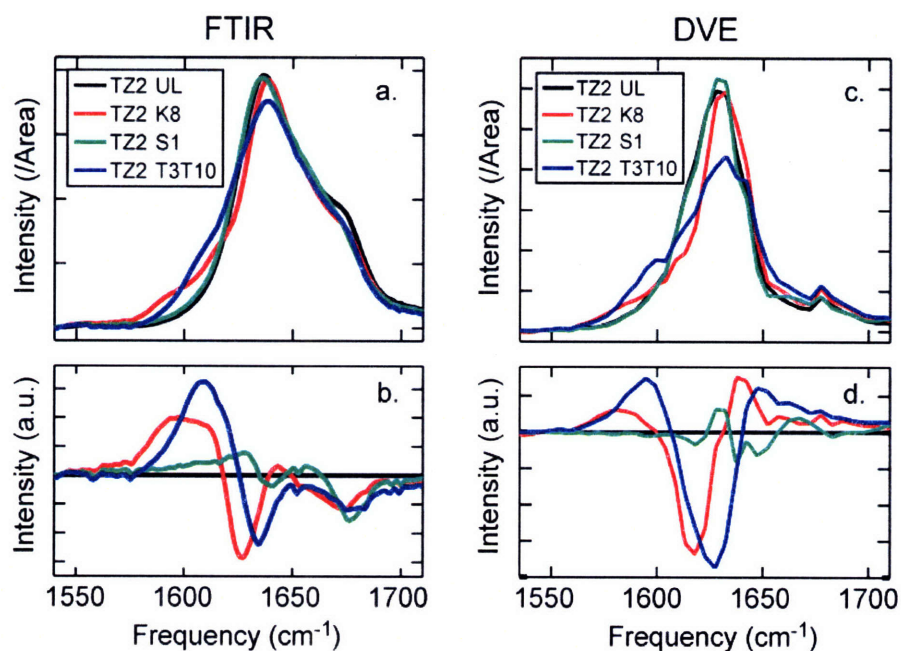
### 7.2.2. Spectroscopic methods

Each peptide sample was probed with FTIR, 2D IR, and DVE spectroscopy over a range of temperatures. For FTIR measurements,  $\sim 25 \mu\text{L}$  of the peptide solution was sandwiched between two 1 mm thick  $\text{CaF}_2$  windows separated by a 50  $\mu\text{m}$  thick Teflon spacer. The temperature of the solution was controlled with the brass sample cell described in Chapter 3. FTIR spectra were collected from 5 to 95 C in 5C increments on a Nicolet 380 FTIR spectrometer (Thermo Electron Corporation) at  $1.0 \text{ cm}^{-1}$  spectral resolution. Each spectrum represents the average of 64 scans. 2D IR spectra were collected as described in Chapter 3, with the polarization geometry all parallel (ZZZZ) unless otherwise noted and  $\tau_2 = 100 \text{ fs}$ . The evolution time,  $\tau_1$ , was scanned to 1.5 ps for the non-rephasing spectrum and 2.0 ps for the rephasing spectrum at 4 fs time steps, and a rectangular window function was used for the Fourier transform. DVE spectra were collected as described previously, with beams 1, 2, and 3 in the parallel polarization geometry, and  $\tau_2 = 100 \text{ fs}$ .

### 7.3. Amide I Spectroscopy of TZ2 at 25 °C

The amide I band is a set of backbone carbonyl stretching modes that can be modelled from a set of interacting local vibrations. Each amide I unit couples to the others via electrostatic coupling (through-space) and kinetic coupling (through-bond) mechanisms. Electrostatic coupling of amide groups across the  $\beta$ -strands leads to eigenstates that are delocalized across the hairpin. The most intense eigenstates contribute to the  $\nu_{\perp}$  band ( $\sim 1635 \text{ cm}^{-1}$ ) introduced in Chapter 4, in which the local amide

groups oscillate in phase across the  $\beta$ -strand and out-of-phase with their nearest neighbors. The symmetry of the  $\nu_{\perp}$  modes leads to a large transition moment that is oriented perpendicular to the  $\beta$ -strands. High frequency modes ( $\sim 1675\text{ cm}^{-1}$ ) are collectively referred to as the  $\nu_{\parallel}$  band, and primarily report on coupling within a single  $\beta$ -strand. Together, the  $\nu_{\perp}$ , and  $\nu_{\parallel}$  bands report directly on the presence of anti-parallel  $\beta$ -sheet secondary structure. At  $\sim 1655\text{ cm}^{-1}$  is the  $\nu_{\text{dis}}$  band that we assign to amide groups in disordered, or random coil-like geometries.



**Figure 7.2** Equilibrium FTIR (a) and DVE (c) spectra are shown for each TZ2 isotopologue at 25 °C and pH 2.5. Each spectrum is baseline corrected with a linear subtraction, and area normalized for comparison. Difference data (b & d) is obtained by subtracting the unlabeled spectrum (TZ2-UL) from each of the spectra.

FTIR and DVE spectra of each TZ2 isotopologue are taken at room temperature and shown in Fig. 7.2. The FTIR peak maximum of the TZ2-UL amide I band is  $1636\text{ cm}^{-1}$ , and there is a prominent shoulder at  $1674\text{ cm}^{-1}$ . The TZ2-UL spectrum is nearly

identical to that at pH 7.0, which was reported in Chapters 4 and 5 and in other previous reports.<sup>2,3</sup> The major difference is that the glutamic acid side-chain carbonyl stretch moves from 1560 cm<sup>-1</sup> (pH 7.0, deprotonated) to ~1710 cm<sup>-1</sup> (pH 2.5, protonated).

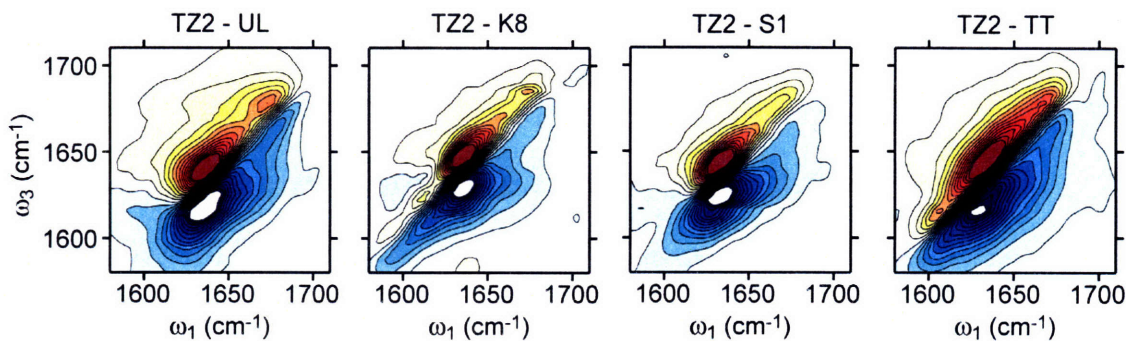
Based on the spectral models in Chapter 4, we interpret the low frequency band as primarily consisting of two nearly degenerate eigenstates with  $\nu_{\perp}$  symmetry. One is localized to the five amide groups in the turn region and the other is localized on the six amide groups near the termini. The high frequency region, or  $\nu_{\parallel}$  band, cannot be assigned to a dominant vibrational mode, but instead consists of several vibrations that have  $\nu_{\parallel}$  symmetry and dipole moments that point parallel to the strand backbone. In the DVE spectrum, the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peak splitting is larger than in FTIR, resulting in better identification and resolution of these  $\beta$ -sheet marker modes.

Isotopic substitution at the K8 amide group leads to a broad increase in intensity from 1585-1615 cm<sup>-1</sup> in the FTIR spectrum, and is hereafter referred to as the  $\nu_{K8}$  band. The band maximum of TZ2-K8 shifts slightly to the blue (1638 cm<sup>-1</sup>), and the width of  $\nu_{\perp}$  peak narrows by 3-4 cm<sup>-1</sup> relative to the TZ2-UL spectrum. Spectral regions that lose intensity compared to the unlabeled spectrum are at the red side of the  $\nu_{\perp}$  band (~1628 cm<sup>-1</sup>) and a broad region from 1650 to 1680 cm<sup>-1</sup>. The width of the distribution of spectral loss at high frequencies reflects the coupling of the unlabeled K8 amide in the TZ2-UL band. The broad (~35 cm<sup>-1</sup>) gain in the TZ2-K8 - TZ2-UL difference spectrum suggests that there is significant structural heterogeneity at the K8 peptide group. The change in the  $\nu_{\perp}$  band suggests that isotopic substitution significantly rearranges the eigenstate distribution. Narrowing of the  $\nu_{\perp}$  band is even more pronounced in the DVE

spectrum of TZ2-K8. The DVE  $\nu_{K8}$  band is less visible, which may be due to interference from the  $\nu_{\perp}$  peak and the limited bandwidth of the excitation laser.

The FTIR spectrum of TZ2-S1 looks qualitatively very similar to the TZ2-UL spectrum. There is an increase to the red of the  $\nu_{\perp}$  band from 1590 to 1625  $\text{cm}^{-1}$ , termed the  $\nu_{S1}$  band. The TZ2-S1  $^{12}\text{C}$   $\nu_{\perp}$  band is red-shifted slightly to 1635  $\text{cm}^{-1}$ , and there is loss from 1670 to 1685  $\text{cm}^{-1}$  and from 1637 to 1650  $\text{cm}^{-1}$  in the difference spectrum. Loss at high frequencies likely corresponds to disorder or a frayed N-terminus, while the mid-frequency loss represents the degree to which the unlabeled S1 amide is coupled to the  $\nu_{\perp}$  band. Changes to the DVE spectrum of TZ2-S1 are nearly imperceptible. There is a slight loss in the  $\nu_{\parallel}$  band, and the  $\nu_{\perp}$  appears to narrow slightly. The  $\nu_{S1}$  band is unresolved in these spectra because the 39  $\text{cm}^{-1}$  frequency shift from the  $^{18}\text{O}$  label moves intensity from the blue side of the amide band to the red side, where it is hidden underneath the  $\nu_{\perp}$  band.

The FTIR spectrum of the dual threonine labeled peptide, TZ2-TT, has been discussed in Chapter 4. Changes to the amide I band are primarily accounted for by a shift of intensity from the  $\nu_{\perp}$  band to the isotope-shifted  $\nu_{TT}$  band. Labeling both of the threonine residues red-shifts the  $\nu_{\perp}$  band that is localized on the mid-strand region of the peptide. In the DVE spectrum of TZ2-TT, the loss of  $\nu_{\perp}$  intensity is even more dramatic than in the FTIR spectrum. The  $\nu_{\perp}$  peak width also increases, with a significant amount of intensity growing in between the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks. This may indicate that the eigenstates of TZ2 are more disordered in the  $^{12}\text{C}$  amide I band of TZ2-TT than in the amide I band of TZ2-UL.



**Figure 7.3** Experimental equilibrium 2D IR spectra for each TZ2 isotopologue at 25°C as described in section 7.2. Each data set shown is taken in the ZZZZ polarization geometry, and contours are plotted at  $\pm 3, 6, 9, 12, 16, 24, 32, 45, 65,$  and 85% of the band maximum.

The TZ2-UL 2D IR spectrum demonstrates the additional information available in 2D IR spectroscopy relative to the one-dimensional spectra shown in Fig. 7.2. By spreading the spectrum out in two dimensions, the diagonal peak separation increases, which helps resolve the  $\nu_{\parallel}$  band as a separate peak rather than a shoulder as in FTIR. The peak splitting between the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks is  $37 \text{ cm}^{-1}$  in the  $\omega_1$  dimension. The 2D widths of the peaks can be used to quantify the homogeneous and inhomogeneous components of the amide I vibrations as shown in Chapter 4. Coupling between the  $\nu_{\perp}$  and  $\nu_{\parallel}$  modes is observed in the positive off-diagonal spectral region as a ridge that extends toward  $\omega_1 = 1636 \text{ cm}^{-1}$  and  $\omega_3 = 1674 \text{ cm}^{-1}$ . In the limit of narrow, non-interfering peaks, the coupling would appear as a cross-peak doublet. In the case of TZ2 peptides, the homogeneous line width of the diagonal peaks interferes with the cross peak doublets giving rise to the ridge structure.

For TZ2-K8, the most striking feature in the 2D IR spectrum is the splitting of the  $\nu_{K8}$  band into two diagonal peaks: one at  $1594 \text{ cm}^{-1}$  ( $\nu_{K8-1}$ ) and the other at  $1612 \text{ cm}^{-1}$  ( $\nu_{K8-2}$ ). Cross peak intensity between the two peaks is undetectable, suggesting that they



represent two structurally distinct sub-ensembles. Coupling to the  $\nu_{\perp}$  band is observed as a ridge extending to  $\omega_1 = 1612 \text{ cm}^{-1}$  and  $\omega_3 = 1650 \text{ cm}^{-1}$ , with most of the intensity above the  $\nu_{K8-2}$  peak. The narrowing of the  $\nu_{\perp}$  band observed in FTIR is also visible in the 2D IR spectrum primarily in the anti-diagonal dimension. The peak splitting between the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks is  $39 \text{ cm}^{-1}$  in the  $\omega_1$  dimension, which is nearly identical to that seen in TZ2-UL.

As expected from the FTIR and DVE spectra, the 2D IR spectrum of TZ2-S1 is very similar to that of TZ2-UL. The major change is in the  $\nu_{\parallel}$  peak, which in TZ2-S1 appears to have decreased significantly relative to  $\nu_{\perp}$ . There is some additional intensity to the red side of the positive  $\nu_{\perp}$  diagonal peak which represents the unresolved  $\nu_{S1}$  band.

In the 2D IR spectrum of TZ2-TT, the  $\nu_{TT}$  band is well-resolved at  $\omega_1 = \omega_3 = 1611 \text{ cm}^{-1}$ , and the splitting between  $\nu_{\perp}$  and  $\nu_{TT}$  in the  $\omega_1$  dimension is  $28 \text{ cm}^{-1}$ . This shift is consistent with the simulations done in Chapter 4. Because the shift is less than the approximately  $40 \text{ cm}^{-1}$  shift expected from  $^{13}\text{C}$  labeling,<sup>4</sup> we conclude that the mid-strand  $\nu_{\perp}$  mode is slightly higher in frequency than  $\nu_{\perp}$  mode at the turn in TZ2-UL. Changes to the  $\nu_{\perp}$  band relative to the other isotopologues are very dramatic, including a diagonal line width increase of approximately  $21 \text{ cm}^{-1}$ . This line width change suggests that the eigenstate distribution of TZ2-TT is much more heterogeneous than in the other peptides, which is consistent with the vibrational modes calculated in Chapter 4.

## 7.4. Simulated Spectra

To better understand the spectral shifts upon isotope labeling, a set of TZ2 conformations was selected to probe the relationship between peptide structure and amide I vibrational frequencies. The peptide structures used in this analysis come from two sources. The first is obtained from the protein data base (PDB ID 1le1; 1<sup>st</sup> Conformer),<sup>5</sup> which was measured using NMR spectroscopy and is hereafter referred to as the native state. The second source of structures comes from a set of Markov states published by Chodera, et al.<sup>6</sup> These Markov states were obtained from a newly developed methodology that automatically decomposes the peptide configurational space explored in an extensive molecular dynamics simulation into a set of kinetically distinct structures. Forty macrostates were identified for TZ2, which are each composed of an ensemble of peptide structures. We present spectral simulations for the first structure in the macrostate ensemble of four Markov states.

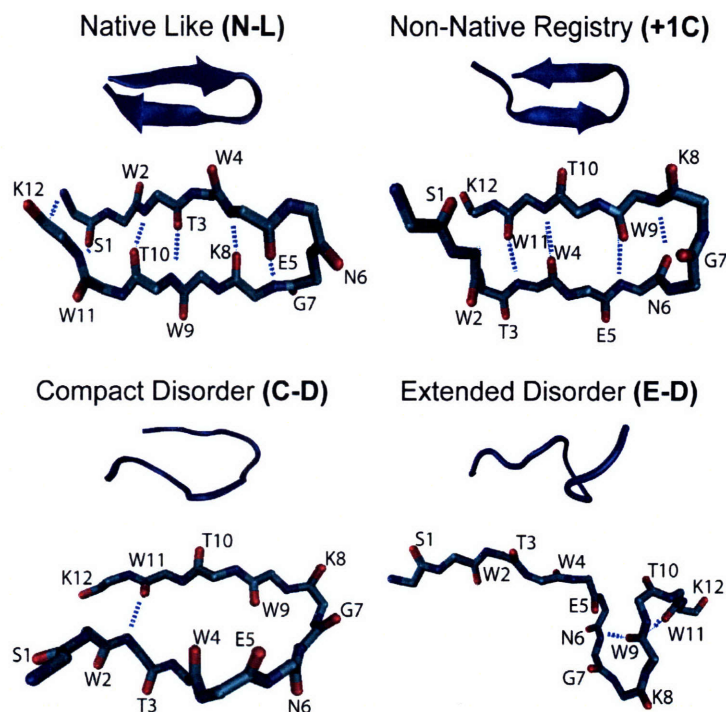
Molecular dynamics simulations were performed with the GROMACS software package using the OPLSAA force field. Simulations were carried out in a 12 Å box of SPC water (spc16). The charge states of the peptide groups were chosen to be consistent with the peptide at the experimental pH = 2.5. The Lys8 and Lys12 side chains and the N-terminus are assigned a charge of +1, and the Glu5 side chain and C-terminus are assigned a charge of 0. The total charge of the peptide is therefore +3, and consequently, three chloride ions were added to the system. The C-terminus of the NMR structure and each of the Markov states are terminated with a carboxylic acid in contrast to the experimental peptides which have C-terminal amides. After a short (200 steps) energy minimization, the peptide bonds were constrained, and the simulation was carried out for

1 ns at 2 fs integration steps. The system configuration was saved every 1 ps for spectral simulations.

For each 1 ps configuration, a Hamiltonian is built with site energies determined from the electrostatic potential on the amide group.<sup>7</sup> Long range couplings are calculated using transition charge coupling,<sup>8</sup> and nearest neighbor couplings are found in an ab initio parameterized look-up table.<sup>9</sup> To simulate the effects of isotope labeling, the Hamiltonian files were adjusted by red-shifting the appropriate amide I site energy. For <sup>13</sup>C labeled amide groups, 42 cm<sup>-1</sup> was subtracted from the appropriate site energy, and for <sup>18</sup>O labels, a 39 cm<sup>-1</sup> red-shift was used. Obtaining the 1D and 2D IR spectra at each time point has been described previously.<sup>10-12</sup> At each time step, the Hamiltonian and amide I dipoles in the site basis are used to produce a set of energies and transition dipoles in the eigenstate basis. For 2D spectra, the anharmonicity of the two-quantum states is 16 cm<sup>-1</sup>. The spectra are produced by convolving the energies and transition dipoles with a Lorentzian function with the width,  $\Gamma = 9 \text{ cm}^{-1}$ . Each of the simulated spectra over the length of the molecular dynamics trajectory is averaged together to produce a spectrum that accounts for disorder of the peptide in a solvated environment.<sup>11</sup>

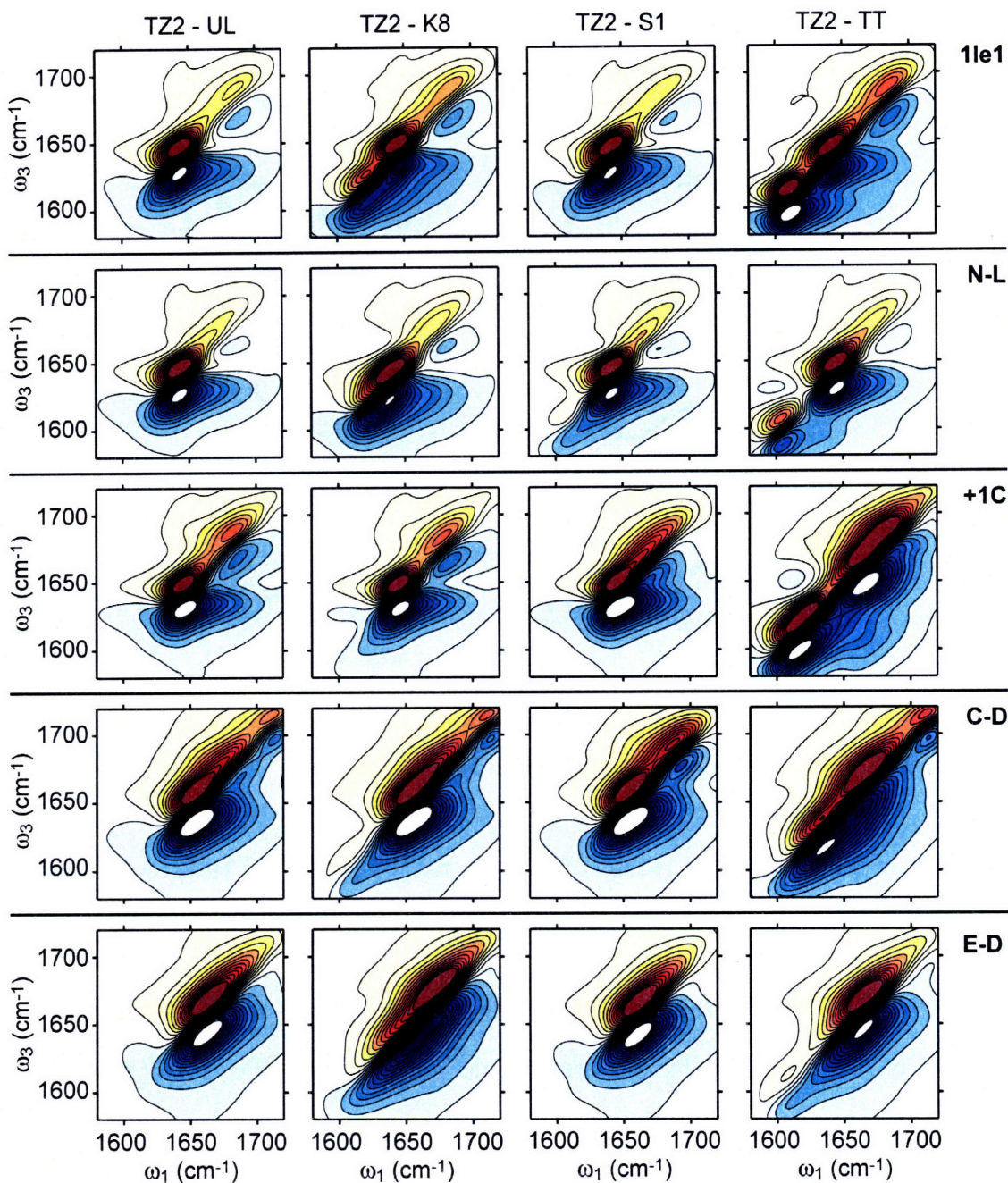
Each configuration chosen for the spectral simulations should be viewed as an ansatz from which structural features can be identified that may play an important role in determining the spectroscopy. The Markov states identified by Chodera et al.<sup>6</sup> are valuable because they represent an exhaustive search and organization of the relevant peptide configuration space. From those states, one configuration is chosen that closely resembles the native state observed in the NMR-derived structure ( $C_{\alpha}$  RMSD = 0.6 Å). It is referred to as the native-like state (N-L, 250851) and is chosen for comparison to the

spectra simulated with the NMR structure. The second structure contains significant  $\beta$ -sheet secondary structure, but has a strand misalignment of +1 residue in the direction of the C-terminus (H-bonds between Trp4 and Trp11, and Asn6 and Trp9). This structure will be referred to as **+1C** (271154). It was chosen to identify spectral markers for misaligned  $\beta$ -strands, which have been implicated as intermediate structures in a reptation-like folding pathway.<sup>13,14</sup> The final two configurations are chosen as possible representations of structure in the denatured peptide ensemble. One of these structures is relatively compact, but with a disordered turn and non-native cross-strand tertiary contacts (**C-D**, 214369). The other structure is much more extended, with a kink in the region near the turn (**E-D**, 64336).



