



Publicly Accessible Penn Dissertations


1-1-2014

Infrared Measurements of Protein Conformational Dynamics

Robert Culik

University of Pennsylvania, rculik@mail.med.upenn.edu

Follow this and additional works at: <http://repository.upenn.edu/edissertations>

 Part of the [Biochemistry Commons](#), [Biophysics Commons](#), and the [Physical Chemistry Commons](#)

Recommended Citation

Culik, Robert, "Infrared Measurements of Protein Conformational Dynamics" (2014). *Publicly Accessible Penn Dissertations*. 1248.
<http://repository.upenn.edu/edissertations/1248>

This paper is posted at Scholarly Commons. <http://repository.upenn.edu/edissertations/1248>
For more information, please contact libraryrepository@pobox.upenn.edu.

Infrared Measurements of Protein Conformational Dynamics

Abstract

The topic of how a protein folds has been a major area of research for several decades; however, important details about this process are still undetermined. Experimental limitations in the study of protein folding are a result of no technique possessing both the necessary spatial and temporal resolution. This Thesis presents several studies conducted with the goal of expanding upon the experimentalist's toolbox, involving new methods of interrogating and/or perturbing protein systems of interest.

The early chapters of this Thesis describe our efforts using established synthetic methods to extend the utility of infrared spectroscopy in the study of protein folding. Specifically, we show that, using the strategy of cysteine alkylation, we can incorporate novel vibrational probes into proteins in a site-specific manner. We also show that combined sidechain mutagenesis, probing at multiple frequencies, and isotopic labeling to obtain secondary structural resolution in infrared studies of protein folding, in the process uncovering details about the folding mechanism of the Trp-cage miniprotein. Similarly, we illustrated the use of thioamides as site-specific reporters of backbone-backbone hydrogen bonding, and applied this functionalization to the Trpzip2 beta-hairpin system to validate its proposed folding mechanism. Further work involved using D-amino acids to interrogate turn regions in proteins, specifically examining Trp-cage folding.

The later chapters of this Thesis are focused on the effects of extrinsic molecules on the structural ordering of proteins. Taking advantage of the lack of tertiary structure of intrinsically disordered proteins, we examined the effect that trifluoroethanol has on protein folding, and found evidence that this cosolvent acts as a nanocrowder. We also introduced the idea of using phototriggered modifications to modify the free energy landscape of folding for a given protein, and demonstrated that a peptide that typically folds in an activated (barrier-containing) manner can be made to fold in a downhill fashion upon irradiation.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Biochemistry & Molecular Biophysics

First Advisor

Feng Gai

Keywords

Dynamics, Infrared, Protein Folding, Spectroscopy

Subject Categories

Biochemistry | Biophysics | Physical Chemistry

INFRARED MEASUREMENTS OF PROTEIN CONFORMATIONAL DYNAMICS

Robert M. Culik

A DISSERTATION

in

Biochemistry and Molecular Biophysics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

Graduate Group Chairperson

Feng Gai

Professor of Chemistry

Kathryn M. Ferguson

Associate Professor of Physiology

Dissertation Committee

S. Walter Englander (Chair)

Jacob Gershon-Cohen Professor of Biochemistry, Biophysics, & Medical Science

Jeffery G. Saven

Associate Professor of Chemistry

Sergei A. Vinogradov

Associate Professor of Biochemistry and Biophysics

Mark Goulian

Edmund J. and Louise W. Kahn Professor of Biology, Physics & Astronomy

Kim A. Sharp

Associate Professor of Biochemistry and Biophysics

Scott H. Brewer

Associate Professor of Chemistry

INFRARED MEASUREMENTS OF PROTEIN CONFORMATIONAL DYNAMICS

COPYRIGHT

2014

Robert Mitchell Culik

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

To my grandfathers, Harold Edward Bellis and Rudolf Wenceslaus Čulík

ACKNOWLEDGMENTS

It is impossible to begin this Thesis without acknowledging the great swathe of people who have helped me get to this point. The most prominent influence in my scientific career thus far has been my thesis advisor, Dr. Feng Gai, whose passion, patience, and ceaseless scientific curiosity have been a great inspiration for me. I am incredibly thankful for his encouragement, his mentorship, and his ability to always believe in me. I also would be remiss not to thank all of the members of my thesis committee, Professors Walter Englander, Jeffery Saven, Sergei Vinogradov, and Mark Goulian, for their useful questions and comments at meetings, and for their efforts to make me realize my full potential. I owe an extra debt of appreciation to Drs. Englander and Saven for their efforts in helping me to receive external funding, and to Dr. Goulian for agreeing to be on my committee after a former committee member left the university. I would also like to thank Dr. Kim Sharp and Dr. Scott Brewer for agreeing to be on my final defense committee; I'd especially emphasize my gratitude to Dr. Brewer, who made the trip to Philadelphia for my defense.

My attempts here at Penn would have been for naught if not for the helpful and supportive community that I have had the good fortune to land in. I am especially grateful to several people for helping to mold me into a better scientist and person. Dr. Arnaldo Serrano has been a mentor, peer, and friend who never flinched at giving me helpful advice or insight, even if the response was something I did not want to hear. Kyle Harpole has been a constant friend and fount of knowledge since the time we first interviewed at grad schools together, and I'm sure without him in my life I'd be a much

blander person. I spent many a night chatting with Dr. Jacob Goldberg, whose unique perspective and strong work ethic left a deep impression on me. I also need to thank Dr. Bob Rarig, who I've shared countless experiences with, none of which haven't contained a hearty chuckle or two. My labmates, both past and present, in the Gai group have been an incredible support net throughout the years, and the camaraderie we have had has been awesome. Thanks very much to Dr. Primit Chowdhury, Dr. Smita Mukherjee, Dr. Julie Rodgers, Dr. Matthias Waegele, Dr. Lin Guo, Dr. Kate Smith-Dupont, Dr. Tom Troxler, Dr. Jian-Xin Chen, Dr. Lev Chuntonov, Dr. Ayanjeet Ghosh, Dr. Jianqiang Ma, Dr. Kwang-Im Oh, Lali Pazos, Chun-Wei Lin, Yi-Ju Chen, Beatrice Markiewicz, Paul Straus, Mary Rose Mintzer, Rachel Roesch, Jeff Rodgers, and Ismail Ahmed. I'd also like to thank the foosball crowd for the stress relief. In addition, I've had many collaborators over the years - I am grateful for their efforts and their desire to work with me. Thanks to Dr. Hyunil Jo for his hard work and his incredibly fast turnaround time and to his advisor, Dr. Bill DeGrado. I'd also like to thank Dr. Michelle Bunagan, Dr. Srinivas Annavarapu, and Dr. Vikas Nanda. I am also grateful to the staffs of the Chemistry and BMB Departments, and to the Chemistry Department as a whole for 'adopting' me and making me feel at home. Finally, I would like to thank my family for all of their love, support, and understanding over these years. My mom, dad, sister, extended family, and girlfriend Nicole have suffered a considerable amount of neglect over the years, and by their grace they still choose to associate with me. All of them have been bedrock for me, and helped me to bounce back time and again in my efforts presented here.

ABSTRACT

INFRARED MEASUREMENTS OF PROTEIN CONFORMATIONAL DYNAMICS

Robert M. Culik

Feng Gai

The topic of how a protein folds has been a major area of research for several decades; however, important details about this process are still undetermined. Experimental limitations in the study of protein folding are a result of no technique possessing both the necessary spatial and temporal resolution. This Thesis presents several studies conducted with the goal of expanding upon the experimentalist's toolbox, involving new methods of interrogating and/or perturbing protein systems of interest.

The early chapters of this Thesis describe our efforts using established synthetic methods to extend the utility of infrared spectroscopy in the study of protein folding. Specifically, we show that, using the strategy of cysteine alkylation, we can incorporate novel vibrational probes into proteins in a site-specific manner. We also show that combined sidechain mutagenesis, probing at multiple frequencies, and isotopic labeling to obtain secondary structural resolution in infrared studies of protein folding, in the process uncovering details about the folding mechanism of the Trp-cage miniprotein. Similarly, we illustrated the use of thioamides as site-specific reporters of backbone-backbone hydrogen bonding, and applied this functionalization to the Trpzip2 beta-hairpin system to validate its proposed folding mechanism. Further work involved using

D-amino acids to interrogate turn regions in proteins, specifically examining Trp-cage folding.

The later chapters of this Thesis are focused on the effects of extrinsic molecules on the structural ordering of proteins. Taking advantage of the lack of tertiary structure of intrinsically disordered proteins, we examined the effect that trifluoroethanol has on protein folding, and found evidence that this cosolvent acts as a nano-crowder. We also introduced the idea of using phototriggers to modify the free energy landscape of folding for a given protein, and demonstrated that a peptide that typically folds in an activated (barrier-containing) manner can be made to fold in a downhill fashion upon irradiation.

TABLE OF CONTENTS

1 Introduction

1.1 Protein Folding.....	2
--------------------------	---

2 Theory

2.1 Protein Folding Thermodynamics.....	12
---	----

2.2 Protein Folding Kinetics	13
------------------------------------	----

3 Methods

3.1 Circular Dichroism Spectroscopy	19
---	----

3.2 Fourier-Transform Infrared Spectroscopy	21
---	----

3.3 Laser-Induced Temperature-Jump Spectroscopy	22
---	----

3.4 Infrared Spectroscopy of Proteins.....	23
--	----

3.5 Using Phototriggers to Initiate Conformational Changes in Proteins.....	25
---	----

4 Selective Incorporation of Nitrile-Based Infrared Probes into Proteins via

Cysteine Alkylation

4.1 Introduction.....	30
-----------------------	----

4.2 Experimental Section	30
--------------------------------	----

4.3 Results and Discussion	33
----------------------------------	----

4.4 Conclusion	35
----------------------	----

4.5 Original Publication.....	36
-------------------------------	----

5 Achieving Secondary Structural Resolution in Kinetic Measurements of Protein Folding: A Case Study of the Folding Mechanism of Trp-cage

5.1 Introduction.....	44
5.2 Experimental Section.....	46
5.3 Results and Discussion	47
5.4 Conclusion	51
5.5 Original Publication	52

6 Using Thioamides to Site-Specifically Interrogate the Dynamics of Hydrogen Bond Formation in β -Sheet Folding

6.1 Introduction.....	66
6.2 Experimental Section.....	68
6.3 Results and Discussion	71
6.4 Conclusion	75
6.5 Original Publication.....	76

7 Using D-Amino Acids to Delineate the Mechanism of Protein Folding: Application to Trp-Cage

7.1 Introduction.....	85
7.2 Experimental Section.....	87
7.3 Results and Discussion	88
7.4 Conclusion	91

7.5 Original Publication	93
8 Experimental Validation of the Role of Trifluoroethanol as a Nano-Crowder	
8.1 Introduction.....	100
8.2 Experimental Section.....	103
8.3 Results and Discussion	103
8.4 Conclusion	108
9 Light-Triggered Modulation of the Free Energy Landscape: Using Azobenzene Isomerization to Create a Downhill Folder	
9.1 Introduction.....	120
9.2 Experimental Section.....	123
9.3 Results and Discussion	125
9.4 Conclusion	128
10 Future Directions	
.....	138
Bibliography	
.....	141

LIST OF TABLES

4.1	Band position (ν) and full width at half maximum ($\Delta\nu$) of the C \equiv N stretching vibration of various probes in H ₂ O and THF.....	37
5.1	Name and sequence of the Trp-cage variants studied.....	57
5.2	Unfolding thermodynamic parameters of the Trp-cage peptides	60
6.1	Trpzip 2 variant relaxation, folding, and unfolding times at 25 °C	80
6.2	Unfolding thermodynamic parameters obtained from global fitting for each Trpzip 2 variant ($T_R = 25$ °C)	81
7.1	Unfolding thermodynamic parameters and folding/unfolding times at 30 °C of TC5b and its D-amino acid mutants.....	94
8.1	Fractional helicity (f_H) approximations of pKID in TFE/water mixtures.....	116
8.2	Relaxation rates of pKID in TFE/water mixtures at 15 °C.....	117
8.3	Relaxation rates of LEA in TFE/water mixtures at 25 °C.	118

LIST OF FIGURES

4.1	Schematic overview of synthetic routes converting cysteine to vibrational probes of interest.....	38
4.2	The C≡N stretching bands of W3C-CN obtained in the presence (open circles) and absence (open triangles) of CaM.....	39
4.3	The C≡N stretching band of the nitrile-labeled human calmodulin-like protein	40
4.4	The C≡N stretching bands of the four cysteine derivatives, as indicated, obtained in water (open triangles) and THF (open circles)	41
4.5	The C≡N stretching bands of A8C-CN obtained in the absence (open triangles) and presence (open circles) of CaM	42
5.1	Structure of the Trp-cage (PDB code: 1L2Y), showing the α -helix (red), the 3_{10} -helix (blue), the polyproline region (green), and the sole tryptophan (orange).....	53
5.2	A representative FTIR difference spectrum of ^{13}C -TC10b between 65.0 °C and 25.0 °C.....	54
5.3	Representative T -jump induced conformational relaxation traces of ^{13}C -TC10b in response to a T -jump from 5 to 10 °C, probed at different frequencies as indicated	55
5.4	Conformational relaxation rate constants (solid symbols) of ^{13}C -TC10b obtained with a probing frequency of 1580 cm^{-1} (red), 1612 cm^{-1} (green), and 1664 cm^{-1} (blue), respectively. The blue open triangles represent the	

	relaxation rates of the fast kinetic phase observed at 1664 cm^{-1} . The black open symbols represent the global folding (circle) and unfolding (square) rates of the protein	56
5.5	Far-UV CD spectra of the Trp-cage peptides studied, as indicated. These data were collected at $1.0\text{ }^{\circ}\text{C}$	58
5.6	CD thermal melting curves (open circles) of the Trp-cage peptides studied, as indicated. Smooth lines represent the best fits of these data to a two-state model.....	59
5.7	Difference FTIR spectra of TC5b in the amide I' region, which were generated by subtracting the FTIR spectrum obtained at $4.0\text{ }^{\circ}\text{C}$ from those measured at higher temperatures (the highest temperature was $82.2\text{ }^{\circ}\text{C}$).....	61
5.8	Temperature dependence of the slow relaxation rate (black), folding rate (blue), and unfolding rate (red) of TC10b-R16K.....	62
5.9	Temperature dependence of the slow relaxation rate (black), folding rate (blue), and unfolding rate (red) of TC10b-P19A.....	63
5.10	Temperature dependence of the slow relaxation rate (black), folding rate (blue), and unfolding rate (red) of TC5b	64
6.1	A cartoon representation of the β -hairpin structure of Trpzip-2c with the BB-HBs shown (dotted lines). BB-HBs that are perturbed using O-to-T mutation are shown in blue	77
6.2	CD thermal melting curves of the Trpzip-2c peptides, as indicated. Smooth lines are global fits of these data to a two-state model discussed in the text.....	78

6.3	Arrhenius plot of the relaxation (black), folding (blue), and unfolding (red) rate constants of Trpzip-2c and mutants	79
6.4	Far-UV CD spectra of the Trpzip-2c peptides at 1 °C, as indicated	82
6.5	A representative <i>T</i> -jump induced relaxation trace of Trpzip-2c in response to a temperature change from 16.2 to 20.7 °C monitored at 1630 cm ⁻¹ . Shown in the inset are the conformational relaxation kinetics of E5/TE in response to a <i>T</i> -jump, from 12.6 to 20.4 °C. The smooth lines in red are single-exponential fits of these data.....	83
7.1	The NMR structure of Trp-cage (PDB 1L2Y). The glycine residue that was mutated to various D-amino acids in the current study is highlighted	95
7.2	Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GDQ. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type.....	96
7.3	Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GDN. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type.....	97
7.4	Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GDA. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type.....	98
8.1	CD wavelength and temperature melt spectra of pKID in aqueous solutions of different TFE percentages, as indicated	110

8.2	Relaxation rate constants of pKID versus temperature for different TFE solutions, as indicated	111
8.3	Representative trace of the relaxation kinetics of the pKID peptide in a 30% TFE solution in response to a temperature jump from 5.7 °C to 11 °C probed at 1630cm ⁻¹ . The smooth line represents the best fit of this curve to a single exponential function.....	112
8.4	CD wavelength and temperature melt spectra of LEA in aqueous solutions of different TFE percentages, as indicated.....	113
8.5	Relaxation rate constants of LEA versus temperature for different TFE solutions, as indicated	114
8.6	Figure 8.6: Representative trace of the relaxation kinetics of the LEA peptide in a 40% TFE solution in response to a temperature jump from 3.8 °C to 8.4 °C probed at 1664 cm ⁻¹ . The smooth line represents the best fit of this curve to a single exponential function.	115
9.1	Cartoon representation of azobenzene-substituted Trp-cage.....	130
9.2	Absorption spectra of azobenzene-incorporated Trp-cage 10b before irradiation with 355 nm light, after irradiation at 355 nm for 5 minutes, and after further irradiation at 420 nm for 5 minutes	131
9.3	CD wavelength spectra of 10b-azob when dark-equilibrated and after irradiation at 355 nm for 5 minutes.....	132

9.4	A kinetic trace of 10b-azob probed at 1630 cm^{-1} in response to irradiation with a 355 nm pulse of light	133
9.5	A kinetic trace of 10b-azob probed at 1680 cm^{-1} in response to irradiation with a 355 nm pulse of light	134
9.6	CD wavelength spectra of helix-azob when dark-equilibrated and after irradiation at 355 nm for 5 minutes, as indicated.....	135
9.7	A kinetic trace of helix-azob probed at 1680 cm^{-1} in response to irradiation with a 355 nm pulse of light. Inset: A kinetic trace of helix-azob probed at 1630 cm^{-1} in response to irradiation with a 355 nm pulse of light	136

CHAPTER 1

Introduction

1.1 Protein Folding

With the discovery of proteins and the emergence of techniques to visualize structures of these macromolecules, a natural follow-up question quickly arose: how do proteins adopt these complicated, three-dimensional structures?[1] The study of how proteins fold essentially began in the 1970s after the pioneering work of Christian Anfinsen,[2,3] who showed that a protein can reversibly fold in a test tube without the presence of any other molecules other than solvent. These findings indicated several important points about folding: the folded state is the lowest in terms of free energy, there is exchange between the folded and unfolded states with access only to thermal energy and the solvent, and the folded structure of a given protein is encoded in its primary (amino acid) sequence. Interestingly, when one considers all of the possible conformations in a given protein, as Levinthal famously pointed out, it would take an incredible amount of time to fold if there was no energetic bias towards the native state.[4] Therefore, folding came to be understood in the framework of pathways, typically pictorialized using a one-dimensional free energy landscape with wells depicting observable states and barriers between them.[5,6] Identifying additional folding states (intermediates) and their connectivities, as well as obtaining a structural picture of folding transition states are major research topics in the contemporary study of protein folding.[7-24] Though there have been great advances in these areas, which will be touched on throughout this Thesis, there are several questions that still remain. One example of a persisting problem in protein folding is whether discovered intermediates are on-pathway or not, and what the functional purpose of an off-pathway intermediate could be.[25] Other studies have focused on

generalizing the characteristics of folding transition states.[26,27] For example, a great deal of evidence now suggests that most transition state structures of small proteins are near-native, containing ~80% of the native structure of the protein of interest.[28] Further work is now being conducted to understand the nature of the denatured state, and how different these species are compared to the random coils we often envision them as.[29] Other work is focused on the nature of initial protein structuring upon addition to solvent: does secondary structure form first before any appreciable hydrophobic collapse, or does the hydrophobic effect dominate the earliest events in protein folding?[30,31]

Regardless of the interpretation of the folding mechanism, however, the prevailing assumption has been that proteins are globular, well-structured molecules in their native state in solution. Efforts to expand folding principles beyond such viewpoints began thirty years ago with the study of membrane proteins, whose structures generally are disordered in solution until association with a lipid membrane.[32,33] Since then, the binding of disordered peptides to membranes and their subsequent folding into helices and insertion into the bilayer have been discovered.[34,35] Much still remains to be studied with regards to disordered proteins that do not rely on membranes to fold.

Recent studies suggest that intrinsically disordered proteins (IDPs), proteins that are unable to fold independently into a well-defined tertiary structure under physiological conditions, seem to actually be much more prevalent in nature than previously surmised.[36] For example, recent estimates based on bioinformatics showed that disordered proteins or proteins containing large disordered regions (≥ 30 residues) may make up as much as 30% of some eukaryotic genomes.[37] These IDPs include many

DNA binding and cell-signaling proteins, and often are found in regulatory pathways, including transcription and translation. The amino acid composition of these proteins follows a trend: IDPs tend to have very low numbers of hydrophobic residues, but contain many charged residues.[38] As hydrophobic collapse is often thought of as a major driving force in protein folding, and repulsions arising from like charges can keep a protein's subunits well-separated, one can see how disorder persists in such proteins. Why, then, do we have IDPs? Preliminary evidence suggests that the flexibility of IDPs allows for multiple binding partners and higher levels of regulation via variable binding affinities.[38,39] IDPs are found with greater frequency in increasingly complex organisms (bacteria → yeast → humans), allowing a greater degree of regulation with each additional level of organismal intricacy.[37] It has also been proposed that the increased binding surface area of IDPs allows for faster association with ligands (the fly-casting hypothesis),[40,41] although the increased hydrodynamic radius could slow molecular diffusion and hence binding.[41] The latter would argue that the increased binding rate observed for IDPs is due to a more favorable free energy change for binding, stemming from the coupled folding process.

While the specifics of IDP binding have not completely been determined yet, another mystery is how IDPs fold. Though it is true that some IDPs never adopt a fully folded structure, even upon binding, many IDPs do fold into an ordered conformation upon binding to a ligand.[39] One can imagine two extreme mechanisms by which an IDP folds: either the IDP associates first with its binding partner and then folds (i.e., induced folding) or the IDP folds first and then associates with its binding partner. The

latter has been referred to as conformational selection, as binding shifts the folding equilibrium toward the folded state. Recent experiments have shown both mechanisms are possible,[42-44] but much work is needed to fully elucidate exactly how folding takes place at the molecular level and what predisposes a given IDP for one mechanism versus another. Additionally, and perhaps most importantly, it should be noted that IDPs can fold via a mechanism that is a combination of these two extremes, where only partial folding is required before binding.[39]

Efforts to address the outstanding questions in protein folding involve both experimental and computational studies. Early measurements of protein folding were limited to denaturant mixing methods, probed either via tryptophan fluorescence or circular dichroism. Details obtained with these experiments were limited, partially due to the inability to observe events occurring faster than milliseconds, and partially due to the lack of structural resolution with these techniques. Over the years, fast-initiation techniques, such as rapid mixing,[45,46] pressure-jump spectroscopy,[47] and temperature-jump spectroscopy[48-50] have extended the effective temporal resolution of folding measurements to the nanosecond to microsecond timescale, where some of the earliest events in protein structural formation occur. Similarly, techniques such as triple resonance nuclear magnetic resonance (NMR) spectroscopy allow for atomistic structural resolution in folding studies, and in some cases even can be used to determine the structures of sparsely populated excited states of proteins.[51] Interestingly, despite all of the technological advances over the past four decades, there is still a tradeoff in experimental techniques of spatial and temporal resolution; in other words, no

contemporary technique is universally applicable to folding questions ranging a great span of time or length scales. To overcome this limitation, one has to devise alternative means to uncover important details. Computational studies of protein folding began with simple lattice models, combined with theory from polymer physics and statistical mechanics, which despite their simple nature were often able to capture many of the characteristics of protein chains. From these elementary frameworks arose the idea that free energy landscapes of protein folding should be representative of additional degrees of freedom, and ultimately led to the ‘folding funnel’ concept when the energetics of folding were plotted as a function of overall conformational entropy.[52-55] Contemporary simulations have increased the complexity of their systems, often employing the full structures of proteins of interest in explicit solvent and simulating for aggregates of time of up to milliseconds.[56] While it is worth mentioning that researchers have devised many ways to perform such a large amount of calculations, for example by using supercomputers,[56-57] employing distributed computing,[58] or using computers dedicated to solving these mathematical problems, often with the help of graphics processor units, it is beyond the scope of this Thesis to delve further into the details of this work. Here, it is sufficient to note that the contemporary problems of protein folding can be approached with equivalent rigor from an experimental or theoretical framework, and often the conclusions obtained from one method can be well-complemented by work from the other end of the spectrum.

In this Thesis, the overarching goal is to expand upon the experimentalists’ toolbox for studying protein folding, either by introducing new methods or by

functionalizing the molecules studied to provide additional information than what can normally be obtained using infrared spectroscopy. For example, chapters 4, 6, and 7 involve chemical modifications of proteins to either directly or indirectly increase the structural resolution of infrared measurements, whereas chapters 5 and 8 utilize new advances with our transient infrared setup that allow us to investigate a miniprotein's folding mechanism and the nature of a cosolvent's effects on protein stability.

Chapter 2 is a summary of the theory used in the doctoral research presented in this Thesis. A general overview of protein folding thermodynamics and kinetics is laid out, as well as an outline for analytical methods used to study folding transition states.

Chapter 3 involves a discussion of the methods and molecular modifications used in the following research. This includes a discussion of circular dichroism (CD), Fourier transform infrared spectroscopy, temperature jump transient IR measurements, and vibrational probes.

Chapter 4 demonstrates a new method to incorporate aromatic nitrile probes into peptides and proteins through cysteine sidechains using an alkylation reaction. We distinguish between several model compounds to identify an ideal candidate for incorporation in terms of extinction coefficient, sensitivity to the electrostatic environment, and ease of reaction: *para* cyano benzyl bromide. We further go on to test the incorporation in a model peptide, and assay probe utility by comparing its vibrational transition when exposed in water and when buried in complex with a larger protein. Finally, we show incorporation of this probe in a protein, revealing the range of applications for this method.

In Chapter 5, we take advantage of isotope labeling and sidechain vibrational transitions to develop a method to use multiple probes and probing frequencies to reveal kinetics of different secondary structural elements. We apply this technique to a well-studied miniprotein, Trp-cage, and observe kinetics indicative of a hidden intermediate. Further analysis identifies this intermediate as the native state minus the folded 3_{10} helix, and the rate-limiting step of folding as being the formation of the alpha helix. By increasing the effective structural resolution of the typical T -jump experiment, we are able to observe folding events that would be otherwise invisible, even in so-called ‘simple’ model systems that are well-studied both experimentally and computationally.

In Chapter 6, we demonstrate an approach to site-specifically perturb a backbone-backbone hydrogen bond. Since thioamides are almost isosteric as native oxoamides, but the sulfur atom in a thioamide is a weaker hydrogen bond acceptor than the oxygen in an oxoamide, this functionalization represents a facile way to perturb a single hydrogen bond and assess its energetic contributions to folding. We tested this method by incorporating thioamide amino acids individually into several positions in a Trpzip beta-hairpin, and found that the backbone-backbone hydrogen bond next to the turn is important in forming the transition state, consistent with previous studies.

Chapter 7 presents an application of using D-amino acids to dissect a protein’s folding mechanism. Since D-amino acids have access to areas of the Ramachandran plot that otherwise are only accessible to glycine residues, they represent a unique avenue to interrogate the importance of glycines in proteins and their frequent placement in turn regions of structures. To assess if this method is practical, we substitute an important

glycine in the Trp-cage model peptide with several different D-amino acids to gain further insight into the peptide's folding mechanism. We observe that the C-terminus of the alpha helix in Trp-cage, where this particular glycine is found, is not fully formed in the transition state.

In Chapter 8, we investigate the mechanism of how the cosolvent trifluoroethanol (TFE) affects protein stability. Specifically, TFE has long been known and used to stabilize helical secondary structure – additionally, it is well-documented that TFE destabilizes tertiary structure. While there are several explanations for these observations, in general it is difficult to separate the effects on secondary and tertiary structure. One way around this problem is by using IDPs, which may have some amount of secondary structure, but lack any significant tertiary structure. We perform kinetic measurements on two different IDP systems, one which forms a helix-turn-helix with increasing TFE concentration and one that forms a single helix, to examine if there are differences in relaxation rates between the two. Interestingly, we find that in the case of the single helix there is no change on relaxation rate with increasing TFE concentration, however for the helix-turn-helix we observe a decrease in relaxation rate at intermediate (~30 % v/v) TFE concentrations, which goes away at higher (~50% v/v) TFE concentrations. We believe that these observations are consistent with recent computational findings suggesting that cosolvents act as ‘mini-crowders’.

Chapter 9 presents a method to trigger ultrafast, barrierless folding with light. The use of phototriggers to initiate folding from a constrained, known state allows researchers to ask specific questions about the nature of the folding landscape that otherwise could

not be addressed experimentally. While this technology is not new, here we add another dimension to the technique by using the phototrigger to constrain structural elements formed at the transition state of folding, allowing us to significantly populate the transition state upon irradiation. We tested this idea with the Trp-cage model peptide, whose folding mechanism and transition state structure we have previously characterized. By inserting an azobenzene crosslinker into the alpha helix of Trp-cage, we effectively prevent its formation, and the folding of the rest of the molecule, unless triggered with light. Kinetic folding measurements after triggering reveal dynamics that have a time constant of ~100 ns at room temperature, which is an order of magnitude faster than the folding rate of wild-type Trp-cage at that temperature. The ability to take a protein that folds in an activated (barrier-containing) manner and convert it into a downhill folder upon irradiation is useful, as it not only validates our understanding of the molecule's dominant folding mechanism, but also allows us to infer useful details about the protein's free energy landscape, such as the landscape roughness.

CHAPTER 2

Theory

2.1 Protein Folding Thermodynamics

To quantitatively investigate the folding of proteins, one first must examine the theoretical framework of protein folding thermodynamics. Folding thermodynamics have been extensively studied in the past.[59-61] While there is a great deal of information worth understanding with regards to examining the equilibria between different protein conformational states, here it is sufficient to explain a few simple concepts.

To begin, it has been commonly observed that a predominant number of small proteins and peptides fold in a two-state manner, where molecules only exist in either a folded (F) or unfolded (U) state. In this scenario, a chemical equilibria equation relating the two states is simply:



with a forward (unfolding) reaction rate constant, k_u , and backward (folding) rate constant, k_f . The equilibrium between these two states is a function of the Gibbs free energy of unfolding (ΔG_u , which is equivalent to the difference in free energy of the unfolded and folded states, respectively), which can be expressed as:

$$\Delta G_u(T) = \Delta H_R^\circ + \Delta C_p(T - T_R) - T(\Delta S_R^\circ + \Delta C_p \ln(T/T_R)) \quad (2.2)$$

where ΔH° , ΔS° , and ΔC_p are the enthalpy, entropy, and heat capacity of unfolding, respectively, and R is a reference point, typically chosen to be the melting point (50% folded, 50% unfolded) of the system of interest. Equation 2.2 can be further detailed with additional variables such as pressure or denaturant, however the work in this thesis does not require consideration of these other factors, and therefore they have been purposely neglected for the sake of simplicity.

2.2 Protein Folding Kinetics

As found for folding thermodynamics, it is useful to apply the two-state folding scenario to interpret folding kinetics. As understood from equation 2.1, at equilibrium the magnitude of the change in the folded population as a function of time, $\frac{d[F](t)}{dt}$, is equivalent to the change in the unfolded population as a function of time, $\frac{d[U](t)}{dt}$, otherwise expressed as:

$$\frac{d[F](t)}{dt} = -\frac{d[U](t)}{dt} \quad (2.3)$$

$\frac{d[F](t)}{dt}$ can also be written in terms of k_f and k_u , as follows:

$$\frac{d[F](t)}{dt} = k_f[U](t) - k_u[F](t) \quad (2.4)$$

Here, it is helpful to relate [F] and [U] with the following equation:

$$[F] + [U] = [B] \quad (2.5)$$

where [B] is the total population. With this relation, Equation 2.4 can now be integrated to yield an equation describing the folded population as a function of time:

$$[F](t) = [B] \frac{k_f}{k_f + k_u} + C e^{-(k_f + k_u)t} \quad (2.6)$$

where C is a constant of integration. Equation 2.6 is useful because examination of its form reveals that at equilibrium [F] equals $[B] * k_f / (k_f + k_u)$. Additionally, it shows that any deviation from equilibrium will have an exponential relaxation back to equilibrium with a rate constant equal to the sum of the folding and unfolding rate constants. Using Equation 2.4 it is easy to relate the equilibrium constant (K_{eq}) to the folding and unfolding rate constants:

$$K_{eq} = \frac{[U]}{[F]} = \frac{k_u}{k_f} \quad (2.7)$$

Thus, for a given system under a common set of conditions, if the equilibrium constant and relaxation rate (k_r) are both known, one can determine k_f and k_u individually.

We can further interpret results from relaxation kinetic measurements if we extrapolate values for k_f and k_u . For example, if one knows the folding rate for a given protein, one can estimate values for the barrier height (ΔG^\ddagger) of the reaction using the Arrhenius equation:

$$k_f = Ae^{-\Delta G^\ddagger/k_B T} \quad (2.8)$$

where A is a constant often referred to as an attempt frequency and k_B is the Boltzmann constant. Precisely defining and evaluating A has been a major research focus of many labs in the past, the full details of which will not be discussed here. Contemporarily, the folding community typically interprets barrier heights in the framework of Kramer's equation:

$$k = \frac{A}{n(T)} e^{-\Delta G^\ddagger/k_B T} \quad (2.9)$$

where $n(T)$ is a frictional term related to the viscosity of the solvent.[62] Addition of this term to the Arrhenius equation was helpful in studies examining the viscosity-dependence of folding of proteins. Kramer's theory also provides a way to interpret data on downhill folding, where the barrier to folding is absent or relatively small in comparison to $k_B T$ (0.6 kcal/mol at room temperature); such analysis allows one to estimate the roughness of the energy landscape, a parameter which is otherwise rather difficult to investigate.

Additionally, methods have been developed to understand the nature of the transition state. Perhaps the most famous of these approaches is phi-value analysis, where

strategic (typically sidechain) mutations are made in a protein in an effort to assess whether some interaction found in the native (i.e. folded) state of the protein is also present in the transition state.[26] To address this question, the equilibrium constants of unfolding for both the mutant and wild-type proteins are measured. Here, it is helpful to point out the relationship between the free energy of unfolding and the equilibrium constant:

$$\Delta G_u = -RT \ln K_{eq} \quad (2.10)$$

Using Equation 2.10, we can determine the change in stabilization of the native state due to the above mutation by examining the ratio of the equilibrium constants of unfolding of the mutant and wild-type as follows:

$$\Delta \Delta G_u = -RT \ln \frac{K_m}{K_{wt}} \quad (2.11)$$

where $\Delta \Delta G_u$ is the change in stability of the native state due to mutation, K_m is the equilibrium constant of the mutant peptide, and K_{wt} is the equilibrium constant of the wild-type peptide. We can analogously estimate the change in stabilization of the transition state due to the above mutation by examining the ratio of the folding rates of the mutant and wild-type as follows:

$$\Delta \Delta G^\ddagger = -RT \ln \frac{k_m}{k_{wt}} \quad (2.12)$$

where $\Delta \Delta G^\ddagger$ is the change in barrier height due to mutation, k_m is the folding rate of the mutant peptide, and k_{wt} is the folding rate of the wild-type peptide. The phi-value (Φ) is defined as:

$$\Phi = \frac{\Delta \Delta G^\ddagger}{\Delta \Delta G_u} \quad (2.13)$$

In an ideal situation, a mutation that affects a native interaction in the transition state would yield a phi-value of 1, whereas a mutation that perturbs an interaction not present in the transition state would yield a value of 0. Thus, with multiple individual mutations throughout a model system, one could, in theory, identify all interactions/structure present in the transition state of folding, something which otherwise is very difficult to assess experimentally at the present time. In practice, however, there are several pitfalls when using this analysis. For one, this method assumes that stabilizing interactions in proteins are pairwise and additive. In reality, this assumption is not correct in all cases. Another premise is that an interaction present in the transition state has the same energetic contribution as in the native state. This may not hold true, as some structures are not completely formed in transition states,[63] and sometimes mutations can alter the dominant transition state,[64,65] leading to the comparison of two different states entirely. Both of these examples could potentially result in a fractional phi-value. Similarly, a mutation in a protein might cause some non-native interaction that results in different changes in the free energy of unfolding for the transition state and the native state. Finally, and perhaps most importantly, this method assumes that mutations do not affect the free energy of the unfolded state. Characterizing protein unfolded states is a popular topic in contemporary protein folding studies, and future measurements of energetic effects on unfolded states of mutant proteins would be very useful in expanding upon the information obtained via this analysis.[29] Ultimately, the caveats of phi-value analysis are realized in experimental results that lead to phi-values that are fractional, negative, or much larger than one – interpretation of these values is rather difficult, and

typically requires additional measurements or other forms of analysis. Despite these limitations, phi-value analysis has been successfully applied to many two-state folding peptides and proteins, allowing researchers to identify structural elements of transition states and better understand folding mechanisms. For completeness, a brief commentary on an analog of phi-value analysis, termed psi-value analysis is worth mentioning. Psi-value analysis utilizes bi-histidine mutations in the protein of interest. Taking advantage of the metal-binding nature of the bi-histidine structure, folding measurements can be performed as a function of metal concentration and, using an analysis similar to phi-value analysis, a psi-value can be obtained, which informs on whether or not the mutated region of the protein is structured in the folding transition state.[27,66] There are many examples of applications of psi-value analysis to the study of protein folding mechanisms, but expounding upon them in great detail is beyond the scope of this Thesis.

CHAPTER 3

Methods

3.1 Circular Dichroism (CD) Spectroscopy

Circular dichroism is a useful and popular technique to assay secondary structural content in proteins. CD measurements send right and left circularly polarized light into a sample compartment, and the difference in absorption of one polarization over the other (the ellipticity), due to the chirality of the sample, gives rise to the resulting spectrum. For proteins, the typical wavelength region examined by CD measurements is the far-UV (190-250 nm), where the peptide backbone electronic transitions $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ absorb. Typically, alpha helices have two minima at 222 and 209 nm, corresponding to $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ transitions. The lower wavelength transition is sensitive to helix length, where a shorter helix will result in a blueshift in the minimum. In contrast, beta-sheets have one minimum ($n \rightarrow \pi^*$) at ~ 215 nm, whereas random coil conformations have a single minimum ($n \rightarrow \pi^*$) at ~ 195 nm. Proteins containing mixtures of secondary structures have spectra that are difficult to interpret due to the fact that the helical transitions have a larger extinction coefficient than transitions of the other structural elements, and therefore signal from alpha-helices tend to dominate such spectra. Despite this, efforts have been made to quantify amounts of secondary structure directly from CD wavelength spectra using software such as K2D2, which compare CD spectra of proteins with unknown secondary structure contents to a library of spectra of proteins with known secondary structures.[67] Some sidechains, most notably the aromatic ones, give rise to signals in CD wavelength spectra – tryptophan, for example, typically has a positive ellipticity at a wavelength of ~ 220 nm from its B_b transition – however, the extinction of these transitions is relatively low compared to backbone transitions.[68] An exception to

this is observed when coupling exists between two or more aromatic sidechains in a given molecule. The best example of this is found in the Trpzip model peptide system, a beta-hairpin that contains four tryptophans in close proximity to one another. The excitonic coupling of the B_b transition of these tryptophans gives rise to a very unique CD wavelength spectrum where the sidechain absorptions dwarf those of the backbone, resulting in a minimum at ~212 nm and a maximum at 227 nm due to splitting. The CD temperature melt, where ellipticity measurements at one wavelength are monitored over a range of temperatures, is a useful method for determining equilibrium constants of unfolding as a function of temperature, which is necessary for extrapolating folding and unfolding rates from relaxation kinetics (see Chapter 2) measured using temperature-jump experiments. The wavelength chosen for this experiment corresponds to the peak ellipticity value in a given protein's CD wavelength spectrum (222 nm for a helical protein, etc.). For a two-state folder, we can interpret the ellipticity measured at a particular temperature ($\theta(T)$) to be the sum of the signals arising from the folded ($\theta_f(T)$) and unfolded ($\theta_u(T)$) states, multiplied by their respective populations and normalized by the total population:

$$\theta(T) = \frac{[U]\theta_u(T) + [F]\theta_f(T)}{[U] + [F]} \quad (3.1)$$

which can be rearranged using the equilibrium constant of unfolding to be:

$$\theta(T) = \frac{\theta_f(T) + K_{eq}\theta_u(T)}{1 + K_{eq}} \quad (3.2)$$

Additionally, $\theta_f(T)$ and $\theta_u(T)$ are assumed to vary linearly with temperature:

$$\theta_f(T) = a + bT \quad (3.3)$$

$$\theta_u(T) = c + dT \quad (3.4)$$

where a, b, c, and d are all constants. Through the use of Equations (2.2) and (3.2) to fit temperature melt data, one can obtain values for all thermodynamic parameters.

3.2 Fourier-Transform Infrared (FTIR) Spectroscopy

FTIR is a popular technique for measuring the infrared absorption of samples over a broad spectral range (typically from 500-4000 cm^{-1}). To accomplish this, a broadband source (e.g. a globar) is sent into an interferometer, and the resulting interferogram is sent through the sample. A Fourier transform of the resulting signal gives rise to a transmission spectrum of the sample, which can easily be converted to absorbance using Beer's law:

$$\text{OD} = -\log \frac{I}{I_0} \quad (3.5)$$

where I is the transmittance intensity of the sample at a given wavenumber, and I_0 is the transmittance intensity of a reference at the same wavenumber. For our sample cells, we have two separate compartments, one for the protein sample, and one for the buffer background. For our measurements, the transmittance through buffer alone is used as the reference unless otherwise specified. Our FTIR setup also has temperature control ($\sim 4-70$ $^{\circ}\text{C}$) via a water bath connected to the sample cell holder. By collecting absorption spectra of a sample over a range of temperatures and subtracting the absorption spectrum at the lowest measured temperature, one can obtain a difference spectrum that shows how the IR spectrum changes as a function of temperature. This is useful for determining probing frequencies in temperature-jump studies, as well as estimating the magnitude of signals obtained in temperature-jump measurements.

3.3 Laser-Induced Temperature-jump (*T*-jump) Spectroscopy

To obtain measurements of folding/unfolding kinetics, one must perturb a system's equilibrium and follow the resulting dynamics. While there have been many methods over the years that vary how the system is taken out of equilibrium or how the dynamics are measured, one practical limiting factor applies to all experimental methods: a measurement can have either very high spatial or temporal resolution, but not both. Efforts to improve the limiting resolution of different methods (including the measures taken in this Thesis) continue, but to date no technique by itself is able to examine every aspect of the folding of a given protein. One of the non-equilibrium techniques used in our lab is laser-induced *T*-jump spectroscopy.[48-50] In our setup, a nanosecond Nd:Yag laser, whose 1064 nm output is shifted to 1.9 μm via the stimulated raman effect, is used to excite the vibrational overtone of D_2O , which relaxes on a picosecond timescale and transfers the energy to heat. The magnitude of a *T*-jump is dependent on several parameters:

$$\Delta T(r, z) = \frac{k}{\rho C_v} I(r, z) \quad (3.6)$$

where r is the beam radius, z is the distance the beam has traveled through the sample in the direction of propagation, k is the extinction coefficient of the sample (10.1 cm^{-1} for bulk D_2O at 1.9 μm), ρ is the density of the solution, and C_v is the heat capacity of the solution.[69] By using relatively small pathlengths (50 μm) and a vibrational transition with a weak extinction coefficient, we are able to have excitation light with roughly the same intensity throughout the pathlength of sample cells, allowing us to prevent a temperature gradient from resulting. In addition, the probe laser used to monitor protein-

related infrared transitions has a smaller beam diameter than the pump laser, which allows us to probe a spatial region of the heated sample that is relatively homogeneous in temperature. FTIR measurements of D₂O as a function of temperature allow us to use a calibration curve relating change in absorbance to the magnitude of a *T*-jump with a specified initial temperature and probing frequency. By comparing the change in absorbance of buffer in a temperature jump experiment to the calibration curve described, we can estimate the final temperature of the excited region.

3.4 Infrared Spectroscopy of Proteins

The use of infrared spectroscopy predominantly focuses on the use of the amide I vibrational transition, which corresponds largely to the antisymmetric stretching of the carbonyl in the polypeptide backbone. This particular vibration is useful because of its location in an otherwise uncongested region of the infrared spectrum, and is sensitive to the local environment and to coupling with other amide I chromophores.[70] The latter property results in the amide I transition having different spectral signatures depending on the secondary structures present in the system of interest.

Like all of the contemporary techniques commonly used to study protein conformational dynamics, there is a tradeoff between temporal and spatial resolution. For example, X-ray crystallography is currently unparalleled in its ability to determine structural details with atomistic resolution, however the resulting measurements are static snapshots of biomolecules. In the other extreme, one of the benefits of using infrared spectroscopy is the fact that the oscillations of chromophores occur on the femtosecond to picosecond timescale, affording incredible temporal resolution, however the

vibrational transitions typically examined for proteins largely lack any appreciable spatial resolution. Several methods exist to circumvent these limitations; two of the most popular are the use of extrinsic probes and isotopic labeling. Extrinsic probes are desirable because their vibrational transitions occur in otherwise unpopulated regions of the infrared spectrum, such as the nitrile stretching transition of *para*-cyanophenylalanine in H₂O (2237 cm⁻¹), and they are sensitive to the local environment. A great number of studies have been devoted to identifying suitable vibrational functional groups and probes for infrared spectroscopic measurements of proteins, which have been reviewed extensively in the literature.[71] While early efforts at identifying probes focused on molecules that could be incorporated into peptides via solid-phase peptide synthesis, more recent studies have examined the insertion of vibrational probes into proteins using amber codon suppression and molecular replacement strategies.[72,73] Though a great amount could be written about the development and application of these techniques, the topic is beyond the scope of this Thesis.

An alternative method to circumvent the limited structural resolution of infrared measurements of proteins is to use isotopic labeling.[74] Specifically, by replacing an atom in the chromophore of interest with a heavier variant (i.e. replacing the ¹²C and/or ¹⁶O in the amide I vibrator to ¹³C and/or ¹⁸O), the labeled chromophore will vibrate at lower frequency with respect to the unlabeled chromophores, as one would expect in the case of a classical vibrator that followed Hooke's law. In the case where both the carbonyl carbon and oxygen are labeled, the resulting vibrational transition will be red-shifted by ~70 cm⁻¹. Since this technique requires the labeling of specific positions and

not entirely throughout the peptide backbone, it is not readily applicable in larger protein systems (>45 amino acids) that must be expressed. However, there are numerous examples in the literature of using isotopic labeling to study the folding and local environment of small model peptides.[75,76]

3.5 Using Phototriggers to Initiate Conformational Changes in Proteins

An increasingly utilized way to perturb protein conformational equilibria is to employ a molecular trigger, which is synthetically incorporated into the system of interest. Molecular triggers for folding have been a topic of study for decades. There are multiple types of molecular triggering methods, which are dependent on the type of functional moieties that are present; examples of triggering methods include mechanical triggering,[77] triggering with current,[78] and phototriggering.[79] Due to the large number of molecules that change conformations upon absorbing light of a particular wavelength, as well as the ease with which triggering can be achieved, phototriggers are the most popular way to switch protein conformations. There are many examples of phototriggers, and the use of each is typically determined based on several parameters. For example, most of the time the desirable spectral absorbance range for phototriggers falls between 310 and 800 nm. This range is practical in nature: protein sidechain and backbone transitions absorb below 310 nm, and above 800 nm there is not (generally) sufficient energy to generate a conformational transition large enough to significantly affect protein structure. Another important consideration is whether the trigger will be reversible or not. If the trigger is reversible, one must ensure that the reverse reaction does not occur on the same timescale as (or before) the desired molecular event that will

be observed. If the trigger is irreversible, one must work with large amounts of sample and typically a flow cell apparatus to ensure that there is always phototriggering occurring during a given experiment. An additional consideration is the photochemical yield of the reaction. In general a high photochemical yield is desirable to maximize signal, however for scenarios where one is using an irreversible phototrigger with a limiting amount of sample, some compromise between signal strength and rate of sample conversion is applied. Further, triggering should occur on a faster timescale than the molecular event to be observed. Finally, the trigger must be able to be synthetically incorporated into the protein of interest. Over the past few decades, a substantial number of molecules have been examined as potential phototriggers. A practical choice for a phototrigger comes from natural elements in proteins: disulfide bonds. It is well known that disulfides can be cleaved with ultraviolet light (270 nm), however the free radicals formed often undergo geminate recombination, which leads to new disulfide bond formation.[80] The radicals can also take part in side reactions with protein sidechains or the backbone, potentially disrupting protein native structure. Additionally, normal disulfide bonds cleave on a microsecond timescale, which would interfere with folding measurements of model peptide/protein systems, since their conformational dynamics also typically occur then. One solution to this is to synthesize aryl disulfides, which are known to cleave in less than one picosecond.[79] To combat geminate recombination, scientists have found that insertion of disulfides into proteins in geometries that strain the bond greatly reduces disulfide bond reformation after cleavage.[80] This can be tricky, however, since ensuring that a disulfide bond is in a strained configuration requires some

trial and error. Another alternative is to insert a tryptophan proximal to the disulfide bond, and by exciting the tryptophan, an electron from the tryptophan will transfer to the disulfide, breaking the bond irreversibly.[81] There are several other examples of irreversible phototriggers, including tetrazines, hydrazines, and dimethoxybenzoins. Of the list of irreversible phototriggers, tetrazines are currently the most popular due to their fast timescales of reaction (picoseconds), a relatively high photochemical yield (~30%), and their excitation wavelengths (355 nm). There are several examples of using tetrazine to study conformational rearrangements in protein systems.[82,83]

Reversible phototriggers have focused on isomerization reactions of molecules, largely stemming from the original investigation of photoexcited stilbene isomerization by Hochstrasser and coworkers.[84] A natural variant of stilbene is the azobenzene moiety, where the stereocenter carbons are replaced with nitrogens. The photoisomerization of azobenzene has been extensively studied; the conversion of the thermally equilibrated *trans* conformation is believed to proceed to the *cis* conformation upon excitation via diffusion on the excited energy landscape followed by movement through a conical intersection to nuclear coordinates in the *cis* geometry.[85] This photochemistry also occurs on the picosecond timescale, while the reverse reaction thermally occurs on a much slower (seconds to minutes) timescale. Many chemical functionalizations of the azobenzene moiety have been investigated to facilitate easy incorporation into proteins and to alter the excitation wavelength for photoisomerization. A great number of studies using this variant of phototriggers apply a di-iodoacetamide azobenzene, which can be inserted into proteins easily using cysteine alkylation strategies

with the cysteine sidechain.[86-90] Here, the excitation wavelength is 355 nm, and the chemical yield is typically ~50%. Though this probe is highly desirable, Woolley and coworkers have recently developed azobenzene variants that have excitations in the visible region of the electromagnetic spectrum as well, and efforts to identify and design other reversible phototriggers still continue.

CHAPTER 4

Selective Incorporation of Nitrile-Based Infrared Probes into Proteins via Cysteine Alkylation

4.1 Introduction

The C≡N (nitrile) stretching vibration has recently emerged as a valuable infrared (IR) probe of the conformation and local environment of biological molecules[91-115] due to its sensitivity to various factors,[103,104,107] such as local electric field and hydrogen bonding interactions.[116] For example, it has been used to probe peptide insertion into membranes,[96] protein-ligand interactions,[115] and the dehydration status of an antimicrobial peptide encapsulated in reverse micelles.[98] For chemically synthesizable peptides, site-specific incorporation of a nitrile group is readily achievable through the use of nitrile-derivatized non-natural amino acids, such as cyanoalanine (Ala_{CN}) and *p*-cyanophenylalanine (Phe_{CN}). For proteins that cannot be chemically synthesized, however, selective incorporation of a nitrile moiety is rather difficult. Although it has been shown that Phe_{CN} can be incorporated into proteins by using an orthogonal tRNA-synthetase pair,[72,117] the techniques involved are time-intensive and available to only a handful of laboratories worldwide. Thus, it would be quite helpful to develop both an alternative and easier method for selective incorporation of different nitrile moieties into proteins. Here, we show that cysteine alkylation and arylation reactions can be used for such a purpose.

4.2 Experimental Section

Synthesis and Purification of Materials

Cysteine derivatives: The four model nitrile probes (**1**, **2**, **3**, and **4**) were synthesized from N-acetyl cysteine using previously reported procedures. All the prepared compounds show satisfactory NMR (¹H and ¹³C) and ESI-MS data, which will be reported elsewhere.

The IR samples of these probes were prepared by directly dissolving lyophilized solids in either Millipore water or 99.5+% spectroscopic grade tetrahydrofuran (Acros Organics) and the corresponding concentrations were estimated by weight.