**Figure 7.4** Structures of Markov states that were used for spectral simulations. Structures were provided by William Swope and have been published in Ref. 6.

The simulated 2D IR spectra of TZ2-UL for **N-L**, **+1C**, and **1le1**, each display two diagonal peaks and a Z-shape to the contours. This shape is characteristic of  $\beta$ -sheet secondary structure and is similar to the experimental TZ2-UL 2D IR spectrum. For the disordered states, the **E-D** band is stretched along the diagonal axis. The anomalous peak at  $1710\text{ cm}^{-1}$  is localized on the S1 amide group. For the **C-D** structure, the amide band I is very similar to that of a peptide in random coil geometry. Simulated FTIR spectra of TZ2-UL for each structure are shown in the Appendix (Fig 1.a.1). The simulated spectra of the isotope labeled peptides will be discussed below.



**Figure 7.5** Simulated 2D IR spectra of TZ2 isotopologues for the NMR structure and four Markov states. Methods and notation are described in text. For each spectrum, 27 contours are plotted from  $\pm 75\%$  of the band maximum.

In the experimental spectra (Section 7.3) it was observed that the  $^{13}\text{C}$  isotope label in the TZ2-K8 gives rise to two peaks explicitly resolved in the 2D IR spectrum (Fig.

7.3). The  $\nu_{K8-1}$  peak maximum is located at  $\omega_1 = 1599 \text{ cm}^{-1}$ , and the  $\nu_{K8-2}$  peak maximum is at  $\omega_1 = 1616 \text{ cm}^{-1}$ . The amide I band maximum, assigned as the  $^{12}\text{C } \nu_{\perp}$  band, is located at  $\omega_1 = 1637 \text{ cm}^{-1}$ , leading to a peak splitting of 38 and 21  $\text{cm}^{-1}$  for  $\nu_{K8-1}$  and  $\nu_{K8-2}$ , respectively. The intensity of  $\nu_{K8-1}$  is 58% of the  $\nu_{K8-2}$  peak, which in turn is 15% of the amide I maximum intensity. The observation that the two peaks are uncoupled in the 2D IR spectrum suggests that they are reporting on two distinct structural ensembles.

For each of the five simulations, there is a strong correlation between the orientation of the K8 amide group and the frequency splitting of the isotope-shifted peak. The NMR structure and N-L state each have a K8 amide group oriented with the carbonyl oxygen pointing toward the opposite  $\beta$ -strand, which hydrogen bonds to the W4 amide proton. For both of these structures the dihedral angles around the K8 alpha carbon are  $\Phi = -140^\circ$ ,  $\Psi = -170^\circ$ , and the  $\nu_{K8}$  peak is split from the amide I maximum by 20 to 25  $\text{cm}^{-1}$ . In each of the other structures (+1C, C-D, and E-D) the carbonyl oxygen points toward the solvent and has  $\Phi$ ,  $\Psi$  values on the order of  $-70^\circ$  and  $-40^\circ$ . The  $\nu_{K8} - \nu_{\perp}$  amide I peak splitting for each of the corresponding spectra is greater than 40  $\text{cm}^{-1}$ . Based on these observations, we conclude that the  $\nu_{K8-1}$  peak represents peptides with non-native turns, and the high frequency,  $\nu_{K8-2}$  peak reports on structures with native, type I'  $\beta$ -turns.

The experimental FTIR, 2D IR, and DVE spectra do not resolve a peak corresponding to the  $^{18}\text{O}$  label in the serine group of TZ2-S1. Observable spectral changes involve a loss of intensity at high frequencies ( $\sim 1675 \text{ cm}^{-1}$ ), and a small ( $\sim 2 \text{ cm}^{-1}$ ) red-shift of the  $\nu_{\perp}$  band. The  $\nu_{\parallel}$  band of the TZ2-S1 2D IR spectrum, which has a resolvable maximum in the TZ2-UL spectrum, now appears as a ridge to the blue of the

$\nu_{\perp}$  band in the diagonal dimension. These spectral changes are assigned to a range of structures explored by the N-terminal serine. First, changes to the red side of the  $\nu_{\perp}$  band are assigned largely to serine groups that are coupled to the amide groups at the C-terminus of the hairpin. This is seen in the **1le1** and **N-L** structures, for which the S1 carboxyl oxygen is hydrogen bound to the W11 amide proton. Strong coupling across the  $\beta$ -strand leads to a low frequency  $\nu_{\perp}$ -like mode delocalized primarily on amide groups S1 and W11, which shifts upon  $^{18}\text{O}$  isotope-labeling. Second, high frequency intensity loss in TZ2-S1 amide I spectra is assigned to disordered serine amide groups. These disordered amides are predicted to have large site energies relative to the rest of the peptide, and are weakly coupled to the other amide groups. Markov states that display this combination of high site energies and weak coupling are **+1C**, **C-D**, and **E-D**. In each of these states, the serine residue is essentially in a random coil geometry, and is too far from the rest of the peptide to experience strong electrostatic coupling. The S1 vibrational frequency for each system moves from very high energies ( $\geq 1700 \text{ cm}^{-1}$ ) in the unlabeled spectrum to frequencies just to the blue of the  $\nu_{\perp}$  band for the  $^{18}\text{O}$  label. The simulated TZ2-S1 FTIR spectra for **1le1** and **E-D** are shown in the appendix (Fig. 7.a.2) to illustrate the patterns in amide I spectral loss and gain upon S1  $^{18}\text{O}$  labeling

As observed in Fig 7.3 and in Chapter 4, the amide I band of dual  $^{13}\text{C}$  labeled TZ2-TT, contains intensity that is red-shifted from the amide I maximum ( $\nu_{\text{TT}}$ ). It is observed as a shoulder in the FTIR spectrum and a resolvable peak in the 2D IR spectrum. The splitting of the 2D IR  $\nu_{\perp}$  and  $\nu_{\text{TT}}$  positive diagonal peaks is  $26 \text{ cm}^{-1}$  in  $\omega_1$ , and the  $\nu_{\text{TT}}$  intensity is 40% of the  $^{12}\text{C}$   $\nu_{\perp}$  intensity. In the native configuration of TZ2, the threonine carboxyl group on each  $\beta$ -strand is hydrogen bonded to the threonine amide proton on the



opposite  $\beta$ -strand. From the NMR structure, the dihedral angles of each threonine are  $\Phi = -126^\circ$ ,  $\Psi = 147^\circ$ , and their relative orientation leads to an electrostatic coupling of  $7.0 \text{ cm}^{-1}$ . The  $\nu_{\text{TT}} - \nu_{\perp}$  splitting in the simulated 2D IR spectrum of the NMR structure is  $30 \text{ cm}^{-1}$ , which agrees with the experimental peak splitting. The  $\nu_{\text{TT}}$  intensity, however, is much too high relative to  $\nu_{\perp}$ . For Markov state N-L, the dihedral angles of the two threonine residues are  $\Phi = -134^\circ$ ,  $\Psi = 135^\circ$  (T3) and  $\Phi = -134^\circ$ ,  $\Psi = 119^\circ$  (T10), and the calculated coupling between them is  $6.5 \text{ cm}^{-1}$ . In the simulated 2D IR spectrum the  $\nu_{\text{TT}} - \nu_{\perp}$  peaks are split by  $40 \text{ cm}^{-1}$ , but the relative intensity is much closer to the experimental value. Because the structural similarity between **1le1** and N-L is high ( $0.6 \text{ \AA}$  backbone atom RMSD) such large spectral differences are surprising. In each of the other Markov states the  $\Phi$ ,  $\Psi$  angles are similarly dispersed, and the  $\nu_{\text{TT}} - \nu_{\perp}$  peak splitting ranges from  $30$  to  $50 \text{ cm}^{-1}$ . This large variation in  $\nu_{\text{TT}}$  peak position and intensity suggests that the experimentally observed  $\nu_{\text{TT}}$  band is significantly broadened by structural heterogeneity. Fast structural fluctuations are also expected to play a significant role in the homogeneous width because a conformational transition from the **1le1** structure and the N-L structure could easily occur on a picosecond time scale.

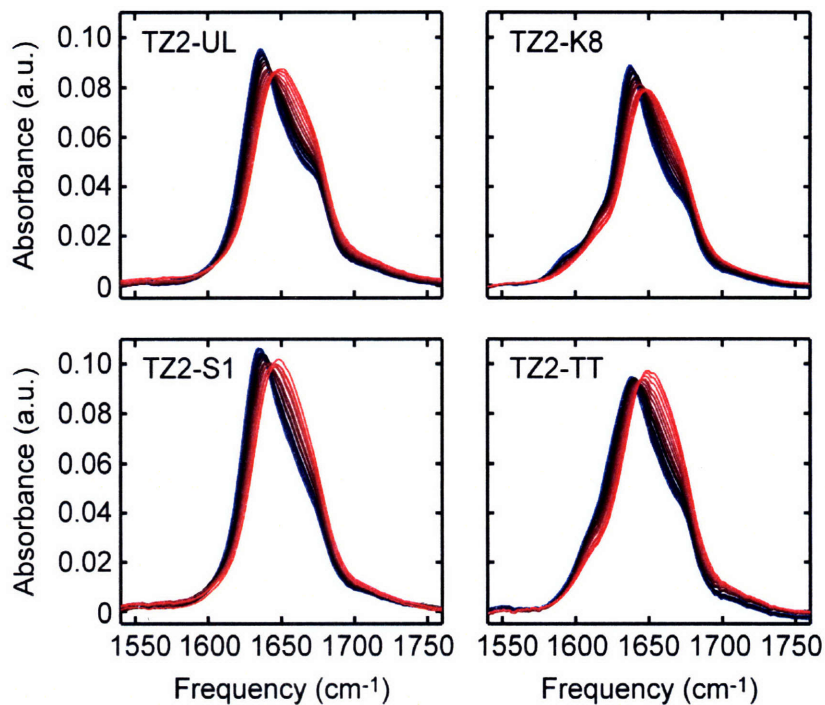
In the +1C TZ2-TT spectrum, the  $\nu_{\text{TT}}$  peak is very intense relative to the  $\nu_{\perp}$  peak, and the  $\nu_{\text{TT}} - \nu_{\perp}$  peak splitting is  $50 \text{ cm}^{-1}$ . The reason for the large deviation is that the T3 and T10 amides are more strongly coupled to  $^{12}\text{C}$  amides. This leads to an isotope shifted peak that contains 40% of the total spectral intensity rather than the 17% expected for 2 out of 12 labeled amides. Because these large shifts and intensities are not observed in the experimental spectra, we conclude that out-of-registry structure is not significantly populated in the equilibrium conditions observed here.

## 7.5. Thermal Denaturation

Effects of thermal denaturation on the infrared spectra of TZ2-UL have been described in Chapter 5. The major result of that work was to show that the intensity of the two anti-parallel  $\beta$ -sheet modes decreases with increasing temperature, and that the splitting between them decreases. These spectral changes, however, do not lead to a completely random coil-like spectrum at the highest accessible temperatures. The  $\nu_{\perp}$  and  $\nu_{\parallel}$  modes at high temperatures indicate that significant cross-strand native contacts are present in the thermally disordered state. This section revisits those observations and describes the nature of the thermal melting transition using the infrared spectra of each isotopologue presented above.

### 7.5.1. FTIR

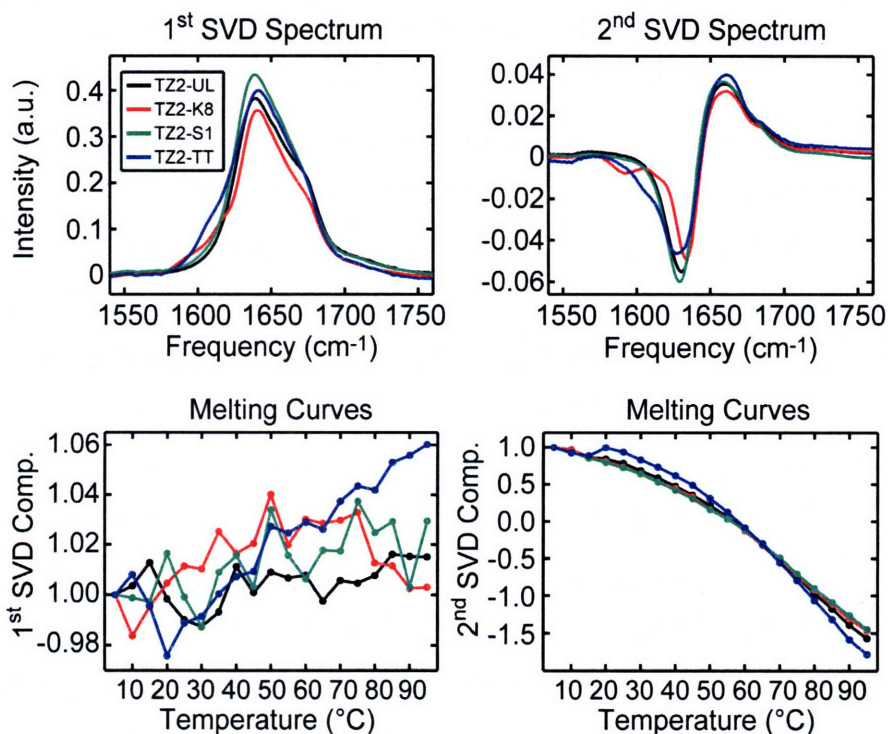
Thermal denaturation of TZ2-UL at pH 7.0 using FTIR spectroscopy has been reported in a number of previous publications, including Chapter 5 of this work. Here, the pH of the solution is 2.5. The amide I spectrum of the low pH solution is nearly identical to that at high pH, and changes to the  $\nu_{\perp}$  and  $\nu_{\parallel}$  bands with temperature are the same for both pH solutions. For example, the 2<sup>nd</sup> SVD component melting curve of the pH 2.5 solution is identical to the pH 7.0 solution presented in Chapter 5. It displays a broad melting transition with  $T_m > 65^{\circ}\text{C}$  (Fig. 7.7).



**Figure 7.6** Temperature dependent FTIR spectra of each TZ2 isotopologue taken from 5 to 95 °C in 5 °C increments (pH = 2.5). Spectrum line color shows progression from 5 °C (blue) to 95 °C (red).

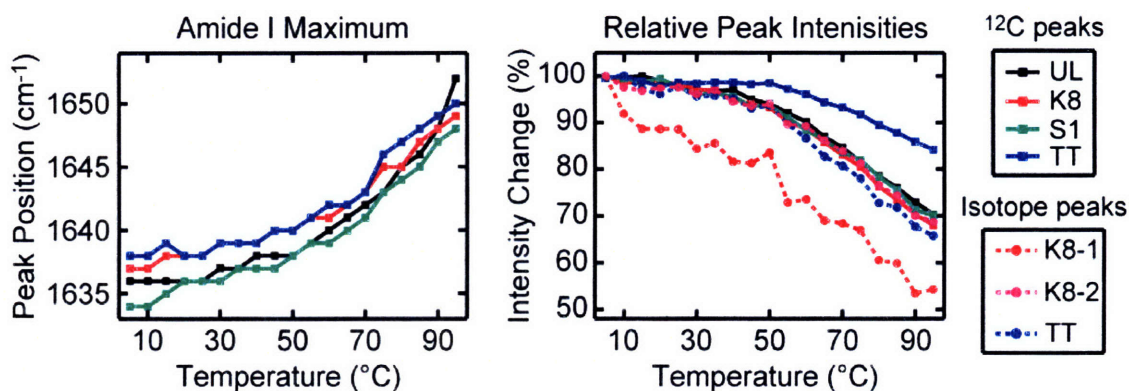
Changes to the FTIR amide I band of the isotope-labeled hairpins are shown in Figs. 7.6 and 7.7. To characterize the melting behaviour, singular value decomposition (SVD) is carried out for the temperature-dependent FTIR data. The SVD analysis produces a series of basis states, or component spectra, that describe the spectral changes. For each TZ2 system, the FTIR data is well described with the first two spectral components as indicated by the relative magnitude of the component weighting factors - the value of the 3<sup>rd</sup> component is an order of magnitude less than the 2<sup>nd</sup> component. In a limiting two-state case, the 1<sup>st</sup> component spectrum represents the average of the temperature dependent spectra, and the 2<sup>nd</sup> component spectrum represents the changes to the spectra with temperature.

The 1<sup>st</sup> and 2<sup>nd</sup> component spectra are shown in Fig. 7.7.a and 7.7.b. The values of these components as a function of temperature are shown in Figs. 7.7.c and 7.7.d, and are normalized to the value at 5°C. The component values are the weight of each component spectrum needed to reconstruct the measured data at each temperature. The value of the 1<sup>st</sup> component of each system varies less than 5% over the temperature range, indicating that the total absorbance of the amide I band is not sensitive to thermal variations. The 2<sup>nd</sup> component values are commonly used to describe the thermodynamics of the melting transition. Here the 2<sup>nd</sup> component melting curves of each isotopologue are observed to be very similar, with the transition midpoints all greater than 65°C. The curvature of each system is slight, but matches the melting curves obtained with previous IR, Fluorescence, and UV-CD measurements.<sup>2,3,5</sup>



**Figure 7.7** SVD analysis of the temperature-dependent FTIR spectra shown in Fig. 7.6. The SVD component spectra are not scaled, and the melting curves are normalized to the 5 °C value.

In addition to SVD characterization of the temperature-dependent amide I' spectrum, it is also useful to observe specific spectral regions within the band. First, the  $^{12}\text{C } \nu_{\perp}$  band is observed during the melting transition, with the frequency of the amide I peak maximum plotted as a function of temperature for each TZ2 system (Fig. 7.8). In each case, the two  $^{12}\text{C } \beta$ -sheet modes show similar melting behaviour. The peak position stays relatively constant until  $\sim 35^{\circ}\text{C}$ , after which there is a steady blue-shift of the band up to  $90^{\circ}\text{C}$ . The total frequency shift,  $\Delta\omega = \omega_{\text{max}}(95^{\circ}\text{C}) - \omega_{\text{max}}(5^{\circ}\text{C})$ , is  $16 \text{ cm}^{-1}$  for TZ2-UL,  $14 \text{ cm}^{-1}$  for TZ2-S1, and  $12 \text{ cm}^{-1}$  for TZ2-K8 and TZ2-TT.



**Figure 7.8** (Left) The frequency of the FTIR amide I maximum is plotted for each TZ2 system. The resolution in the vertical frequency dimension is  $1 \text{ cm}^{-1}$ , which is defined by the FTIR instrument resolution during data collection. (Right) TZ2 FTIR peak intensities are plotted vs. temperature. The  $^{12}\text{C}$  peak intensities are obtained by averaging the absorbance around the peak positions found in the left panel ( $\nu \pm \sim 5 \text{ cm}^{-1}$ ). The  $\nu_{\text{K8-1}}$  trace corresponds to frequencies from  $1588$  to  $1598 \text{ cm}^{-1}$ , the  $\nu_{\text{K8-2}}$  trace is averaged from  $1612$  to  $1619 \text{ cm}^{-1}$ , and the  $\nu_{\text{TT}}$  trace corresponds to frequencies from  $1603$ - $1611 \text{ cm}^{-1}$ . Each trace is normalized to the peak intensity at  $5^\circ\text{C}$ .

Another metric of the thermally-induced FTIR spectral changes is the intensity of specific bands with temperature. For comparison, the intensities of the  $^{12}\text{C}$ ,  $\nu_{\perp}$  peaks and the isotope labeled peaks are all plotted in Fig. 7.8. Intensities are obtained by averaging the absorbance around the peak positions ( $\nu \pm \sim 5 \text{ cm}^{-1}$ ), and normalizing by the value at  $5^\circ\text{C}$ . The TZ2-UL, TZ2-K8, and TZ2-S1  $^{12}\text{C}$   $\nu_{\perp}$  peak intensities have profiles that are consistent with the 2<sup>nd</sup> SVD melting curves seen in Fig. 7.7. In the case of TZ2-K8, the two  $\nu_{\text{K8}}$  peaks show distinct melting behaviour. The  $\nu_{\text{K8-1}}$  peak undergoes the most drastic change with temperature, decaying linearly with increasing temperature. In contrast, the higher frequency  $\nu_{\text{K8-2}}$  peak decays identically with the TZ2-K8  $^{12}\text{C}$   $\nu_{\perp}$  peak. For TZ2-TT, the  $^{12}\text{C}$   $\nu_{\perp}$  peak decay is much slower with temperature, and the  $^{13}\text{C}$   $\nu_{\perp}$  peak decays similar to, or slightly faster than, the TZ2-UL peak.

## 7.5.2. 2D IR

Changes to the 2D IR spectra of TZ2-UL have been described in Chapter 5, and peak shifts and intensity changes are consistent with those described in the FTIR section. Here we focus on two unique observables of the 2D IR spectrum: cross peaks and the 2D line widths. Cross peaks between the  $\nu_{\perp}$  and  $\nu_{\parallel}$  bands of TZ2-UL are observed as a ridge extending towards  $\omega_1 = 1636 \text{ cm}^{-1}$  and  $\omega_3 = 1675 \text{ cm}^{-1}$  (Fig. 7.9). The strength of the cross-peak ridge and the splitting between the two modes decreases with increasing temperature, which is a signature of disruption of  $\beta$ -sheet secondary structure.

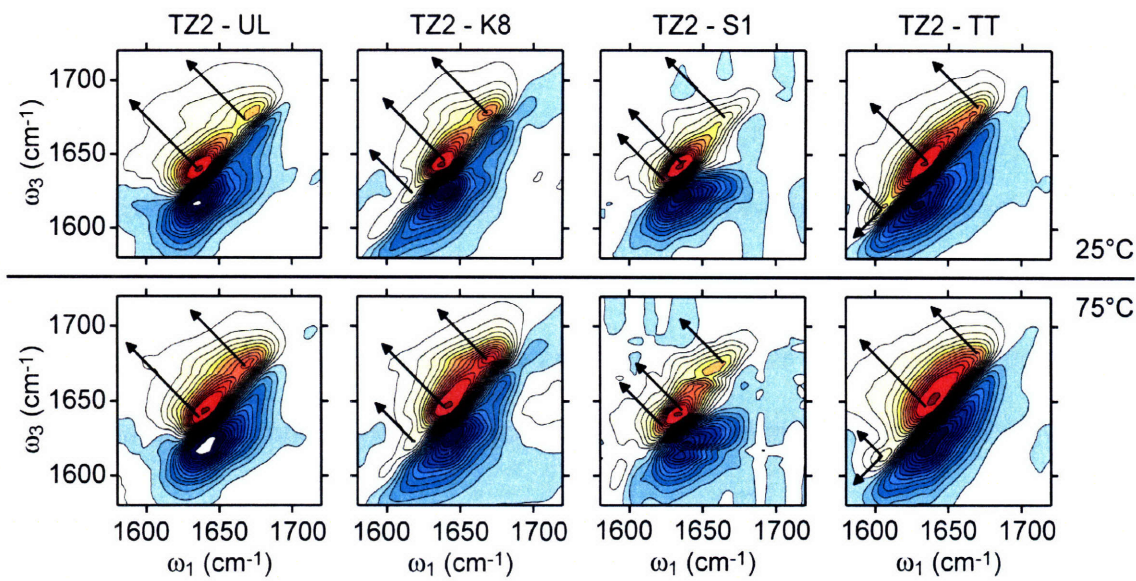
In the temperature-dependent 2D IR spectra of TZ2-K8 (Fig. 7.9) the  $^{12}\text{C}$   $\nu_{\perp}$  and  $\nu_{\parallel}$  bands both shift towards band center and the cross peak between them tracks the changes. At high temperatures the cross peak ridge is still visible, indicating that the  $^{12}\text{C}$   $\nu_{\perp}$  eigenstates are still present in the ensemble. The ridge extending from the  $^{12}\text{C}$   $\nu_{\perp}$  band to the  $^{13}\text{C}$   $\nu_{\text{K8}}$  band also persists although this is difficult to quantify because of the interference from the wings of the positive diagonal  $^{12}\text{C}$   $\nu_{\perp}$  peak. The relative intensity of the two  $\nu_{\text{K8}}$  peaks does change with temperature: The  $\nu_{\text{K8-2}}$  peak gains intensity relative to the  $\nu_{\text{K8-1}}$  peak. This is consistent with the FTIR peak intensities.