Mastoparan-X mutants: The two mutants of mastoparan-X (sequence: INWKGIAAMAKKLL), W3C and A8C, were synthesized on an Argonaut Quest 210 synthesizer (Argonaut Technologies) at 0.092 mmol scales using Fmoc-CLEARTM Amide resin (substitution level - 0.46 mmol/g). Activation of Fmoc-amino acids (5-fold excess) was achieved with 0.95 equiv (relative to the amino acid) excess of HBTU in the presence of 10 equiv of DIPEA in DMF (HPLC grade, Aldrich). Side chain deprotection and simultaneous cleavage from the resin were carried out using a mixture of TFA/thioanisole/ethanedithiol/anisole (90:5:3:2, v/v) at room temperature, for 3 hours. After filtration, most of the solvent was then evaporated using a stream of N₂. The crude peptide was obtained by ether precipitation, which was then purified to homogeneity by reverse-phase chromatography (Agilent 1100 Series) with a C18 preparative column (Vydac) using a constant 10% buffer A (1% TFA in Millipore water), and a linear gradient of buffer B (100% acetonitrile) and buffer C (100% Millipore water). The mass of all peptides was verified using a Bruker Ultraflex III mass spectrometer.

Purified MpX mutants (W3C and A8C) were then used to perform the cysteine alkylation reaction with 4-bromomethyl benzonitrile under the conditions specified in Figure 4.1. Purification was performed as described above for the cysteine derivatives. The IR samples were prepared by directly dissolving lyophilized peptide solids into 50 mM HEPES buffer (pH 7). For the peptide binding experiments, the buffer also contains 30

mM CaCl₂ and 2 mM CaM. The final concentration of the A8C-CN was approximately 2 mM, estimated by Trp absorbance at 280 nm using an extinction coefficient of 5,500 M⁻¹cm⁻¹. The final concentration of the W3C-CN was estimated to be between 1-2 mM based on the absorbance at 214 nm using an extinction coefficient of 15,521 M⁻¹cm⁻¹.

Calmodulin variants: Human CaM was used in the MpX binding experiments, which has the following sequence:

ADQL TEEQI AEFKE AFSLF DKDGD GTITT KELGT VMRSL GQNPT EAELQ
DMINE VDADG NGTID FPEFL TMMAR KMKDT DSEEE IREAF RVFDK DGNGY
ISAAE LRHVM TNLGE KLTDE EVDEM IREAD IDGDG QVNYE EFVQM MTAK

A his-tagged version of human calmodulin-like protein CALM3, which contains a single cysteine residue, was used to test the cysteine alkylation reaction in proteins. The CALM3 gene was obtained from Invitrogen (cloned into pEXP5-NT/TOPO vector) and has the following sequence:

SGSHH HHHHG SSGEN LYFQS LMADQ LTEEQ VTEFK EAFSL FDKDG DGCIT
TRELG TVMRS LGQNP TEAEL RDMMS EIDRD GNGTV DFPEF LGMMA
RKMKD TDNEE EIREA FRVFD KDGNG FVSAA ELRHV MTRLG EKLS D EEVDE
MIRAA DTDGD GQVNY EEFVR VLVSK

Both CaM and CALM3 were expressed in *E. coli* BL21(DE3) pLysS cells (Novagen) at 37° and purified on an ACTA FPLC system (Amersham Biosystems) over a Ni-NTA column (Qiagen) in the case of CALM3, and twice over a phenyl sepharose HiTrap column in the case of CaM. The purity of the proteins was assessed by SDS-PAGE, MALDI and ESI-MS.

Cysteine alkylation of CALM3: 3 mL of a buffered (50 mM HEPES, pH 7.0, 100 mM NaCl) solution of CALM3 (100 μ M) was first diluted with a NH_4HCO_3 buffer (15 mL, 100 mM, pH = 8.4). Then, an aliquot of an aqueous TCEP solution (0.12 mL, 4.2 mM) was added to the above solution and stirred for 1 hour at room temperature, which was followed by addition of an aliquot of 4-bromomethyl benzonitrile DMF solution (0.25 mL, 2.5 mM). The reaction mixture was then stirred overnight at room temperature. Reaction progress was monitored by LC-MS. The small-molecule reagents were removed using a desalting column 10DG (BioRad). The solution obtained was then dialyzed twice into 50 mM HEPES buffer (pH 7.4) using a 10 kD dialysis cassette (Thermo Scientific).

FTIR Measurements

FTIR spectra were collected on a Nicolet Magna-IR 860 spectrometer using 1 cm^{-1} resolution and a CaF_2 sample cell that was divided into two compartments with a Teflon spacer (52 μm). For all reported spectra, a background was subtracted.

4.3 Results and Discussion

We tested the feasibility of the proposed method on four cyanobenzyl derivatives (Figure 4.1), based on the consideration that the oscillator strength and stark tuning rate of aromatic nitriles are normally larger than those of alkyl nitriles.[93,94] As shown (Figure 4.1), these model probes can be quite easily attached to the cysteine sidechain via either thiol alkylation or arylation.[118-120] Similar to that of Phe_{CN} ,[93] the $\text{C}\equiv\text{N}$ stretching frequency of these nitrile derivatives in water is found to be in the range of 2233 - 2241 cm^{-1} (Table 4.1 and Figure 4.4). Since any interactions that decrease/increase the electron density of the $\text{C}\equiv\text{N}$ bond will result in an increase/decrease in the $\text{C}\equiv\text{N}$ stretching

vibrational frequency of nitriles,[121] these results can be understood qualitatively in the context of the effect of an activating substituent on the cyanobenzyl ring (i.e., sulfur versus methylene and *para* versus *ortho* position with respect to the nitrile group). Furthermore, in comparison with those obtained in water, the C≡N stretching bands of these probes in tetrahydrofuran (THF) show a 7-8 cm⁻¹ shift toward lower frequency and also a concomitant decrease in the bandwidth by approximately a factor of two (Table 4.1), demonstrating the potential utility of these aromatic nitriles as local environmental probes.

Moreover, it is interesting to note that in aqueous solution the C≡N stretching bandwidths of **2** and **4** are noticeably larger than those of **1** and **3**. This finding is consistent with the study of Waegele *et al.*,[122] which showed that the C≡N stretching vibration of an aromatic nitrile can be influenced by direct interactions between the nitrile group and solvent molecules and also indirectly by solvation status of the aromatic ring. In other words, the larger bandwidth of **2** and **4** arises most likely from their asymmetric molecular shape (with respect to the nitrile group), which leads to a more heterogeneous solvation of the respective aromatic ring and hence a broader vibrational transition.

Considering the fact that the synthesis of **1** and **2** involves much milder conditions than those used in the synthesis of **3** and **4** and that the extinction coefficient of the C≡N stretching vibration of **1** and **3** is about an order of magnitude larger than that of **2** and **4** (Figure 4.4), only probe **1** is used in the subsequent proof-of-principle tests involving peptides and proteins.

First, the method of cysteine alkylation is applied to two cysteine mutants of mastoparan-X (MpX), W3C and A8C. These mutants are chosen because upon association with calmodulin (CaM) the sidechains of Trp3 and Ala8 of MpX are known to situate inside the peptide-protein binding groove and, as a result, become less solvent-exposed.[123] It is found that both peptides are efficiently labeled by *p*-cyanobenzyl bromide with a >80% yield (determined by LC-MS) under the conditions specified in Figure 4.1 (the corresponding nitrile-containing peptides are referred to hereafter as W3C-CN and A8C-CN). The FTIR spectra of W3C-CN (Figure 4.2) and A8C-CN (Figure 4.5) also support the site-specific incorporation of probe **1** into these peptides, as their C≡N stretching bands in aqueous solution are centered at about 2236.5 cm⁻¹ but shift to lower wavenumbers upon binding to CaM, as expected. Second, the applicability of the cysteine alkylation reaction to proteins is tested by applying it to human calmodulin-like protein CALM3, which contains a unique cysteine residue. As shown (Figure 4.3), the IR spectrum of the reaction product confirms the successful incorporation of probe **1** into the protein of interest and the corresponding yield was estimated to be >50%.

4.4 Conclusion

In summary, we have demonstrated a post-translational method allowing site-specific incorporation of nitrile-based IR probes into peptides and proteins via cysteine alkylation or arylation. Because this method involves relatively routine and mild reaction conditions, we expect that it will find wide application in biophysical studies of proteins.

Acknowledgement

We thank the NIH (GM-065978) and the NSF (DMR05-20020) for funding.

4.5 Original Publication

This Chapter has been reprinted from Biochemistry, Hyunil Jo*, Robert M. Culik*, Ivan V. Korendovych, William F. DeGrado and Feng Gai, (2010) 49, 10354-10356. DOI: 10.1021/bi101711a (* equal contributions), with permission from the American Chemical Society.

Probe	1	2	3	4
$\nu(\text{cm}^{-1}, \text{H}_2\text{O})$	2236.6	2240.3	2233.7	2237.9
$\Delta\nu(\text{cm}^{-1}, \text{H}_2\text{O})$	11.8	14.4	11.2	16.4
$\nu(\text{cm}^{-1}, \text{THF})$	2228.5	2232.3	2226.8	2229.6
$\Delta\nu(\text{cm}^{-1}, \text{THF})$	7.4	7.8	6.1	7.1

Table 4.1: Band position (ν) and full width at half maximum ($\Delta\nu$) of the $\text{C}\equiv\text{N}$ stretching vibration of various probes in H_2O and THF.

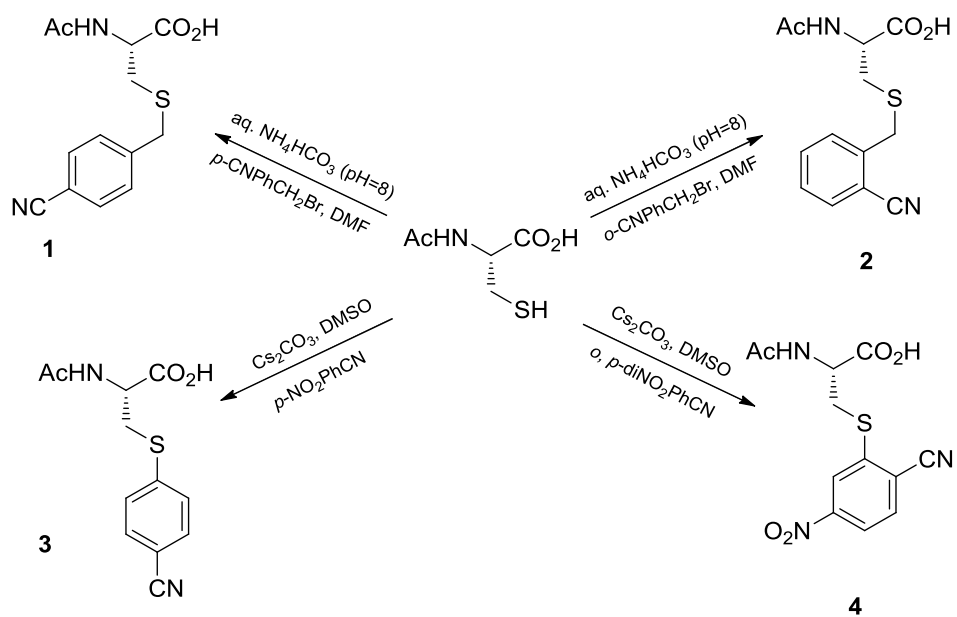


Figure 4.1: Overview of synthetic routes converting cysteine to vibrational probes of interest.

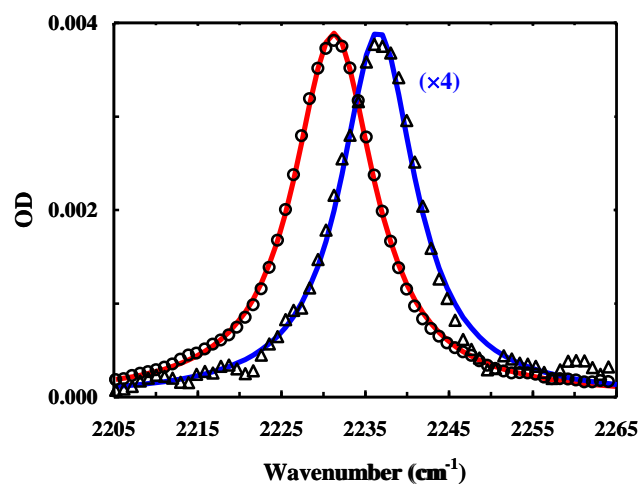


Figure 4.2: The C≡N stretching bands of W3C-CN obtained in the presence (open circles) and absence (open triangles) of CaM (50 mM HEPES buffer, pH 7.4, 30 mM CaCl₂). For the latter case, the concentrations of W3C-CN and CaM were estimated to be 1-2 mM. Lines are respective fits of these data to a Lorentzian function with the following parameters: for W3C-CN $\nu = 2236.5 \text{ cm}^{-1}$ and $\Delta\nu = 10.7 \text{ cm}^{-1}$, for W3C-CN/CaM $\nu = 2231.2 \text{ cm}^{-1}$ and $\Delta\nu = 11.7 \text{ cm}^{-1}$.

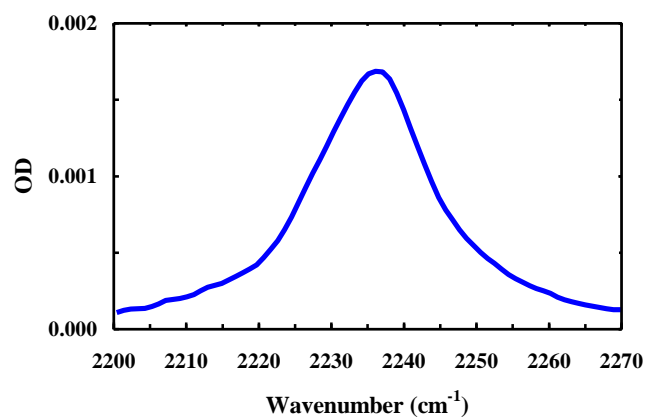


Figure 4.3: The C≡N stretching band of the nitrile-labeled human calmodulin-like protein (approximately 1-2 mM in 150 mM HEPES buffer, pH 7.4). The peak position of this band is consistent with the fact that the labeled cysteine residue is exposed to solvent.

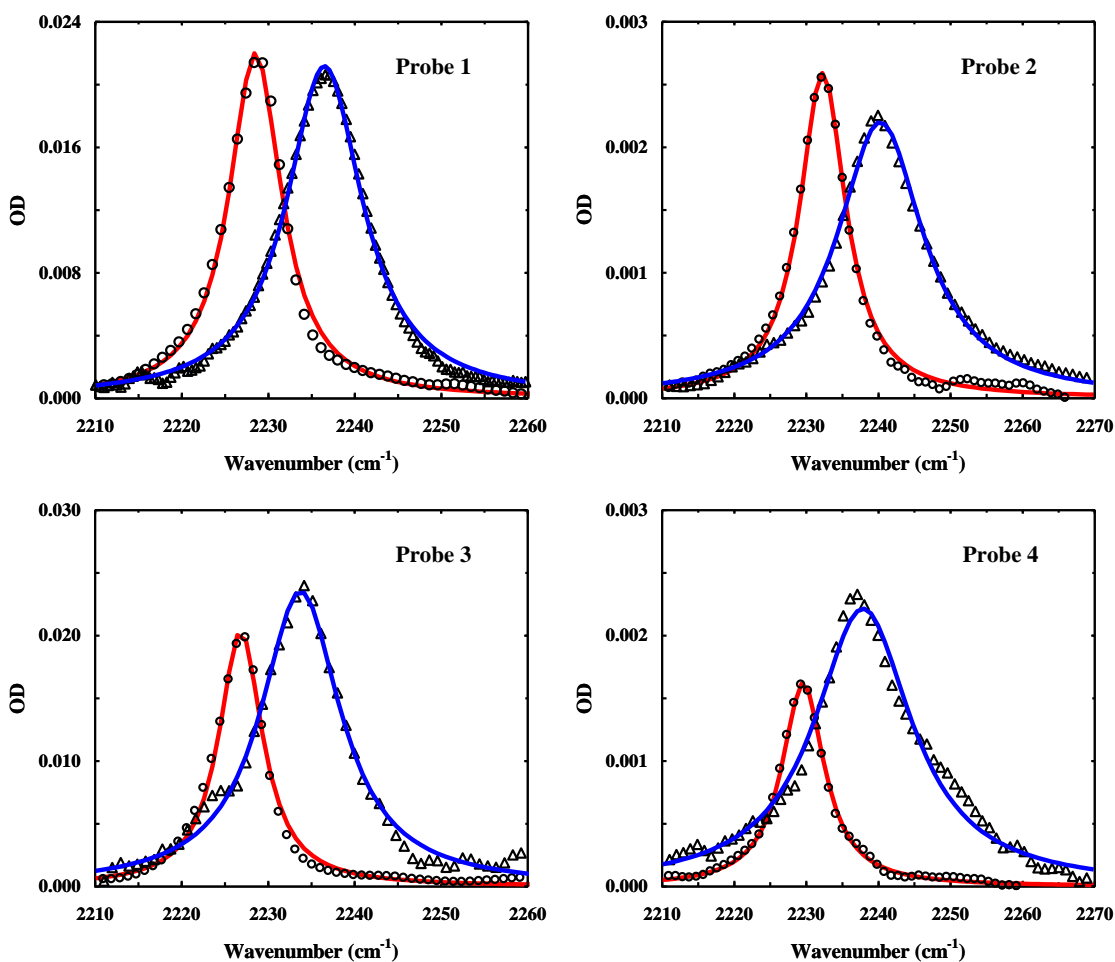


Figure 4.4: The C≡N stretching bands of the four cysteine derivatives, as indicated, obtained in water (open triangles) and THF (open circles). Lines are respective fits of these data to a Lorentzian function and the fitting parameters are given in Table 1. The sample concentrations were in the range of 5-10 mM, estimated by weight.

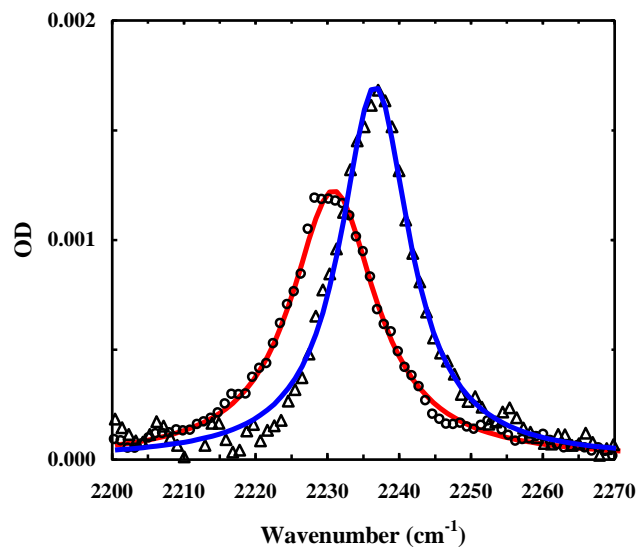


Figure 4.5: The C≡N stretching bands of A8C-CN obtained in the absence (open triangles) and presence (open circles) of CaM (50 mM HEPES buffer, pH 7.4, 30 mM CaCl₂). The concentrations of A8C-CN and CaM were estimated to be about 1-2 mM. Lines are respective fits of these data to a Lorentzian function with the following parameters: for A8C-CN $\nu = 2236.6 \text{ cm}^{-1}$ and $\Delta\nu = 11.7 \text{ cm}^{-1}$, for A8C-CN/CaM $\nu = 2230.8 \text{ cm}^{-1}$ and $\Delta\nu = 14.4 \text{ cm}^{-1}$.

CHAPTER 5

Achieving Secondary Structural Resolution in Kinetic Measurements of Protein

Folding: A Case Study of the Folding Mechanism of Trp-cage

5.1 Introduction

Protein folding kinetics are often measured by monitoring the change of a single spectroscopic signal, such as the fluorescence of an intrinsic fluorophore or the absorbance at a single frequency within an electronic or vibrational band of the protein backbone. While such an experimental strategy is easy to implement, the use of a single spectroscopic signal can leave important folding events undetected and overlooked. Herein, we demonstrate, using the mini-protein Trp-cage as an example, that the structural resolution of protein folding kinetics can be significantly improved when a multi-probe and multi-frequency approach is used, thus allowing a more complete understanding of the folding mechanism.

Trp-cage is a 20-residue mini-protein designed by Andersen and coworkers.[124] Among the many Trp-cage variants (the name and sequence of the Trp-cage peptides studied here are listed in Table 5.1), TC5b is the most studied, both experimentally and computationally. As shown (Figure 5.1), the folded structure of Trp-cage consists of three secondary structural elements: an α -helix from residues 2-8, a 3_{10} -helix consisting of residues 12-14, and a polyproline region spanning residues 17-19, which together generate a hydrophobic cage housing the peptide's sole tryptophan residue. Because of its small size and fast folding rate, Trp-cage has been an extremely popular model for computational studies of protein folding dynamics.[125-165] However, experimental investigations of the folding kinetics and mechanism of Trp-cage remain scarce. Using a temperature-jump (*T*-jump) fluorescence technique, Hagen and coworkers[166] showed that TC5b folds in about 4 μ s at room temperature, while an infrared (IR) *T*-jump study

by Bunagan *et al.* indicated that the P12W mutant of TC5b, or Trp²-cage, folds even faster.[167] In both cases, single-exponential relaxation kinetics were observed, suggesting that folding proceeds in a two-state manner. On the other hand, equilibrium unfolding studies provided evidence suggesting the existence of folding intermediates corresponding to a compact denatured state[168,169] and a partially folded state with maximal thermal stability of 20 °C.[170] Moreover, a large number of different folding pathways have been observed in computer simulations, including, for instance, the formation of an early intermediate where the hydrophobic core is bisected by the D9-R16 salt-bridge,[171] and the concurrent formation of the α -helix and the hydrophobic core,[142,150,151] among others.

Generating a conclusive experimental verification of these previous simulation results poses a great challenge to experimentalists, because the kinetic techniques commonly used in protein folding studies offer relatively low structural resolution. To overcome this limitation and to provide new insights into the folding mechanism of Trp-cage, we seek to use a multi-probe approach to dissect the folding kinetics of individual local structural elements of the native fold. To this end, we measure *T*-jump induced conformation relaxation kinetics[172] at well-chosen frequencies in the amide I' region of the protein that report the absorbance changes of the α -helix, the 3_{10} -helix, the unfolded structural ensemble, as well as the Asp sidechain. Separation of the α -helix IR signal from those arising from other structural motifs is facilitated by using the following Trp-cage sequence: DA*YA*QWLKDGGPSSGRPPPS (hereafter referred to as ¹³C-TC10b), where A* represents ¹³C=O labeled alanine, whose amide I' frequency is known to red-

shift by about 40 cm^{-1} from that of the unlabeled helical amides.[74,75,173] Andersen and coworkers have shown that this sequence, which is referred to as TC10b in their study, yields a more stable Trp-cage fold and is therefore a better model for both experimental and computational studies.[174] In addition, we employ several well-chosen mutations and ϕ -value analysis[26] in order to determine the structural elements formed in the folding transition state.

5.2 Experimental Section

The Trp-cage peptides were synthesized on a PS3 automated peptide synthesizer (Protein Technologies, MA) using Fmoc-protocols, purified by reverse-phase chromatography, and identified by matrix assisted laser desorption ionization mass spectroscopy. Trifluoroacetic acid (TFA) removal and H-D exchange were achieved by multiple rounds of lyophilization.

CD spectra and thermal melting curves were obtained on an Aviv 62A DS spectropolarimeter (Aviv Associates, NJ) with a 1 mm sample holder. The peptide concentration was in the range of 30-50 μM in 50 mM phosphate D_2O buffer solution (pH* 7).

Fourier transform infrared (FTIR) spectra were collected on a Magna-IR 860 spectrometer (Nicolet, WI) using a home made, two-compartment CaF_2 sample cell of 56 μm . [172] The detail of the *T*-jump IR setup has been described elsewhere.[172] The only difference is that in the current study a quantum cascade (QC) mid-IR laser (Daylight Solutions, CA) was used to probe the *T*-jump induced conformational relaxation kinetics, which significantly improved the signal-to-noise ratio of the kinetic data. The peptide

samples used in the IR measurements were prepared by directly dissolving lyophilized solids in 50 mM phosphate D₂O buffer (pH* 7) and the final peptide concentration was between 1-2.5 mM.

5.3 Results and Discussion

As shown (Figures 5.5 and 5.6 and Table 5.2), the thermal unfolding properties of the Trp-cage variants studied here, determined via circular dichroism (CD) spectroscopy, are in quantitative agreement with those reported in the literature.[167,174] For example, the thermal melting temperature (i.e., T_m) of ¹³C-TC10b is determined to be 55.0 ± 1.0 °C, comparing well with 56 °C for TC10b reported by Andersen and coworkers.[174]

In comparison with that of TC5b (Figure 5.7), the FTIR difference spectrum of ¹³C-TC10b (Figure 5.2) indicates that the negative spectral feature at ~ 1615 cm⁻¹ is due to the ¹³C-labeled Ala residues, thus uniquely reporting the thermal melting of the α -helical segment within the protein. The negative peak at ~ 1646 cm⁻¹ arises from the loss of unlabeled helical amides. The apparent blue-shift and lower intensity of the unlabeled helical amide I' band in the difference spectrum, in comparison with that observed for unlabeled Trp-cage, is due to spectral overlapping with the amide I' band of ¹³C=Os in the thermally denatured state.[173] On the other hand, the positive spectral feature arises from ¹²C=Os in the thermally unfolded state of ¹³C-TC10b. In addition, the negative feature at ~ 1586 cm⁻¹ is due to the absorbance change of the deprotonated Asp sidechain (i.e., $\nu_{as}(\text{COO}^-)$)[70] in response to protein unfolding. Since the salt bridge formed between the sidechains of D9 and R16 is a key structural determinant of the Trp-cage

stability and fold,[175] we believe that this spectral feature provides an excellent IR marker for probing the global folding/unfolding kinetics of the cage structure.[176]

As shown (Figure 5.3), the T -jump induced conformational relaxation kinetics probed at both 1580 and 1612 cm^{-1} can be adequately described by a single-exponential function and the corresponding rate constants, as indicated (Figure 5.4), are indistinguishable from each other within the limit of experimental errors. Interestingly, however, when probed at 1664 cm^{-1} , a frequency where both the 3_{10} -helix and disordered conformation are known to absorb,[177] the T -jump induced conformational relaxation kinetics can only be fit by two exponentials with amplitudes of opposite sign (Figure 5.3). As indicated (Figure 5.4), the rate constant of the positive (and slower) kinetic phase is also identical, within experimental uncertainty, to those measured at 1580 and 1612 cm^{-1} . Therefore, we attribute this kinetic phase to the global folding-unfolding transition of the Trp-cage structure. Consequently, we assign the fast phase, whose amplitude decreases with time, to the local unfolding of the 3_{10} -helix.

The assignment of the fast kinetic phase observed at 1664 cm^{-1} to T -jump induced conformational relaxation of the 3_{10} -helix is consistent with several lines of evidence. First, it has been shown that 3_{10} -helices absorb in the 1660 cm^{-1} region.[177,178] Second, the full amplitude of this phase decreases with increasing final temperature (for the same T -jump amplitude) and becomes practically undetectable when the final temperature is higher than ~ 20 °C (Figure 5.4). This result is consistent with the work of Asher and coworkers[170] as well as Day *et al.*,[164] both of which showed that the unfolding of a structural element that likely includes the 3_{10} -helix occurs at a temperature that is much

lower than the thermal melting temperature of the cage structure. Third, the relaxation rate of this kinetic phase is on the order of hundreds of nanoseconds, comparable to that observed for short α -helices.[173,179-182] Fourth, many molecular dynamics simulations carried out at 300 K[127,132,171,182] fail to reproduce the native 3_{10} -helix in the NMR structure determined at 285 K,[124] which suggests that this structural element is only stable at low temperatures (<25 °C). Finally, our finding is in accord with the computational study of Bolhuis and coworkers,[183] which showed that every unfolding trajectory in their molecular dynamics simulations begins with unfolding of the 3_{10} -helix. Moreover, the T -jump induced relaxation kinetics of both TC5b[166] and Trp²-cage[167] obtained at 1664 cm^{-1} also contain this fast kinetic phase (data not shown), indicating that it is not unique to ^{13}C -TC10b but rather reports the conformational relaxation of the 3_{10} -helix in each case. What is more interesting, however, is that for TC5b this negative phase is detectable only at final temperatures below ~ 12 °C, whereas for Trp²-cage the temperature range within which this phase is detectable is similar to that of ^{13}C -TC10b. Since the T_m of Trp²-cage is almost identical to that of ^{13}C -TC10b, but is approximately 15 °C higher than that of TC5b,[167] these results suggest that while the 3_{10} -helix can fold/unfold independently, its stability is to some extent affected by the stability of the cage. Similar to the observation that a nearby structural constraint can stabilize the helical structure of very short peptides,[184] the above correlation most likely reflects the constraining effect of the cage on the 3_{10} -helix.

The fact that the relaxation rates obtained at 1580 cm^{-1} and 1612 cm^{-1} are identical indicate that the α -helix and the cage are formed at the same rate (Figure 5.4). However,

these results alone are insufficient to establish whether the D9-R16 salt bridge is formed early, as suggested by many molecular dynamics simulations,[135,148,150,152,171,182] or on the downhill side of the major folding free energy barrier. To provide additional insights into the folding transition state of Trp-cage, we further conducted ϕ -value analysis.

Since the stability of the 3_{10} -helix is sufficiently low compared to that of the cage structure, the folding rates of the cage are obtained by analyzing the corresponding relaxation rates and CD thermal melting curves using a two-state model.[184] We first compare the folding rates of TC10b and its mutant R16K. As shown (Table 5.2 and Figure 5.8), while this mutation decreases the thermal melting temperature of the cage by more than 9 °C, the folding rate of the resultant peptide (i.e., TC10b-R16K) at 25 °C is $(1.9 \pm 0.4 \mu\text{s})^{-1}$, which, in comparison to the folding rate of $(1.6 \pm 0.3 \mu\text{s})^{-1}$ of the parent at the same temperature (Figure 5.4), leads to a ϕ -value of 0.1 ± 0.15 . This result indicates that the D9-R16 salt bridge has not been formed when folding reaches the transition state. Similarly, we find that the ϕ -value of the P19A mutant of TC10b is also essentially 0.0 ± 0.1 at 25 °C (Figure 5.9), indicating that the folding transition state of Trp-cage is not stabilized by interactions involving P19 and that the hydrophobic cage is formed at a later stage of the folding process. On the other hand, we find that the cage folding rate of TC5b at 25 °C is $(3.7 \pm 0.3 \mu\text{s})^{-1}$ (Figure 5.10). This leads to a ϕ -value of 1.16 ± 0.15 , indicating that the α -helix is fully formed in the transition state. Thus taken together, our ϕ -value results depict a Trp-cage folding mechanism wherein the formation of the α -helix directs folding toward the native state. In other words, those interactions

that stabilize the cage structure are only fully developed at the native side of the major folding free energy barrier. This folding mechanism is consistent with several simulations[142,150,151] and is further supported by the fact that monomeric α -helices can fold in 1-2 μ s.[185,186]

5.4 Conclusion

In summary, we demonstrate that much improved structural resolution can be achieved in protein folding kinetics studies using IR *T*-jump spectroscopy. This method combines several strategies: (a) using isotopically labelled amide groups to assess the conformational relaxation kinetics of a specific secondary structural element, (b) using sidechain absorption to probe the relaxation kinetics of a specific long-range tertiary interaction, and (c) scanning the probing frequencies across the amide I' band of the protein backbone to reveal relaxation events that occur with different rates. For Trp-cage, we find that the 3_{10} -helix unfolds at a temperature much lower than the global unfolding temperature of the cage structure, which is similar to the notion that protein folding occurs via step-wise assembly of structural foldons.[187] Using ϕ -value analysis, we further show that only the α -helix is formed in the folding transition state, which is in disagreement with most previous simulation studies.

Acknowledgement

We thank the National Institutes of Health (GM-065978, RR01348 and GM-008275) for funding. R.M.C. is a Structural Biology Training Grant Fellow.

5.5 Original Publication

This Chapter has been reprinted from *Angewandte Chemie International Edition*, Robert M. Culik*, Arnaldo L. Serrano*, Michelle R. Bunagan and Feng Gai, (2011) 50,10884-10887. DOI: 10.1002/anie.201104085 (* equal contributions), with permission from John Wiley and Sons.

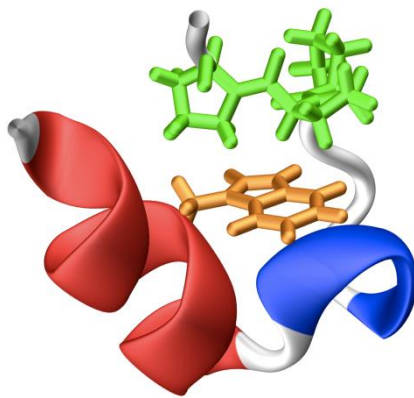


Figure 5.1: Structure of the Trp-cage (PDB code: 1L2Y), showing the α -helix (red), the 3_{10} -helix (blue), the polyproline region (green), and the sole tryptophan (orange).

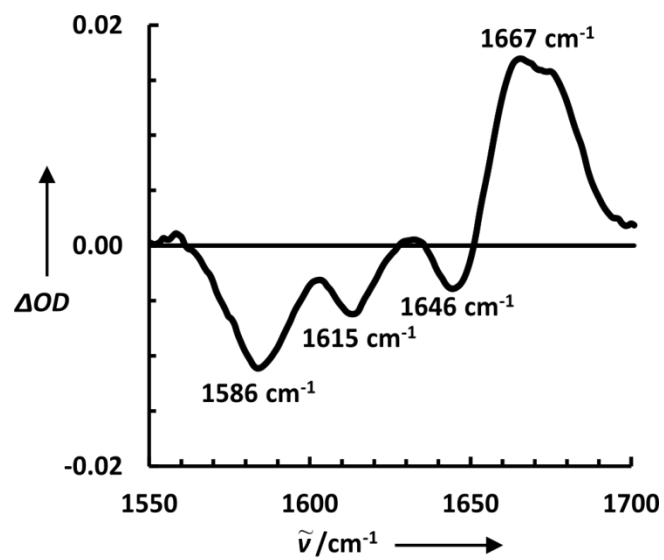


Figure 5.2: A representative FTIR difference spectrum of ^{13}C -TC10b between $65.0\text{ }^{\circ}\text{C}$ and $25.0\text{ }^{\circ}\text{C}$.

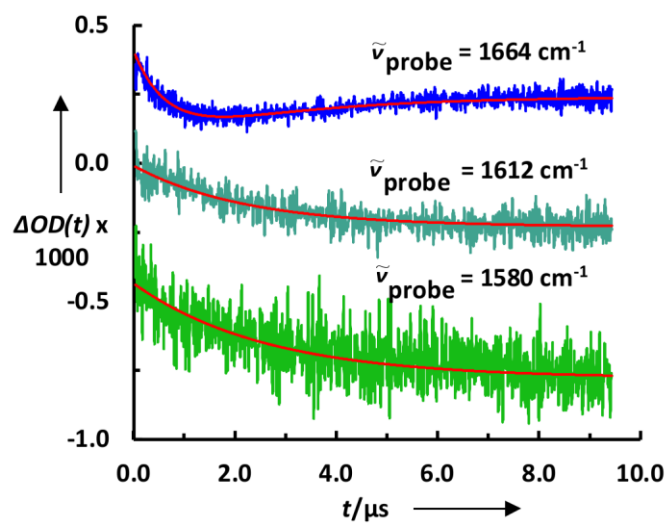


Figure 5.3: Representative T -jump induced conformational relaxation traces of ^{13}C -TC10b in response to a T -jump from 5 to 10 °C, probed at different frequencies as indicated. The smooth lines are the corresponding fits of these data to either a single-exponential (for 1580 cm^{-1} and 1612 cm^{-1}) or a double-exponential function (for 1664 cm^{-1}) and the resulting rate constants are given in Figure 4. For easy comparison, these data have been offset.

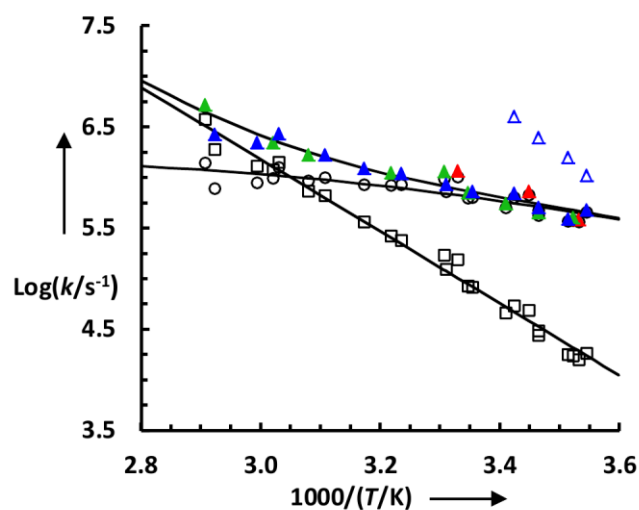


Figure 5.4: Conformational relaxation rate constants (solid symbols) of ^{13}C -TC10b obtained with a probing frequency of 1580 cm^{-1} (red), 1612 cm^{-1} (green), and 1664 cm^{-1} (blue), respectively. The blue open triangles represent the relaxation rates of the fast kinetic phase observed at 1664 cm^{-1} . The black open symbols represent the global folding (circle) and unfolding (square) rates of the protein.

Name	Sequence
TC5b	NLYIQWLKDGGPSSGRPPPS
TC10b	DAYAQWLKDGGPSSGRPPPS
¹³ C-TC10b	DA*YA*QWLKDGGPSSGRPPPS
TC10b-R16K	DAYAQWLKDGGPSSGKPPPS
TC10b-P19A	DAYAQWLKDGGPSSGRPPAS
Trp ² -cage	NLYIQWLKDGGWSSGRPPPS

Table 5.1: Name and sequence of the Trp-cage variants studied or discussed. A* represents ¹³C=O labeled alanine.

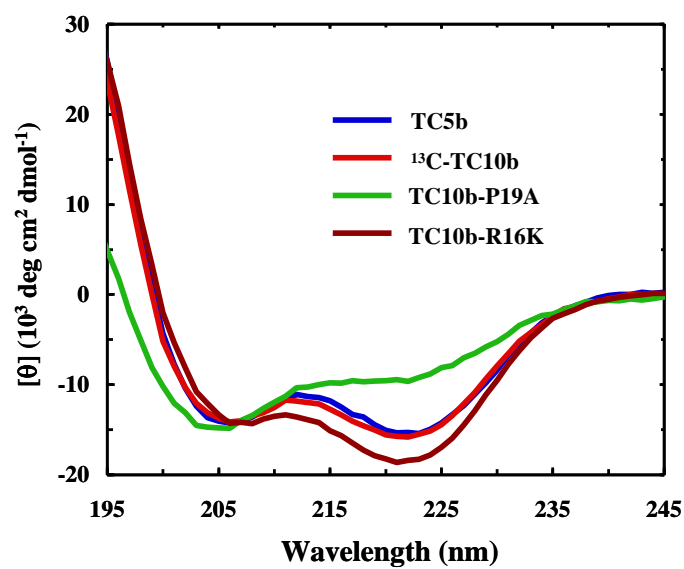


Figure 5.5: Far-UV CD spectra of the Trp-cage peptides studied, as indicated. These data were collected at 1.0 °C.

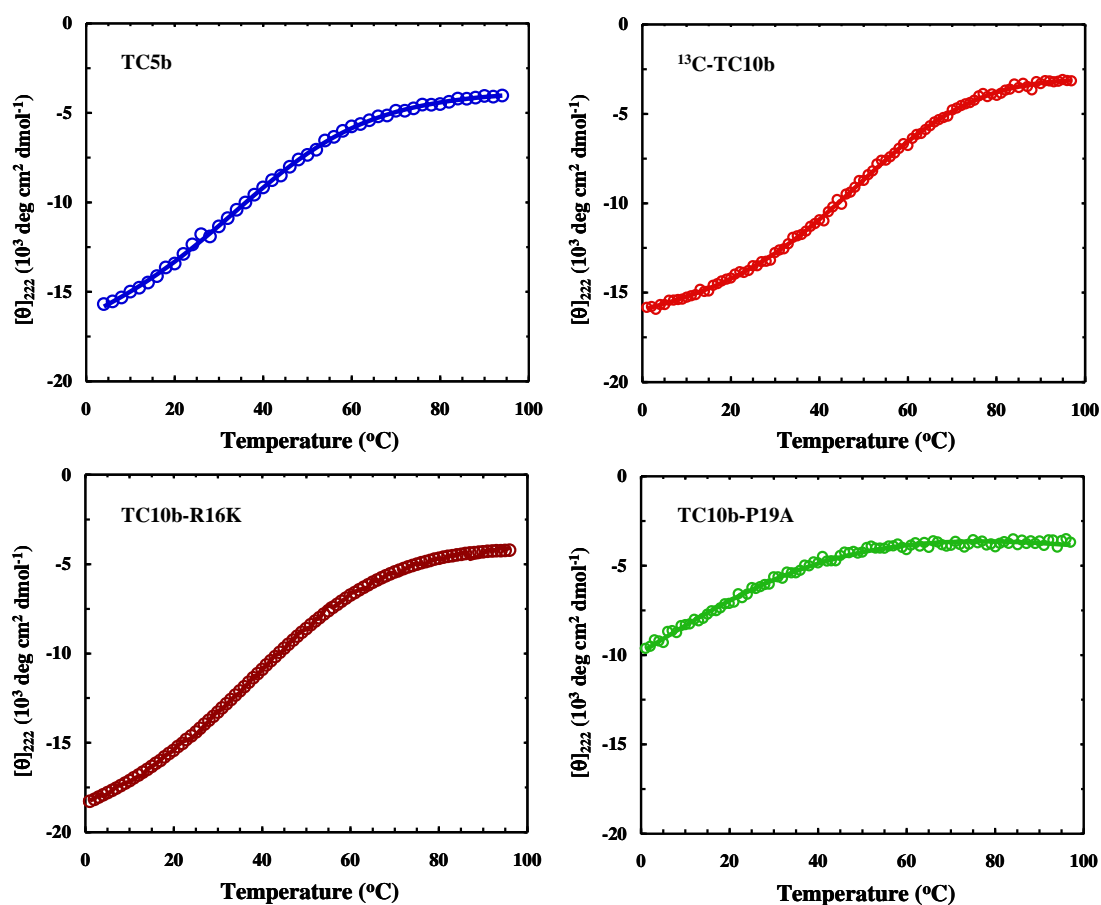


Figure 5.6: CD thermal melting curves (open circles) of the Trp-cage peptides studied, as indicated. Smooth lines represent the best fits of these data to a two-state model. For TC10b and its mutants, the slopes of the folded and unfolded CD baselines were treated as global fitting parameters. The resultant unfolding thermodynamic parameters are given in Table S2.

	TC5b	13C-TC10b	TC10b-R16K	TC10b-P19A
ΔH_m (kcal mol ⁻¹)	13.5	13.9	11.4	6.0
ΔS_m (cal K ⁻¹ mol ⁻¹)	42.7	42.2	35.8	21.0
ΔC_p (cal K ⁻¹ mol ⁻¹)	23.3	42.1	36.0	42.0
T_m (°C)	43.2	55.0	45.8	11.0

Table 5.2: Unfolding thermodynamic parameters of the Trp-cage peptides.

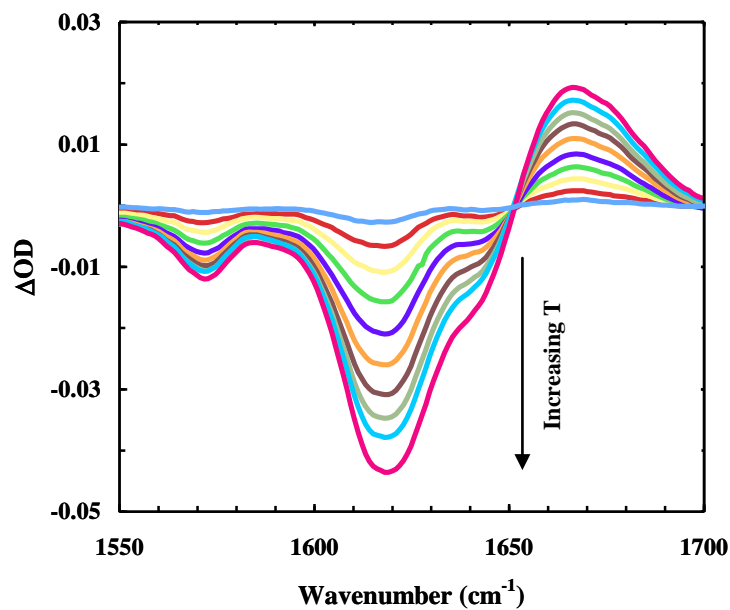


Figure 5.7: Difference FTIR spectra of TC5b in the amide I' region, which were generated by subtracting the FTIR spectrum obtained at 4.0 °C from those measured at higher temperatures (the highest temperature was 82.2 °C).

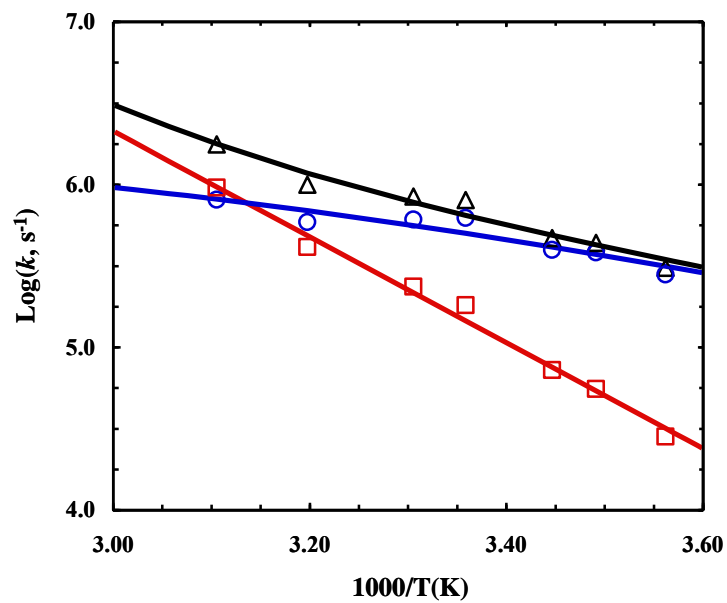


Figure 5.8: Temperature dependence of the slow relaxation rate (black), folding rate (blue), and unfolding rate (red) of TC10b-R16K. The probing frequency was 1664 cm^{-1} . The folding time at $25 \text{ }^\circ\text{C}$ was determined to be $1.9 \pm 0.4 \text{ } \mu\text{s}$, which yields a ϕ -value of 0.1 ± 0.15 compared to that of TC10B.

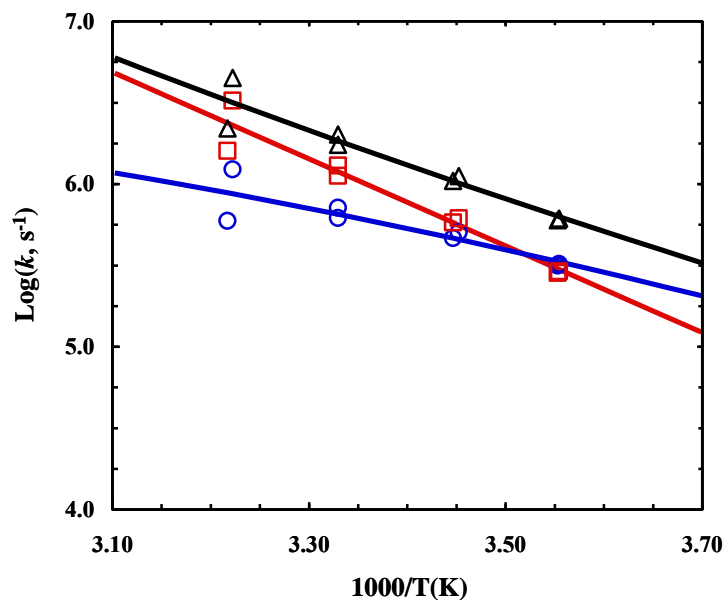


Figure 5.9: Temperature dependence of the slow relaxation rate (black), folding rate (blue), and unfolding rate (red) of TC10b-P19A. The probing frequency was 1664 cm^{-1} . The folding time at $25 \text{ }^\circ\text{C}$ was determined to be $1.6 \pm 0.4 \text{ } \mu\text{s}$, which, in comparison to that of TC10B, gives rise to a ϕ -value of 0.0 ± 0.1 .

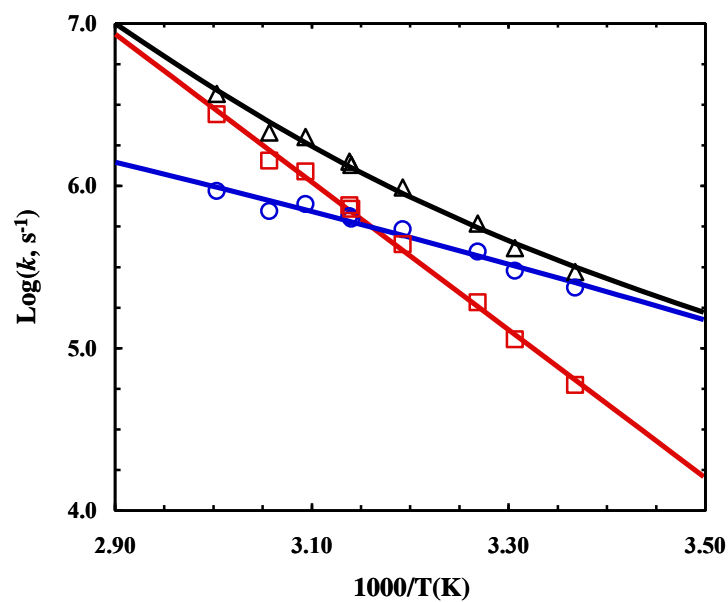


Figure 5.10: Temperature dependence of the slow relaxation rate (black), folding rate (blue), and unfolding rate (red) of TC5b. The probing frequency was 1630 cm^{-1} . The folding time at $25\text{ }^{\circ}\text{C}$ was determined to be $3.7 \pm 0.3\ \mu\text{s}$, which gives rise to a ϕ -value of 1.16 ± 0.15 compared to that of TC10b.

CHAPTER 6

Using Thioamides to Site-Specifically Interrogate the Dynamics of Hydrogen Bond Formation in β -Sheet Folding

6.1 Introduction

Folded (globular) proteins are characterized by two important structural features: a hydrophobic core consisting of well packed sidechains and an intricate network of backbone-backbone hydrogen bonds (hereafter referred to as BB-HBs). The latter exclusively determines the secondary structure content of the protein. Thus, in order to provide a comprehensive understanding of how a protein folds, one needs to determine the order and kinetics of its sidechain packing, as well as the temporal sequence of its BB-HB formation. Because site-directed mutagenesis is relatively easy and straightforward, almost all previous protein folding kinetic studies have relied on ϕ -value analysis[188] through sidechain perturbation to infer the underlying folding mechanism. In comparison, perturbing the energetics of an individual BB-HB is more difficult and, as a result, only a few experimental studies[189-193] have been performed, for example, using the technique of amide-to-ester (hereafter referred to as A-to-D) mutation to directly assess the role of BB-HB formation in the folding dynamics of proteins. Since esters are sterically different from amides and replacement of an amide with an ester completely eliminates a BB-HB, an A-to-D mutation could affect molecular packing and thus complicate interpretation of the experimental findings. Herein, we demonstrate an alternative approach, which only reduces the strength of the targeted BB-HBs, for mechanistic studies of protein folding.