Temperature-induced changes to the 2D IR spectra of TZ2-S1 (Fig. 7.9) are very similar to those of the unlabeled peptide. The most striking difference induced by the isotopic substitution is in the shape and temperature profile of the  $\nu_{\parallel}$  band. The intensity of the positive diagonal peak at  $\omega_1=1675 \text{ cm}^{-1}$  is less than that of TZ2-UL, which also affects the cross-peak ridge along  $\omega_3=1675 \text{ cm}^{-1}$ . While difficult to resolve, it appears that this ridge remains at high temperatures like in the other systems. The  $\nu_{\text{S1}}$  band is not

directly observable in the TZ2-S1 2D IR spectrum, and the  $^{12}\text{C}$   $\nu_{\perp}$ ,  $\nu_{\text{S1}}$  cross peak ( $\omega_1=1620\text{ cm}^{-1}$ ,  $\omega_3=1640\text{ cm}^{-1}$ ) is not prominent enough to quantify the thermally-induced intensity changes.

Thermal changes the TZ2-TT 2D IR spectra deviate significantly from the other TZ2 isotopologues. From Fig. 7.9, it is clear that the cross-peak ridge is difficult to resolve at low temperatures, and effectively disappears at higher temperatures. The  $\nu_{\parallel}$  band, which is also poorly resolved at low temperatures, disappears at high temperatures. Both of these features, of course, are strongly affected by the rise of the  $\nu_{\text{dis}}$  band. However, given that the peptide structural ensembles are the same for each system, it seems clear that the amide I eigenstate distribution is significantly disrupted for the dual threonine isotopic substitution. The  $^{12}\text{C}$   $\nu_{\perp}$  red-shift, on the other hand, does match the analogous changes in the other hairpin systems.



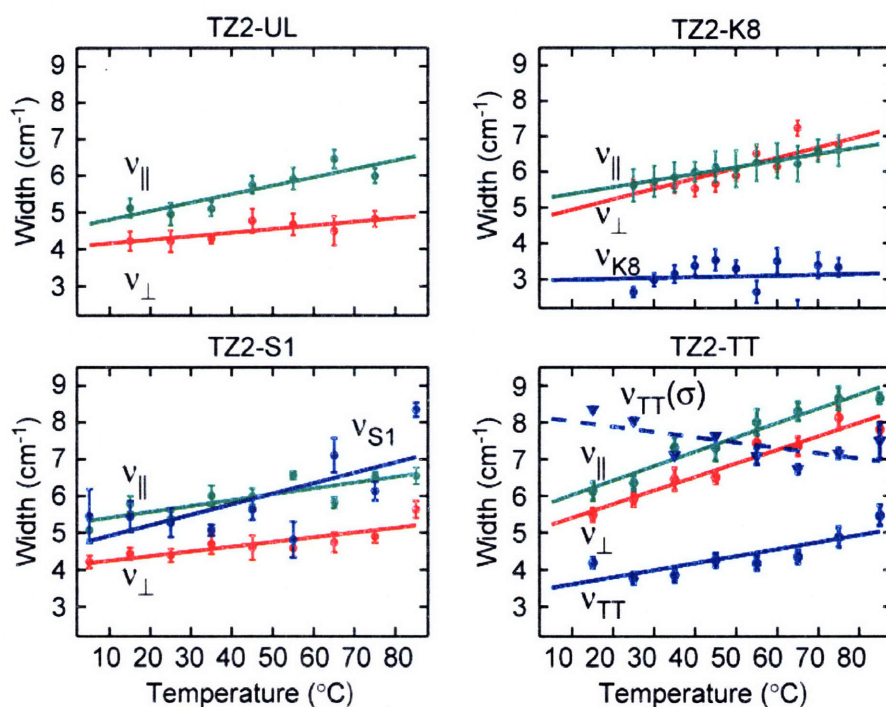


**Figure 7.9** 2D IR spectra of each TZ2 isotopologue are shown at 15 °C (top row) and 75 °C (bottom row). Black lines show the anti-diagonal slices analyzed in Fig. 7.10.

The 2D IR amide I line widths report on disorder within the vibrational modes, and are observed to broaden as the temperature is raised. Generally, diagonal line broadening arises from inhomogeneous mechanisms such as static structural disorder. For TZ2, this increase is difficult to quantify because the  $\nu_{\perp}$ ,  $\nu_{\parallel}$  splitting is on the same order as the line width. Line-broadening in that dimension is further complicated by the frequency shifts of the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks. Changes to the homogeneous width, in contrast, can be quantified by measuring the half-width at half maximum (HWHM) of the positive diagonal peaks in the anti-diagonal dimension. Thermally-induced changes to the anti-diagonal line-widths are assigned primarily to increasing interactions between the solvent and the peptide, but also include fast ( $\sim 1$ -2 ps) structural fluctuations that increase with temperature. In practice, the anti-diagonal slices are fit to the Lorentzian function.

$$\frac{2A}{\pi} \frac{\Gamma}{4(\omega - \omega_0)^2 + \Gamma^2} \quad (7.1)$$

In the function above,  $A$  represents the area under the curve,  $\omega_0$  is the center frequency, and  $\Gamma$  is the full-width at half maximum. In Fig. 7.10, the HWHM ( $\Gamma/2$ ) of the fitted function is plotted versus temperature for each of the TZ2 diagonal peaks. Each data point represents the average width of six diagonal slices around the arrows shown in Fig. 7.9. Error bars are obtained from the standard deviation of the averaged widths.



**Figure 7.10** For each TZ2 isotopologue, the width of the anti-diagonal slice through the positive diagonal peaks is plotted as a function of temperature. For TZ2-TT (bottom right), the diagonal width of the  $\nu_{TT}$  peak is also shown. Each data point represents the  $\Gamma/2$  value of a Lorentzian function (Equation 7.1) fit to the (anti-) diagonal slices. Solid lines represent linear fits to the HWHM vs. temperature, with the slopes reported in Table 7.1.

For TZ2-UL, the HWHM grows with temperature for both of the  $\beta$ -sheet modes in the 2D IR spectrum. The width of the  $\nu_{||}$  band is  $\sim 1.0 \text{ cm}^{-1}$  larger than that of the  $\nu_{\perp}$  band

at 15 °C, and the slope of the line width increase versus temperature of  $\nu_{\parallel}$  is twice that of the  $\nu_{\perp}$  band (Table 7.1). Line-width changes to the TZ2-K8 spectra are consistent with those of the unlabeled hairpin (Fig. 7.10). The slope of the line fit to the HWHM of the  $^{12}\text{C}$   $\nu_{\perp}$  band ( $0.029\pm 0.006\text{ cm}^{-1}/^{\circ}\text{C}$ ) is twice as large as the corresponding slope of TZ2-UL ( $0.010\pm 0.003\text{ cm}^{-1}/^{\circ}\text{C}$ ). For the  $\nu_{\text{K8-1}}$  and  $\nu_{\text{K8-2}}$  peaks, the HWHM measurement is noisier due to lower signal strength. However, the line width of the high frequency peak,  $\nu_{\text{K8-2}}$ , does not display a significant slope with temperature, indicating that the homogeneous broadening does not increase significantly with temperature. The limited resolution of the  $\nu_{\text{K8}}$  peaks prevents a quantitative description of the diagonal line-width. However, in the spectra shown in Fig. 7.14 there does not seem to be any clear sign of peak broadening with temperature. Together, these observations support the idea that two bands are reporting on sub-ensembles of peptides, for which the population may change with temperature, but the linewidth properties remain relatively constant. For TZ2-S1, the line width changes with temperature match those of TZ2-UL (Fig. 7.10). The slope of the  $\nu_{\parallel}$  HWHM versus temperature is lower than that of TZ2-UL, and is nearly identical to the TZ2-S1  $\nu_{\perp}$  HWHM slope. A frequency channel to the red of the  $\nu_{\perp}$  band is chosen to measure the linewidth of the  $\nu_{\text{S1}}$  band. However, the noise in the width of the diagonal peak in that region is large, and does not yield a strong correlation.

Changes to the anti-diagonal line widths of the  $^{12}\text{C}$   $\nu_{\perp}$  and  $\nu_{\parallel}$  bands are much more pronounced for TZ2-TT (Fig. 7.10). The slope of the  $\nu_{\perp}$  HWHM vs. temperature ( $0.037\pm 0.004\text{ cm}^{-1}/^{\circ}\text{C}$ ) is more than double that of the TZ2-UL and TZ2-S1. The  $\nu_{\perp}$  and  $\nu_{\parallel}$  widths are nearly identical, as well as their increase with temperature. The shape of the  $\nu_{\text{TT}}$  peak changes drastically with temperature, going from a diagonally elongated

peak to a symmetrically distributed peak as the temperature is raised. From 15 to 75 °C, the anti-diagonal  $\nu_{\text{TT}}$  HWHM increases more than  $1.0 \text{ cm}^{-1}$ , and the  $\nu_{\text{TT}}$  diagonal width decreases by a similar amount over the same temperature range.

These data show that in the region of T3 and T10, a significant aspect of the thermally-induced disorder is increasing structural fluctuations that produce frequency shifts correlated on a picosecond time-scale. A limited range of configurations is available to the peptide on this time scale, which precludes structural events like the breaking of intra-peptide hydrogen bonds or the insertion of solvent molecules between the  $\beta$ -strands. One possible description of the fluctuations giving rise to the band changes is small fluctuations around the native state. For example, the spectral simulations shown Section 7.4 show that small structural changes can have a large effect on the frequency and intensity of the  $\nu_{\text{TT}}$  peak (e.g. compare 1le1 to N-L in Fig. 7.5). If the T3 and T10 amides do not retain their native geometry, then the temperature-dependent homogeneous broadening could be assigned to increasing solvent-peptide interactions as  $\beta$ -strand separation grows.

System	Fitted Slope (cm <sup>-1</sup> /°C)					
	$\nu_{\perp}$	$\nu_{\parallel}$	$\nu_{K8}$	$\nu_{S1}$	$\nu_{TT} (\Gamma)$	$\nu_{TT} (\sigma)$
TZ2-UL	0.010±0.003	0.023±0.006				
TZ2-K8	0.029±0.006	0.019±0.002	0.002±0.009			
TZ2-S1	0.013±0.002	0.016±0.004		0.030±0.010		
TZ2-TT	0.037±0.004	0.039±0.004			0.019±0.005	-0.014±0.007

**Table 7.1** The rate of line broadening with temperature is shown for diagonal peaks in the TZ2 2D IR spectra. Values shown represent the slope of the line fit to the 2D IR peak width versus temperature shown in Fig. 7.10.

### 7.5.3. DVE

The temperature-dependent DVE spectra shown in this section are consistent with the observations from the FTIR and 2D IR spectra. Amide I peak shifts and intensity changes are well described with linear absorption spectroscopy, and higher order observables such as vibrational anharmonicity, anharmonic coupling, and line-broadening sources are well described with 2D IR spectroscopy. The aim of this section is to describe how the thermally-induced spectral changes in the FTIR and 2D IR data map onto DVE spectra. This will aid our interpretation of the T-jump DVE data, which will be used to describe the folding kinetics of TZ2.

Each set of temperature-dependent DVE data is shown in Fig 7.11 from 10 to 85°C. One feature that varies greatly from system to system is the temperature-dependent intensity of the DVE spectrum. This feature is captured explicitly in the melting curves of the 1<sup>st</sup> component spectrum shown in Fig. 7.12, which essentially describes the intensity of the average spectrum over the temperature range. A fraction of this intensity change can be attributed to laser intensity noise, which fluctuates during the collection of

the data set. However, for each set of spectra below, the intensity changes were found to be largely reversible. This surprising variation between TZ2 isotopes highlights the complex relationship between the homodyne third order signal (DVE), and the heterodyned 2D IR signal which retains the line-shape and phase information necessary to interpret the spectroscopy.

There are three conceptual classes of effects that determine the intensity profile of the temperature dependent DVE spectrum. The first class is analogous to intensity changes observed in a linear, absorptive spectroscopy (i.e. FTIR). These spectral trends are directly related to thermally-induced structural changes and transition frequencies shifting from folded peptide vibrations to unfolded peptide vibrations. Examples of this are peak shifts and amplitude changes like the loss/shift of the  $\nu_{\perp}$  peak and the growth of the  $\nu_{\text{dis}}$  band. These effects drive the observed shifting of the DVE  $\nu_{\perp}$  and  $\nu_{\parallel}$  bands towards the band center at high temperatures.

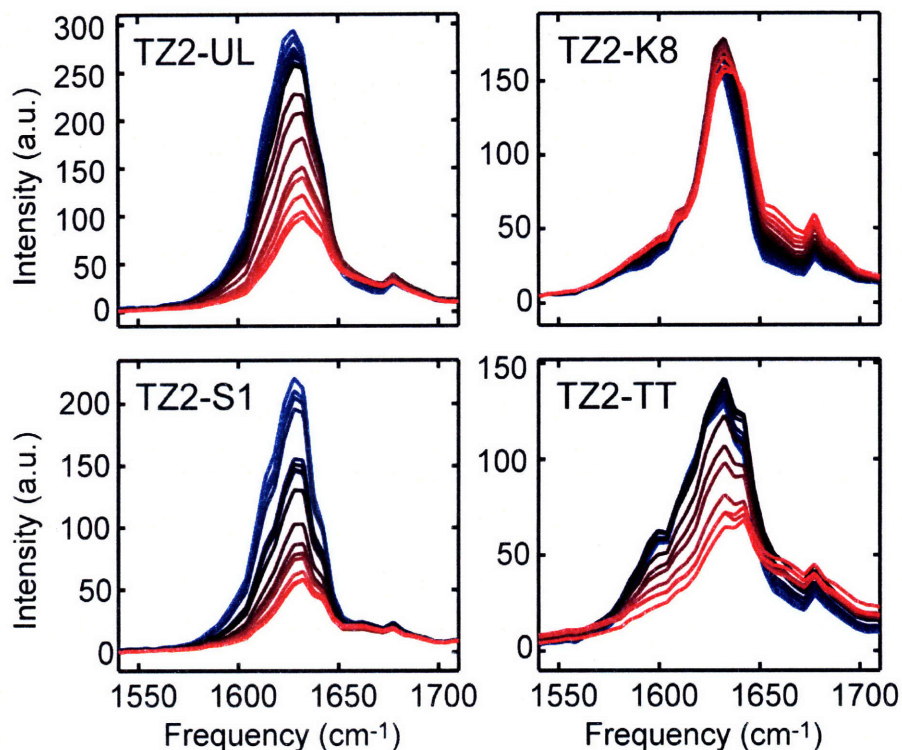
The second class of changes results from the additional observables that are directly related to the 2D IR spectrum through the projection relationship defined in Chapter 6. The DVE signal can be reproduced from the complex 2D IR spectrum,  $S_{2D} = \text{Re}[S_{2D}] + i \text{Im}[S_{2D}]$  by a projection onto the detection axis  $\omega_3$ :

$$S_{DVE}(\omega_3) = \int d\omega_1 |S_{2D}(\omega_1, \omega_3)|^2 \quad (7.2)$$

One major contribution to the DVE signal from this projection relationship is the cross peak intensity, which is best understood as the projection of the cross-peak ridges onto the intensity of the diagonal peaks. This is a major reason for the increased resolution of the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peaks in the DVE spectrum relative to the FTIR spectrum.

Loss of the cross peak intensity is expected as the peptide denatures, which leads to loss of intensity in the DVE peaks. Another major effect in this class of changes is from the projection of the 2D line shapes onto the detection ( $\omega_3$ ) axis. For a congested amide I band, any increase in 2D line width (heterogeneous or homogeneous) will lead to decreasing DVE intensity.

The final class of changes is solely due to the temperature dependence of the buffer solution. For  $D_2O$  there is a well-known transmission increase with temperature. Because DVE is a background-free measurement that is non-linearly proportional to the excitation field intensity, this effect will have a significant (positive) effect on the observed spectral signal. One way to conceptualize this is to imagine a system that does not undergo any structural changes as the temperature is raised. In this case there will be a steady increase in signal resulting from the higher number of excitation photons interacting with the sample.



**Figure 7.11** Equilibrium thermal denaturation DVE spectra for each TZ2 isotopologue from 10 (blue) to 85 °C (red) in 5 °C increments.

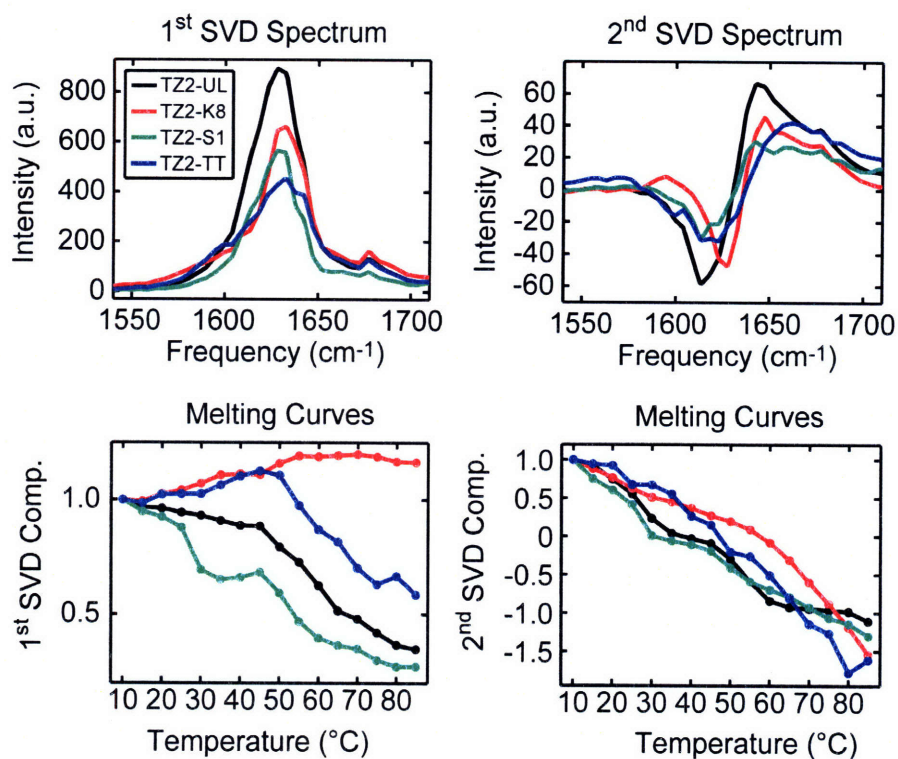
Each of the effects discussed above play a role in the temperature-dependent DVE spectra shown in Fig. 7.11. For TZ2-UL, the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peak splitting decreases as the temperature is raised and the intensity of the  $\nu_{\perp}$  peak drops significantly. The intensity drop is assigned primarily to thermally-induced line broadening, and the peak shifts are assigned to the same mechanism ascribed to the FTIR and 2D IR peak shifts. The TZ2-S1 data display very similar spectral changes upon thermal denaturation.

Thermal changes to the TZ2-K8 spectra are distinct in that the intensity of the spectrum stays relatively flat throughout the thermal melt, and actually displays a steady increase. The spectral region around the two  $\nu_{K8}$  peaks grows steadily with temperature. While it is difficult from the DVE spectrum to resolve the low frequency peak from the



high frequency peak, FTIR and 2D IR spectra show that the low frequency peak decreases while the high frequency peak increases. It is also known that the 2D line width of the band stays constant (Fig. 7.10) and that the population giving rise to the band does not change. Therefore, the primary cause of the increasing DVE signal in this spectral region is the contribution of the temperature dependent solvent transmission, which in the absence of line width and total population changes, will increase homodyne signal with temperature.

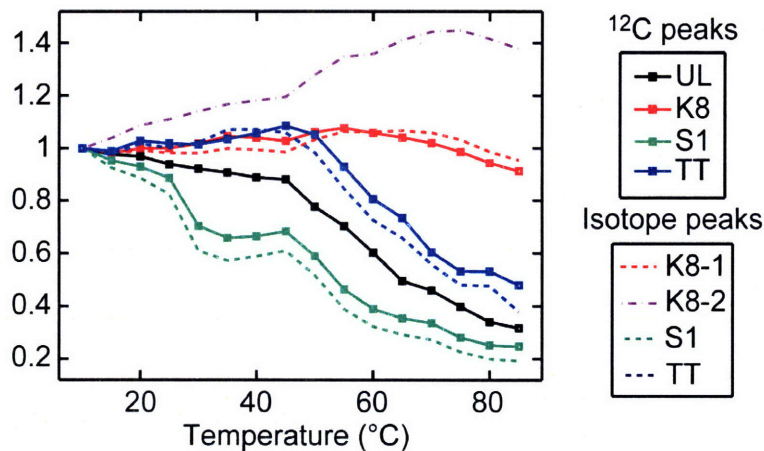
In the case of TZ2-TT, the thermally-induced spectral changes follow the same trends as TZ2-S1 and TZ2-UL. However, there is a small signal increase that is seen in the 1<sup>st</sup> component melting curve from 10 to 50°C. The reason for this increase is not precisely known, but may result from the delayed melting in specific regions of the peptide. For example, it is known from FTIR (Figs. 7.8) that the loss of the  $\nu_{\text{TT}}$  peak is faster than that of the  $^{12}\text{C } \nu_{\perp}$  peak. The DVE spectra show this as a rise in temperature induced by the increased solvent transmission until after 50 °C, at which point the structural effects dominate as the peptide unfolds.



**Figure 7.12** SVD analysis of the temperature-dependent DVE spectra shown in Fig. 7.11. The SVD component spectra are not scaled, while the melting curves are normalized to the 10 °C value.

The SVD analysis shown in Fig. 7.12 is valuable for sorting out two types of changes: intensity shifts, which are captured by the 1<sup>st</sup> component spectrum, and band profile changes, which are captured by the 2<sup>nd</sup> component spectrum. Each second component spectrum shows the  $^{12}\text{C } \nu_{\perp}$  loss, and the rise of the  $\nu_{\text{dis}}$  band. The melting profiles are also similar, which is expected from the similar plots of the FTIR melting curves (Fig. 7.7). Looking at the relative intensity of individual bands is not as informative as for FTIR. The  $\nu_{\text{S1}}$ , high-frequency  $\nu_{\text{K8}}$ , and  $\nu_{\text{TT}}$  band intensities follow their respective  $^{12}\text{C } \nu_{\perp}$  intensities almost exactly (Fig. 7.13). The only deviation is the

high-frequency  $\nu_{K8}$  band that shows a slightly faster increase with temperature compared to the TZ2-K8  $^{12}\text{C}$   $\nu_{\perp}$  band.