Among the existing methods[189-193] for BB-HB mutations, replacing a backbone amide unit with a thioamide represents a distinctly advantageous approach to modulate the strength of targeted BB-HBs because thioamides are not only weaker hydrogen-bond

acceptors but also sterically very similar to their oxoamide counterparts.[194] Previous studies[195,196] suggested that an oxoamide-to-thioamide (hereafter referred to as O-to-T) mutation would decrease the protein's stability by about 1.6 kcal/mol, a value that is appropriate for ϕ -value analysis.[197] Furthermore, advances in synthetic chemistry have made it relatively easy to synthesize thioamide amino acids, especially those containing aliphatic sidechains,[198-200] making O-to-T mutational studies of protein folding more practical.[196]

The applicability of thioamides in the conformational study of helical proteins has recently been established.[196,197,201] Using β -hairpins as an example, herein we extend the utility of O-to-T mutations to interrogate the dynamic role of BB-HB formation in β -sheet protein folding. While there are a large number of experimental studies on the folding mechanism of β -hairpins,[202-208] a direct assessment of interstrand hydrogen bond formation in the folding transition state of β -hairpins has, to the best of our knowledge, never been done before. Specifically, we chose to study a variant of well-studied β -hairpins, tryptophan zippers (Trpzips), due to the large body of experimental and computational research[209-230] on their folding thermodynamics, kinetics, and mechanisms. As shown (Figure 6.1), this Trpzip variant (Trpzip-2c following Keiderling and coworkers,[224] sequence: NH₂-AWAWENGKWAWA-CONH₂) folds into an antiparallel β -sheet structure that is stabilized by several interactions, including six BB-HBs, among which three are perturbed in the current study by individually substituting Ala1, Ala10, and Glu5 with their thioamide derivatives, i.e.,

thioalanine (TA) and thioglutamate (TE), and the corresponding mutants are hereafter referred to as A1/TA, A10/TA, and E5/TE, respectively.

6.2 Experimental Section

Synthesis of Thioamide Amino Acids

Fmoc-Ala-diamino-nitrobenzene ((S)-(9H-fluoren-9-yl)methyl (1-((2-amino-5-nitrophenyl)amino)-1-oxopropan-2-yl)carbamate)

This compound was prepared from Fmoc-Alanine-OH and diaminonitrobenzene using HCTU/DIEA following the method of Dawson and coworkers. Specifically, a mixture of diisopropylethylamine (11.45 mL, 66 mmol) and diaminonitrobenzene (5.06 g, 33 mmol) was added to a DMF (100 mL) solution of Fmoc-Alanine-OH (10.0 g, 33 mmol) and HCTU (12.5 g, 88.7 mmol) at room temperature. After 1.5 h stirring, this reaction mixture was poured into a saturated KCl solution (250 mL). After addition of Et₂O (200 mL) to the mixture, the precipitated product was filtered and washed with water and Et₂O. The yellow solid (12.6 g, 83%) obtained after drying the precipitate under vacuum was used for the next step of synthesis without further purification.

Fmoc-thiocarbonyl-Ala-nitrobenzotriazole ((S)-(9H-fluoren-9-yl)methyl (1-(6-nitro-1H-benzo[d][1,2,3]triazol-1-yl)-1-thioxopropan-2-yl)carbamate)

This compound was prepared from the crude Fmoc-Ala-diamino-nitrobenzene following existing procedures. Briefly, anhydrous Na₂CO₃ (0.935 g, 8.8 mmol) was added to a suspension of P₂S₅ (3.92 g, 8.8 mmol) in anhydrous THF (200mL) and the resulting mixture was stirred at room temperature until a clear solution was obtained. The solution

was cooled to 0 °C, and then the Fmoc-Ala-diamino-nitrobenzene (6.0 g, 13.45 mmol) obtained from the previous step was added to it. The mixture was stirred for 2 h and filtered. The filtrate was concentrated under reduced pressure and the residue was dissolved in ethyl acetate and washed with 5% NaHCO₃. After drying over Na₂SO₄, the crude mixture was concentrated again under reduced pressure to yield the orange, oil-like product (3.4 g, 55%), which was subsequently dissolved in 95% AcOH (10 mL) and cooled down to 0 °C. NaNO₂ (760 mg, 11 mmol) was then added portionwise and stirred at 0 °C for 1.5 h. Precipitation of the product was accomplished by addition of ice-cooled water (60 mL) to the above mixture; the precipitate was then filtered and washed with ice-water. The crude product was further dried by lyophilization under protection from light to yield a yellow-orange solid (1.37 g, 60%).

Fmoc-thiocarbonyl-Glu-nitrobenzotriazole was prepared similar to Fmoc-Glu(OtBu)-OH except that the final product was isolated by extraction with dichloromethane instead of precipitation.

The Trpzip-2c peptides were synthesized on a PS3 automated peptide synthesizer (Protein Technologies, MA) using Fmoc-protocols, except for incorporation of Fmoc-thioalanine-nitrobenzotriazole or Fmoc-thioglutamate-nitrobenzotriazole, which was added manually (0.4 mmol scale) to the deprotected, resin-bound peptide chain in 4 ml of dimethylformamide and coupled for 45 minutes. Thioamide peptides were cleaved in a modified cocktail of trifluoroacetic acid, water, and triisopropylsilane (14:5:1, v/v). Peptide products were further purified by reverse-phase chromatography, and identified

by matrix assisted laser desorption ionization mass spectroscopy. Trifluoroacetic acid (TFA) removal and H-D exchange were achieved by multiple rounds of lyophilization.

All peptide samples were prepared in 20 mM phosphate buffer solution (pH 7) and the peptide concentrations were in the range of 30-100 μ M for CD and 1-2 mM for IR measurements. The details of all spectroscopic measurements, including the T -jump IR setup, have been described elsewhere.[172]

Global Fitting of the CD Thermal Unfolding Curves to a Two-State Model

Because of their decreased stabilities, the CD unfolding curves of the Trpzip-2c mutants lack well defined folded baselines. Thus, to better estimate their unfolding thermodynamics, we globally fit the CD denaturation curves of the wild type Trpzip-2c and mutants to the following two-state model:

$$[\theta](i, T) = \frac{[\theta]_F(i, T) + K_{eq}(i, T) \times [\theta]_U(i, T)}{1 + K_{eq}(i, T)} \quad (6.1)$$

where $[\theta](i, T)$ is the temperature-dependent mean residue ellipticity for peptide i , and $[\theta]_F(i, T)$ and $[\theta]_U(i, T)$ are the corresponding folded and unfolded baselines, respectively. In the current study, both $[\theta]_F(i, T)$ and $[\theta]_U(i, T)$ were assumed to be linearly dependent on temperature:

$$[\theta]_F(i, T) = m + n \times T \quad (6.2)$$

$$[\theta]_U(i, T) = p(i) + q \times T \quad (6.3)$$

where m , n , and q are constants and were treated as global fitting parameters, whereas $p(i)$ was peptide dependent and treated as a local fitting parameter. In this model, we assume that the folded baseline offset is the same for all peptides. This implies that the final folded state of each Trpzip variant is similar enough that CD measurements would give the same signal readout for each variant under ~100% folded conditions. Based on the excitonic coupling argument in our paper, we propose that thioamide mutation does not significantly perturb the final native structure of the Trpzip peptides, and thus we believe that this is a reasonable assumption. In eq. (6.1), $K_{eq}(i, T)$ is the folding equilibrium constant of peptide i , which is determined by the free energy change (ΔG) between the unfolded and folded states, namely,

$$K_{eq}(i, T) = \exp(-\Delta G(i, T) / RT) \quad (6.4)$$

with

$$\Delta G(i, T) = \Delta H_R(i) + \Delta C_p \cdot (T - T_R) - T \cdot [\Delta S_R(i) + \Delta C_p \cdot \ln(T / T_R)] \quad (6.5)$$

where $\Delta H_R(i)$ and $\Delta S_R(i)$ are the enthalpy and entropy changes for peptide i at the reference temperature T_R , which was set to 298.15 K in the current study. In addition, we have assumed that the heat capacity change of the folding transition, ΔC_p , is temperature and peptide independent.

6.3 Results and Discussion

The thioamide amino acids were synthesized using established methods described in detail above. As shown (Figure 6.4), at 1.0 °C both the wild type and thioamide mutants of Trpzip-2c show the characteristic far-UV CD spectrum of Trpzips,[209] with a distinctive positive band centered at 227 nm that arises from the excitonic coupling

between the B_{\parallel} transitions of tryptophan residues.[231] In comparison to that of the wild type, however, the mean residue molar ellipticities of the mutants at 227 nm are lower, indicating that the O-to-T mutation in all cases decreases the β -hairpin stability, as expected. Moreover, because the underlying excitonic coupling is sensitive to both the distance and orientation of the tryptophan residues in the folded state, the fact that the 227 nm band only changes its intensity upon O-to-T mutations suggests that thioamide incorporation does not significantly perturb the β -hairpin structure. Following our previous studies,[232,233] we further quantify the unfolding thermodynamics of these Trpzip-2c peptides by globally fitting all the CD thermal melting curves obtained at 227 nm to a two-state model[224] (see the Experimental Section above for details), wherein the folded CD baselines and ΔC_p for unfolding of the mutants are assumed to be the same as those of the wild type. Because the CD thermal unfolding curves of the thioamide mutants lack folded baselines, such a stringent global fitting constraint is necessary to enable us to best estimate the folding/unfolding thermodynamics of the mutants. As indicated (Figure 6.2 and Table 6.2), the thermal melting temperature (T_m) of the wild-type Trpzip-2c is quantitatively consistent with that reported by Keiderling and coworkers,[224] whereas those of the thioamide mutants show different degrees of decrease, depending on the position of the BB-HB that is perturbed. It is clear that the energetic destabilization arising from thioamide mutation is greatly reduced when a BB-HB close to the terminal region is perturbed, consistent with several previous studies indicating that the ends of β -hairpins are frayed.[227,234,235]

To determine how these O-to-T mutations change the folding and unfolding rates of the β -hairpin, we further measured the relaxation rates of these peptides in response to a laser-induced temperature-jump (*T*-jump)[236] using time-resolved infrared (IR) spectroscopy.[237]⁵⁰ As shown (Figure 6.5), the relaxation kinetics of these peptides can be well described by a single-exponential function, consistent with previous studies.[238-240] As shown (Figure 6.3), the folding rate constant of wild type Trpzip-2c, determined from the corresponding relaxation rate constant (k_R) and folding thermodynamics, shows a modest dependence on temperature in the temperature range of the experiment. Specifically, the folding rate constant at 25.0 °C is determined to be $(3.2 \pm 0.7 \mu\text{s})^{-1}$, comparable to the $(2.5 \mu\text{s})^{-1}$ measured for Trpzip2.[238] Since Trpzip2 and Trpzip-2c share the same turn sequence, this result alone provides additional evidence supporting our previous conclusion that the turn sequence plays a key role in determining the folding rate of β -hairpins.[239-241]

As indicated (Figure 6.3 and Table 6.1), the folding rates of A1/TA and A10/TA are almost identical within experimental uncertainty to that of the wild type. In contrast, their unfolding rates at any given temperature are significantly faster than that of the wild type. Taken together, these results indicate that the reduced stability of these mutants arises almost exclusively from a smaller unfolding free energy barrier. In other words, the corresponding native interstrand hydrogen bonds perturbed by the O-to-T mutation are formed at the downhill side of the folding free energy barrier.

More interestingly, the E5/TE mutant, which is designed to decrease the strength of the first BB-HB next to the turn region of the peptide, has a more dramatic effect on both the

relaxation and folding rates of the β -hairpin (Figure 6.3 and Figure 6.5). For example (Table 6.1), at 25 °C its relaxation rate constant is decreased to about $(17.5 \mu\text{s})^{-1}$, compared to $(3.0 \mu\text{s})^{-1}$ for the wild type, whereas its folding rate is slowed down by nearly a factor of 32. While this mutation also affects the unfolding rate of the β -hairpin, it leads to an increase in the unfolding rate constant only by a factor of approximately 4. Thus, these results together provide strong evidence indicating that the BB-HB immediately next to the turn region stabilizes both the folding transition state and the native state. While such a folding mechanism has been observed in simulations,[220] the current study provides direct experimental insights into the mechanistic detail of BB-HB formation in β -hairpin folding. The comparatively smaller but significant increase in the unfolding rate further suggests that either this BB-HB is not entirely native-like in the transition state ensemble or the corresponding thioamide mutation induces a local distortion to the transition state conformation.

Thus, taken together our findings demonstrate that the O-to-T mutation method is a useful technique for site-specifically probing the dynamics of hydrogen bonding formation and indicate that for β -hairpins the BB-HBs located beyond the turn region stabilize the folded state only by increasing the unfolding free energy barrier. In other words, they are formed not in but after the folding transition state. On the other hand, the BB-HB directly next to the β -turn is formed in the transition state, potentially acting as a staple to hold the turn region together for further propagation of folding down the two strands. This mechanistic picture is entirely consistent with our previous notion that for β -hairpin folding the turn formation is the rate limiting step and that the native

hydrophobic cluster is only formed at the downhill side of the major folding free energy barrier, when folding begins from thermally denatured states.[239,240,242] It is worth pointing out, however, that the O-to-T mutation strategy described here is insensitive to probing alternative folding pathways that, for example, involve an initial collapsing step arising from sidechain-sidechain interactions.

6.4 Conclusion

We demonstrate that thioamide mutation provides a site-specific means to interrogate the role of backbone-backbone hydrogen bonds in protein folding dynamics. Application of this method to a model β -hairpin allows us to pinpoint which hydrogen bond is formed in the folding transition state, yielding a conclusion fully supported by several previous studies.[220,239-241] We believe that this method is a useful addition to the existing experimental toolkit used in protein folding studies and will find many important applications.

Acknowledgement

We thank the National Institutes of Health (GM-065978 and GM-008275) for funding.

R.M.C. is an NIH Ruth Kirschstein Predoctoral Fellow.

6.5 Original Publication

This Chapter has been reprinted from the Journal of the American Chemical Society,

Robert M. Culik, Hyunil Jo, William F. DeGrado and Feng Gai (2012) 134, 8026-8029.

DOI: 10.1021/ja301681v, with permission from the American Chemical Society.

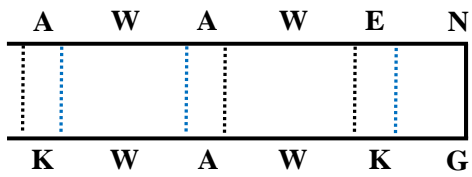


Figure 6.1: A cartoon representation of the β -hairpin structure of Trpzip-2c with the BB-HBs shown (dotted lines). BB-HBs that are perturbed using O-to-T mutation are shown in blue.

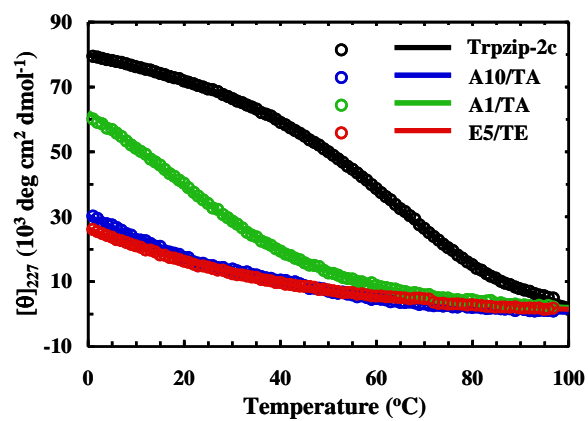


Figure 6.2: CD thermal melting curves of the Trpzip-2c peptides, as indicated. Smooth lines are global fits of these data to a two-state model discussed in the text.

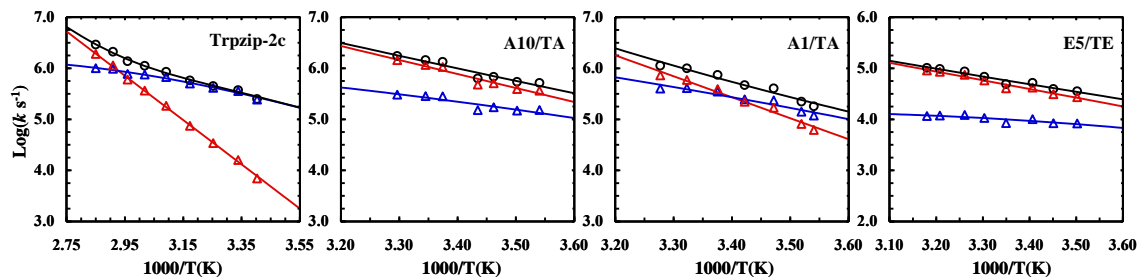


Figure 6.3: Arrhenius plot of the relaxation (black), folding (blue), and unfolding (red) rate constants of Trpzip-2c and mutants.

Peptide	τ_R (μs)	τ_f (μs)	τ_u (μs)
Trpzip-2c	3.0 ± 0.7	3.2 ± 0.7	79.4 ± 18
A10/TA	0.8 ± 0.2	3.8 ± 0.5	1.0 ± 0.5
A1/TA	1.3 ± 0.2	3.0 ± 0.6	2.4 ± 0.6
E5/TE	17.5 ± 2.5	100.6 ± 14	21.2 ± 3.0

Table 6.1: Relaxation, Folding, and Unfolding Times at 25 °C.

Peptide	ΔH_R (kcal mol ⁻¹)	ΔS_R (cal mol ⁻¹ K ⁻¹)	ΔC_P (cal mol ⁻¹ K ⁻¹)	T_m (°C)
Trpzip-2c	13.9	40.2	52.7	68.4
A10/TA	5.9	22.4	52.7	-20
A1/TA	9.6	32.7	52.7	21.0
E5/TE	5.3	20.8	52.7	-37

Table 6.2: Unfolding thermodynamic parameters obtained from the aforementioned global fitting for each peptide ($T_R = 25$ °C).

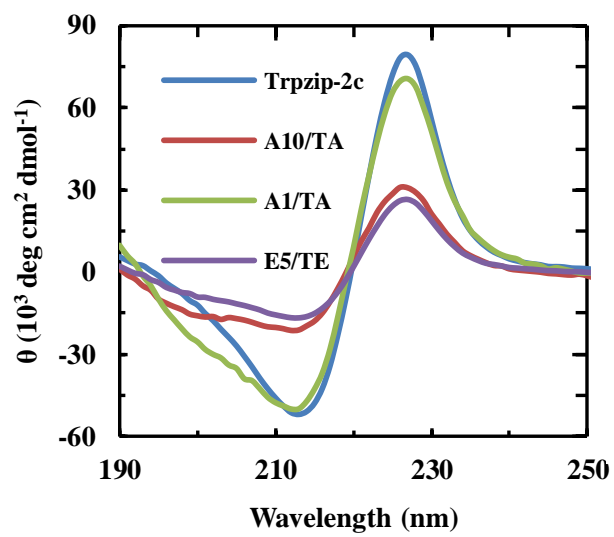


Figure 6.4: Far-UV CD spectra of the Trpzip-2c peptides at 1 °C, as indicated.

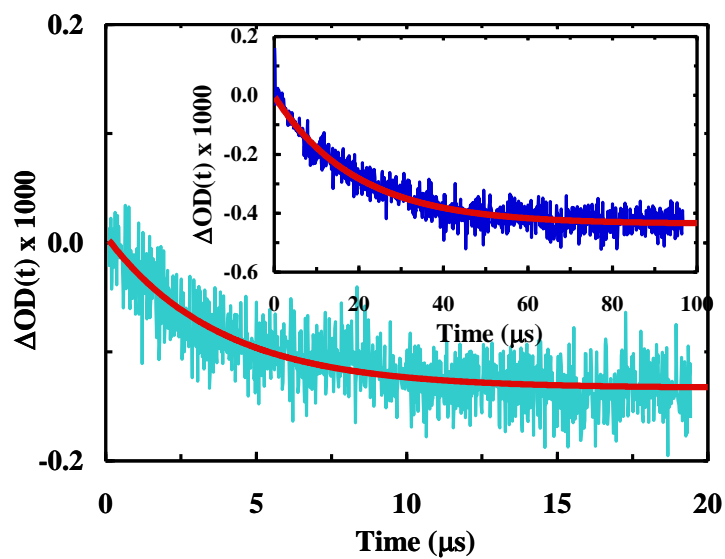


Figure 6.5: A representative T -jump induced relaxation trace of Trpzip-2c in response to a temperature change from 16.2 to 20.7 °C monitored at 1630 cm^{-1} . Shown in the inset are the conformational relaxation kinetics of E5/TE in response to a T -jump, from 12.6 to 20.4 °C. The smooth lines in red are single-exponential fits of these data, and the resultant rate constants are reported in Figure 3. It is apparent that at the same final temperature, the conformational relaxation of the mutant occurs much slower than that of the wild type.

CHAPTER 7

**Using D-Amino Acids to Delineate the Mechanism of Protein Folding: Application to
Trp-Cage**

7.1 Introduction

Loops and turns in proteins not only ensure connectivity of the polypeptide chain, but they also, in many cases, play an important role in directing the process of folding. A survey of the available protein structures in the Protein Data Bank (PDB) indicated that the most frequently found amino acids in loops are glycine (Gly), serine (Ser), aspartate (Asp), asparagine (Asn), and proline (Pro).[243] In particular, Gly is frequently found to be at the C-terminus of α -helices in proteins, serving as a secondary structure breaker and directing the main chain in a new direction. This is due to the fact that Gly can extensively populate backbone conformations that are normally forbidden for other L-amino acids, as seen in the Ramachandran plot. For the same reason, however, there is a relatively large conformational entropy penalty associated with restricting a Gly residue in a small region of phi and psi conformations. Therefore, Gly has recently become a popular target in protein design to improve the stability of proteins of interest via mutations that could alleviate this entropy penalty.[244] For example, it has been shown that Pro, which has low backbone conformational entropy, can in some cases replace Gly and, as a result, further stabilize the native fold.[245] A different, but perhaps more elegant and more effective approach, is to replace Gly with various D-amino acids,[244,246-248] the enantiomeric relative of L-amino acids, which can occupy conformations on the Ramachandran plot that otherwise only glycine could. For example, D-alanine (D-Ala), which has no sidechain rotamer preferences and hence a smaller entropy penalty upon folding, has been used, in a number of cases, to increase the

stability of proteins,[246,248] to reconstitute an active and specific ion channel,[247] and to create novel folds.[249]

In a more recent study,[250] Rodriguez-Granillo *et al.* have shown that D-Ala, D-Asn, and D-Gln all show a stabilizing effect, although to a different extent, on the folding free energy of the mini-protein Trp-cage (Figure 7.1), when replacing the Gly residue at position 10 (Specifically, the TC5b variant, sequence: NLYIQWLKDGGPSSGRPPPS, was used). Interestingly, both D-Ala and D-Gln were found to significantly stabilize the native fold, raising the thermal melting temperature (T_m) of Trp-cage by about 20 °C, whereas D-Asn only increased the T_m by about 10 °C. This observation was interpreted as due to a combination of effects, including the difference in sidechain-backbone hydrogen-bonding interaction ability and conformational entropy of these D-amino acids. The hydrogen bonding ability of the D-Gln sidechain was found to be greater than that of D-Asn, while the sidechain conformational entropy penalties for both D-Gln and D-Asn are larger than for D-Ala. While this study demonstrated the effectiveness, especially when placed at the C-terminus of α -helices, of several D-amino acids to alleviate the intrinsic conformational limitations of naturally occurring L-amino acids, and offered a thermodynamic rationalization of the stabilizing effect of these D-amino acids on the Trp-cage fold, it was unable to yield a mechanistic and dynamic view of how this is achieved, i.e., through an increase in the folding rate, a decrease in the unfolding rate, or even a change in the folding mechanism. Herein, we use a laser-induced temperature-jump (T -jump) infrared (IR) technique to study the conformational dynamics of several previously characterized D-amino acid mutants of TC5b, including 5b-G10DA, 5b-

G10DN, and 5b-G10DQ, aiming to elucidate the kinetic role of these D-amino acids. In addition, Culik et al. have recently put forward a folding mechanism for Trp-cage, based on extensive spectroscopic studies and amino acid mutations,[237] in which the transition state of the main folding event, i.e., the formation of the cage structure, involves the consolidation of the α -helix, from amino acids 2-9. Since the abovementioned D-amino acids are expected to provide further stabilization to the α -helix,[246,248] studying the folding kinetics of these D-amino acid mutants will help improve our understanding of the folding mechanism of this extremely popular protein folding model.[124-131,133-135,137-141,143-150,152-158,164,166,168,170,171,174,175,183]

7.2 Experimental Section

Peptides were synthesized using standard solid-phase Fmoc methods at the Tufts University Core Facility (<http://www.tucf.org>). All peptides were acetylated at the N-terminus and amidated at the C-terminus. Peptides were purified using reversed-phase high-performance liquid chromatography (HPLC), and peptide identities were verified using mass spectrometry. Trifluoroacetic acid (TFA) removal and H-D exchange were achieved by multiple rounds of lyophilization. Fourier transform infrared (FTIR) spectra were collected on a Magna-IR 860 spectrometer (Nicolet, WI) using a homemade, two-compartment CaF₂ sample cell of 56 μ m. The details of the *T*-jump IR setup have been described elsewhere.[236] For both static and time-resolved IR measurements, the peptide sample was prepared by directly dissolving lyophilized solids in 20 mM phosphate D₂O buffer (pH* 7) and the final peptide concentration was between 1-2 mM.

7.3 Results and Discussion

A number of studies have shown that the short 3_{10} -helix in the Trp-cage fold is much less stable than the global cage structure[164,170,237,] and its folding can proceed independently and with a faster rate.[237] Therefore, in the following discussion, we focus only on the effect of the corresponding D-amino acid mutations on the two-state folding and unfolding kinetics of the cage, which consists of several key structural elements: an α -helix, a polyproline stretch, and a hydrophobic core that encloses a tryptophan residue (Figure 7.1). In addition, all the thermodynamic parameters used to analyze the T-jump-induced relaxation kinetics are taken from a previous study[250] and are summarized in Table 7.1.

5b-10GDQ. As indicated (Table 7.1), substitution of Gly10 in TC5b with D-Gln significantly improves Trp-cage stability, increasing its T_m from 45.9 to 69.3 °C. It has been hypothesized and subsequently computationally verified that this stabilizing effect of D-Gln arises from, among other factors, its ability to form hydrogen bonds between the backbone amine of D-Gln and the backbone carbonyl of Leu7, and another between the backbone carbonyl of D-Gln and the backbone amine of Ser13. From a kinetic perspective, however, the higher stability of *5b-10GDQ*, compared to that of the wild type, could result from either a faster folding rate or a slower unfolding rate, or both. Since the CD spectrum of *5b-10GDQ* indicates that it has a higher helical content than the wild type,[250] it is reasonable, based on the notion that the α -helix is populated in the folding transition state of the cage,[237] to assume that the α -helix stabilizing effect of D-

Gln would lower the free energy of the transition state and, as a result, speed up the rate of folding. As shown (Figure 7.2), however, the folding rate of 5b-10GDQ is, within our experimental uncertainties, identical to that of the wild type, indicating that the stabilizing interactions stemming from the D-Gln mutation only start to develop at the downhill side of the major folding free energy barrier. In other words, these interactions act primarily to decrease the free energy of the folded state, making the protein unfold slower as a result of an increased unfolding free energy barrier, as observed (Figure 7.2). This picture is entirely consistent with earlier studies[126,129,134,135,137,143,164,237] that indicated that the developments of several key structural elements that involve residues near this part of the sequence, e.g., the formation of the D9-R16 salt-bridge and the burial of the Trp6 sidechain, occurs after the folding transition state. Furthermore, this result is in accord with the observation that stabilizing sidechain-sidechain and/or sidechain-backbone interactions can decrease the unfolding rate of α -helices in isolation.[186]

5b-10GDN. Similar to D-Gln, D-Asn was also found to form backbone-backbone hydrogen bonds with Leu7 and Ser13. However, its stabilizing effect is less pronounced compared to D-Gln (Table 7.1). For example, the T_m of 5b-10GDN was determined to be 56 °C, only 10 °C higher than the wild type. This improvement in stability has been attributed to the difference in backbone hydrogen bonding ability of D-Asn, in comparison to that of D-Gln,[250] due to the difference in sidechain-backbone sterics.[245] Despite this thermodynamic difference between D-Gln and D-Asn however, we expect that these amino acid substitutions would exert similar impacts on the

folding/unfolding kinetics of the cage because of their structural similarity. Indeed, as shown (Figure 7.3), substitution of Gly10 with D-Asn does not, within our experimental uncertainty, change the folding rate of TC5b; instead it decreases the unfolding rate. Previous studies have shown that the α -helix of TC5b in isolation is unfolded.[175] Thus, folding of this α -helix in the absence of other native stabilizing interactions would place the system at a higher free energy position, as concluded previously.[129,137,143,164,237] However, such a free energy ‘penalty’ is warranted and necessary, as it prepares and ensures the polypeptide chain to rapidly evolve toward the native state. Since D-Gln and D-Asn are able to act, in the current case, as a C-terminal helix cap, one could have assumed that both would be able to lower the folding free energy barrier, making the protein fold faster. The fact that this is not the case suggests that these mutations involve interactions that stabilize not only the α -helix, through capping, but also other parts of the cage structure. This is consistent with previous MD simulations of these mutants, which found that backbone fluctuations in the D-amino acid mutants are reduced, especially from residues 9 to 15.[250] However, these ‘nonlocal’ interactions are expected to be intimately coupled to the structural evolution of the cage, not the α -helix. Thus, the native rotameric state of these D-amino acids in the corresponding mutants would be expected to be significantly populated only after the major folding free energy barrier.

5b-10GDA. As shown (Table 7.1), D-Ala is almost as effective as D-Gln in stabilizing the Trp-cage fold at position 10. One of the factors this has been attributed to is that D-Ala

has less sidechain conformational entropy than both D-Gln and D-Asn, which highlights the well-known importance of entropy-enthalpy compensation in protein folding and design. In light of the above discussion and the fact that D-Ala does not have a sidechain rotamer preference, one might expect that this mutation would lead to an increase the folding rate, in addition to a decrease in the unfolding rate. As shown (Figure 7.4), indeed, while the unfolding rate of 5b-10GDA is significantly slower than TC5b, similar to that observed for 5b-10GDQ and 5b-10GDN, its folding rate also shows a marginal, but detectable, increase. These results thus confirm the above hypothesis that, unlike D-Gln and D-Asn, D-Ala helps to stabilize, although only to a small extent, the folding transition state or the nascent α -helix. A simple calculation indicates that at 30 °C the D-Ala mutation stabilizes the folding transition state by approximately 0.33 kcal/mol, suggesting that D-Ala already becomes partially native in the transition state, again owing to the relatively small number of configurations the residue can sample. Thus, taken together, the results obtained from these D-amino acids demonstrate their utility in helping reveal certain fine details of the folding mechanism of proteins of interest.

7.4 Conclusion

Many previous studies have elegantly demonstrated the versatility of D-amino acids in protein design, in particular to increase the conformational stability of the targeted folds. However, what is lacking in our current understanding of the effectiveness of specific D-amino acids in this regard is their kinetic mechanism of action. Herein, we use the miniprotein Trp-cage as a model to examine the role of three D-amino acids, i.e., D-Gln, D-Asn, and D-Ala in altering the folding/unfolding kinetics of this protein, based on an

earlier study that showed that these D-amino acids can significantly increase the stability of Trp-cage, when replacing the Gly residue in the loop region at position 10. Our results show that these mutations almost exclusively reduce the unfolding rate of the Trp-cage, although the D-Ala mutation also shows a marginal increase in the folding rate. In light of the ability of these D-amino acids to stabilize the α -helix in this miniprotein via a C-capping effect, these results indicate that this α -helix, while being populated in the folding transition state as suggested by a previous study, is nucleated at its N-terminus when passing through the folding free energy barrier.

Acknowledgements

We thank the National Institutes of Health (GM-065978, RR01348, GM-008275, and GM-089949) for funding. R.M.C. is an NIH Ruth Kirschstein Predoctoral Fellow.

7.5 Original Publication

This Chapter has been reprinted from Chemical Physics, Robert M. Culik, Srinivas Annavarapu, Vikas Nanda and Feng Gai (2013) 422, 131-134. DOI: 10.1016/j.chemphys.2013.01.021, with permission from Elsevier Science.

Peptide	ΔH_m (kcal mol ⁻¹)	ΔS_m (cal mol ⁻¹ K ⁻¹)	ΔC_p (cal mol ⁻¹ K ⁻¹)	T_m (°C)	τ_f (μ s)	τ_u (μ s)
TC5b	12.9	40.4	72.1	45.9	3.0	7.6
5b-10GDA	17.1	50.4	53.6	66.6	1.7	33.4
5b-10GDQ	16.5	48.2	63.9	69.3	2.5	44.0
5b-10GDN	15.8	47.8	91.6	56.4	3.7	25.5

Table 7.1: Unfolding thermodynamic parameters of TC5b and its D-amino acid mutants, adapted from ref 250. Also listed are the folding and unfolding times of these peptides at 30 °C.

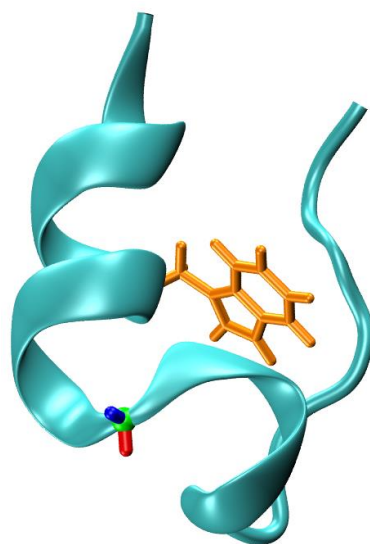


Figure 7.1: The NMR structure of Trp-cage (PDB 1L2Y). The glycine residue that was mutated to various D-amino acids in the current study is highlighted.

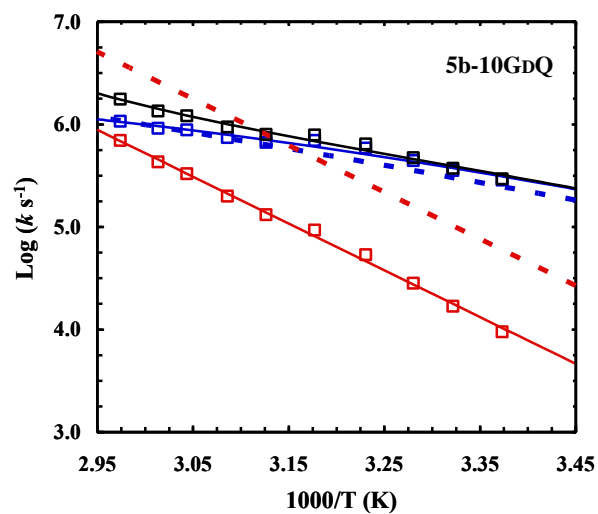


Figure 7.2: Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GdQ. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type, adapted from ref 237.

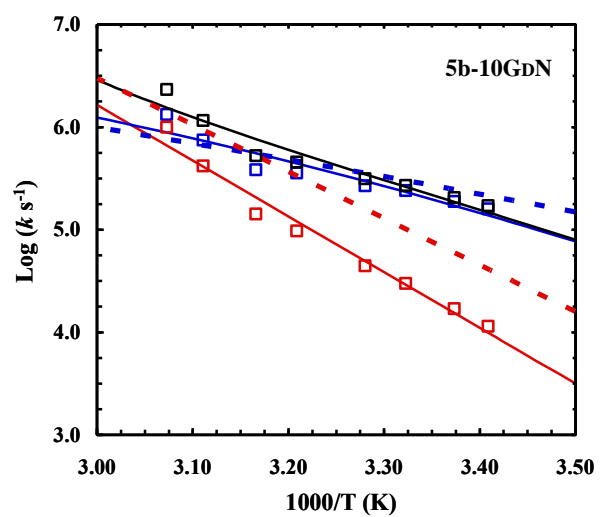


Figure 7.3: Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GDN. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type, adapted from ref 237.

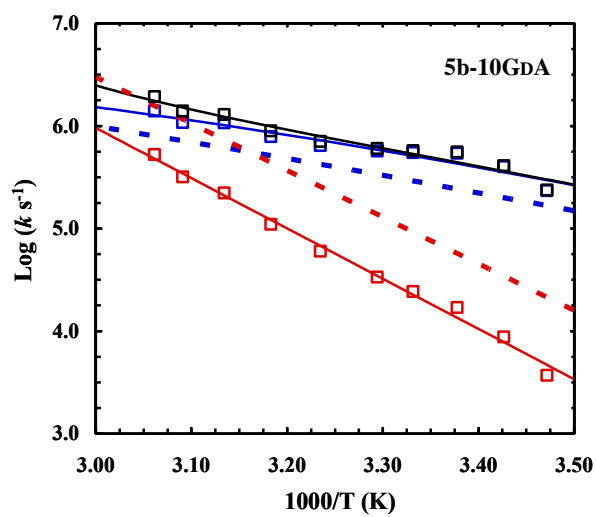


Figure 7.4: Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GdA. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type, adapted from ref 237.

CHAPTER 8

Experimental Validation of the Role of Trifluoroethanol as a Nano-Crowder

8.1 Introduction

While there are many ways to experimentally perturb protein stability and/or dynamics, perhaps one of the most common is through the use of cosolvents. Since the formation of native protein structure relies on protein-bath interactions, cosolvents can affect the protein folding pathway by altering the characteristics of the solvent. One advantage of using cosolvents rather than other methods to modify a system's equilibrium state, such as changing the temperature or pressure of the sample, is the ease with which cosolvents can be incorporated into various investigational methods without expensive or elaborate changes in the experimental setup. There are several examples of cosolvents used in protein folding studies, such as the denaturants guanidine hydrochloride (GdnHCl) and urea, the viscosogens glycerol and polyethylene glycol, and the stabilizers hexafluoroisopropanol (HFIP) and trifluoroethanol (TFE). Of the alcohols that stabilize protein secondary structure, TFE is one of the most commonly used due to its high potential for stabilizing structure relative to other alcohols. Although there have been great efforts to understand how various cosolvents perturb protein conformational preferences, the mechanisms of action for several denaturants and stabilizers are still in contention. Herein, we aim to gain insight into the effect of TFE on protein structure and dynamics.

The stabilizing ability of TFE has been studied extensively both experimentally and computationally.[253-265] One of the predominant views of TFE's mechanism of action is that it more favorably surrounds the protein than water, effectively leading to desolvation of the protein backbone.[254] This effect is thought to be the result of the

lower dielectric constant of TFE, which causes it to preferentially cluster around the protein backbone. However, at low concentrations, this hydrophobic interaction is not enough to lead to disruption of buried hydrophobic residues. Instead, protein-solvent hydrogen bonds are replaced with intramolecular hydrogen bonds, to lower the free energy of the system.[236,260,284,285] Conversely, other studies indicate that rather than stabilize the folded state, TFE acts to destabilize the random coil state by further structuring the solvent. This scenario also increases the folded population by effectively lowering the free energy barrier.[254,286] Despite its stabilizing ability on the secondary structural level, TFE has also been shown to cause denaturation in larger proteins. Specifically, the amphiphilic TFE molecule is capable, at larger volume percentages, of exposing and interacting with hydrophobic sidechains, thereby leading to a decrease in tertiary structure.[287]

Previously, Dobson *et al.* performed folding measurements as a function of TFE percentage. They showed that the folding rates for a set of globular proteins in various water-TFE solutions follow a chevron-like trend.[259] The interpretation for these results was that at low TFE percentages, folding rates are increased due to stabilization of native hydrogen bonding groups, whereas at higher percentages, folding rates are decreased in a similar manner as is found with denaturants, as determined by a high correlation between *m*-values of TFE and GdnHCl.[259] To complicate this interpretation, several studies have shown that the amphiphilic nature of TFE also results in molecular aggregation at certain volume percentages.[257,265] Dynamic light scattering and NMR measurements as well as MD simulations find that TFE clusters together with increasing propensity up

to 30% v/v, above which it mixes more homogeneously.[256, 257, 265] Taken together, these findings suggest that TFE could act as a molecular crowder thus increasing folding rates at certain percentages, similar to previous findings for the stabilizing cosolvent trimethylamine N-oxide (TMAO). Interestingly, viscosity profiles indicate that the viscosity of TFE/water mixtures doubles from 0% to 60% TFE.[288] Such a drastic increase in solvent viscosity could also have notable impacts on the folding rates of proteins in these solutions.

In order to gain insight into the effect of viscosity and aggregation on protein folding in TFE/water solutions, we have examined the folding of two intrinsically disordered peptides (IDPs) in different amounts of TFE. IDPs are ideal candidates for this study, because they lack appreciable tertiary structure when isolated in buffer, simplifying our interpretations. Here, we examine the phosphorylated kinase inducible domain (pKID) peptide [289] and the late embryogenesis abundant (LEA) peptide.[290] We chose these two systems because both have folded states that are alpha helical, however pKID is a helix-turn-helix, whereas LEA is a single monomeric helix. Previous experiments have shown that crowding effects have only a small effect on folding dynamics of monomeric alpha helices, whereas more extended systems experience more of a change.[266] Our hypothesis is that if TFE indeed behaves as a molecular crowder, there will be a TFE-dependent effect on folding dynamics for pKID that will not be seen in the folding dynamics of LEA. Additionally, previous studies have shown that relaxation rates obtained from kinetic measurements of peptides in highly viscous solutions are decreased in a manner consistent with Kramer's rule.[291, 292] We would

expect a similar effect if the viscosity of the cosolvent solution changed the dynamics of the system.

8.2 Experimental Section

Deuterated TFE was purchased from Cambridge Isotope Labs (MA) and stored in a drybox upon opening. Peptides were synthesized on a PS3 automated peptide synthesizer (Protein Technologies, MA) using Fmoc-protocols, purified by reverse-phase chromatography, and identified by matrix assisted laser desorption ionization (MALDI) mass spectroscopy. Phosphorylated serine was incorporated into pKID via the modified amino acid Fmoc-Ser(HPO₃Bzl)-OH. H-D exchange of IR samples was achieved by multiple rounds of lyophilization. CD spectra and thermal melting curves were obtained on an Aviv 62A DS spectropolarimeter (Aviv Associates, NJ) with a 1 mm sample holder. The peptide concentration was in the range of 50-60 μ M in H₂O and various concentrations of TFE (pH 7). Fourier transform infrared (FTIR) spectra were collected with 1 cm⁻¹ resolution on a Magna-IR 860 spectrometer (Nicolet, WI) using a homemade, two-compartment CaF₂ sample cell of 56 μ m. The details of the *T*-jump IR setup has been described elsewhere.[236] The peptide samples used in IR measurements were prepared by directly dissolving lyophilized solids in D₂O/deuterated TFE solutions (pH* 7) and the final peptide concentration was between 1-2 mM.

8.3 Results and Discussion

We chose pKID (sequence: DSVTDSQKRREILSRRPS*YRKILNDLSSDAPG-CONH₂, here S* is phosphoserine) as our model system because previous steady state CD measurements of pKID show that there is a significant increase in helicity of the peptide

upon addition of TFE up to ~30% v/v.[267] As shown (Figure 8.1A), the secondary structure of pKID is highly dependent on the concentration of TFE present in the solution, in qualitative agreement with earlier studies. As the percentage of TFE in the solution increases, there is a distinct red shift in the absorbance minimum from 200 nm to 208 nm. Additionally, with increasing TFE percentage, a local minimum at 222 nm appears, indicative of α -helical secondary structure. There is little change in the CD wavelength spectra beyond 30% TFE, which suggests that there is no further increase in helical content after this point. CD thermal melting experiments were also performed at all TFE percentages (Figure 8.1B). While there is negligible change in ellipticity at 222 nm at 0% TFE over the range of measured temperatures, with increasing amounts of TFE a larger change is observable, and at 15% TFE the transition appears to be cooperative. As expected, beyond 15% TFE the transition is no longer cooperative, to the extent that at 50% TFE the melt is essentially linear. This type of uncooperative transition has been seen in several other studies that have used TFE to induce helical structure.[256] Due to the lack of baselines in these temperature-dependent measurements, as well as the changing nature of the melts themselves as a function of TFE percentage, no quantitative analysis was performed to extract additional information from this data. We do, however, use the absolute ellipticity at 222 nm as an estimate of the fraction helix formed (Table 8.1). Previous studies of pKID without its binding partner have estimated that the peptide alone in buffer has marginal folding in its two helices (α A and α B), on the order of 10-15%, respectively.[267, 289] At 40% TFE, helical conformations in α A and α B increase to 30-40%, respectively. Since the peptide ellipticity does not appreciably further

decrease beyond 30% TFE, we assume that the final helical amounts of α A and α B are ~30-40%. To more accurately determine the fractional helicity of pKID at different trifluoroethanol percentages, we used the method of Baldwin,[253] which defines the fractional helicity (f_H) as:

$$f_H = \frac{[\theta]_{222} - [\theta]_C}{[\theta]_H - [\theta]_C} \quad (8.1)$$

where $[\theta]_H$ is defined as:

$$[\theta]_H = -44000 * \left(1 - \frac{2.5}{n}\right) + 100 * T \quad (8.2)$$

and $[\theta]_C$ is defined as:

$$[\theta]_C = 640 - 45 * T \quad (8.3)$$

and $[\theta]_{222}$ is the mean residue ellipticity of the protein at 222 nm, $[\theta]_H$ is the mean residue ellipticity of a 100% helical peptide at 222 nm, n is the number of residues that could potentially be in helical conformations, and T is the temperature in Celsius. By applying this analysis to our CD data for pKID in 0% and 30% TFE, we obtain fractional helicity values of 16% and 43%, respectively, at 1 °C. These values are in quantitative agreement with those obtained in previous studies.[267,289]

To understand the effect of TFE on peptide dynamics, we measured conformational relaxation rates of pKID in various concentrations of TFE using laser-induced temperature jump (T -jump) IR spectroscopy. As shown (Figure 8.2 and Table 8.2), at low TFE percentages there is a small but measurable decrease in relaxation rates up to 5% TFE, which is no longer observed upon increasing the TFE percentage to 15%. This small change might be a viscosity effect due to the increased viscosity of the

solvent. As previously reported,[288] the dynamic viscosity of TFE/water mixtures is relatively linear from 0% to 40% TFE. From these results, by increasing the solvent concentration from 0% to 5% TFE, the viscosity changes from 0.89 cP to 1.00 cP. Eaton *et al.* showed that the rate of α -helix formation exhibited a $\eta^{-0.6}$ indicating that larger changes in relaxation rates should be seen for lower percentages of TFE.[293] From this estimate, we would expect only a 1.07 ratio between the rates for 0% TFE and 5% TFE; however, our data corresponds to a 1.34 ratio. Therefore, we hypothesize that this decrease might be a microviscosity effect rather than a macroviscosity effect. Previous NMR experiments of a tetrapeptide in TFE/water solutions provided measurements of the diffusion of the tetrapeptide versus TFE percentage.[294] Using these results, we can estimate that microscopic viscosity effects should only increase the relaxation rates of pKID by a 1.09 ratio when going from a 0% TFE solution to a 5% TFE solution, which also does not fully explain the decrease observed. This observation might be explained by the fact that TFE is known to crowd around peptides so that there is a higher local concentration of TFE in the first hydration layer of the peptide than there is in the bulk solution. This phenomenon might cause altered viscosities felt by the peptide than these simple calculations can explain.

After the initial viscosity effect on pKID folding, relaxation rates appear to increase up to TFE percentages of 15%. These rates then remain stable, within experimental error, up to 30% TFE. Interestingly, at higher TFE percentages (Figure 8.2 and Table 8.2) there is a distinct decrease in the rate of pKID relaxation – for example, the relaxation rates of pKID in 50% TFE are slower than those at 30%. From this data

alone it is hard to conclude what this decrease in relaxation rate is due to, however it is well-known that the behavior of TFE varies as a function of concentration in solution. For example, TFE is known to form aggregates or micelles at concentrations at or around 30% v/v, whereas at higher concentrations this behavior disappears.[257,265] Therefore, one possibility is that the aggregates formed by TFE could be acting as molecular crowders, in a manner similar to that recently proposed by Thirumalai and coworkers for TMAO.[268] If this were the case, then folding rates would increase under crowding conditions, whereas unfolding rates would remain relatively the same, effectively increasing relaxation rates. Thus, if this crowding effect of TFE disappears due to higher TFE percentage (50%), we would expect relaxation rates to be slower. It has previously been shown that crowding does not affect the folding kinetics of peptides that have local, nearby structural interactions (i.e. single alpha helices), as diffusive motion over relatively small length scales will not be hampered by the excluded volumes of other molecules in solution.[266] In contrast, we assume that the pKID helix-turn-helix is a large enough system that crowding does affect its folding kinetics. To test our hypothesis that TFE is indeed acting as a nano-crowder, we performed similar measurements using the LEA peptide (sequence: AADGAKEKAGEAADGAKEKAGE-CONH₂) as a control.

Unlike the folded helix-turn-helix conformation of pKID, the folded state of LEA is that of a monomeric α -helix. CD measurements of LEA also reveal the sensitivity of the peptide to trifluoroethanol. While at 0% TFE the CD wavelength spectrum of LEA has signatures indicative of random coil conformations, at higher TFE percentages there is an increase in helical content up to ~40% TFE (Figure 8.4A), after which there is no

appreciable change. Using the fractional helicity analysis described above, we obtain a value of 36.1% for LEA in 40% TFE at 1 °C. Notably, temperature melt experiments (Figure 8.4B) of LEA show no loss of cooperativity with increasing TFE percentage. We interpret this as specific to LEA's folding conditions – this peptide folds upon dehydration of its exposed surfaces – however, a complete investigation of this is beyond the scope of the present paper.

Kinetic measurements for LEA yielded a different result than that obtained for pKID. Specifically, the relaxation rates of LEA at different TFE percentages were indistinguishable within experimental error (Figure 8.5 and Table 8.3). This is in agreement with our hypothesis that TFE is acting as a nanocrowder, since we expected that crowding would have no effect on the folding kinetics of a monomeric alpha helix. Taken together, our results using two different types of IDPs, one containing long-range structure and one without, demonstrate that TFE has the ability to act as a crowding agent at intermediate volume percentages 15-30% v/v.

8.4 Conclusion

There are a great number of studies in the literature dedicated to understanding and taking advantage of the stabilizing effect of TFE on protein secondary structure. However, most efforts to interrogate the mechanism of TFE action use proteins that contain appreciable amounts of tertiary structure, potentially complicating analysis. Here, we use two types of IDPs to investigate TFE effects on relaxation kinetics as a function of volume percentage. Our results indicate that at intermediate volume percentages (15-30%), TFE acts as a molecular crowder, increasing relaxation rates for molecules with long-range

interactions. In agreement with this conclusion, relaxation rates for a peptide whose native state is a monomeric helix are unaffected by this concentration-dependent TFE mode of action, consistent with previous crowding studies.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978). R.M.C. is an NIH Ruth Kirschstein Predoctoral Fellow (GM-008275).

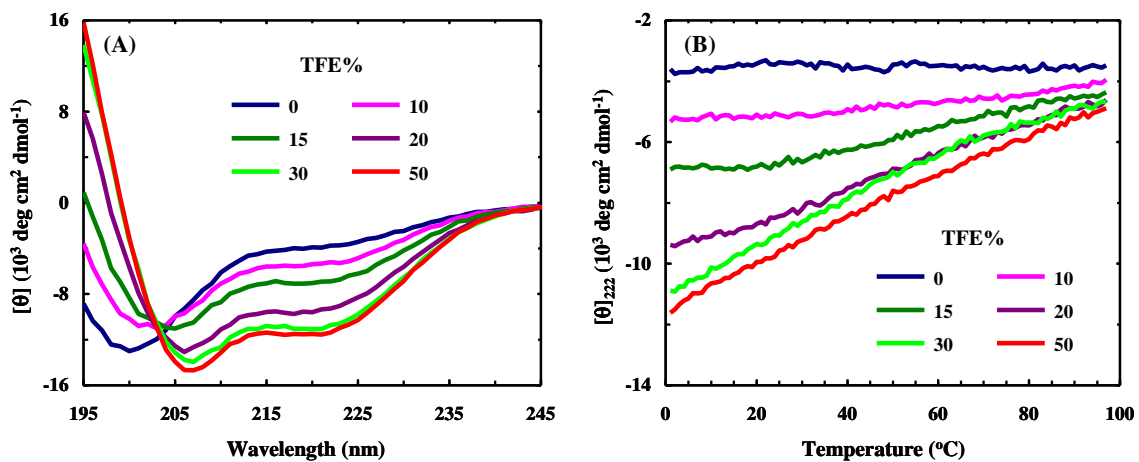


Figure 8.1: (A) CD spectra of pKID collected at 1 °C and in aqueous solutions of different TFE percentages, as indicated. (B) The corresponding CD thermal melting curves of these samples at 222 nm.