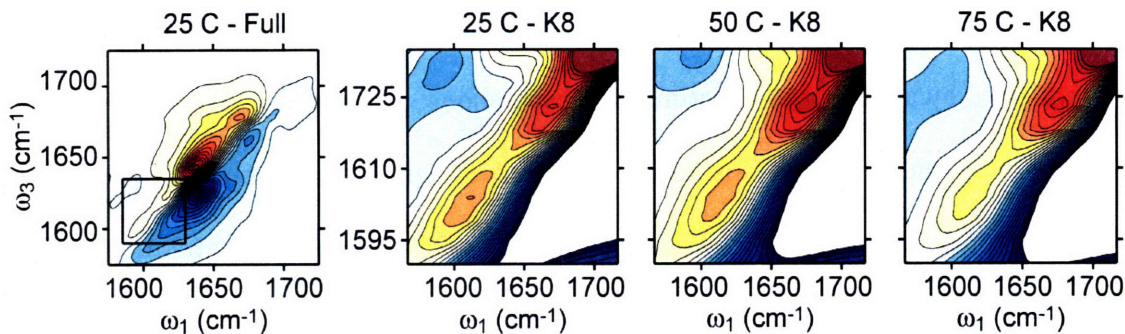
**Figure 7.13** DVE peak intensities are plotted versus temperature. The data points are obtained by averaging the intensity around each peak position ( $\nu \pm \sim 5 \text{ cm}^{-1}$ ). The  $\nu_{K8-1}$  trace corresponds to the low frequency peak ( $1588\text{-}1598 \text{ cm}^{-1}$ ) and the  $\nu_{K8-2}$  peak is the high frequency  $^{13}\text{C}$ -shifted peak ( $1612\text{-}1619 \text{ cm}^{-1}$ ). The  $\nu_{TT}$  trace is obtained by averaging the frequency range from  $1603\text{-}1611 \text{ cm}^{-1}$  in the TZT-TT DVE spectra. Each trace is normalized to the peak intensity at  $5^\circ\text{C}$ .

## 7.6. Thermal Response of Isotope-shifted Bands.

### 7.6.1. TZ2-K8

Thermally-induced spectral changes to the  $\nu_{K8}$  peaks can be seen more clearly in Fig. 7.14. As the temperature of the system increases, the spectrum responds both to structural rearrangements and shifting populations between native and non-native ensembles. For TZ2-K8, the peak positions of  $\nu_{K8-1}$  and  $\nu_{K8-2}$  do not change significantly throughout the temperature titration (Fig. 7.14), and the width of the peaks also remains constant. The only trend correlated with temperature is the relative intensity of the peaks.

As seen in the FTIR data (Figs 7.6 & 7.8), the  $\nu_{K8-1}$  peak intensity drops linearly with temperature, while the  $\nu_{K8-2}$  peak intensity follows the same trend as the  $^{12}\text{C } \nu_{\perp}$  peak. The 2D IR peak intensities are consistent with the FTIR data, with  $\nu_{K8-2}$  decaying at the same rate as  $^{12}\text{C } \nu_{\perp}$ , and the  $\nu_{K8-1}$  peak decaying linearly with temperature relative to  $\nu_{K8-2}$ . The spectral simulations above assigned the  $\nu_{K8-2}$  peak to K8 amide groups in a native, type I' turn and the  $\nu_{K8-1}$  peak to K8 amide oxygen exposed to solvent. Based on these assignments, we conclude that the  $\nu_{K8-2}$  peak intensity decay is similar to that of the rest of the peptide because it is reporting on peptide unfolding from the native to a denatured state. The  $\nu_{K8-2}$  FTIR intensity profile as well as the  $^{12}\text{C } \nu_{\perp}$  intensity and 2<sup>nd</sup> SVD component melting curve are not rigorously sigmoidal, but their shape is consistent with an activated barrier crossing with a melting temperature,  $T_m$ , of 65 to 75 °C. The intensity profile of the  $\nu_{K8-1}$  peak, on the other hand decays linearly with temperature, which is not consistent with a two-state folding model.



**Figure 7.14** Zoom in of temperature-dependent 2D IR spectra of TZ2-K8 shown for the region highlighted in the left plot. Contours lines in the right three spectra are plotted from 20% to -20% of the amide I maximum at the respective temperature.

With the assignments above, one might expect that if the two peaks report on native ( $\nu_{K8-2}$ ), and non-native ( $\nu_{K8-1}$ ) ensembles then as the peptide denatures, intensity will flow from the high frequency peak to the low frequency peak. However, the data show an opposite trend. Specifically, the  $\nu_{K8-1}$  peak actually decays relative to the  $\nu_{K8-2}$  peak with increasing temperature. To rationalize this we propose that the structures giving rise to the two peaks do not interconvert significantly at the temperature and solvent conditions in these experiments. This would mean that the  $\beta$ -turn of TZ2 ( $\nu_{K8-2}$ ) does not transition from the native state to a random coil, nor are there any other large changes in its dihedral angles. Instead, it is decoupling from the  $^{12}\text{C}$  amide groups due perhaps to larger fluctuations and a slightly larger distance from the other amide groups. The presence of the  $\nu_{K8-2}$  peak at high temperatures and the absence of a high temperature baseline indicate that there is a significant population of peptides with native-like type I' turns. An intact turn at high temperatures is consistent with an unfolding transition that preferentially disrupts the amide groups near the N and C termini. This fraying of the

terminal regions of the peptide has been observed in previous simulations, and has been implicated by a number of experimental observations.

In recent work, a  $^{13}\text{C}$  isotope was introduced into the G7 amide group of TZ2.<sup>15</sup> The isotope-shifted peak was well-resolved in the 2D IR spectrum at room temperature, and the spectral simulations were consistent with a solvent exposed amide group. In addition, long waiting-time 2D IR data ( $\tau_2 = 1$  ps) of this peptide showed evidence for a number of solute solvent substates.<sup>16</sup> These observations are consistent with the assignments of the  $\nu_{\text{K8}}$  peaks made above. While only one diagonal peak is observed for the G7-labeled peptide, the G7 amide group in each of the Markov states above shows similar solvent exposure. It is never observed to form a strong intra-peptide hydrogen bond like the natively structured K8 amide, and consequently does not show a large variation in diagonal peak frequency.

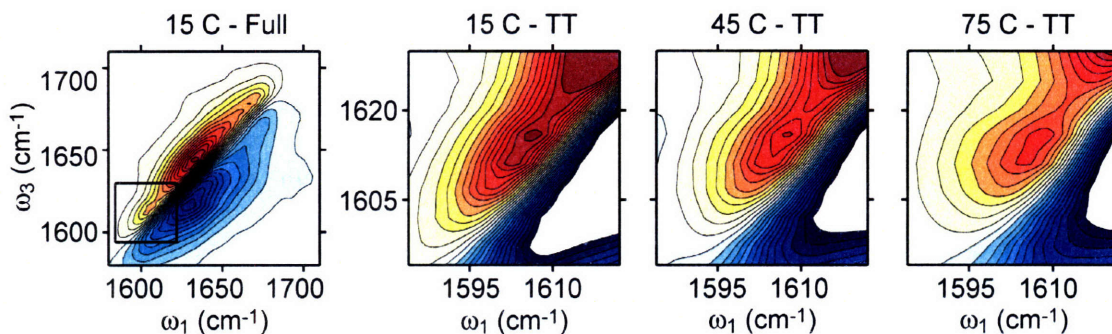
### **7.6.2. TZ2-S1**

In each TZ2-S1 spectrum, the  $\nu_{\text{S1}}$  peak is not well-resolved, and most of the isotope-induced spectral changes are observable in the  $\nu_{\perp}$  and  $\nu_{\parallel}$  bands. The intensity gain in the TZ2-S1 – TZ2-UL difference spectrum upon isotope labeling is consistent with the range of frequencies and intensities observed in the spectral simulations. Based on these observations, we conclude that the N-terminal serine group is free to explore a large configuration space, especially compared to the other labeled residues. The isotope-shifted peaks in the spectral simulations are done on structures that should be seen as limiting cases. The NMR structure, for example, may over-emphasize the degree to which the serine amide group is coupled across the  $\beta$ -strand. Similarly, the disordered

Markov states may over-estimate the blue-shift of the S1 site energy, which would explain why the calculated  $\nu_{S1}$  energies over  $1720\text{ cm}^{-1}$  do not contribute significantly to the spectroscopy. These observations lead to a picture in which the peptide is preferentially disordering at the N and C-termini as described above.

### 7.6.3. TZ2-TT

Thermally-induced spectral changes to the  $\nu_{TT}$  peak are emphasized in Fig. 7.15. Similar to the  $\nu_{K8}$  peaks, the maximum of the positive diagonal  $\nu_{TT}$  peak does not shift appreciably with temperature. Instead, the  $\nu_{TT} - \nu_{\perp}$  peak splitting increases from  $26\text{ cm}^{-1}$  at  $15\text{ }^{\circ}\text{C}$  to  $34\text{ cm}^{-1}$  at  $85\text{ }^{\circ}\text{C}$ , which mainly tracks the blue-shift of the  $\nu_{\perp}$  peak. The relative intensity of the isotope-shifted peak also drops, going from 40% of the band maximum at  $15\text{ }^{\circ}\text{C}$  to 20% at  $85\text{ }^{\circ}\text{C}$ . This decrease in intensity is partially ascribed to a loss of coupling between isotope-labeled amide groups and the  $^{12}\text{C}$  amides. The anti-diagonal width of the peak increases substantially with temperature, while the diagonal width decreases (Fig. 7.10). This indicates that the  $\nu_{TT}$  peak is going from a diagonally elongated peak at low temperatures to a nearly symmetric peak in  $\omega_1$  and  $\omega_3$  at high temperatures.



**Figure 7.15** Zoom in of temperature-dependent 2D IR spectra of TZ2-TT shown for the region highlighted in the left plot. Contours lines in the right three spectra are plotted from 40% to -40% of the amide I maximum at the respective temperature.

The observation of decreasing intensity and increasing homogeneous broadening for  $\nu_{TT}$  indicates that there are two types of temperature-induced structural disorder in the mid-strand region of TZ2. The first disordering pathway is the transition of the peptide backbone away from its native geometry. This results in less vibrational coupling between the amide groups, and hence a lower participation of the labeled amides in the excitonic vibrations.<sup>17</sup> The second disordering pathway is increasing picosecond-scale structural fluctuations that lead to correlated frequency shifts. The two sources of this are small orientational fluctuations between the T3 and T10 residues in native-like geometries, and increasing fluctuations of disordered peptide groups. These disordering pathways are consistent with the fraying transition described in previous hairpin simulations.<sup>18-20</sup> At low temperatures, the mid-strand region around T3 and T10 is relatively rigid, and at high temperatures it is significantly more flexible.



## 7.7. Conclusions

In this chapter, temperature-dependent spectral changes to the amide I band of TZ2 were studied at equilibrium. Three isotope substitutions were introduced to probe the spectral response of the  $\beta$ -turn (TZ2-K8), the N-terminus (TZ2-S1), and the mid-strand region (TZ2-TT) of the hairpin structure. From these data, specific conclusions were made regarding the thermal denaturation pathway. The  $\nu_{\perp}$  ( $\sim 1636\text{ cm}^{-1}$ ) and  $\nu_{\parallel}$  bands ( $\sim 1675\text{ cm}^{-1}$ ) report generally on the presence of  $\beta$ -sheet secondary structure, and the  $\nu_{\text{dis}}$  band ( $\sim 1650\text{ cm}^{-1}$ ) corresponds to amide groups in a disordered state. As the temperature is raised, the  $\beta$ -sheet structure partially denatures, but retains some native contacts at high temperatures. The  $^{13}\text{C}'$  label at the K8 amide group in TZ2-K8 gives rise to two bands that identify unique turn geometries present from high to low temperatures. At high temperatures, structural disorder within the two turn configurations does not increase significantly. The  $\beta$ -turn is therefore described as well-ordered over the range of experimentally accessed temperatures. For TZ2-S1, an  $^{18}\text{O}$  label at the S1 amide group provides evidence for an N-terminus that explores a range of configurations. These vary from structures in which the S1 amides are hydrogen bonded to the W11 amide to structures in which there is no interaction with the other peptide amides. Thermal disordering at the N-terminus is described as an increase in the range of configurations explored by the S1 peptide group. Finally, data from TZ2-TT ( $^{13}\text{C}'$  labels at T3 and T10) show that the mid-strand backbone contacts are statically disordered at low temperatures. At high temperatures, the labeled amides denature specifically by exploring a larger configurational space at a faster time scale.

A number of recent reports have observed that the folding pathway and free energy landscape of TZ2 are heterogeneous.<sup>1,2,6,13,14</sup> Molecular dynamics simulations have produced free energy surfaces from 255 to 350 K that predict an energy minimum 2.5 Å ( $C_{\alpha}$  RMSD) away from the native state.<sup>2</sup> A significant barrier separates it from the native minimum, and the relative energy of the two minima shifts as the temperature increases. The existence of the  $\nu_{K8}$  peaks is direct evidence for two populated states that are separated by barrier. One of these states (corresponding to  $\nu_{K8-1}$ ) is populated by peptides with non-native turns, and the other (corresponding to  $\nu_{K8-2}$ ) is populated by structures with native  $\beta$ -turns. No spectral signatures of inter-conversion between the states are observed, which is further evidence of a significant barrier.

The observations of this chapter are consistent with a peptide that at low temperatures populates at least two distinct conformational states (native and non-native turn). As the temperature increases, disorder increases at the termini and gradually the T3-T10 backbone region of the peptide samples a larger configurational space on a faster time scale. This disordering transition can be described as hairpin fraying, which has been proposed as a mechanism for hairpin unfolding.<sup>13,18</sup> In the following chapter, the transient thermal unfolding of TZ2 will be presented for each of the isotopologues shown here.

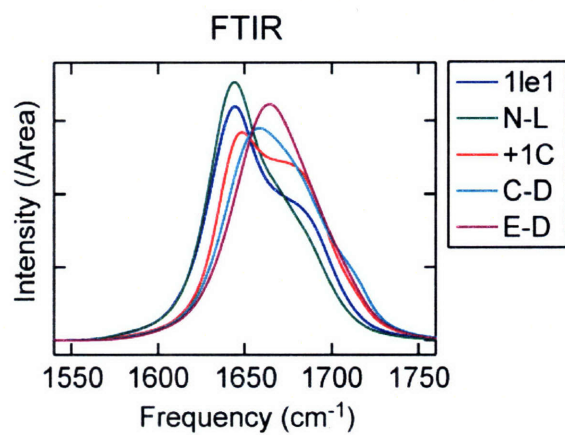
## 7.8. References

- (1) Yang, W. Y.; Gruebele, M. J. *Am. Chem. Soc.* 2004, 126, 7758-7759.
- (2) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. J. *Mol. Biol.* 2004, 336, 241-251.
- (3) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* 2004, 101, 4077-4082.
- (4) Decatur, S. M. *Acc. Chem. Res.* 2006, 39, 169-175.
- (5) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. *Proc. Natl. Acad. Sci. USA* 2001, 99, 9081-9081.
- (6) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. *J. Chem. Phys.* 2007, 126, 155101.
- (7) Bour, P.; Keiderling, T. A. *J. Am. Chem. Soc.* 1993, 115, 9602-9607.
- (8) Hamm, P.; Woutersen, S. *Bull. Chem. Soc. Jpn.* 2002, 75, 985-988.
- (9) Ham, S.; Cha, S.; Choi, J.-H.; Cho, M. *J. Chem. Phys.* 2003, 119, 1451-1461.
- (10) Cheatum, C. M.; Tokmakoff, A.; Knoester, J. *J. Chem. Phys.* 2004, 120, 8201-8215.
- (11) Ganim, Z.; Tokmakoff, A. *Biophys. J.* 2006, 91, 2636-2646.
- (12) Smith, A. W.; Tokmakoff, A. *J. Chem. Phys.* 2007, 126, 045109.
- (13) Wei, G.; Mousseau, N.; Derremaux, P. *Proteins* 2004, 56, 464-474.
- (14) Pitera, J. W.; Haque, I.; Swope, W. C. *J. Chem. Phys.* 2006, 124, 141102.
- (15) Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. *J. Phys. Chem. B* 2006, 110, 7545-7555.

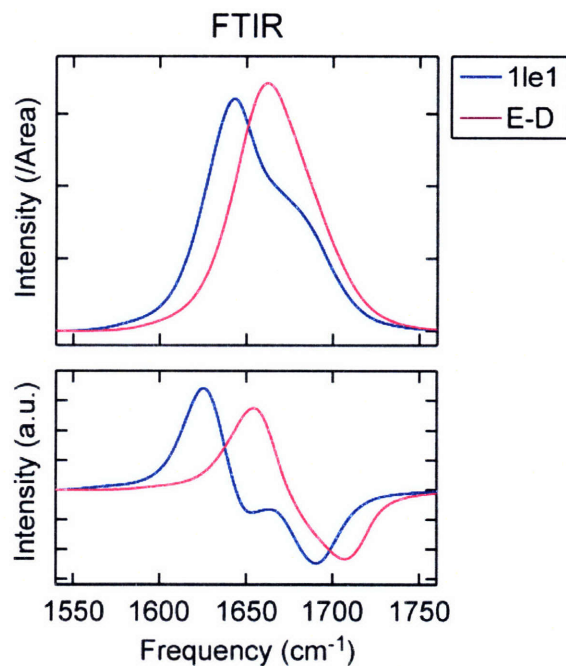
- (16) Kim, Y. S.; Hochstrasser, R. M. *J. Phys. Chem. B* 2007, 111, 9697-9701.
- (17) Kubelka, J.; Keiderling, T. A. *J. Am. Chem. Soc.* 2001, 123, 6142-6150.
- (18) Bolhuis, P. G. *Biophys. J.* 2005, 88, 50-61.
- (19) Bolhuis, P. G. *Proc. Nat. Acad. Sci., USA* 2003, 100, 12129-12134.
- (20) Wei, G. H.; Mousseau, N.; Derreumaux, P. *Proteins* 2004, 56, 464-474.

## 7.9. Appendix

### 7.9.1. Simulated FTIR spectra



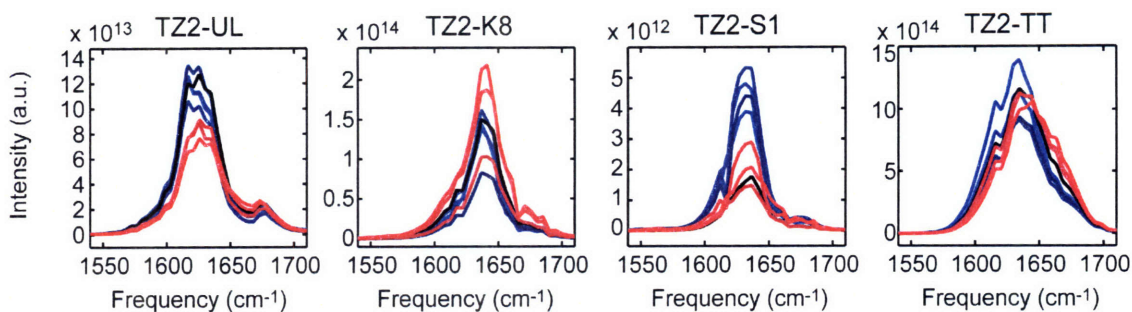
**Figure 7.a.1** Simulated FTIR spectra of unlabeled TZ2-UL for the NMR structure and four Markov states. Methods and notation are described in Section 7.4.



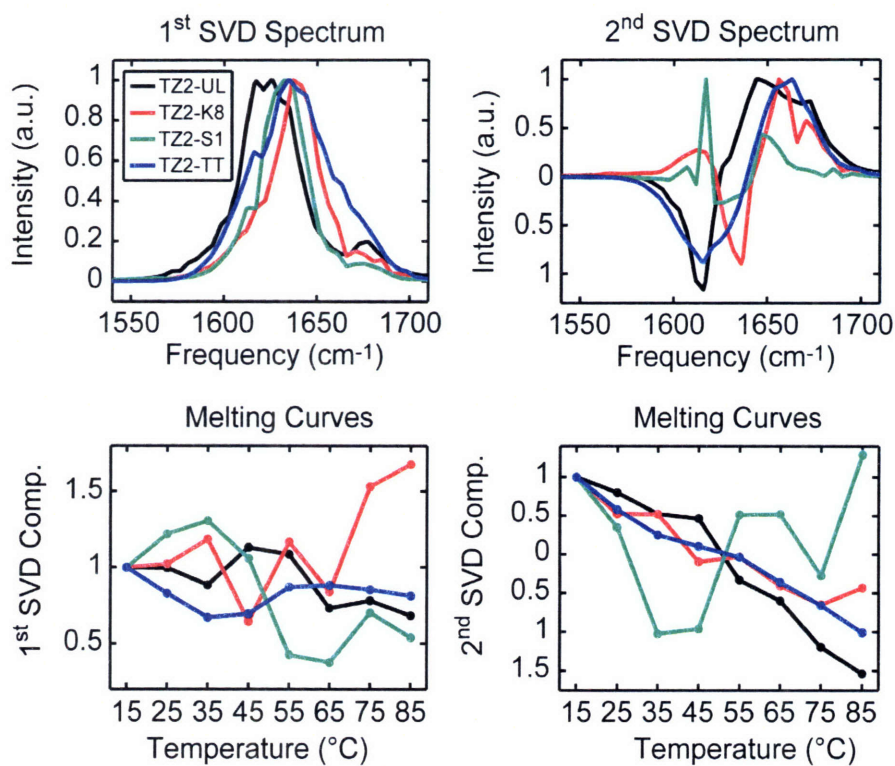
**Figure 7.a.2** (Top) Simulated FTIR spectra of TZ2-S1 for the NMR structure and E-D. (Bottom) Difference spectra obtained by subtracting the simulated TZ2-UL spectra (Fig. 7.a.1.) from the TZ2-S1 spectra in the top panel.

### 7.9.2. DVE from 2D IR

Major features of the TZ2 DVE spectra can be reproduced by numerically calculating the DVE spectrum from the experimentally-obtained 2D IR spectra using Equation 7.2. The signal-to-noise of the reconstructed DVE spectra is lower than the directly measured DVE spectra, but the trends with temperature are consistent within the error limits. These trends are observed in the SVD analysis of the reconstructed spectra (Fig. 7.a.4)



**Figure 7.a.3** Equilibrium thermal denaturation of DVE spectra reconstructed from 2D IR data taken from 15 to 85 °C in 10 °C increments.



**Figure 7.a.4** SVD analysis of the temperature dependent (reconstructed) DVE spectra shown in Figure 7.a.3. The SVD component spectra are not scaled, while the  $n^{\text{th}}$  component melting curves are normalized to the 15 °C value.

# Chapter 8

## Transient Unfolding of TZ2

### 8.1. Introduction

In Chapter 7 the thermal denaturation of TZ2 was investigated by measuring the amide I spectrum of the peptide at equilibrium over a range of temperatures. Specific structural disordering events were observed and summarized in the conclusion of Chapter 7. In this chapter, the thermal unfolding of TZ2 is studied with temperature-jump (T-jump) methods that characterize the response of the peptide to a fast perturbation. The experiments described below produce a 10 to 20 °C increase to the solvent temperature within 10 ns. Following the T-jump, the amide I band of TZ2 is then probed to observe spectral changes as the peptide relaxes to the higher temperature environment. Using this technique, kinetic rates can be assigned to the structural disordering events described in the previous chapter. Two T-jump probes are used to measure the TZ2 amide I band. First, DVE spectroscopy is used to sample the delay time following the T-jump laser pulse from nanoseconds to milliseconds, and to assign rates to the observed spectral changes. Second, 2D IR spectra are collected at each decade of delay time from 10 ns to



10  $\mu\text{s}$  following the T-jump to provide a more detailed description of disordering transition.