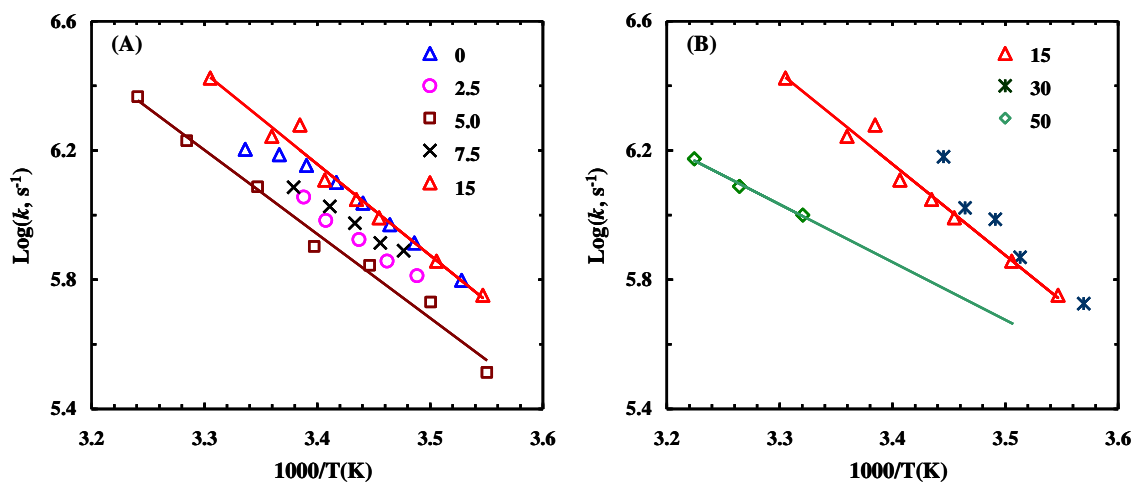


Figure 8.2: Relaxation rate constants of pKID versus temperature for different TFE solutions, as indicated.

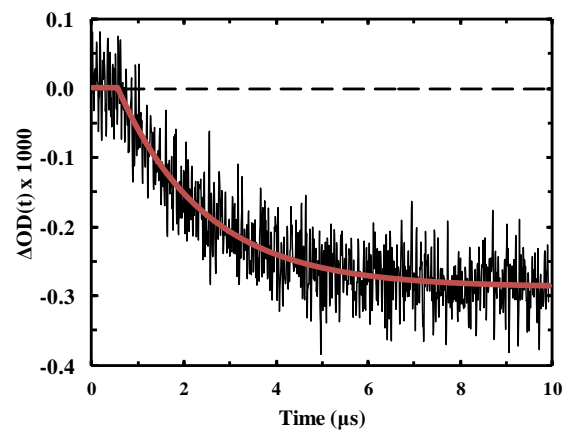


Figure 8.3: Representative trace of the relaxation kinetics of the pKID peptide in a 30% TFE solution in response to a temperature jump from 5.7 °C to 11 °C probed at 1630cm^{-1} . The smooth line represents the best fit of this curve to a single exponential function with a relaxation time constant of $1.8 \mu\text{s}$.

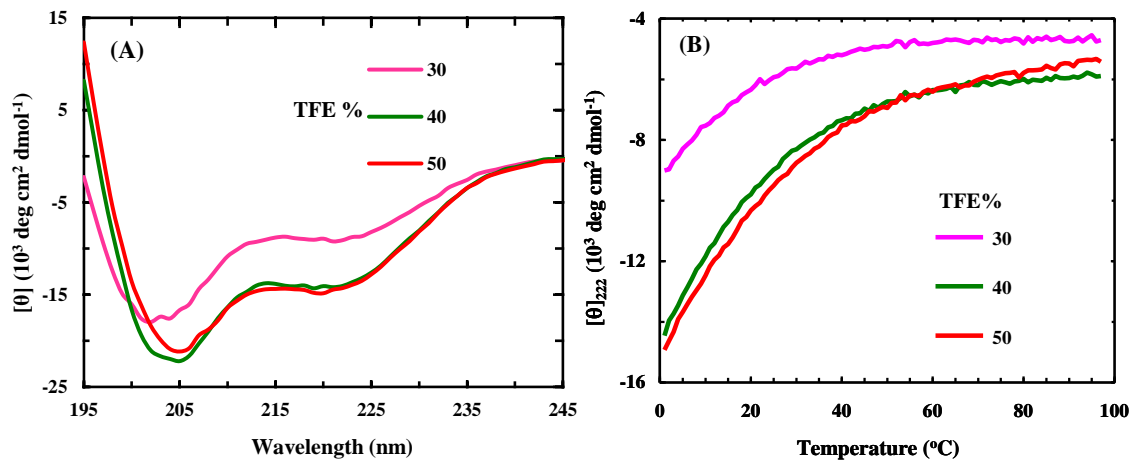


Figure 8.4: (A) CD spectra of LEA collected at 1 $^{\circ}\text{C}$ and in aqueous solutions of different TFE percentages, as indicated. (B) The corresponding CD thermal melting curves of these samples at 222 nm.

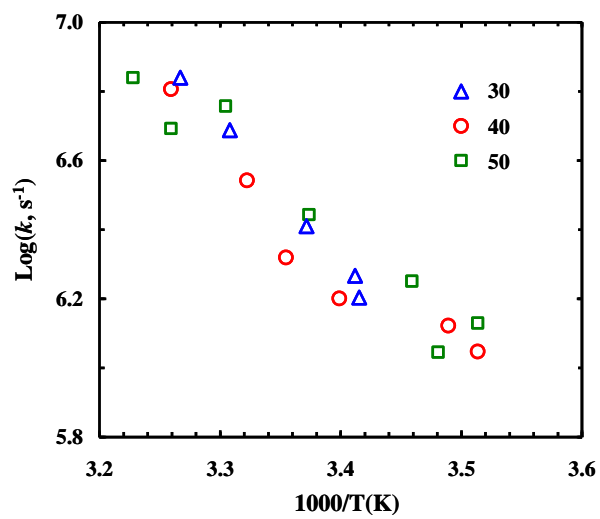


Figure 8.5: Relaxation rate constant of LEA versus temperature for different TFE solutions, as indicated.

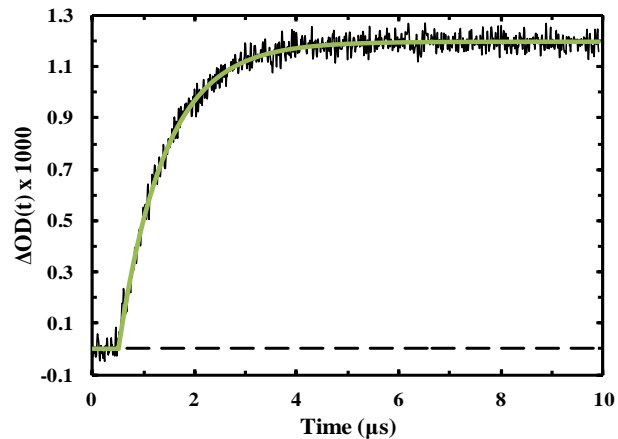


Figure 8.6: Representative trace of the relaxation kinetics of the LEA peptide in a 40% TFE solution in response to a temperature jump from 3.8 °C to 8.4 °C probed at 1664 cm^{-1} . The smooth line represents the best fit of this curve to a single exponential function with a relaxation time constant of 880 ns.

TFE %	f_H (%)
0	16
10	22
15	28
20	37
30	43
50	45

Table 8.1: Fractional helicity (f_H) approximations of pKID in TFE/water mixtures.

TFE %	T_f (°C)	τ_R (μs)
0	15.5	1.07 ± 0.09
2.5	15.7	1.4 ± 0.1
5	17	1.43 ± 0.08
7.5	16.2	1.22 ± 0.09
15	16.3	1.0 ± 0.3
30	16.1	1.4 ± 0.1
50	16.1	2.4 ± 0.5

Table 8.2: Relaxation rates of pKID in TFE/water mixtures at 15 °C.

TFE %	T_f (°C)	τ_R (μs)
30	24.2	0.21 ± 0.03
40	24.7	0.36 ± 0.09
50	23.1	0.18 ± 0.05

Table 8.3: Relaxation rates of LEA in TFE/water mixtures at 25 °C.

CHAPTER 9

Light-Triggered Modulation of the Free Energy Landscape: Using Azobenzene Isomerization to Create a Downhill Folder

9.1 Introduction

Despite the great amount of research dedicated to the study of protein folding, there are still many unresolved issues. For instance, questions remain regarding the nature of the unfolded state, the importance of off-pathway folding intermediates, and how nascent polypeptides fold out of the tunnel of the ribosome. Of the contemporary topics in protein folding, one of the greatest continuing debates is the nature of folding transition states. For example, for a given protein, what is/are the transition state(s) and is it appropriate to think about transition states in terms of a one-dimensional free energy coordinate?

There is a substantial body of literature dedicated to studying folding transition states. One approach dedicated to trying to understand and identify structural elements of folding transition states borrowed ideas from enzymologists – the method of developing transition state analogs, trapping enzyme structures in some intermediate conformation that can then be crystallized or otherwise characterized, was quickly adapted. Examples of trapping structure include the plethora of cross-linking strategies that, after ensuring molecules' native folded conformations weren't disrupted, were used to identify interactions present in the transition state.[27] Hydrogen exchange or other forms of kinetic isotope effect experiments are also very useful in identifying the structure of transient states, and do not require sophisticated mutation strategies as well.[187] Perhaps the most famous methods to probe transition state structure are phi and psi analysis. Phi value analysis posits that the energetic consequences of a given mutation in a protein that folds in a two-state manner fall in between two extremes. At one extreme, the energy of

the folded state changes, but there is no change in the energy of the transition state. At the other extreme, the change in energy of the folded state exactly matches the change in energy of the transition state. The first extreme would indicate that an interaction perturbed by the mutation is not present in the transition state ($\phi = 0$), while the second extreme would indicate that the interaction is present in the transition state ($\phi = 1$).[26] Psi value analysis takes advantage of bi-histidine mutations that can coordinate zinc ions, since the addition of zinc can be used to nucleate local structure. Folding measurements taken as a function of zinc concentration can then be used to identify regions where this zinc-dependent nucleation changes the folding rate.[27] While these two methods especially have contributed greatly to our knowledge of folding transition states, the next major hurdle for researchers was to use this information to design mutations to increase the stability and/or folding rate of biologically important proteins, or to generalize some design principles for creating *de novo* proteins. Although the latter of these two directions has not been achieved yet, there are many examples of proteins which have been improved upon after knowledge of folding transition states was obtained. Another idea for applying our understanding of protein transition states is to design molecular switches that could be used to switch proteins to fold or unfold with some external stimulus.

A great deal of effort has been made by the scientific community to develop molecules that conformationally rearrange in response to a triggering event. There are several ways to trigger conformational changes in molecules, however the most commonly employed is phototriggering, due to the straightforward use of a light source to effect structural change and the great library of molecules that rearrange upon

illumination.[269-273] A phototrigger is useful if it meets several criteria. First, its excitation should occur somewhere between 310 and 800 nm; lower wavelengths begin to also excite protein sidechain and/or backbone transitions, and higher wavelengths are typically not sufficient for exciting trigger molecules to undergo some molecular rearrangement. Additionally, there should be an appreciable photochemical yield to ensure sufficient signal for transient measurements. Furthermore, the triggering itself should occur on a timescale that is significantly faster than the dynamics that are desired to be probed. Also, a choice needs to be made about whether or not the triggering is reversible. Here, we use an acetamide azobenzene phototrigger to study conformational rearrangements in the Trp-cage model system.

Acetamide azobenzene is a desirable phototrigger because of its excitation at 355 nm, its photochemical yield of ~50%, and its picosecond timescale of photochemistry.[85] The reaction is a reversible isomerization around the central nitrogen stereocenter; the thermally populated *trans* configuration (>95% at room temperature) changes to *cis*, with a change in the end-to-end distance of ~3.5 Å.[274] It should be noted that the reverse reaction occurs spontaneously on a timescale of minutes, although it can be made to proceed on a picosecond timescale with irradiation at 430 nm. Since the conformational dynamics of interest occur on a nanosecond to microsecond timescale, we can effectively assume that the azobenzene trigger has fully isomerized before folding dynamics occur, and we can ignore any appreciable amount of *cis*→*trans* during the measurements.

We chose the Trp-cage model system for our measurements because of the extensive number of studies previously conducted for this peptide, which have led to a fairly detailed understanding of the molecule's folding mechanism. For example, our lab has previously shown that the transition state of folding for the Trp-cage involves formation of the α -helix.[237] Since the acetamide azobenzene is an ideal crosslinker for helices, we made cysteine mutations within the Trp-cage alpha helix to allow for azobenzene incorporation.[275-283] In this instance, we inserted the azobenzene at a length where in the *trans* configuration, it would not allow the helix to fold, whereas the *cis* conformer would allow it. In this scenario, upon irradiation we expect that if indeed helix formation is the rate limiting step of folding for Trp-cage, we could nucleate folding with light, allowing the molecule to fold in a downhill manner subsequently.

9.2 Experimental Section

Di-iodoacetamide azobenzene was acquired from our collaborators in the Andrew Woolley lab. The Trp-cage peptides were synthesized on a PS3 automated peptide synthesizer (Protein Technologies, MA) using Fmoc-protocols, purified by reverse-phase chromatography, and identified by matrix assisted laser desorption ionization (MALDI) mass spectroscopy. Incorporation of the acetamide azobenzene into Trp-cage was accomplished by following the methods established by Woolley. Briefly peptide (1.13 mM) was dissolved in Tris·Cl buffer (pH 8) containing Tris(carboxyethyl)phosphine (TCEP, 1.13 mM) and left to incubate at room temperature for two hours. Afterwards, di-iodoacetamide azobenzene was dissolved in DMSO and added to the peptide solution to a final concentration of 0.66 mM. The mixture was placed under foil in a hood with the

lights out and allowed to stir for 10 minutes. Afterwards, 56 μ l of a 10 mM solution of di-iodoacetamide azobenzene was added to the mixture, and this was allowed to stir in the dark for 10 minutes. Following this, another 56 μ l of 10 mM di-iodoacetamide azobenzene was added to the mixture and allowed to stir in the dark for 10 minutes. Then, the mixture was exposed to light (the lights in the hood were turned on), and the mixture was stirred for another 10 minutes. The mixture was purified using reverse-phase chromatography, and a second round of MALDI spectroscopy was performed to identify the pure labelled compound. Trifluoroacetic acid (TFA) removal and H-D exchange were achieved by multiple rounds of lyophilization.

UV-Vis spectra were collected on a Lambda 25 UV-Vis spectrometer (Perkin Elmer, MA) before sample exposure to 355 nm light, after 5 min. exposure to 355 nm light (\sim 7 mW), and again after 5 min. exposure to 430 nm light to convert the azobenzene molecules predominantly back to *trans*. Spectra were collected at room temperature using a 1 cm quartz cuvette and an integration time of 0.5 s/nm. Sample concentrations were \sim 10 μ M.

For steady-state measurements, samples were irradiated at room temperature with 355 nm (or 430 nm) light using a Fluorolog 3.10 spectrofluorometer (Jobin Yvon Horiba, NJ) with 2 nm resolution and a 1 cm quartz sample holder. Irradiation was always performed for 5 minutes.

CD wavelength spectra were obtained on an Aviv 62A DS spectropolarimeter (Aviv Associates, NJ) with a 1 mm sample holder. The peptide concentration was in the range of 35 μ M in a 20% trifluoroethanol/80% D₂O solution.

The transient IR setup used is similar to the *T*-jump IR setup described previously. The only difference is that here, the beamsplitter in the setup is removed and the entirety of the energy of the 1064 nm pump laser is sent into a harmonic generator that contains two sets of crystals, allowing for the generation of both the second (532 nm) and third (355 nm) harmonic. The second harmonic and fundamental (1064 nm) were separated from the desired 355 nm light with a pair of dichroic mirrors, and the 355 nm light was sent into the sample compartment. The IR cell used for these experiments has only one compartment and a spacer of 400 μm thickness. Samples were dissolved in a 20% trifluoroethanol/80% D_2O solution and concentrated to give an OD at 355 nm of 0.25 (35 μM). Transient measurements were performed at probing frequencies of 1630 and 1680 cm^{-1} , and the sample holder was moved constantly during measurement to ensure that each pump laser pulse hit a different spot on the cell than the previous pulse.

9.3 Results and Discussion

Previous studies have shown that by using an *i, i+7* spacing for azobenzene insertion in an α -helix, the *trans*-equilibrated sample will have the peptide backbone in an extended geometry that prevents helix formation.[86] Upon photoexcitation, however, the formation of the *cis* isomer does permit helical conformations. Following this reasoning we designed a mutant Trp-cage sequence with cysteines placed at an *i, i+7* spacing (sequence: CAYAQWLCDGGPSSGRPPPS) and incorporated acetamide azobenzene into it (Figure 9.1). After MALDI analysis confirmed azobenzene incorporation, we performed UV-Vis measurements of samples before and after irradiation with 355 nm light. As seen (Figure 9.2), the π - π^* transition of *trans* acetamide azobenzene at ~ 367 nm

has a significant decrease in intensity upon irradiation with 355 nm light, whereas there is a gain in absorbance at ~258 nm, which corresponds to the π - π^* of *cis* acetamide azobenzene.[86] For completeness, we also performed back-irradiation measurements, where after 5 minutes of irradiation with 355 nm light to promote *trans* to *cis* isomerization, samples were irradiated for 5 minutes at 430 nm to promote *cis* to *trans* isomerization. As shown (Figure 9.2), there is a recovery of the *trans*-azobenzene absorption spectrum, although it is not completely recovered, presumably because the fluorimeter light does not illuminate the entire cell. After convincing ourselves that the azobenzene-incorporated Trp-cage (hereafter referred to as 10b-azob) behaves similarly to other azobenzene-containing molecules, and that the spontaneous back-reaction of *cis* to *trans* isomerization is slow compared to the folding rate of Trp-cage (1 μ s), we performed CD measurements of the sample before and after irradiation with 355 nm light.

The CD spectrum of the dark-equilibrated sample (Figure 9.3) indicates that the molecules are mostly disordered. In contrast, the spectrum of the same sample after irradiation with 355 nm light (Figure 9.3) has signatures (minima at 209 and 222 nm) suggesting the formation of a substantial amount of helical structure. By comparing these two spectra to the wild-type 10b CD spectrum, we can approximate the percent helicity measured under the two conditions of this experiment. We estimate that the dark equilibrated 10b-azob has ~20% helical content, whereas the irradiated sample has ~70% helical content. To probe the kinetics of this large conformational rearrangement, we performed UV-pump IR-probe kinetic measurements on 10b-azob.

We first probed 10b-azob at 1630 cm^{-1} , where helical content is known to absorb. If indeed we are populating more helical conformations upon irradiation with 355 nm light in this system, as observed in our steady state CD measurements, we would expect a decrease in transmittance in our observed kinetics. Indeed, the measured dynamics show a decrease in transmittance, with kinetics that are best fit to a double-exponential (Figure 9.4). Interestingly, the time constant for the slower phase is $\sim 1\text{ }\mu\text{s}$, which is the measured folding time for 10b at room temperature. The other time constant, however, is $\sim 100\text{ ns}$, an order of magnitude faster. We interpret this fast component as corresponding to a population of molecules present at the folding transition state that proceed to the native state via downhill folding. To further confirm this hypothesis, we probed at 1680 cm^{-1} , where random coil configurations are expected to absorb. At this probe frequency, we observe an increase in transmittance with time, as expected if the photoisomerization were to produce more helical conformations. The kinetics observed at this frequency also are biphasic (Figure 9.5), with time constants within experimental error of those observed at 1630 cm^{-1} . We therefore put forth the hypothesis that insertion of the azobenzene crosslinker into Trp-cage 10b is able to modify its free energy landscape, creating an additional unfolded state that has a negligible folding barrier.

To confirm this, we inserted this crosslinker into a truncate of the 10b sequence containing only the alpha helix (sequence: CAYAQWLCD, hereafter referred to as helix-azob). If our hypothesis is correct, we would not expect to observe the 100 ns component in this truncate system. The CD spectrum of dark-equilibrated helix-azob is indicative of beta-sheet structure (Figure 9.6), however the intensity of this signature is low

considering the peptide concentration in the measurement ($\sim 50 \mu\text{M}$), suggesting that the molecule is largely disordered. Upon irradiation with 355 nm light, there is an increase in helical content as expected, although there is also an atypical minimum present at 235 nm. We speculate that this arises from some interaction of the peptide with the azobenzene moiety when the linker is in the *cis* conformation that is not possible when the linker is in the *trans* conformation.

Transient IR measurements of helix-azob probed at either 1630 or 1680 cm^{-1} yield relaxation kinetics that fit well to a single exponential function, with time constants of 1 and $1.1 \mu\text{s}$, respectively (Figure 9.7). Assuming that the folded structure of this peptide is similar to that of the alpha helix in 10b-azob, we conclude that the fast component seen in transient measurements of 10b-azob folding corresponds to downhill folding of the molecule. This is an interesting finding because it 1) further validates the previously-proposed folding mechanism of the Trp-cage and 2) suggests that phototrigger crosslinkers can be used to switch a protein from an activated (i.e. barrier-containing) to downhill folder.

9.4 Conclusion

While extensive effort has gone into identifying the structures of folding transition states of peptides and proteins, aside from the ability to further stabilize these proteins and to obtain generic protein design strategies, there have not been many examples of using this knowledge to actively change the nature of a protein's folding, e.g. to take a two-state folder and make it fold in a downhill manner. Here, we insert an azobenzene phototrigger, which isomerizes from a *trans* to *cis* configuration upon irradiation with

355 nm light, into the alpha helix of Trp-cage, the formation of which has previously been shown to be the rate-limiting step in Trp-cage folding. Steady-state measurements of the azobenzene-containing Trp-cage show a significant increase in helical content upon phototriggering, and transient IR measurements reveal biphasic kinetics, which correspond to the folding rate of the wild-type Trp-cage and folding rates that are an order of magnitude faster, suggesting a population of the Trp-cage molecules are folding in a downhill manner upon irradiation. Control experiments on a truncate of Trp-cage containing just the alpha helix with the azobenzene incorporated confirm that this fast component is not a result of the phototrigger itself or due to helix folding independent of the rest of the molecule.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978). R.M.C. is an NIH Ruth Kirschstein Predoctoral Fellow (GM-008275).

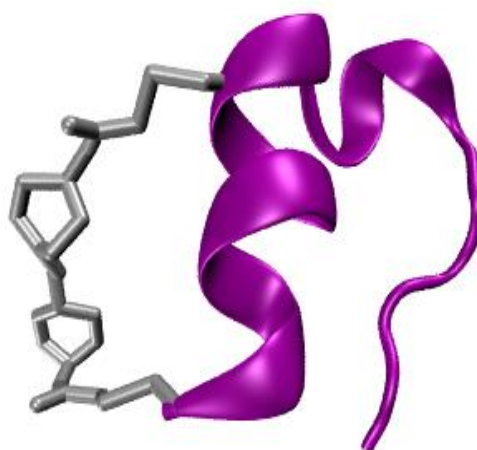


Figure 9.1: Cartoon representation of azobenzene-substituted Trp-cage. The azobenzene moiety is shown in grey. Structure adapted from the PDB deposit 1L2Y.

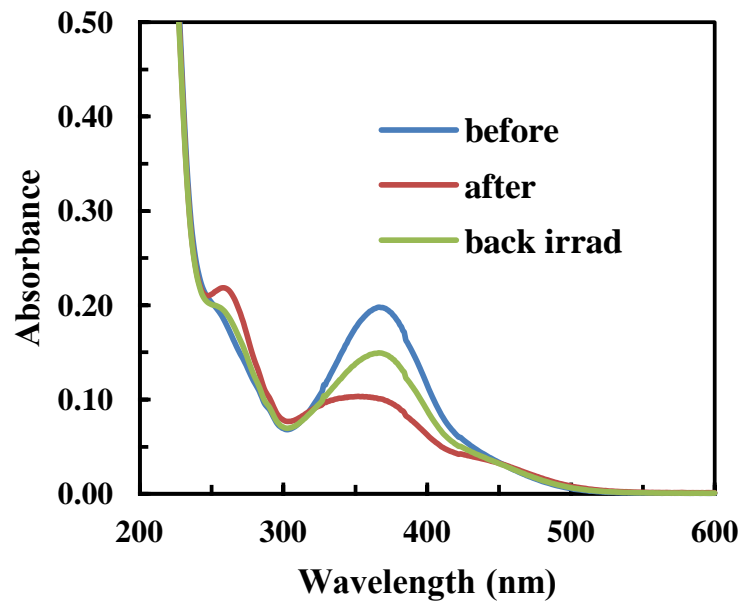


Figure 9.2: Absorption spectra of azobenzene-incorporated Trp-cage 10b before irradiation with 355 nm light, after irradiation at 355 nm for 5 min., and after further irradiation at 420 nm for 5 min, as indicated.

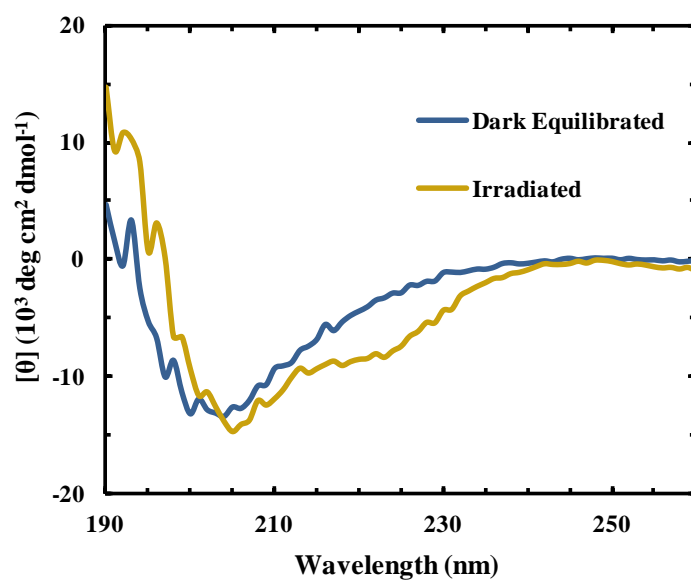


Figure 9.3: CD wavelength spectra of 10b-azob when dark-equilibrated and after irradiation at 355 nm for 5 minutes, as indicated.