Using T-jump DVE spectroscopy, a kinetic rate is assigned to the structural events giving rise to the spectral shifts observed in Chapter 7. For example, the decay of the TZ2-UL  $\nu_{\perp}$  band from a starting temperature of 45  $^{\circ}\text{C}$  is measured to be  $0.97\pm 0.12$   $\mu\text{s}$ , which is assigned to the disordering of  $\beta$ -sheet secondary structure. The rise time of the  $\nu_{\text{dis}}$  peak is found to be  $0.90\pm 0.07$   $\mu\text{s}$ , which signifies the concomitant gain of disordered conformations as the peptide ensemble relaxes to the higher temperature environment. Under identical conditions, the TZ2-K8  $\nu_{\text{K8}}$  band decays with a  $1.08\pm 0.58$   $\mu\text{s}$  time scale, which is assigned to intensity loss as the K8 amide decouples from the  $^{12}\text{C}$  amide modes. Finally, the TZ2-TT  $\nu_{\text{TT}}$  band intensity decays with a  $0.70\pm 0.09$   $\mu\text{s}$  time scale compared to the  $0.82\pm 0.12$   $\mu\text{s}$  TZ2-TT  $\nu_{\perp}$  band decay. This unfolding rate is assigned to the transition of the mid-strand region of the peptide from a statically disordered ensemble to a larger distribution of inter-converting structures. Each of these measured time scales are very similar, ranging from 0.7 to 1.08  $\mu\text{s}$ . Our transient folding experiments, therefore, do not show any evidence for a preferential ordering of local structural events at  $T_0 = 45$   $^{\circ}\text{C}$ .

Temperature-jump 2D IR spectroscopy is used to explore the transient disordering events by observing diagonal peak broadening and intensity shifts as well as loss or gain in the off-diagonal cross peaks. In the transient 2D IR data, the  $\nu_{\perp}$  band of each system undergoes anti-diagonal line broadening on the order of 0.5-1.0  $\text{cm}^{-1}$ , which is consistent with the equilibrium temperature-dependent 2D IR spectra. The time scale of this broadening is  $\sim 1$   $\mu\text{s}$ , which is the rate at which the peptide becomes more dynamically

disordered. The line width of the  $\nu_{K8}$  peak corresponding to native peptide turns does not increase during the T-jump experiment. This is evidence that the turn region remains relatively static throughout the unfolding transition. In contrast, the line width of the  $\nu_{TT}$  band increases substantially, demonstrating that the unfolding of the mid-strand region can be described as increasingly fast structural fluctuations. For both TZ2-K8 and TZ2-TT, cross peak intensity between the  $\nu_{\perp}$  band and the isotope-shifted band is observed to decay with the same rate as the other spectral features. Loss of cross-peak intensity signifies structural rearrangements that decouple the isotope-labeled amide groups from the rest of the peptide. Based on the conclusions from the diagonal peaks and the other spectral observations, the decay of the  $\nu_{K8} - \nu_{\perp}$  cross peak is interpreted as disordering away from the K8 amide group, while the  $\nu_{TT} - \nu_{\perp}$  cross peak decay is interpreted as disordering in both the  $^{12}\text{C}$  and  $^{13}\text{C}$  amide groups.

These observations of the transient unfolding of TZ2 support a mechanism that is heterogeneous in the sense that the disordering transition goes from a small set of structures to a larger set of structures through a distribution of folding pathways. No evidence is seen for a preferential ordering of events that was proposed in research reviewed in Chapter 1.<sup>1,4</sup> However, the N-terminus is observed to be more disordered than the mid-strand region of the hairpin, and the mid-strand region is more disordered than the  $\beta$ -turn. This scenario can be described as hairpin fraying, which has been proposed for other systems.<sup>2,5</sup> Together, these experiments represent a novel approach to observe residue level structure during the transient thermal denaturation of polypeptides.

## 8.2. Methods

Temperature-jump DVE and 2D IR methods are described in detail in Chapter 3. DVE and 2D IR data were collected at a starting temperature,  $T_0$ , of 45 °C, with an approximately 20 °C T-jump using the retro-reflecting mirror described in Chapter 3. T-jump DVE data were also collected at three starting temperatures (25, 35, and 45 °C) at  $\Delta T \sim 10$  °C. The T-jump DVE delay time  $\tau$  was stepped from -10 ns to 0.9 ms with four data points per decade. At each delay, data was collected in-phase with an optical chopper for 5000 shots and 180° out-of-phase for 2000. In the figures shown throughout this chapter, the change in the DVE signal with T-jump delay time,  $\Delta S(\tau, \omega)$ , is defined as:

$$\Delta S(\tau, \omega) = \frac{S(\tau, \omega) - S(\tau_{eq}, \omega)}{S(\tau_{eq}, \omega_{max})}. \quad (8.1)$$

In equation 8.1, the DVE spectrum collected at each T-jump delay,  $\tau$ , is defined as  $S(\tau, \omega)$ , and  $S(\tau_{eq}, \omega)$  is the average of the DVE spectra collected from the 48<sup>th</sup>, 49<sup>th</sup> and 50<sup>th</sup> DVE probe pulses following the T-jump. The delay time for these pulses is at least 48 ms +  $\tau$ , at which point the bath and peptide have relaxed to the equilibrium temperature,  $T_0$ . The difference DVE signal is normalized to the maximum of the equilibrated spectrum,  $S(\tau_{eq}, \omega_{max})$ , to compare percentage change from system to system. For specific spectral regions,  $\Delta S(\tau, \omega)$  is fit to a single exponential function:

$$y(\tau) = y_0 - A \cdot e^{-\tau/\tau_{obs}} \quad (8.2)$$

Because the DVE signal,  $S_{\text{DVE}}$ , scales as the square of the population,  $N^2$ , the rate of decay goes as  $N \propto A^2 e^{-2k\tau}$ , so that  $\tau_{\text{obs}} = 1/2k_{\text{obs}}$ . However, for small changes in signal, the observed rate is taken to be  $k_{\text{obs}}^{-1} = \tau_{\text{obs}}$ .<sup>6,7</sup> In some previous studies of TZ2, a two-state kinetic model was used to extract the folding and unfolding rates.<sup>8</sup> Here, we choose not to make this approximation because the melting curves from the infrared spectra do not fit well to a two-state model. Therefore, the observed folding rate,  $k_{\text{obs}}^{-1}$ , is used for the data analysis presented in this chapter. This is consistent with two recent treatment of TZ2 folding kinetics.<sup>9,10</sup>

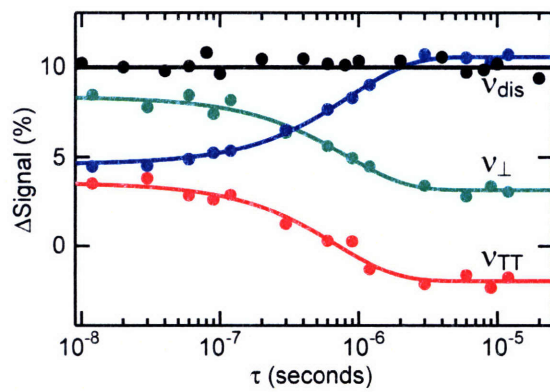
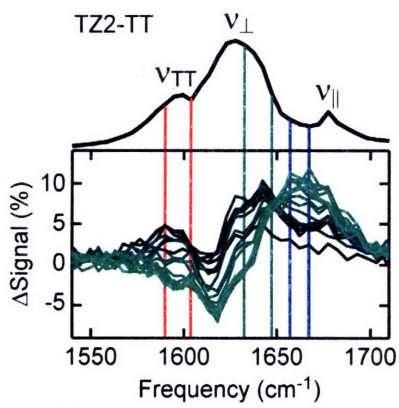
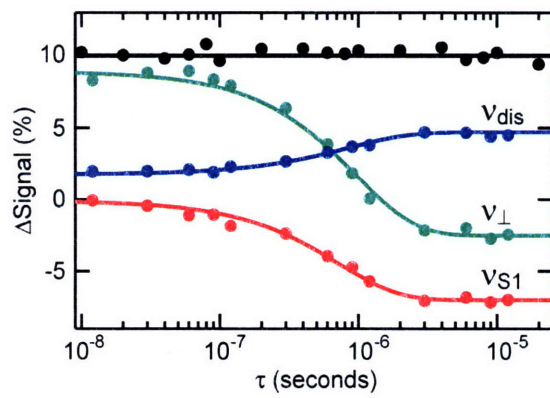
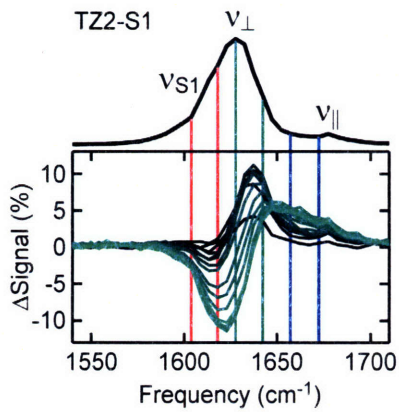
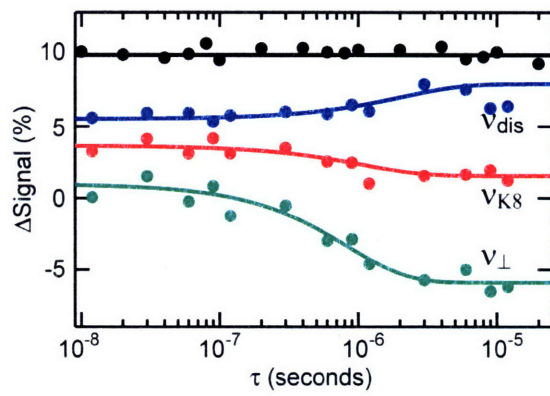
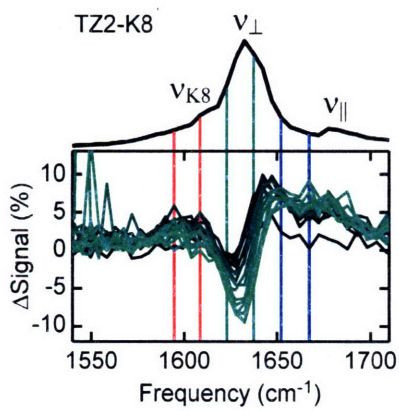
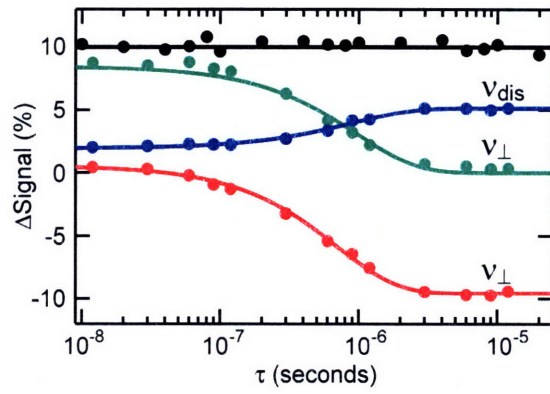
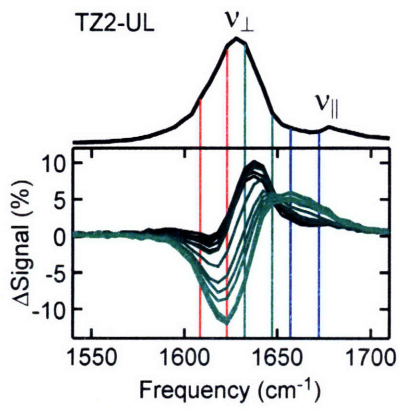
For T-jump 2D IR data, the evolution time,  $\tau_1$ , was collected out to 1.5 and 1.2 ps for rephasing and non-rephasing geometries respectively. Each  $\tau_1$  time point was averaged for 5000 shots leading to a 2.5 hour collection time for one 2D IR spectrum at a single  $\tau$  delay time. Each T-jump 2D IR spectrum consists of at least four 2.5 hour averages. The 2D IR difference spectra are displayed similarly to the DVE difference data, with the difference signal normalized by the peak maximum of the equilibrated spectrum.

## 8.3. Transient Folding Results

### 8.3.1. T-jump DVE at $T_0 = 45$ °C

The time-dependant DVE difference data of TZ2-UL are shown in Fig. 8.1 (top row). During the range of time delays shown in the figure, the transient temperature profile is constant at approximately 63 °C. After 10 ns of delay time, the difference

spectrum is largely positive, with a band shape that resembles the equilibrium spectrum. This positive change is assigned primarily to the rise in solvent transmission directly following the T-jump. The transient signal stays relatively flat until around 100 ns, at which point the  $\nu_{\perp}$  peak intensity begins to drop exponentially and then levels out at  $\sim 10$   $\mu\text{s}$ . There are two contributions to the  $\nu_{\perp}$  peak decay. The first is the  $\nu_{\perp}$  blue-shift as the  $\beta$ -sheet is partially disrupted, and the second is loss of intensity associated with the 2D IR line shape broadening. These two mechanisms cannot be resolved separately with DVE spectroscopy, but they will be revisited in the discussion of the T-jump 2D IR data. The fitted time scale for the  $\nu_{\perp}$  decay is between 0.70 and 0.97  $\mu\text{s}$  depending on which part of the band is observed. The time scale difference across the band may represent the unique folding rates for different  $\beta$ -sheet configurations, however the differences are too small to draw definite conclusions. The  $\nu_{\text{dis}}$  peak intensity increases with time, which is assigned to the population increase of disordered amide groups expected as the peptide relaxes to the higher temperature environment. The fitted time constant for this change is  $0.90 \pm 0.07$   $\mu\text{s}$ .



**Figure 8.1** (Previous page) (Left column) Transient difference DVE spectra of each TZ2 isotopologue is shown from 1 ns (black) to 3  $\mu$ s (green). The calibrated T-jump,  $\Delta T$ , for TZ2-UL and TZ2-TT is 18.5  $^{\circ}$ C, and for TZ2-K8 and TZ2-S1 is 18.2  $^{\circ}$ C. The equilibrium temperature,  $T_0$ , for each data set is 45  $^{\circ}$ C. Difference spectra are obtained as described in Section 8.2, with the  $\Delta$ Signal defined as the difference DVE spectrum normalized by the equilibrium band maximum. The top trace is the equilibrated DVE spectrum obtained by averaging the 48-50<sup>th</sup> pulses following the T-jump. (Right column) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles, and solid lines represent single-exponential fits to the data from 12 ns to 12  $\mu$ s. Fitted parameters are displayed in Table 8.1. The black circles represent a sample transient absorbance trace and the solid black line is an exponential fit to those data, which are scaled to show the temporal profile of the solvent temperature.

Changes to the TZ2-K8 spectrum are observed in the time-dependent DVE difference spectra (Fig. 8.1, 2<sup>nd</sup> row). The total DVE signal is three to four times smaller than that of TZ2-UL, and so the noise in the difference signal is more significant. At  $\tau = 10$  ns, the difference spectrum is positive due to the solvent transmission increase, with most of the intensity in the  $\nu_{K8}$  and  $\nu_{dis}$  band regions. The  $\nu_{\perp}$  peak intensity decays exponentially with a  $0.84 \pm 0.19$   $\mu$ s time constant, which agrees with the TZ2-UL time scale within the margins of error. The  $\nu_{dis}$  band also increases, but the noise makes it difficult to assign a time constant. Consequently, the value shown in Table 8.1 (2.1  $\mu$ s) has a wide margin of error ( $\pm 1.15$   $\mu$ s). After the initial increase of the  $\nu_{K8}$  band, the signal decays with a  $1.08 \pm 0.58$   $\mu$ s time constant. Analogous to the other bands, the  $\nu_{K8}$  decay is assigned to peptide structural changes upon unfolding.

For TZ2-S1, the spectral changes (Fig. 8.1, 3<sup>rd</sup> row) are very similar to those of the unlabelled peptide. The difference signal starts positive and develops at a  $\sim 1$   $\mu$ s time scale. Spectral changes are assigned to the same mechanism described for TZ2-UL. The

$\nu_{\perp}$  peak decays with a  $0.96 \pm 0.11 \mu\text{s}$  fitted time scale and the disordered band increases with a  $0.86 \pm 0.13 \mu\text{s}$  time constant. In an effort to assign the kinetics of the S1 isotope-labelled amide group, the frequency channels around the unresolved  $\nu_{\text{S1}}$  band are shown in red. The fitted time constant is  $0.72 \pm 0.08 \mu\text{s}$ , which is equal to the similar trace of the TZ2-UL data making it difficult to assign the  $\nu_{\text{S1}}$  band dynamics. This lack of a distinguishable time constant is consistent with a serine residue that is highly disordered and very loosely coupled to the rest of the peptide. As discussed in Chapter 7, a disordered N-terminus is consistent with the TZ2-S1 amide I spectrum as well as molecular dynamics simulations.

The time-dependent DVE difference data of TZ2-TT are shown in Fig. 8.1 (bottom row). The total DVE signal of this compound is lower than that of TZ2-UL, and the noise of the difference signal is comparable to the TZ2-K8 data. As in the other systems, the difference spectrum at 10 ns is positive due to the solvent transmission increase. The  $^{12}\text{C}$   $\nu_{\perp}$  peak intensity decays exponentially with a  $0.82 \pm 0.12 \mu\text{s}$  time constant, which agrees with the  $^{12}\text{C}$   $\nu_{\perp}$  peak decay of the other systems. The time constant of the  $\nu_{\text{dis}}$  band increase is  $0.86 \pm 0.07 \mu\text{s}$ , which is also similar to the other data sets. Decay of the  $\nu_{\text{TT}}$  peak occurs with a  $0.70 \pm 0.09 \mu\text{s}$  time constant. This loss is assigned to structural disordering around the T3 and T10 amide groups, primarily in the form of increasing homogeneous and heterogeneous line broadening. The decay of the  $\nu_{\text{TT}}$  peak is slightly faster than the  $\nu_{\perp}$  peak, but the margin of error is too large to convincingly distinguish them. This suggests that unfolding of the mid-strand region of the hairpin occurs on a similar time scale as the other peptide disordering events.



Another way to compare the T-jump DVE spectra of the isotope-labeled peptides is with singular value decomposition of the DVE spectra as a function of time. This was done for each system from 3 ns to 3  $\mu$ s, and the time constants for the 2<sup>nd</sup> component decay are included in Table 8.1. The variation in the time constant is small between the systems, suggesting that the dominant features of the transient signal are similar from system to system.

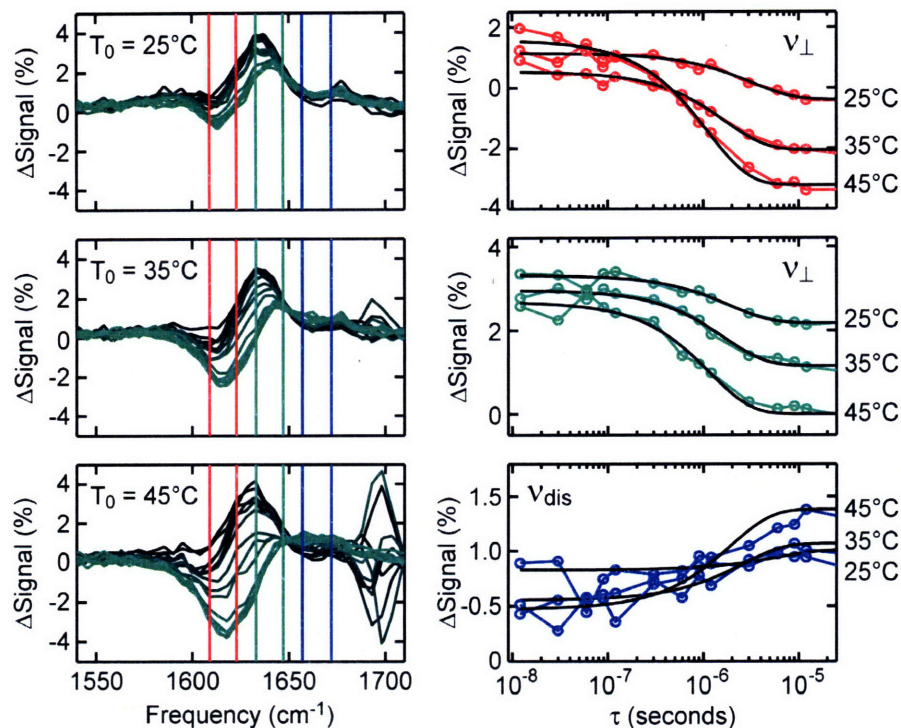
System	T <sub>0</sub> (°C)	$\Delta T$ (°C)	$k_{obs}^{-1}(\mathbf{v}_{\perp})^*$	$k_{obs}^{-1}(\mathbf{v}_{dis})^*$	$k_{obs}^{-1}(\mathbf{v}_{iso})^*$
TZ2-UL	45	18.5	0.97±0.12	0.90±0.07	0.70±0.03
TZ2-K8	45	18.2	0.84±0.19	2.10±1.15	1.08±0.58
TZ2-S1	45	18.2	0.96±0.11	0.86±0.13	0.72±0.08
TZ2-TT	45	18.5	0.82±0.12	0.86±0.07	0.70±0.09

**Table 8.1** Results of single exponential fits to data shown in Figs. 8.1. Error values for  $\tau_1$  represent the error in fit to Equation 8.2. \*All values are reported in  $\mu$ s.

### 8.3.2. T-jump DVE at variable T<sub>0</sub>

T-jump DVE data of TZ2 were also measured at three starting temperatures (T<sub>0</sub> = 25, 35 and 45 °C) with smaller T-jumps than in the last section ( $\Delta T$  = 8.1, 9.7, and 9.5 °C, respectively). The variation in  $\Delta T$  is small enough to compare the folding rates at each temperature, which is shown in Fig. 8.2 for TZ2-UL at T<sub>0</sub> = 25, 35 and 45 °C. The DVE difference spectra are presented using Equation 8.1, and the frequency channels used to construct the transient traces are consistent with those in Fig. 8.1. At 25 °C, the negative amplitude on the red side of the TZ2-UL  $\mathbf{v}_{\perp}$  peak is lower than at the higher starting temperatures. This suggests that the  $\mathbf{v}_{\perp}$  peak experiences less blue-shift upon peptide unfolding. As this blue shift is a sign of backbone disordering, we conclude that the T-jump induced structural changes at low temperatures is much smaller than at higher

temperatures, which is consistent with the melting curves observed in Chapter 7. This is borne out in the exponential fits to the transient data (Fig. 8.2, left column), where the amplitude of the decay increases with temperature. Changes to the  $\nu_{\text{dis}}$  band are very small at  $T_0 = 25$  and  $35$  °C, so fitted time constants are less reliable at low temperatures. The exponential decay of the  $\nu_{\perp}$  signal at  $T_0 = 45$  °C and  $\Delta T = 9.5$  °C is 70% slower than the decay of the band in Fig. 8.2 with  $T_0 = 45$  °C and  $\Delta T = 18.5$  °C.



**Figure 8.2** (Left) Transient difference DVE spectra of TZ2-UL shown from 1 ns to 3  $\mu$ s at 25, 35, and 45  $^{\circ}$ C. Difference spectra are obtained as described in Section 8.2, with the  $\Delta$ Signal defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu$ s. Time constants from the transient fits are shown in Table 8.2. The intensity fluctuations at 1700  $\text{cm}^{-1}$  are from anomalous electronic noise.

For each of the TZ2 isotopologues, the  $T_0$ -dependent data is shown in the Appendix (Figs. 8.a.1 - 8.a.2). For each system, the  $v_{\perp}$  DVE difference band with temperature is similar to that of TZ2-UL. At lower temperatures the negative amplitude on the red side of the  $v_{\perp}$  peak is less significant, and the amplitude of the change rises with temperature. The exponential decay of each  $^{12}\text{C}$   $v_{\perp}$  signal at  $T_0 = 45^{\circ}\text{C}$  and  $\Delta T = 9.5^{\circ}\text{C}$  is about 70% slower than the decay of corresponding band in Fig. 8.1 with  $\Delta T$

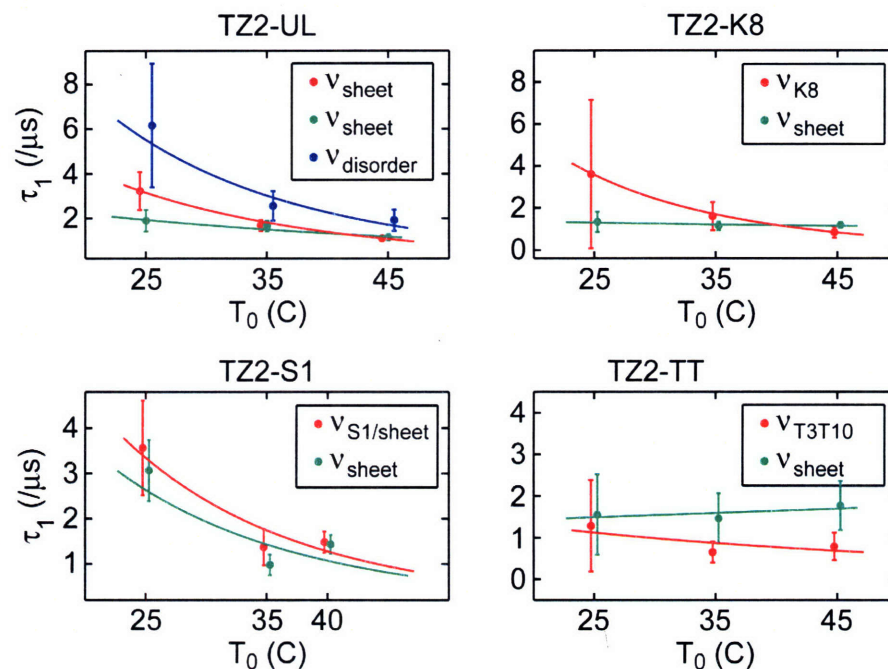
equal to at least 18.2 °C. For TZ2-K8, the changes to the  $\nu_{K8}$  band also grow in with increasing  $T_0$ . At 25 °C, the  $\nu_{K8}$  difference signal is low enough to prevent an accurate exponential fit, but at 45 °C the time constant is 0.86  $\mu$ s. For TZ2-TT, the amplitude of the  $\nu_{TT}$  difference band grows in with increasing  $T_0$ . At 25 °C, the  $\nu_{TT}$  difference signal is too low to accurately fit, but at 45 °C the time constant is 0.80  $\mu$ s. This is similar to the rate found with  $\Delta T = 18.5$  °C, but it is more than twice as fast as the  $\nu_{\perp}$  band decay.

System	$T_0$ (°C)	$\Delta T$ (°C)	$k_{obs}^{-1}(\nu_{\perp})^*$	$k_{obs}^{-1}(\nu_{dis})^*$	$k_{obs}^{-1}(\nu_{iso})^*$
TZ2-UL	25	8.1	1.91±0.48	6.17±2.77	3.23±0.85
	35	9.7	1.67±0.23	2.56±0.66	1.69±0.24
	45	9.5	1.17±0.14	1.92±0.48	1.12±0.11
TZ2-K8	25	8.1	1.36±0.48	0.69±0.32	3.62±3.53
	35	9.7	1.16±0.20	--	1.62±0.67
	45	9.5	1.19±0.14	--	0.86±0.26
TZ2-S1	25	8.1	3.07±6.76	--	3.57±1.05
	35	9.7	0.98±0.23	0.89±0.48	1.37±0.40
	45	9.5	1.44±0.21	1.74±0.74	1.49±0.24
TZ2-TT	25	8.1	1.56±0.96	0.74±0.22	1.29±1.10
	35	9.7	1.47±0.61	1.61±1.71	0.66±0.25
	45	9.5	1.78±0.59	0.58±0.19	0.80±0.33

**Table 8.2** Summary of single exponential fits to the  $T_0$ -dependant T-jump DVE data.  $\Delta T$  is the calibrated rise in temperature. \*All time scales are shown in  $\mu$ s.

For the T-jump DVE data at low temperatures, there is evidence of a time-scale separation between the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peak decay rates. At  $T_0 = 25$  °C, the TZ2-UL  $\nu_{dis}$  time constant is  $\sim 3$  times larger than the  $\nu_{\perp}$  time constant. This separation of time scales is evidence for heterogeneous folding that disappears as the temperature is raised. This trend was also observed in an isotope-labelled trpzip peptide, TZ2C, using T-jump IR

absorbance spectroscopy.<sup>9</sup> It was interpreted as evidence of a kinetic intermediate that is not accessed at higher temperatures.



**Figure 8.3** Summary of time constants for TZ2 T-jump data at  $T_0 = 25, 35,$  and  $45\text{ }^\circ\text{C}$  ( $40\text{ }^\circ\text{C}$  for TZ2-S1). Each set of data points is the average of several channels similar to Fig 8.1. The solid lines represent the fit of the kinetics to the Arrhenius equation. The activation energies are shown in Table 8.3.

Using the Arrhenius rate equation, it is possible to extract an approximate activation energy from the fitted rate constants. For TZ2, this is done for each of the bands described above. For TZ2-UL, the activation energies have modest fitting errors and the energies are comparable to those for TZ2C.<sup>9</sup> For the isotope labelled species, the rates used to find the activation energy have relatively large errors, so some of the values in Table 8.3 are approximate.

System	$E_a(v_{\perp})^*$	$E_a(v_{\text{dis}})^*$	$E_a(v_{\text{iso}})^*$
UL	19±5	46±12	41±5
K8	5±4	--	57±3
S1	47±36	--	50±21
TT	-5±6	--	19±19

**Table 8.3** Summary of activation energies obtained by fitting the temperature-dependant rates to the linear form of the Arrhenius equation:  $\ln(k) = -E_a/(RT) + \ln(A)$ . \*Values are in units of kJ/mol.

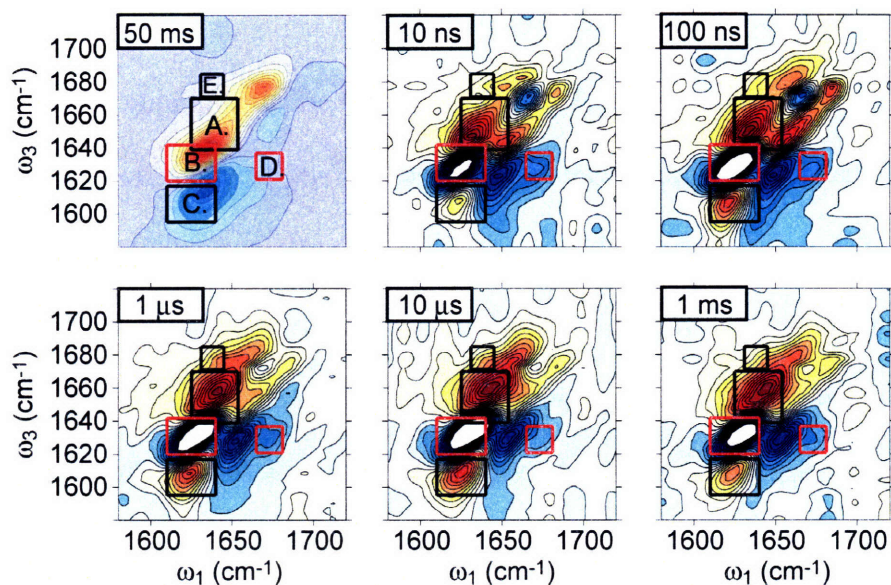
### 8.3.3. T-jump 2D IR

After an initial T-jump perturbation, 2D IR spectra were collected for each of the TZ2 isotopologues as the peptide relaxes to a high temperature environment. T-jump 2D IR spectroscopy resolves the effects of structural rearrangements on the 3<sup>rd</sup> order signal, and allows us to assign the DVE signal changes. The focus of this section will be to observe the 2D peak-shifts and line-width changes during this process.

The 2D IR difference data shown in Fig. 8.4 shows a number of peaks demarcated in the top left panel. Peak **A** is a positive peak that represents gain on the blue side of the positive  $v_{\perp}$  diagonal peak. As the delay time,  $\tau$ , increases from 10 ns to 10  $\mu$ s, the intensity of peak **A** grows from 1.8 to 3.2% of the equilibrated amide I maximum, and the peak maximum shifts from a  $\omega_1 = 1631$  to  $1645 \text{ cm}^{-1}$ . These changes are assigned to an increase in the population of unfolded molecules, which are correlated with loss on the red side of the  $v_{\perp}$  peak (peak **B**). The changes also correlate with the increasing negative intensity on the blue side of the  $v_{\perp}$  overtone band. The time scale of this loss matches the  $\sim 1 \mu$ s decay time observed in the DVE  $v_{\perp}$  peak. Boxes **D** and **E** mark the expected position of cross peaks between the  $v_{\perp}$  and  $v_{\parallel}$  bands. While there is no resolvable

maximum, the spectral intensity in the center of those regions can potentially be used to quantify the loss of vibrational coupling upon thermal unfolding. For TZ2-UL, the absolute change in the cross peak regions increases with temperature. However, shifts in the diagonal peaks interfere with the cross peak intensity and make it difficult to quantify the loss in  $\nu_{\perp}$  -  $\nu_{\parallel}$  vibrational coupling upon transient thermal denaturation.

In addition to the spectral shifts and intensity loss described above, the T-jump 2D IR difference spectra are sensitive to the broadening line shape. This is observed in the diagonal peaks as the peptide relaxes to a higher-temperature environment, and is seen as additional intensity on the wings of the equilibrium peaks. A simpler way to visualize and quantify these changes is to add the difference spectrum to the equilibrated spectrum and then measure the anti-diagonal line widths. Using this approach, we find that the  $\nu_{\perp}$  and  $\nu_{\parallel}$  peak widths increase by 7%, or about  $0.5 \text{ cm}^{-1}$  (Fig. 8.6) at delay times greater than  $1 \mu\text{s}$ . The time scale of this line-width increase is slower than the 10 ns rise in the solvent temperature, but may be faster than the  $1 \mu\text{s}$  time scale of the folded to unfolded structural transition. This suggests that the thermally induced structural fluctuations are not immediately accessed with the solvent temperature, but instead involve  $>10 \text{ ns}$  conformational changes.

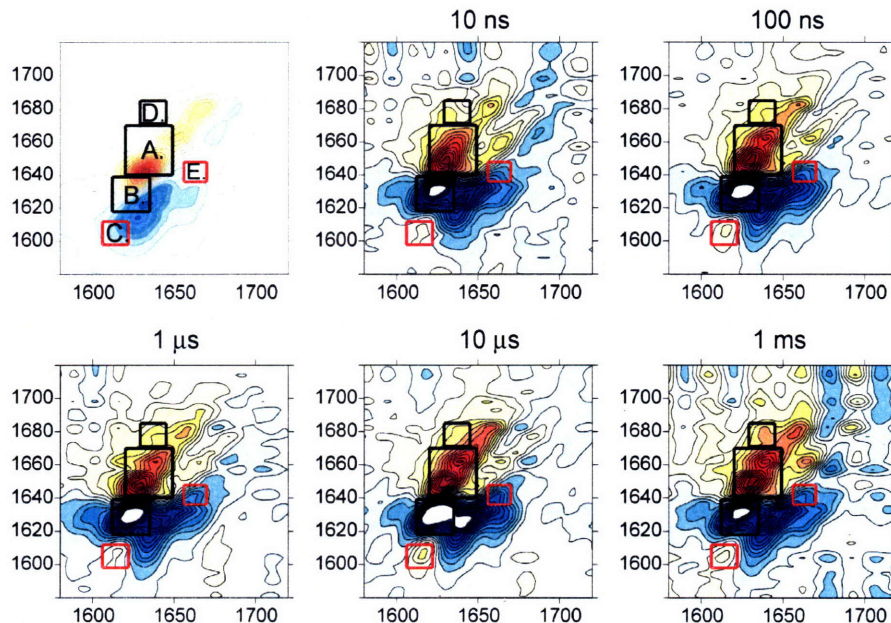


**Figure 8.4** T-jump 2D IR data for TZ2-UL at  $T_0 = 45$  °C, and  $\Delta T = 20.8$  °C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra.

T-jump 2D IR difference spectra were collected at  $T_0 = 10$  °C, where there is limited peptide disordering and the spectral changes are due largely to solvent transmission effects. A comparison to the T-jump 2D IR data at  $T_0 = 45$  °C will help distinguish solvent-induced spectral changes from those reporting on peptide structure. The peaks in the transient 2D IR difference spectra are largely similar to those in Fig. 8.4 except that the intensity is smaller at low temperatures. For example, the maximum changes go from -3% to 2.4% at 10 °C, compared to -6% to +3.2% at 45 °C. This difference is consistent with the variable  $T_0$  T-jump DVE data, which was found to be dramatically smaller at low temperatures. As an example, peak C, which reports on the intensity of the low frequency side of the  $\nu_{\perp}$  overtone, displays the largest disparity from 10 to 45 °C. This difference peak reports specifically on the  $\beta$ -sheet vibrations. The



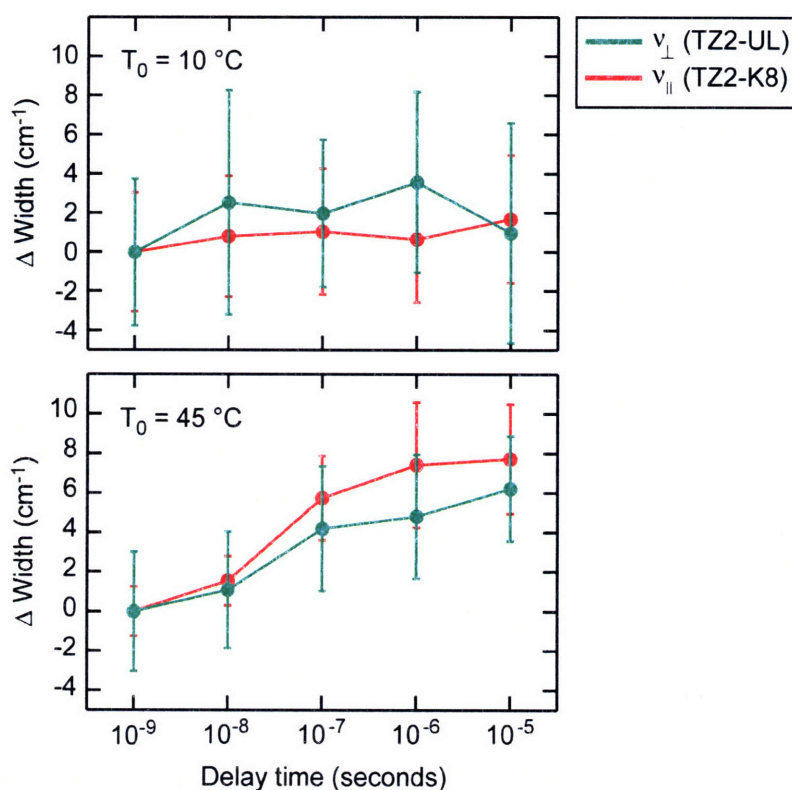
small amplitude of the 2D difference peak **C** is an indication that the  $\beta$ -sheet structure of TZ2 is not disordering substantially at  $T_0 = 10^\circ\text{C}$ .



**Figure 8.5** T-jump 2D IR data for TZ2-UL at  $T_0 = 10^\circ\text{C}$ , and  $\Delta T = 17.3^\circ\text{C}$ . (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra.

In Fig. 8.6, the linewidth changes of the T-jump 2D IR experiment at  $T_0 = 45^\circ\text{C}$  is compared to that of the same measurement done with  $T_0 = 10^\circ\text{C}$ . The data points are obtained by averaging the widths of the peaks at the same frequency channels and in the same way as in Chapter 7 (Section 7.5.2), and error bars are obtained from the standard deviation of the average. In the equilibrium data of Chapter 7, a linear increase to the line width of  $0.010\text{ cm}^{-1}/^\circ\text{C}$  was observed. In the transient 2D IR spectra at  $T_0 = 45^\circ\text{C}$ , there is a gradual (100 ns to 1  $\mu\text{s}$  rise time) increase to the peak width that ends at  $\sim 8\%$  change. The time scale is much longer than the 10 ns rise time of the solvent

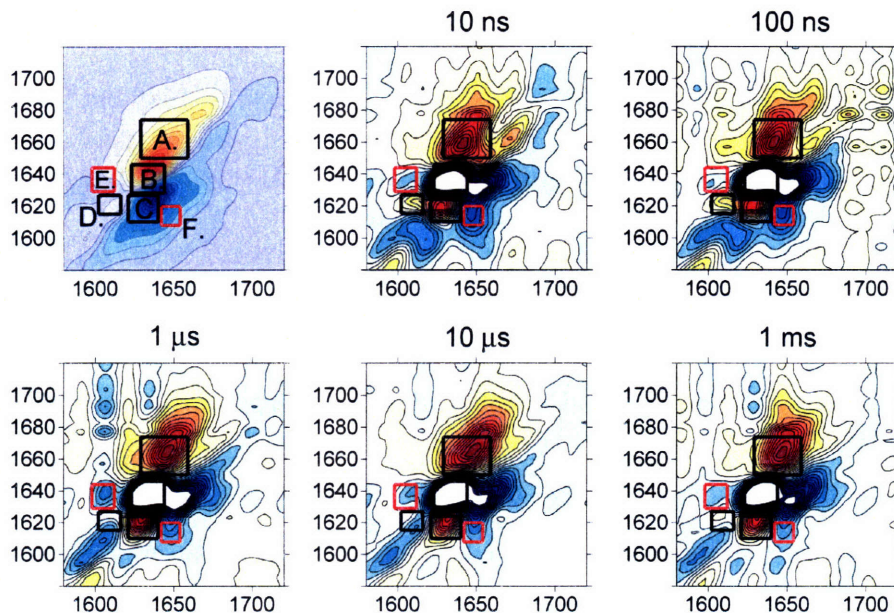
transmission, and is therefore assigned to peptide dynamics. The data at  $T_0 = 10\text{ }^\circ\text{C}$ , on the other hand, is much noisier and appears essentially flat over the same range of time scales. We conclude that the line width increase observed in the equilibrium data is significantly affected by structural disorder. As the temperature increases, the structural contribution to the homogeneous width increases. In the transient data at  $T_0 = 10\text{ }^\circ\text{C}$ , this change may occur too slowly to be seen in the time window of the experiment.



**Figure 8.6** (Top) Line width of the  $v_{\perp}$  and  $v_{\parallel}$  bands of TZ2-UL as a function of T-jump delay time. Each data point represents the  $\Gamma/2$  value of a Lorentzian function (Equation 7.1) fit to the anti-diagonal slices. Error bars are the standard deviation of the fitted widths for six consecutive anti-diagonal slices.

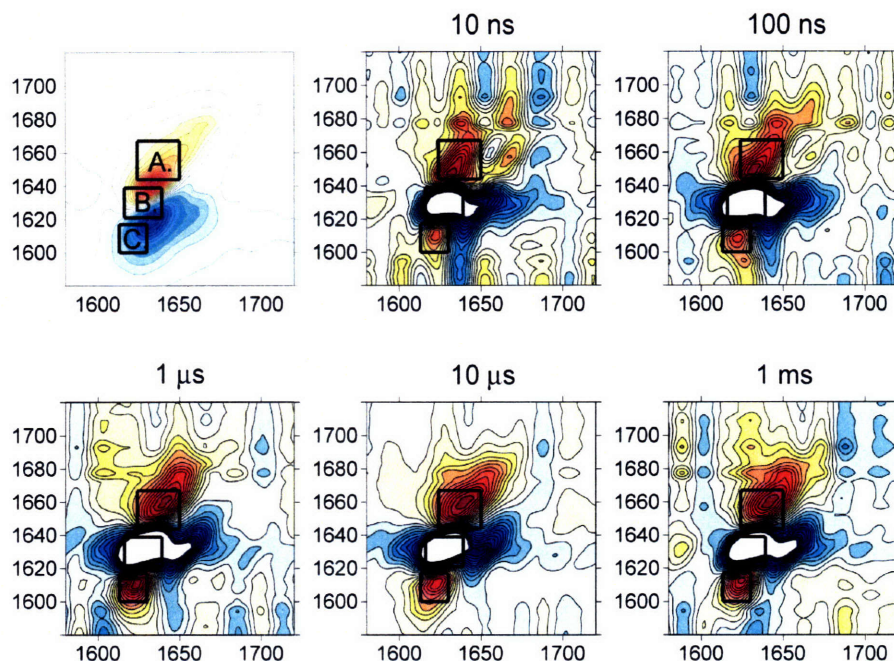
The T-jump 2D IR spectra of TZ2-K8 are shown below in Fig. 8.7. As in TZ2-UL, peak **A** represents the shift of intensity from the positive  $v_{\perp}$  band to the  $v_{\text{dis}}$  band. For

TZ2-K8, the intensity of peak A grows from 2.1 to 3.1% of the equilibrated amide I maximum, and the peak maximum shifts from a  $\omega_1 = 1641$  to  $1649 \text{ cm}^{-1}$  as  $\tau$  goes from 10 ns to 10  $\mu\text{s}$ . The  $\nu_{\text{K8}}$  diagonal difference peak (box D) starts positive (0.5%) at 10 ns, but then decays to negative values (-0.4%) as the delay increases to 10  $\mu\text{s}$ . The  $\nu_{\text{K8}} - \nu_{\perp}$  cross peak (box E) also becomes more negative with time (-0.3 to -0.9%). The correlated spectral changes of these two peaks suggest that the intensity loss is due to a decoupling of the high frequency  $\nu_{\text{K8}}$  peak from the  $\nu_{\perp}$  band. The width of the TZ2-K8,  $^{12}\text{C}$   $\nu_{\perp}$  (Fig. 8.11) and  $\nu_{\parallel}$  bands grows in with a  $\sim 1 \mu\text{s}$  delay time similar to the TZ2 bands. The  $\nu_{\text{K8}}$  T-jump 2D IR data is discussed below (Section 8.4.4)



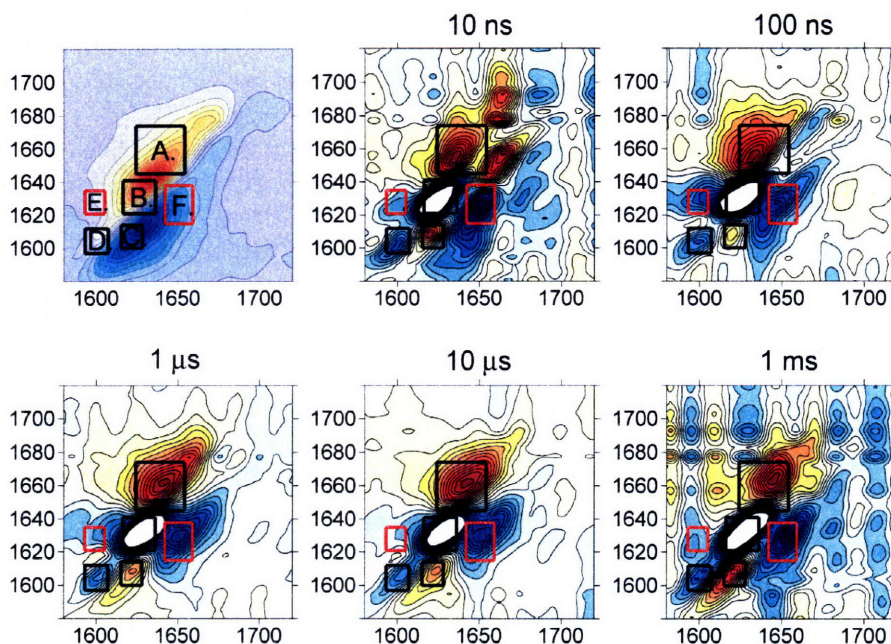
**Figure 8.7** T-jump 2D IR data for TZ2-K8 at  $T_0 = 45\text{ }^\circ\text{C}$ , and  $\Delta T = 20.8\text{ }^\circ\text{C}$ . (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra.

The T-jump 2D IR spectra of TZ2-S1 are shown below in Fig. 8.8. The intensity of peak A grows from 2.2 to 5.1% of the equilibrated amide I maximum, and the peak maximum shifts from  $\omega_1 = 1631$  to  $1644\text{ cm}^{-1}$  as  $\tau$  goes from 10 ns to 10  $\mu\text{s}$ . As in TZ2-UL, this intensity shift represents spectral intensity shifting from the positive  $\nu_\perp$  band to the  $\nu_{\text{dis}}$  band. The spectral changes as a function of delay time are very similar to TZ2-UL. Signatures of the  $\nu_{\text{S1}}$  diagonal peak are indistinguishable from the  $\nu_\perp$  difference peaks.



**Figure 8.8** T-jump 2D IR data for TZ2-S1 at  $T_0 = 45$  °C, and  $\Delta T = 24.7$  °C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra.

The T-jump 2D IR spectra of TZ2-TT are shown below in Fig. 8.9. The intensity of peak **A** grows from 1.9 to 4.9% of the equilibrated amide I maximum, and the peak maximum shifts from  $\omega_1 = 1636$  to  $1644$   $\text{cm}^{-1}$  as  $\tau$  goes from 10 ns to 10  $\mu\text{s}$ . 2D difference peaks **A**, **B** and **F** track the spectral shift from the positive  $\nu_{\perp}$  band to the  $\nu_{\text{dis}}$  band. Peaks **D** and **E** track changes to the  $\nu_{\text{TT}}$  band upon transient thermal denaturation. The intensity of peak **D**, which reports on the loss of the  $\nu_{\text{TT}}$  positive diagonal peak, drops from -1.0 to -1.9% of the equilibrated amide I maximum with increasing delay time. The  $\nu_{\text{TT}}$  difference peak changes are discussed below (Section 8.4.4).



**Figure 8.9** T-jump 2D IR data for TZ2-TT at  $T_0 = 45$  °C, and  $\Delta T = 19.8$  °C. (Top left) Equilibrated 2D IR spectrum obtained by adding the signal from the 49 and 50<sup>th</sup> pulses following the T-jump (8.2). Each additional plot is the 2D difference signal at T-jump delays,  $\tau$ . Boxes shown are used to identify peaks in the 2D difference spectra.

## 8.4. Discussion

### 8.4.1. Transient DVE time scales.

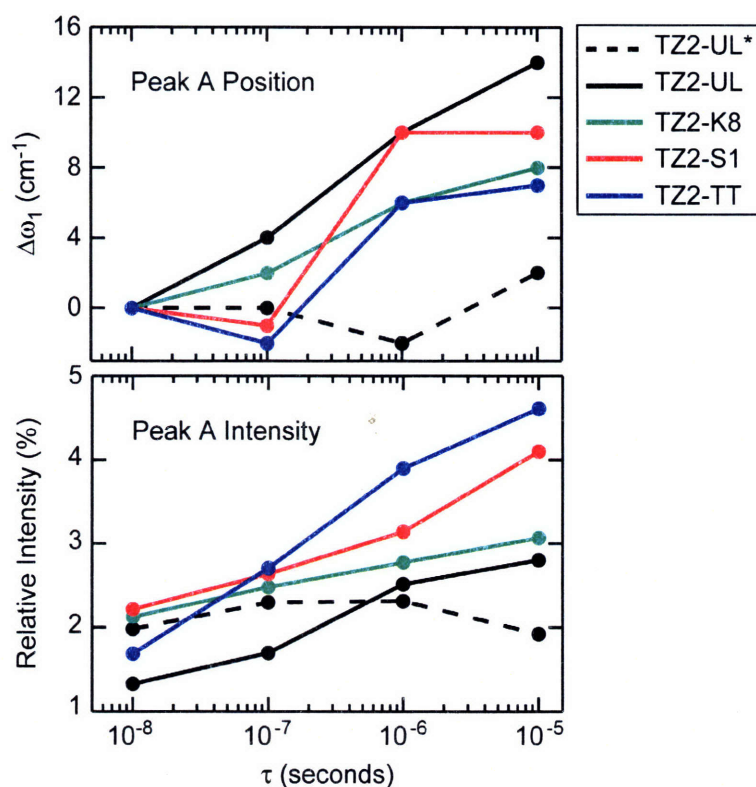
The time scales of the DVE T-jump response at  $T_0 = 45$  °C are remarkably consistent between the various spectral features and between the different isotopologues. Exponential decay times of the  $\nu_{\perp}$  bands are between 0.82 and 0.97  $\mu\text{s}$ , with fitting errors on the order of 15-20%. The exponential rise of the  $\nu_{\text{dis}}$  band for each system is not consistently faster or slower than the respective  $\nu_{\perp}$  decay time. Furthermore, there is no specific time ordering for the turn (K8) or the mid-strand (TT) native contacts based on

the observed rates of the  $\nu_{K8}$  and  $\nu_{TT}$  band decays. For TZ2-UL, TZ2-S1, and TZ2-TT, the decay rate of the lowest frequency channels is about 100 ns faster than the  $\nu_{\perp}$  rates. This might be consistent with a frequency shift of  $\nu_{\perp}$  as the peptide decays through a frayed state. However, the error bars are large enough to prohibit a rigorous time scale separation. This consistency between time scales suggests two possible scenarios. The first is that the transient thermal unfolding of TZ2 does not significantly populate intermediate states. This was the conclusion of Hauser et al. with respect to the TZ2C kinetic rates that were observed to converge at high temperatures. A second possibility is that the time scale reflects population exchange through a heterogeneous distribution of states.

#### 8.4.2. Signatures of $\beta$ -sheet disordering

The 2D IR difference peaks shown in Figs. 8.4, 8.5, and 8.7 to 8.9 each track the peptide spectral response to the T-jump perturbation. The most intense features in the difference spectra are Peaks **A**, **B**, and **C**, which track the thermal response of the  $^{12}\text{C}$  fundamental and overtone transitions. Peak **A** is particularly sensitive to intensity that shifts from the  $\nu_{\perp}$  band to the  $\nu_{\text{dis}}$  band. In Fig. 8.10, the  $\omega_1$  frequency of peak **A** is plotted as the difference between the value at 10 ns and that of longer delay times, ( $\Delta\omega_1(\tau) = \omega_1^{\text{max}}(\tau) - \omega_1^{\text{max}}(\tau = 10\text{ns})$ ). The position of the peak increases by 8-12  $\text{cm}^{-1}$  for each system at  $T_0 = 45\text{ }^{\circ}\text{C}$ . The TZ2-UL spectrum at  $10\text{ }^{\circ}\text{C}$  does not shift appreciably with delay time, which is evidence that the  $\beta$ -sheet structure does not disrupt significantly at this temperature.

The time constant of the frequency shift cannot be fit rigorously with the limited data points. However, it is consistent with an unfolding time of slightly less than 1  $\mu\text{s}$ , and the decay appears to be non-exponential. Within the error of the experiment, this agrees with the time constants found in the T-jump DVE data. In each isotopologue, peak **A** shifts from  $\sim 1640$  to  $\sim 1650$   $\text{cm}^{-1}$ , which corresponds to spectral intensity moving from  $\nu_{\perp}$  to  $\nu_{\text{dis}}$ . Consequently, these shifts are assigned to peptide unfolding.

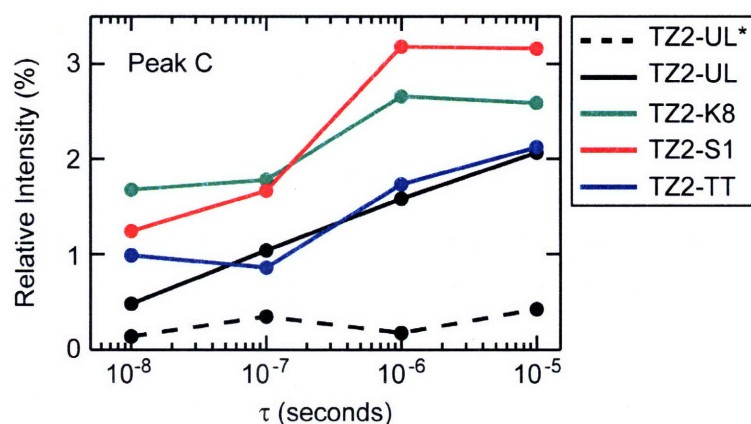


**Figure 8.10** (Top panel) The  $\omega_1$  shift of peak **A** maximum is shown as a function of delay time from 10 ns to 10  $\mu\text{s}$ . The  $\Delta\omega_1$  value is obtained by subtracting the peak max position at 10 ns from each of the following time points. (Bottom panel) Intensity of the peak **A** maximum relative to the equilibrated 2D IR amide I band maximum as a function of T-jump delay time.

For each TZ2 isotopologue, peak **B** represents the decrease of the  $^{12}\text{C}$  diagonal overtone peak. The frequency of peak **B** stays relatively constant at each value of  $t$ , and



the absolute intensity change matches that of peak A. Peak C is located on the red side of the diagonal  $\nu_{\perp}$  overtone band, and therefore reports directly on the loss of  $\beta$ -sheet structure. The frequency of peak C is relatively constant with  $\tau$ , and the time-dependent intensity (Fig. 8.11) is consistent with peak A. For the TZ2-UL data at  $T_0 = 10$  °C, the intensity is very low, and it does not grow substantially with  $\tau$ . This is further evidence that there is little disruption of  $\beta$ -sheet secondary structure at this starting temperature.

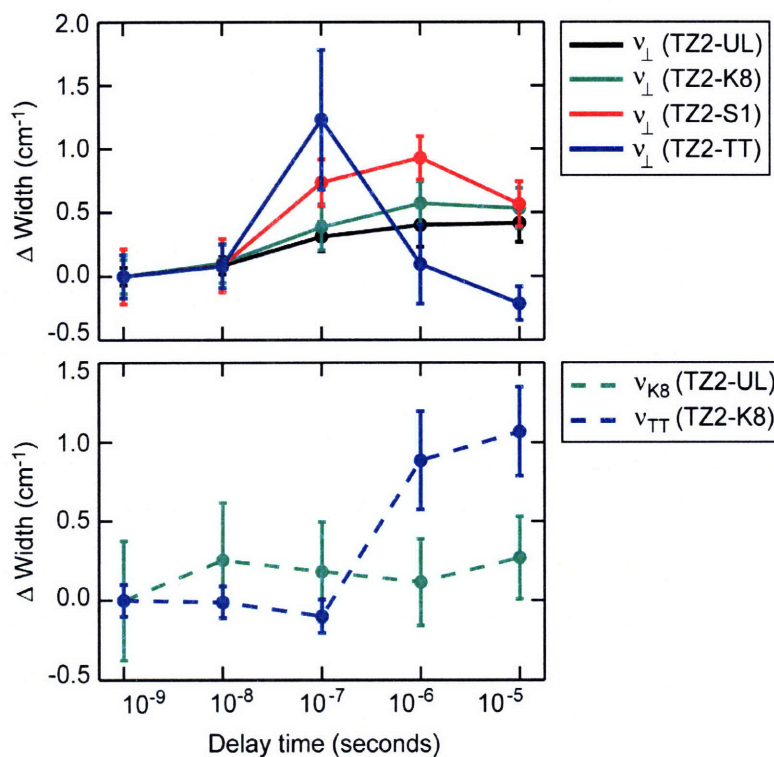


**Figure 8.11** Intensity of the peak C maximum relative to the equilibrated 2D IR amide I band maximum as a function of T-jump delay time from 10 ns to 10  $\mu$ s.

### 8.4.3. Rise time of the homogeneous line width

In Chapter 7, the anti-diagonal line width was interpreted as a probe of thermally-induced structural disorder. As the temperature increases, the width of the 2D IR peaks grows, which is assigned to increasing frequency fluctuations that are correlated on a picosecond time scale. Structural fluctuations consistent with this time scale are small orientational fluctuations of the T3 and T10 amides. Increasing fluctuations can be assigned to a higher degree of peptide backbone flexibility, or could be caused by increasingly solvated amide groups. The line broadening was found to be roughly linear

with temperature, with a fitted slope of 0.01 to 0.04  $\text{cm}^{-1}/^{\circ}\text{C}$ . In a T-jump experiment of  $\Delta T = 20 \text{ cm}^{-1}$ , the slopes predict line width changes on the order of 0.2 to 0.8  $\text{cm}^{-1}$ . The line width changes observed in the T-jump 2D IR spectra are roughly similar to these predicted values. As a function of T-jump delay time, the line broadening shows a sub-microsecond time scale. The fact that it is slower than the  $<10 \text{ ns}$  temperature increase is further evidence to assign these line width changes to increasing structural disorder. For TZ2-UL, TZ2-S1, and TZ2-K8, the  $\nu_{\perp}$  line width reaches the maximum value by 1  $\mu\text{s}$ , suggesting that the width increase occurs either faster than or at the same rate as the shift from  $\nu_{\perp}$  to  $\nu_{\text{dis}}$ .

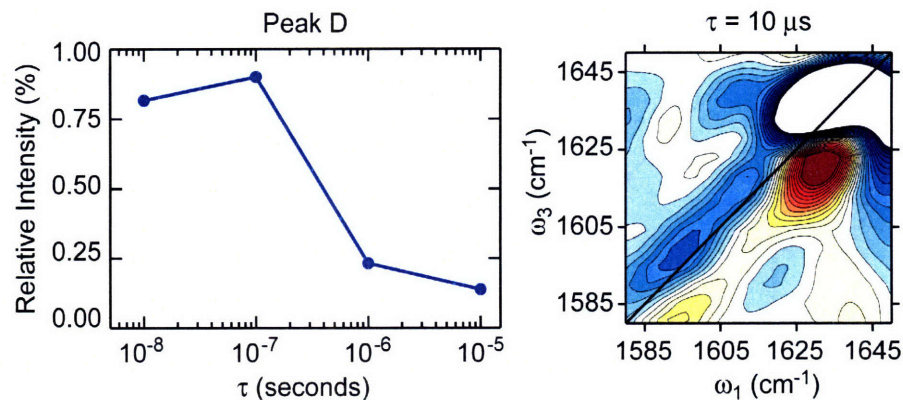


**Figure 8.12** (Top) Line width of the  $\nu_{\perp}$  band of each TZ2 isotopologue as a function of T-jump delay time. Each data point represents the  $\Gamma/2$  value of a Lorentzian function (Equation 7.1) fit to the anti-diagonal slices. The  $\Delta$ width is obtained by subtracting the width of the equilibrated spectrum at each time point. Error bars are the standard deviation of the fitted widths for six consecutive anti-diagonal slices.

#### 8.4.4. Isotope-shifted peaks in T-jump 2D IR

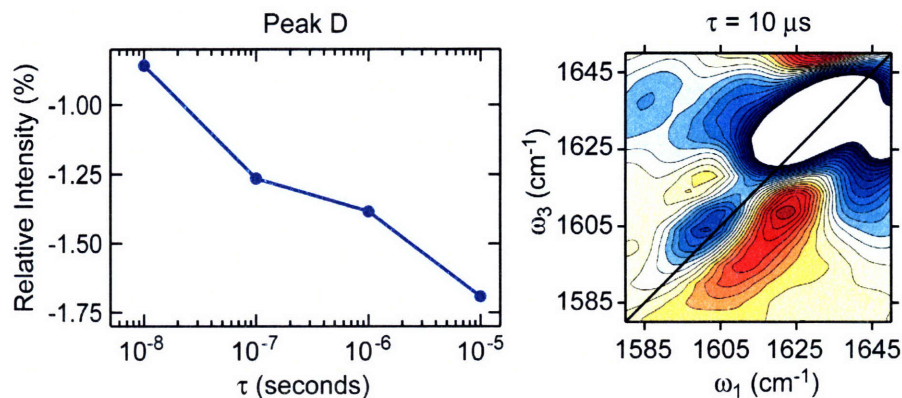
During the T-jump 2D IR experiment, the  $\nu_{K8}$  and  $\nu_{TT}$  bands undergo spectral shifts consistent with the equilibrium data in Chapter 7. The  $\nu_{K8}$  band decay is plotted in Fig. 8.13, which corresponds to peak **D** in Fig. 8.7. The intensity of the  $\nu_{K8-1}$  and  $\nu_{K8-2}$  bands decay on a  $\mu\text{s}$  time scale as expected from the T-jump DVE data. In the 2D difference spectrum at  $10 \mu\text{s}$  (Fig. 8.13), the loss of the diagonal peaks is observed at  $\omega_1 = 1595 \text{ cm}^{-1}$  and  $\omega_1 = 1612 \text{ cm}^{-1}$  corresponding to the two  $\nu_{K8}$  peaks. Also observable in the

difference spectrum is intensity loss in the  $\nu_{\perp}$ - $\nu_{K8-2}$  cross peak region at  $\omega_1 = 1612 \text{ cm}^{-1}$ ,  $\omega_3 = 1635 \text{ cm}^{-1}$ . This loss of cross peak intensity is direct evidence for the loss of coupling between the  $\nu_{K8-2}$  vibration corresponding to the internally hydrogen bonded carbonyl and the  $^{12}\text{C}$   $\nu_{\perp}$  vibrations. Vibrational coupling between the amide groups is based on their relative distance and orientation. As the native orientation is disrupted during thermal unfolding, the coupling decreases and the intensities of the isotope-shifted peaks decrease. The T-jump 2D IR data does not directly distinguish whether the loss of coupling is due to changes in the configuration of K8, the remaining  $^{12}\text{C}$  units, or both. However, based on the temperature-independent width of the  $\nu_{K8-2}$  peak, it is most likely that this loss of coupling reflects increasing disorder of the remaining  $^{12}\text{C}$  peptide groups. Finally, no significant cross peak intensity loss is observed between the  $\nu_{K8-1}$  and  $\nu_{\perp}$  bands. This is consistent with the disordered nature of the K8 amides contributing to the  $\nu_{K8-1}$  peak, for which the vibrational coupling to other amides is expected to be low.



**Figure 8.13** (Left) Intensity of peak **D** as a function of delay time (relative to equilibrated 2D IR maximum), which corresponds to the intensity change of the  $\nu_{\text{K8-2}}$  band. (Right) The  $\nu_{\text{K8}}$  band region of the T-jump difference 2D IR spectrum at  $t = 10 \mu\text{s}$  is plotted with 27 contours from  $\pm 65\%$  of the spectral region maximum.

The intensity of the  $\nu_{\text{TT}}$  peak is plotted versus  $\tau$  in Fig. 8.14. The decay time of the peak is roughly  $1 \mu\text{s}$ , which is consistent with the time constants observed above. In the 2D IR difference spectrum at  $10 \mu\text{s}$ , the negative peak at  $\omega_1 = 1602 \text{ cm}^{-1}$  reflects the loss of  $\nu_{\text{TT}}$  intensity as a function of time. The  $\nu_{\text{TT}}$  difference peak also displays significant positive intensity on each side of the negative peak. This positive intensity represents the broadening of the peak in the anti-diagonal dimension. The significant broadening of the  $\nu_{\text{TT}}$  band during the transient unfolding is evidence that the unfolding mechanism of the T3 and T10 amides is primarily described as increasing structural fluctuations, which is assigned to greater backbone flexibility and increasing interactions with solvent. As for TZ2-K8, the 2D IR difference spectrum shows intensity loss in the cross peak region between the isotope labeled band and the  $\nu_{\perp}$  band ( $\omega_1 = 1602 \text{ cm}^{-1}$ ,  $\omega_3 = 1635 \text{ cm}^{-1}$ ). In the case of TZ2-TT, this loss of cross peak intensity is assigned to disordering of the labeled and unlabeled amide groups.



**Figure 8.14** (Left) Intensity of peak **D** as a function of delay time (relative to equilibrated 2D IR maximum), which corresponds to the intensity change of the  $\nu_{TT}$  band. (Right) The  $\nu_{TT}$  band region of the T-jump difference 2D IR spectrum at  $t = 10 \mu\text{s}$  is plotted with 27 contours from  $\pm 65\%$  of the spectral region maximum.

## 8.5. Conclusion

In this chapter, the folding kinetics of TZ2 have been investigated with T-jump DVE and 2D IR spectroscopy. Temperature-jump DVE spectroscopy was used primarily to assign rates to the transient spectral change. At  $45^\circ\text{C}$ , the decay of the TZ2-UL  $\nu_{\perp}$  band was measured to be  $0.97 \pm 0.12 \mu\text{s}$ , which is assigned to the disordering of  $\beta$ -sheet secondary structure. The rise time of  $\nu_{\text{dis}}$  is found to be  $0.90 \pm 0.07 \mu\text{s}$ . For TZ2-S1, no band could be resolved to assign a folding time to the N-terminus. The TZ2-K8  $\nu_{K8}$  band decays with a  $1.08 \pm 0.58 \mu\text{s}$  time scale, which is assigned to intensity loss as the K8 amide decouples from the  $^{12}\text{C}$  amide modes. Finally, the TZ2-TT  $\nu_{TT}$  band intensity decays with a  $0.70 \pm 0.09 \mu\text{s}$  time scale compared to the  $0.82 \pm 0.12 \mu\text{s}$  TZ2-TT  $\nu_{\perp}$  band decay. Each of these measured time scales are very similar, ranging from 0.7 to  $1.08 \mu\text{s}$ .

This observation of similar folding times across the band is in contrast to findings of Yang et. al., in which probe dependant kinetics were observed in the unfolding of TZZ.<sup>10</sup> Using fluorescence spectroscopy, they observed that the frequency dependence of the transient fluorescence signal grows with increasing temperature. This folding heterogeneity is attributed to a growing population of peptides with a partially denatured hydrophobic core, so that one side of the band probes structures already near the transition state. The major difference between those observations and the data shown in this chapter is that the amide I band is a backbone carbonyl stretch that is sensitive to secondary structure and amide hydrogen bonding. Therefore, it is possible to have folding heterogeneity growing in with temperature at the hydrophobic core, while simultaneously observing consistent folding times across the amide I band. The  $T_0$ -dependent data shown in this chapter do agree with those of an alanine-substituted trpzip peptide, where they observed independent kinetic rates for different features in the amide I band at low temperatures, and a convergence of these rates with increasing temperature.<sup>9</sup>

The fact that the T-jump DVE kinetic rates are similar allows for the possibility that there is a common transition state for each of the structural events giving rise to the DVE changes. For example, the nanosecond phase in the T-jump fluorescence study of Yang et al. was assigned to loosening of the tryptophan packing.<sup>10</sup> It is possible that the backbone disordering giving rise to the amide I spectral changes is dependant on the increased tryptophan mobility. In addition, the  $\mu\text{s}$  phase of the fluorescence kinetics is slower than the rates measured with T-jump DVE spectroscopy, suggesting that the backbone disordering ( $\sim 1 \mu\text{s}$ ) is faster than the breaking up of the tryptophan core (3 to 5

$\mu\text{s}$  at 45 °C).<sup>10</sup> These comparisons lead to a picture in which the unfolding of the backbone and the hydrophobic core are intimately connected. A similar level of cooperativity between the core and the backbone contacts has been observed in molecular dynamics simulations.<sup>11,12</sup>

Temperature-jump 2D IR spectroscopy is used to explore the thermal unfolding transition beyond the kinetic rates discussed above. For the peptides shown here we observe transient disordering events as diagonal peak broadening, diagonal peak intensity shifts, and intensity loss in the off-diagonal cross peaks. In the transient 2D IR data, the  $\nu_{\perp}$  band of each system undergoes anti-diagonal line broadening on the order of 0.5-1.0  $\text{cm}^{-1}$ , which is consistent with the equilibrium temperature-dependent 2D IR spectra. The time scale of this broadening is  $\sim 1 \mu\text{s}$ , which is the rate at which the peptide becomes more dynamically disordered. The line width of the  $\nu_{\text{K8}}$  peak corresponding to native peptide turns does not increase during the T-jump experiment. This is evidence that the turn region remains relatively unperturbed throughout the unfolding transition. In contrast, the line width of the  $\nu_{\text{TT}}$  band increases substantially, demonstrating that the unfolding of the mid-strand region can be described in part as increasingly fast structural fluctuations. For both TZ2-K8 and TZ2-TT, cross peak intensity between the  $\nu_{\perp}$  band and the isotope-shifted band is observed to decay with the same rate as the other spectral features. Loss of cross-peak intensity signifies structural rearrangements that decouple the isotope-labeled amide groups from the rest of the peptide. Based on the conclusions from the diagonal peaks and the other spectral observations, the decay of the  $\nu_{\text{K8}} - \nu_{\perp}$  cross peak is interpreted as disordering toward the fraying ends of the peptide, while the



$\nu_{\text{TT}} - \nu_{\perp}$  cross peak decay is interpreted as disordering in both the  $^{12}\text{C}$  and  $^{13}\text{C}$  amide groups.

Using the data shown in this chapter and in Chapter 7, it is possible to describe specific features of the thermally induced unfolding transition. At equilibrium, TZ2 is slightly disordered at the mid-strand region of the peptide, with the  $\nu_{\text{TT}}$  peak showing signs of homogeneous and inhomogeneous line broadening. The N-terminus is also disordered, and is presumed to sample a larger range of configurations than the mid-strand threonine residues. There are also at least two structural ensembles with unique turn geometries: One in a native, Type I'  $\beta$ -turn, and one in which the K8 amide group is reversed and solvent exposed. During the transient unfolding, the ensemble of peptides with native turns ( $\nu_{\text{K8-2}}$ ) unfolds by increasing the structural disorder about the T3 and T10 residues. The disorder at the N-terminus also presumed to increase, although it cannot be resolved directly. Each of these processes occurs on a  $\sim 1 \mu\text{s}$  time scale, as evidenced in the  $\nu_{\perp}$  peak shifts and the anti-diagonal line-broadening. The  $\nu_{\perp} - \nu_{\text{TT}}$  cross peak intensity drops perhaps as a result of the larger average difference between the two  $\beta$ -strands, which weakens the coupling of the threonine amides to the rest of the peptide. Cross peak intensity between the  $\nu_{\perp}$  and  $\nu_{\text{K8-2}}$  peaks also drops due to the larger average distance to the opposing  $\beta$ -strand. The ensemble of peptides with non-native turns ( $\nu_{\text{K8-1}}$ ) unfolds in a similar way, except that the  $\nu_{\text{K8-1}}$  peak melting curve (observed by FTIR) has a transition temperature less-well-defined than the melting curves of the  $\nu_{\text{K8-1}}$  peak and  $\beta$ -sheet modes. This suggests that the peptides with non-native turns are more disordered to begin with and that the structural transition is less-well defined than for the native turn

ensemble. Evidence of a heterogeneous TZ2 free energy surface has been observed in previous simulations<sup>8,13,14</sup> and experiments.<sup>9,10,14</sup>

These observations of the transient unfolding of TZ2 at  $T_0 = 45$  °C can be rationalized with a mechanism that is heterogeneous in the sense that the disordering transition goes from a small set of structures to a larger set of structures through a distribution of folding pathways. In this study we directly observe the global unfolding of the  $\beta$ -sheet secondary in the  $\nu_{\perp}$  band, and the unfolding of the turn region ( $\nu_{K8}$ ) and mid-strand region ( $\nu_{TT}$ ) of the hairpin. The unfolding pathways have been described above, and do not reveal a time ordering between turn and mid-strand disordering. However, because the natively structured  $\beta$ -turn is not observed to disorder significantly, the unfolding transition can be described as hairpin fraying for the ensemble of peptides with a native turn. A comparison to the rates observed in T-jump fluorescence<sup>10</sup> reveals that the backbone unfolding events are intermediate between the loosening of the tryptophan packing and the breakup of the hydrophobic core. This is consistent with a folding pathway that is a hybrid zipper. Through an assumed microscopic reversibility, our data support the following folding mechanism. First, the turn nucleates, after which the tryptophan side chains form loose tertiary contacts. Next, the peptide backbone folds to the native configuration, following which the hydrophobic core reaches its optimal packing geometry. A similar hybrid zipper mechanism has been observed in simulations of GB1<sub>41-56</sub> folding<sup>3,11</sup> as well as for TZ2 folding simulations.<sup>8</sup>

The mechanism described here is supported by the kinetics of the terminal region of TZ2C observed with isotopes labels and T-jump amide I absorbance.<sup>9</sup> However, the kinetics of the turn region isotope-shifted peak in that work were observed to be

intermediate between terminus and the mid-strand region. This can be rationalized, however, by recognizing that the 2D IR spectra show most of the spectral loss of the turn region peak is actually due to decreasing vibrational coupling to the mid-strand and terminal amides. Furthermore, the turn-region peak in that study is a dual label that will exacerbate the effect of mid-strand disorder on the isotope-shifted peak. This highlights the need for 2D IR spectroscopy to describe the sources of spectral loss that is ambiguous in FTIR: vibrational coupling and structural heterogeneity observed in the 2D line widths. One further comparison is with the unfolding of PG12, which was discussed in Chapter 6. The faster disordering of the PG12 V3 amide group (140 ns) can now be better understood by observing that the hydrophobic side chain contacts of PG12 are much less stable than the four-tryptophan motif of TZ2. Therefore, the backbone disordering rate is not restricted by the tertiary hydrophobic contacts.

In order to completely verify the proposed folding mechanism it is necessary to develop a larger set of TZ2 isotopologues. Specifically, the residues around the turn region should be individually labeled to identify the extent of the region affected by the two K8 amide conformations. The W2 and G7 peptide groups have been labeled previously, and no evidence was seen in the diagonal peaks of the 2D IR amide I band of distinct configurational states.<sup>15</sup> The folding thermodynamics and kinetics of the W2 and G7 isotopologues have not been characterized, but the G7 group has been observed to display distinct solvation states at long waiting times in the 2D IR amide I spectrum.<sup>16</sup> For the isotope labels that do not give rise to a resolvable peak (e.g. S1), dual  $^{13}\text{C}=\text{}^{18}\text{O}$  labels should be used so that the spectral shift is high enough to produce a peak free from interference with the unlabeled amide I band.<sup>17</sup> The advantage of the techniques shown

in this chapter and in Chapter 7 is that they are generalizable to other systems, including small proteins and peptides with mixed-structure domains.<sup>18-20</sup>

## **8.6. Acknowledgements**

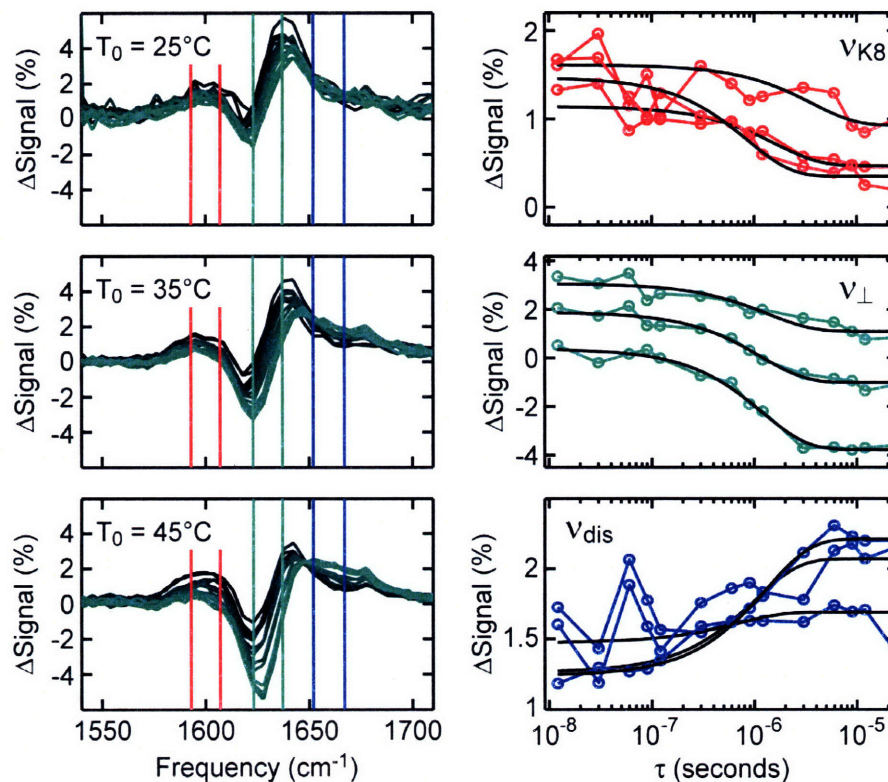
I would like to thank William Swope, John Chodera, and Jed Pitera for their cooperation in providing structures for the TZ2 Markov states. Kevin Jones and Benjamin Dietzek provided regular assistance in the laser lab during the collection of the 2D IR and T-jump experiments. Finally, Joshua Lessing designed and conducted the synthesis for <sup>18</sup>O labeled TZ2-S1, and also conducted the synthesis of TZ2-K8. He also measured the temperature-dependant FTIR spectra shown in Chapter 7.

## 8.7. References

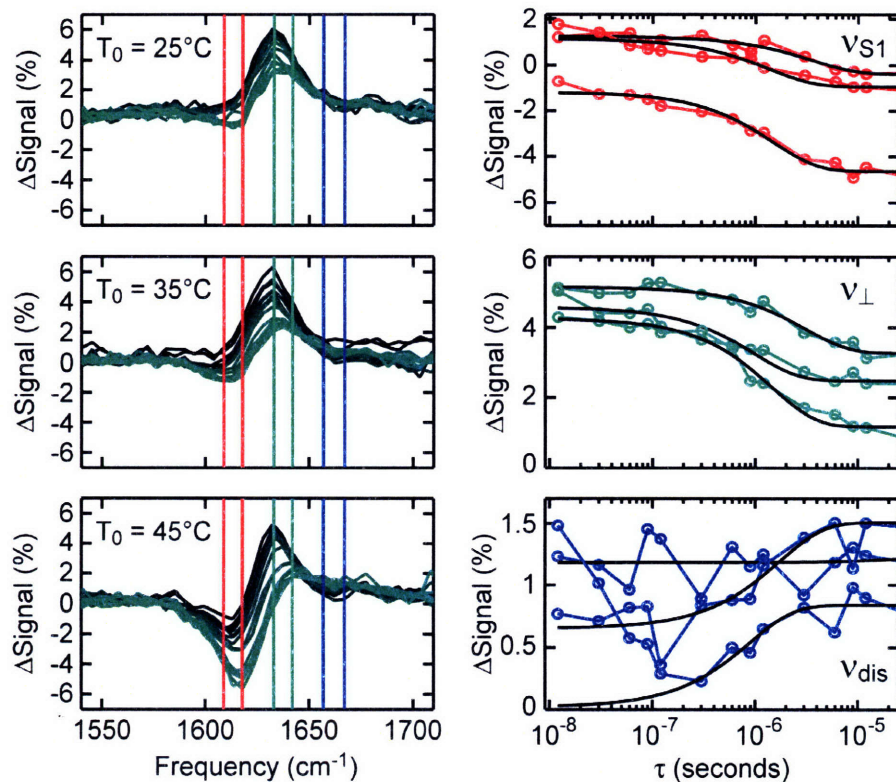
- (1) Pitera, J. W.; Haque, I.; Swope, W. C. *J. Chem. Phys.* 2006, 124, 141102.
- (2) Bolhuis, P. G. *Biophys. J.* 2005, 88, 50-61.
- (3) Wei, G.; Mousseau, N.; Derremaux, P. *Proteins* 2004, 56, 464-474.
- (4) Dyer, R. B.; Maness, S. J.; Peterson, E. S.; Franzen, S.; Fesinmeyer, R. M.; Andersen, N. H. *Biochemistry* 2004, 43, 11560-11566.
- (5) Bolhuis, P. G. *Proc. Nat. Acad. Sci., USA* 2003, 100, 12129-12134.
- (6) Chung, H. S.; Khalil, M.; Smith, A. W.; Ganim, Z.; Tokmakoff, A. *Proc. Natl. Acad. Sci. USA* 2005, 102, 612-617.
- (7) Chung, H. S.; Khalil, M.; Tokmakoff, A. *J. Phys. Chem. B* 2004, 108, 15332-15343.
- (8) Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. *Proc. Natl. Acad. Sci. USA* 2004, 101, 4077-4082.
- (9) Hauser, K.; Krejtschi, C.; Huang, R.; Wu, L.; Keiderling, T. A. *J. Am. Chem. Soc.* 2008, In Press: ja0742151.R2.
- (10) Yang, W. Y.; Gruebele, M. *J. Am. Chem. Soc.* 2004, 126, 7758-7759.
- (11) Tsai, J.; Levitt, M. *Biophysical Chemistry* 2002, 101, 187-201.
- (12) Zhou, R.; Berne, B. J.; Germain, R. *Proceedings of the National Academy of Sciences* 2001, 98, 14931-14936.
- (13) Chodera, J. D.; Singhal, N.; Pande, V. S.; Dill, K. A.; Swope, W. C. *J. Chem. Phys.* 2007, 126, 155101.

- (14) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. J. *Mol. Biol.* 2004, 336, 241-251.
- (15) Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. *J. Phys. Chem. B* 2006, 110, 7545-7555.
- (16) Kim, Y. S.; Hochstrasser, R. M. *J. Phys. Chem. B* 2007, 111, 9697-9701.
- (17) Jaume Torres, A. K., Jonathan M. Goodman, Isaiah T. Arkin, *Biopolymers* 2001, 59, 396-401.
- (18) Fang, C.; Senes, A.; Cristian, L.; DeGrado, W. F.; Hochstrasser, R. M. *Proceedings of the National Academy of Sciences* 2006, 103, 16740-16745.
- (19) Fang, C.; Wang, J.; Kim, Y. S.; Charnley, A. K.; Barber-Armstrong, W.; Smith, A. B., III; Decatur, S. M.; Hochstrasser, R. M. *J. Phys. Chem. B* 2004, 108, 10415-10427.
- (20) Brewer, S. H.; Song, B.; Raleigh, D. P.; Dyer, R. B. *Biochemistry* 2007, 46, 3279-3285.

## 8.8. Appendix

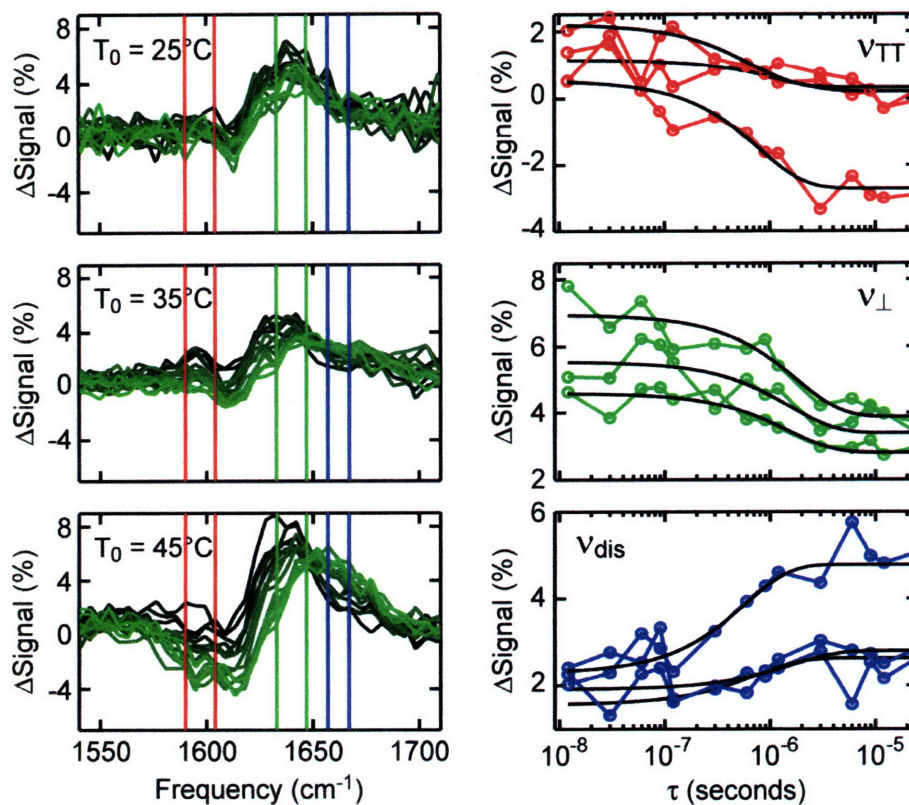


**Figure 8.a.1** (Left) Transient difference DVE spectra of TZ2-K8 shown from 1 ns to 3  $\mu\text{s}$  at 25, 35, and 45  $^\circ\text{C}$ . Difference spectra are obtained as described in Section 8.2, with the  $\Delta\text{Signal}$  defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu\text{s}$ . Time constants from the transient fits are shown in Table 8.2.



**Figure 8.a.2** (Left) Transient difference DVE spectra of TZ2-UL shown from 1 ns to 3  $\mu\text{s}$  at 25, 35, and 45  $^\circ\text{C}$ . Difference spectra are obtained as described in Section 8.2, with the  $\Delta\text{Signal}$  defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu\text{s}$ . Time constants from the transient fits are shown in Table 8.2.





**Figure 8.a.3** (Left) Transient difference DVE spectra of TZ2-UL shown from 1 ns to 3  $\mu\text{s}$  at 25, 35, and 45  $^\circ\text{C}$ . Difference spectra are obtained as described in Section 8.2, with the  $\Delta\text{Signal}$  defined as the difference DVE spectrum normalized by the equilibrium band maximum. (Right) Pixels between each similarly-colored line are binned and averaged to produce the transient DVE traces. Transient traces are displayed as closed circles connected by solid colored lines. Black lines represent single-exponential fits to the data from 12 ns to 12  $\mu\text{s}$ . Time constants from the transient fits are shown in Table 8.2.

# Adam W. Smith

261 Columbia St. · Cambridge MA 02139 · Phone (617)-661-3674 · E-mail [awsmith@mit.edu](mailto:awsmith@mit.edu)

---

- Education**    **Massachusetts Institute of Technology** (Cambridge, MA)  
**Ph.D. degree in Physical Chemistry**, Expected August, 2007. (g.p.a 4.5/5.0)  
**Thesis:** Structure and folding pathways of  $\beta$ -hairpin peptides. **Advisor:** Andrei Tokmakoff
- University of Utah** (Salt Lake City, UT)  
**Honors Degree B.S.**, 2002. (g.p.a. 3.92/4.0; *cum laude*)  
Delivered student address at the University of Utah, College of Science Commencement  
**Honors Thesis:** Optical Spectroscopy of Diatomic Tungsten Carbide. **Advisor:** Michael D. Morse
- Research**    **Massachusetts Institute of Technology** (Cambridge, MA)  
**Advisor: Andrei Tokmakoff**  
**Thesis Title: Probing the Unfolding Transition of  $\beta$ -Hairpin Peptides with Nonlinear Infrared Spectroscopy**
- Conducted solid-phase peptide synthesis including residue-specific isotope labeling.
  - Experienced in a variety of time-resolved, non-linear optical methods including pump-probe spectroscopy, photon echo spectroscopy, and heterodyne-detected 2D IR.
  - Experienced in a variety of equilibrium spectroscopic methods including Raman microscopy, FTIR, UV-CD, fluorescence, and UV-visible absorbance
  - Designed and built a femtosecond regenerative amplifier.
  - Developed and helped build an apparatus to conduct spectroscopic measurements following a nanosecond, laser-induced, temperature jump.
  - Programming skills include MatLab, C++, Visual Basic, Perl, and LabView.
- University of Utah** (Salt Lake City, UT)  
**Advisor: Michael D. Morse**
- Used high-resolution laser spectroscopy to study transition metal carbides.
  - Worked extensively with high vacuum technology
  - Skilled in UV and visible narrow-band laser systems.
- Teaching**    **Tutoring Plus** (Cambridge, MA)  
**Volunteer High School Tutor/Mentor**, 2005 – 2007.  
Spent one evening a week with an economically disadvantaged high school student, primarily to help with homework. Time is also spent developing a friendship and teaching life skills.
- Massachusetts Institute of Technology** (Cambridge, MA)  
**General Chemistry Teaching Assistant**, 2002 – 2003.  
Led a 2hr/week recitation section two semesters. Presented the information found in lectures and textbooks in a way that the students could both conceptualize it and apply it to real problems.

- Awards**
- Morse Travel Grant** (2006) Competitive travel award for students in the MIT Chemistry Department.
  - Lester Wolfe Graduate Fellowship** (2004) awarded to 5-6 Ph.D. students studying spectroscopy at MIT.
  - College of Science Scholarship** full tuition award given to six outstanding undergraduates in the U. of Utah College of Science (2001-02).
  - Departmental Awards** at the University of Utah given to top 5-10 students in each academic year for excellence in chemistry: **Leon Watters Memorial Award** (2002), **Hypercube Scholar** (2002), **Mack Thomas Rozelle Scholarship** (2001), and the **Loyd E. Malm Memorial Award** (2000).
- Publications**
- Ganim, Z.; Chung, H. S.; Smith, A. W.; DeFlores, L. P.; Jones, K. C.; Tokmakoff, A. **Amide I Two-Dimensional Infrared Spectroscopy of Proteins.** *Acc. Chem. Res.* **2008**, (In Press).
  - Smith, A. W., Tokmakoff, A. **Probing local structural events in  $\beta$ -hairpin unfolding with transient nonlinear infrared spectroscopy.** *Angewandte Chemie International Edition* **2007**, 46, 7984.
  - Chung, H. S., Khalil, M., Smith, A. W. & Tokmakoff, A. **Transient 2D IR spectrometer for probing nanosecond temperature-jump kinetics.** *Review of Scientific Instruments* **2007**, 78, 063101.
  - Ganim, Z. Chung, H. S., Smith, A. W., DeFlores, L. P., Jones, K. C. & Tokmakoff, A. **Amide I two-dimensional infrared spectroscopy of proteins.** *Accounts of Chemical Research* (In press)
  - Smith, A. W.; Tokmakoff, A.. **Amide I Two-Dimensional Infrared Spectroscopy of  $\beta$ -Hairpin Peptides** *Journal of Chemical Physics* **2007**, 126, 045109.
  - Smith, A. W., Chung, H. S., Ganim, Z., Tokmakoff, A. **Multidimensional IR Spectroscopy of Site-Specific Hairpin Folding.** *Proceedings of the 15th International Conference on Ultrafast Phenomena, Pacific Grove, CA* **2006**.
  - Smith, A. W.; Chung, H. S.; Ganim, Z.; Tokmakoff, A.; **Residual native structure in a thermally denatured beta-hairpin.** *Journal of Physical Chemistry B* **2005**, 109, 17025.
  - Chung, H. S.; Khalil, M.; Smith, A. W.; Ganim, Z.; Tokmakoff, A.; **Conformational changes during the nanosecond-to-millisecond unfolding of ubiquitin.** *Proceedings of the National Academy of Sciences of the United States of America* **2005**, 102, 612.
  - Sickafoose, S. M.; Smith, A. W.; Morse, M. D.; **Optical spectroscopy of tungsten carbide (WC).** *Journal of Chemical Physics* **2002**, 116, 993.