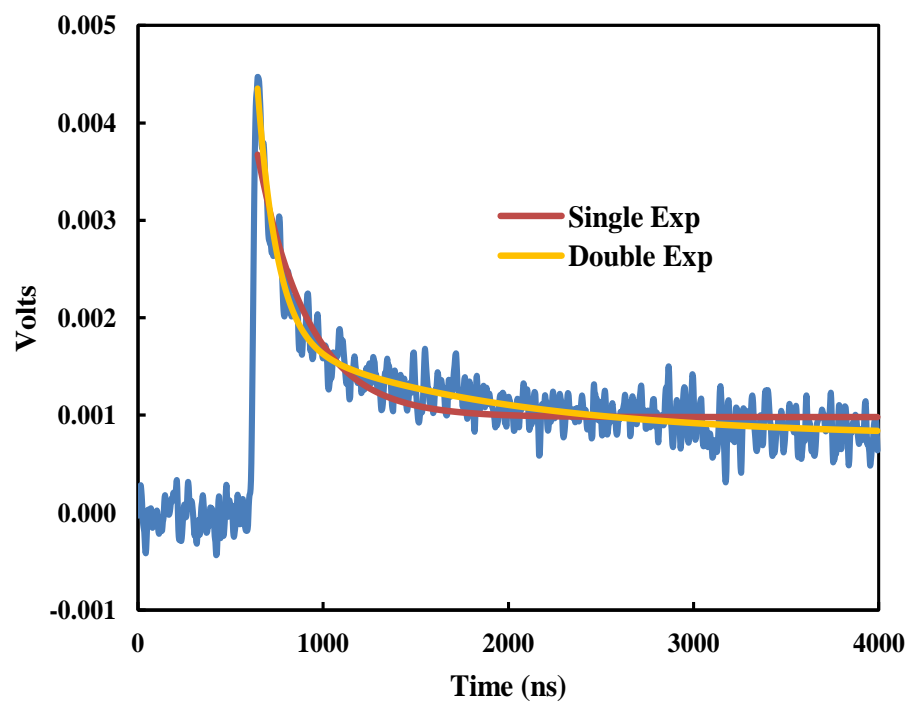


Figure 9.4: A kinetic trace of 10b-azob probed at 1630 cm^{-1} in response to irradiation with a 355 nm pulse of light. The time constants for the double-exponential fit are $\tau_1 = 100\text{ ns}$ and $\tau_2 = 1.2\text{ }\mu\text{s}$.

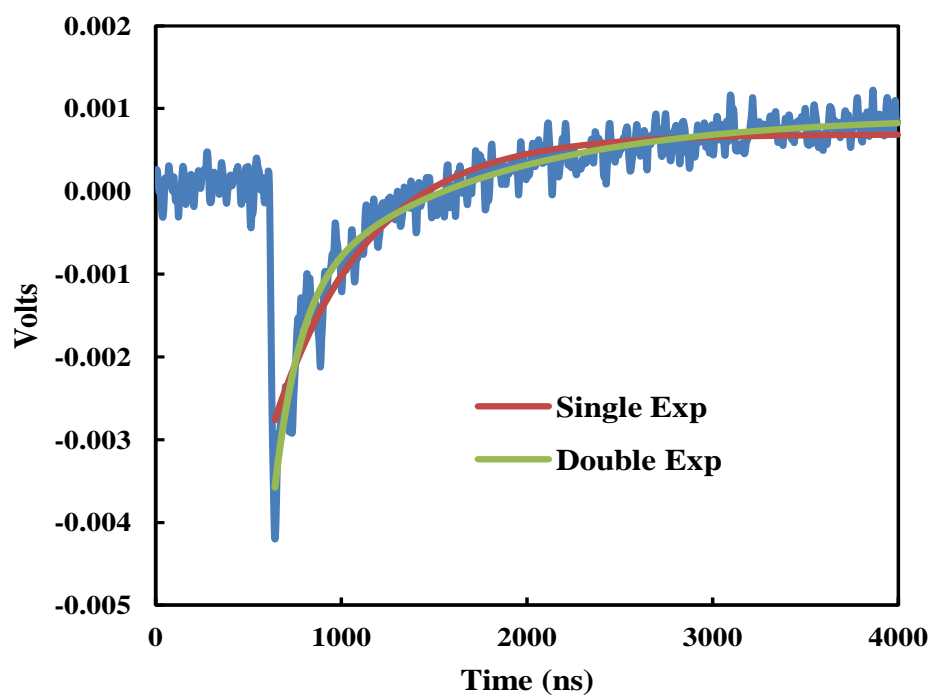


Figure 9.5: A kinetic trace of 10b-azob probed at 1680 cm^{-1} in response to irradiation with a 355 nm pulse of light. The time constants for the double-exponential fit are $\tau_1 = 140\text{ ns}$ and $\tau_2 = 1.1\text{ }\mu\text{s}$.

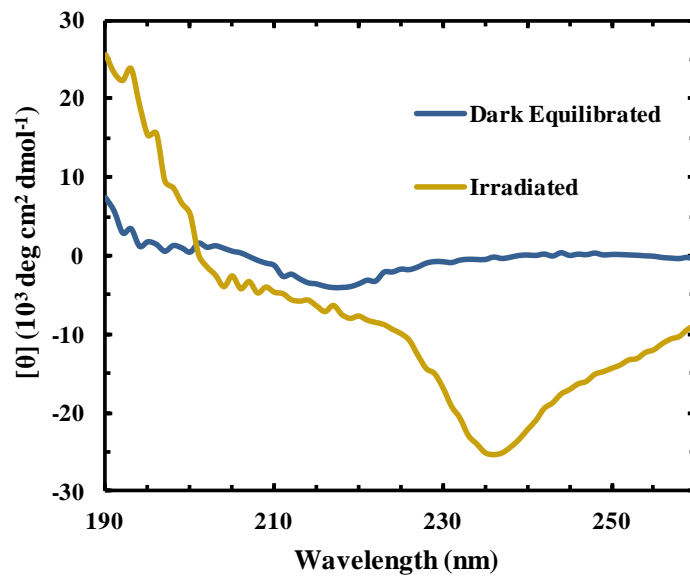


Figure 9.6: CD wavelength spectra of helix-azob when dark-equilibrated and after irradiation at 355 nm for 5 minutes, as indicated.

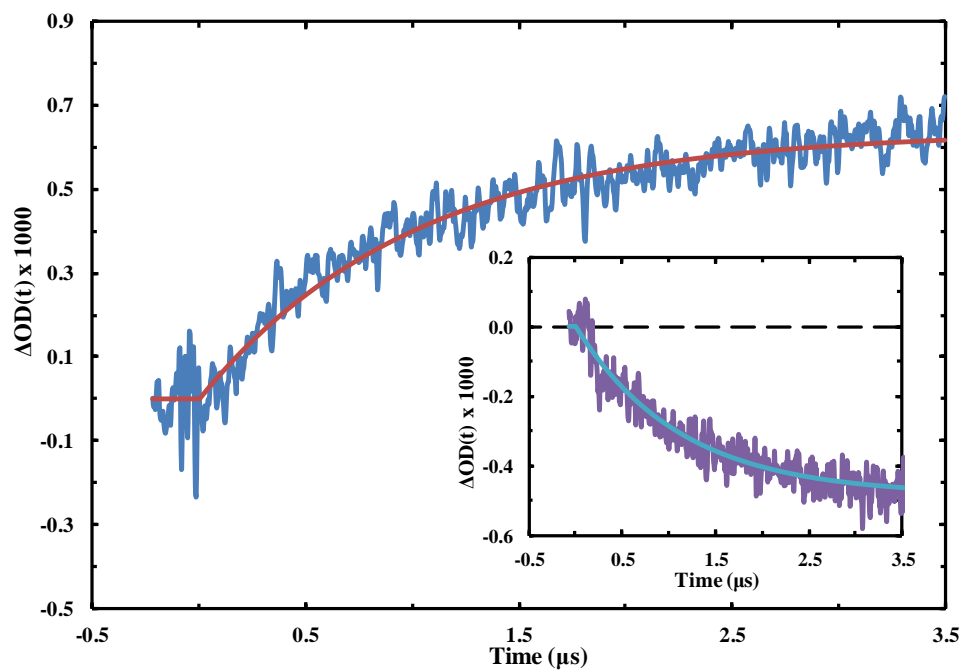


Figure 9.7: A kinetic trace of helix-azob probed at 1680 cm^{-1} in response to irradiation with a 355 nm pulse of light. The time constant for the single-exponential fit is $1.1\text{ }\mu\text{s}$. Inset: A kinetic trace of helix-azob probed at 1630 cm^{-1} in response to irradiation with a 355 nm pulse of light. The time constant for the single-exponential fit is $1.0\text{ }\mu\text{s}$.

CHAPTER 10
Future Directions

In Chapter 4, we demonstrated a method to easily incorporate nitriles into peptides/proteins via cysteine alkylation under relatively mild reaction conditions and identified the paracyanobenzyl moiety as a useful probe of local environment. One point that was not emphasized significantly in the chapter is the larger dynamic range of the paracyanobenzyl moiety in comparison to beta-thiocyanatoalanine, which is the common probe used for insertion into cysteine residues. Thus, it would be useful to insert the probe identified in Chapter 4 into an enzyme active site to estimate the magnitude of the local electric field in a manner similar to studies done in the Webb lab. Current work is being performed to insert this probe in the interior of a protein near a buried charge group in an effort to measure the dynamics and electric field of the local environment.

In Chapter 5, we described a combinatorial approach involving isotopic labeling, sidechain infrared transitions, and mutagenesis to increase the effective structural resolution of infrared kinetic measurements of protein folding. In the process, we identified the dominant folding pathway of the Trp-cage miniprotein, and characterized the structure of its transition state. The work presented in Chapter 9, where the transition state-structure of Trp-cage is targeted for azobenzene incorporation, is a direct result of this study. Another future direction would be to use the strategies mentioned above in larger, more complicated peptide systems, such as LysM or NTL9.

In Chapter 6, we presented the idea of using thioamides as site-specific probes of backbone-backbone hydrogen bonding and demonstrated its application in the Trpzip2 model system, where we showed that the hydrogen bond closest to the turn is present in the transition state, consistent with previous studies. There is wide applicability of this

technique in other studies, most notably as a type of ‘hydrogen bonding phi-value analysis’ to show, backbone hydrogen bond by backbone hydrogen bond, what interactions are present in the transition state. Another interesting application would be to take advantage of the increased hydrogen bond donating ability of the thioamide amine to perform similar measurements that stabilize the native state.

In Chapter 7, we use D-amino acids to investigate the importance of a conserved glycine in the Trp-cage peptide. We find that this glycine, which is at the C-terminus of the Trp-cage alpha helix, is not present in the transition state. A future direction of this work would be to use D-amino acids to stabilize the folding transition state of a peptide – replacement of a hydrophobic residue with a polar D-amino acid in a monomeric helix (e.g. one of the AKA peptides) would presumably stabilize the molecule through sidechain-backbone interactions, and possibly change the folding nucleation site. If this were the case, changing the position of the D-amino acid would allow us to more concretely understand the helix-coil transition, a process that to this day still cannot be perfectly explained.

In Chapter 8, we examined how TFE stabilizes the folded states of proteins using intrinsically disordered proteins as model systems. We found that TFE works, at least in part, by acting as a molecular crowder, which causes a destabilization of the unfolded state. An interesting follow-up to this study would be to assess how TFE affects water dynamics, which could be done using two dimensional infrared (2D-IR) measurements to examine the frequency-frequency correlation function of the amide I vibration as a function of TFE percentage.

In Chapter 9, we use an azobenzene crosslinker as a photolabile constraint to prevent formation of the folding transition state of the Trp-cage peptide, and therefore largely prevent its folding. Upon releasing this structural constraint with light, we find that the peptide has a significantly reduced free energy barrier (i.e. drastically increased folding time), indicating that the transition state is being populated with phototriggering. A future direction of this work might be to use this phototrigger to examine the roughness of the free energy landscape. Since the ‘triggered’ azobenzene Trp-cage mutant folds in a downhill manner, analysis of its folding kinetics can provide a direct estimate of the roughness of the free energy landscape. Comparison of folding rates with those obtained for the dark-equilibrated peptide would also allow for further refinement of the roughness value.

REFERENCES

1. Muirhead, H.; Perutz, M. F. Structure of haemoglobin – A 3-dimensional fourier synthesis of reduced human haemoglobin at 5.5 Å resolution. *Nature* 1963 199 633-638.
2. Schechte, A. N.; Chen, R. F.; Anfinsen, C. B. Kinetics of folding of staphylococcal nuclease. *Science* 1970 167 886-887.
3. Anfinsen, C. B. Principles that govern folding of protein chains. *Science* 1973 181 223-230.
4. Levinthal, C. Are there pathways for protein folding. *Journal de chimie physique et de physico-chimie biologique*. 1968 65 44-&.
5. Hantgan, R. R.; Hammes, G. G.; Scheraga, H. A. Pathways of folding of reduced bovine pancreatic ribonuclease. *Biochemistry* 1974 13 3421-3431.
6. Creighton, T. E. 2-Disulfide intermediates and folding pathway of reduced pancreatic trypsin-inhibitor. *Journal of Molecular Biology* 1975 95 167-199.
7. Garel, J. R.; Nall, B. T.; Baldwin, R. L. Guanidine-unfolded state of ribonuclease A contains both fast-refolding and slow-refolding states. *Proceedings of the National Academy of Sciences of the United States of America* 1976 73 1853-1857.
8. Davis, A.; Parr, G. R.; Taniuchi, H. Kinetic study of the folding of nuclease-B, a possible precursor of staphylococcal nuclease-A. *Biochimica et Biophysica Acta* 1979 578 505-510.

9. Leutzinger, Y.; Beychok, S. Kinetics and mechanism of heme-induced refolding of human alpha-globin. *Proceedings of the National Academy of Sciences of the United States of America* 1981 78 780-784.
10. Kato, S.; Okamura, M.; Shimamoto, N.; Utiyama, H. Spectral evidence for a rapidly formed structural intermediate in the refolding kinetics of hen egg-white lysozyme. *Biochemistry* 1981 20 1080-1085.
11. Ansari, A.; Berendzen, J.; Bowne, S. F.; Frauenfelder, H.; Iben, I. E. T.; Sauke, T. B.; Shyamsunder, E.; Young, R. D. Protein shakes and protein quakes. *Proceedings of the National Academy of Sciences of the United States of America* 1985 82 5000-5004.
12. Nall, B. T. Native or native-like species are transient intermediates in folding of alkaline iso-2 cytochrome-c. *Biochemistry* 1986 25 2974-2978.
13. Udgaonkar, J. B.; Baldwin, R. L. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease-A. *Nature* 1988 335 694-699.
14. Matouschek, A.; Kellis, J. T.; Serrano, L.; Fersht, A. R. Mapping the transition-state and pathway of protein folding by protein engineering. *Nature* 1989 340 122-126.
15. Dill, K. A. Dominant forces in protein folding. *Biochemistry* 1990 29 7133-7155.

16. Ptitsyn, O. B.; Pain, R. H.; Semisotnov, G. V.; Zerovnik, E.; Razgulyaev, O. I. Evidence for a molten globule state as a general intermediate in protein folding. *FEBS Letters* 1990 262 20-24.
17. Honeycutt, J. D.; Thirumalai, D. Metastability of the folded states of globular-proteins. *Proceedings of the National Academy of Sciences of the United States of America* 1990 87 3526-3529.
18. Freire, E.; Murphy, K. P. Molecular-basis of cooperativity in protein folding. *Journal of Molecular Biology* 1991 222 687-698.
19. Weissman, J. S.; Kim, P. S. Reexamination of the folding of BPTI – Predominance of native intermediates. *Science* 1991 253 1386-1393.
20. Jackson, S. E.; Fersht, A. R. Folding of chymotrypsin inhibitor-2 .2. Influence of proline isomerization on the folding kinetics and thermodynamic characterization of the transition-state of folding. *Biochemistry* 1991 30 10436-10443.
21. Schuler, B.; Lipman, E. A.; Eaton, W. A. Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy. *Nature* 2002 419 743-747.
22. Daggett, V.; Fersht, A. The present view of the mechanism of protein folding. *Nature Reviews Molecular Cell Biology* 2003 4 497-502.
23. Kubelka, J.; Hofrichter, J.; Eaton, W. A. The protein folding ‘speed limit’. *Current opinion in structural biology*. 2004 14 76-88.

24. Englander, S. W.; Mayne, L.; Krishna, M. M. G. Protein folding and misfolding: mechanism and principles. *Quarterly Reviews of Biophysics* 2007 40 287-326.
25. Bedard, S.; Krishna, M. M. G.; Mayne, L.; Englander, S. W. Protein folding: Independent unrelated pathways or predetermined pathways with optional errors. *Proceedings of the National Academy of Sciences of the United States of America* 2008 105 7182-7187.
26. A. R. Fersht, A. Matouschek, L. Serrano. The Folding of an Enzyme. 1. Theory of Protein Engineering Analysis of Stability and Pathway of Protein Folding. *Journal of Molecular Biology*. 1992, 224, 771–782.
27. Krantz, B. A.; Sosnick, T. R. Engineered metal binding sites map the heterogeneous folding landscape of a coiled coil. *Nature Structural Biology* 2001 8 1042-1047.
28. Pandit, A. D.; Jha, A.; Freed, K. F.; Sosnick, T. R. Small proteins fold through transition states with native-like topologies. *Journal of Molecular Biology* 2006 361 755-770.
29. Luan, B. W.; Lyle, N.; Pappu, R. V.; Raleigh, D. P. Denatured state ensembles with the same radii of gyration can form significantly different long-range contacts. *Biochemistry* 2014 53 39-47.
30. Crick, S. L.; Jayaraman, M.; Frieden, C.; Wetzel, R.; Pappu, R. V. Fluorescence correlation spectroscopy shows that monomeric polyglutamine molecules form

collapsed structures in aqueous solutions. *Proceedings of the National Academy of Sciences of the United States of America* 2006 103 16764-16769.

31. Yoo, T. Y.; Meisburger, S. P.; Hinshaw, J.; Pollack, L.; Haran, G.; Sosnick, T. R.; Plaxco, K. Small-angle X-ray scattering and single-molecule FRET spectroscopy produce highly divergent views of the low-denaturant unfolded state. *Journal of Molecular Biology* 2012 418 226-236.
32. Wickner, W. Assembly of proteins into membranes. *Science* 1980 210 861-868.
33. Garnier, J.; Gaye, P.; Mercier, J. C.; Robson, B. Structural-properties of signal peptides and their membrane insertion. *Biochimie* 1980 62 231-239.
34. Ladokhin, A. S.; White, S. H. Interfacial folding and membrane insertion of a designed helical peptide. *Biochemistry* 2004 43 5782-5791.
35. Reshetnyak, Y. K.; Andreev, O. A.; Segala, M.; Markin, V. S.; Engelman, D. M. Energetics of peptide (pHLIP) binding to and folding across a lipid bilayer membrane. *Proceedings of the National Academy of Sciences of the United States of America* 2008 105 15340-15345.
36. Tompa, P. Intrinsically unstructured proteins. *Trends in Biochemical Sciences* 2002 27 527-533.
37. Fink, A. L. Natively unfolded proteins. *Current Opinion in Structural Biology* 2005 15 35-41.
38. Uversky, V. N. The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome. *Journal of Biomedicine and Biotechnology* 2010 1-14.

39. Wright, P. E.; Dyson, H. J. Linking folding and binding. *Current Opinion in Structural Biology* 2009 19 31-38.
40. Shoemaker, B. A.; Portman, J. J.; Wolynes, P. G. Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proceedings of the National Academy of Sciences of the United States of America* 2000 97 8868-8873.
41. Huang, Y. Q.; Liu, Z. R. Kinetic advantage of intrinsically disordered proteins in coupled folding-binding process: A critical assessment of the “fly-casting” mechanism. *Journal of Molecular Biology* 2009 393 1143-1159.
42. Sugase, K.; Dyson, H. J.; Wright, P. E. Mechanism of coupled folding and binding of an intrinsically disordered protein. *Nature* 2007, 447, 1021.
43. Onitsuka, M.; Kamikubo, H.; Yamazaki, Y.; Kataoka, M. Mechanism of induced folding: Both folding before binding and binding before folding can be realized in staphylococcal nuclease mutants. *Proteins- Structure, Function, and Bioinformatics* 2008 72 837-847.
44. Kleerekoper, Q. K.; Putkey, J. A. PEP-19, an intrinsically disordered regulator of calmodulin signaling. *Journal of Biological Chemistry* 2009 284 7455-7464.
45. Chan, C. K.; Hu, Y.; Takahashi, S.; Rousseau, D. L.; Eaton, W. A.; Hofrichter, J. Submillisecond protein folding kinetics studied by ultrarapid mixing. *Proceedings of the National Academy of Sciences of the United States of America* 1997 94 1779-1784.

46. Roder, H.; Maki, K.; Cheng, H.; Shastry, M. C. R. Rapid mixing methods for exploring the kinetics of protein folding. *Methods* 2004 34 15-27.
47. Dumont, C.; Emilsson, T.; Gruebele, M. Reaching the protein folding speed limit with large, sub-microsecond pressure jumps. *Nature Methods* 2009 6 515-520.
48. Gilmanshin, R.; Williams, S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. Fast events in protein folding: Relaxation dynamics of secondary and tertiary structure in native apomyoglobin. *Proceedings of the National Academy of Sciences of the United States of America* 1997 94 3709-3713.
49. Dyer, R. B.; Gai, F.; Woodruff, W. H. Infrared studies of fast events in protein folding. *Accounts of Chemical Research* 1998 31 709-716.
50. Gruebele, M.; Sabelko, J.; Ballew, R.; Ervin, J. Laser temperature jump induced protein refolding. *Accounts of Chemical Research*. 1998 31 699-707.
51. Korzhnev, D. M.; Religa, T. L.; Banachewicz, W.; Fersht, A. R.; Kay, L. E. A transient and low-populated protein-folding intermediate at atomic resolution. *Science* 2010 329 1312-1316.
52. Lesk, A. M.; Rose, G. D. Folding units in globular proteins. *Proceedings of the National Academy of Sciences of the United States of America* 1981 78 4304-4308.
53. Dill, K. A.; Fiebig, K. M.; Chan, H. S. Cooperativity in protein-folding kinetics. *Proceedings of the National Academy of Sciences of the United States of America* 1993 90 1942-1946.

54. Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of protein folding: the energy landscape perspective. *Annual Review of Physical Chemistry* 1997 48 545-600.
55. Dill, K. A.; Chan, H. S. From Levinthal to pathways to funnels. *Nature Structural Biology* 1997 4 10-19.
56. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y. B.; Wrighers, W. Atomic-level characterization of the structural dynamics of proteins. *Science* 2010 330 341-346.
57. Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Shaw, D. E. How fast-folding proteins fold. *Science* 2011 334 517-520.
58. Snow, C. D.; Nguyen, N.; Pande, V. S.; Gruebele, M. Absolute comparison of simulated and experimental protein-folding dynamics. *Nature* 2002 420 102-106.
59. Privalov, P. L. Stability of proteins: Small globular proteins. *Advances in Protein Chemistry* 1979 33 167-241.
60. Schellman, J. A. The thermodynamic stability of proteins. *Annual Review of Biophysics and Biophysical Chemistry* 1987 16 115-137.
61. Bechtel, W. J.; Schellman, J. A. Protein stability curves. *Biopolymers*. 1987 26 1859-1877.
62. Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* 1940 7 284-304.

63. Sosnick, T. R.; Dothager, R. S.; Krantz, B. A. Differences in the folding transition state of ubiquitin indicated by phi and psi analyses. *Proceedings of the National Academy of Sciences of the United States of America* 2004 101 17377-17382.
64. Martinez, J. C.; Pisabarro, M. T.; Serrano, L. Obligatory steps in protein folding and the conformational diversity of the transition state. *Nature Structural Biology* 1998 5 721-729.
65. Feng, H. Q.; Vu, N. D.; Zhou, Z.; Bai, Y. W. Structural examination of Phi-value analysis in protein folding. *Biochemistry* 2004 43 14325-14331.
66. Krantz, B. A.; Dothager, R. S.; Sosnick, T. R. Discerning the structure and energy of multiple transition states in protein folding using psi-analysis. *Journal of Molecular Biology* 2004 337 463-475.
67. Perez-Iratxeta, C.; Andrade-Navarro, M. A. K2D2: Estimation of protein secondary structure from circular dichroism spectra. *BMC Structural Biology* 2008 8 25.
68. Woody, R. W. Contributions of tryptophan side-chains to the far-ultraviolet circular-dichroism of proteins. *European Biophysics Journal with Biophysics Letters* 1994 23 253-262.
69. Kubelka, J. Time-resolved methods in biophysics. 9. Laser temperature-jump methods for investigating biomolecular dynamics. *Photochemical and Photobiological Sciences* 2009 8 499-512.

70. A. Barth, C. Zscherp. What vibrations tell us about proteins. *Quarterly Review of Biophysics*. 2002, 35, 369–430.
71. Waegele, M. M.; Culik, R. M.; Gai, F. Site-specific reporters of the local electric field, hydration, structure, and dynamics of biomolecules. *Journal of Physical Chemistry Letters* 2011 2 2598-2609.
72. Schultz KC, Supekova L, Ryu YH, Xie JM, Perera R, Schultz PG. A genetically encoded infrared probe. *Journal of the American Chemical Society* 2006; 128:13984–13985.
73. Connor, R. E.; Tirrell, D. A. Non-canonical amino acids in protein polymer design. *Polymer Reviews* 2007 47 9-28.
74. L. Tadesse, R. Nazarbaghi, L. Walters. Isotopically Enhanced Infrared-Spectroscopy - A Novel Method for Examining Secondary Structure at Specific Sites in Conformationally Heterogeneous Peptides. *Journal of the American Chemical Society*. 1991, 113, 7036–7037.
75. S. M. Decatur, J. Antonic. Isotope-edited infrared spectroscopy of helical peptides. *Journal of the American Chemical Society*. 1999, 121, 11914–11915.
76. Decatur, S. M. Elucidation of residue-level structure and dynamics of polypeptides via isotope-edited infrared spectroscopy. *Accounts of Chemical Research* 2006 39 169-175.
77. Franco, I.; George, C. B.; Solomon, G. C.; Schatz, G. C.; Ratner, M. A. Mechanically activated molecular switch through single-molecule pulling. *Journal of the American Chemical Society* 2011 133 2242-2249.

78. Huang, T.; Zhao, J.; Peng, M.; Popov, A. A.; Yang, S. F.; Dunsch, L.; Petek, H. A molecular switch based on current-driven rotation of an encapsulated cluster within a fullerene cage. *Nano Letters* 2011 11 5327-5332.
79. Volk, M.; Kholodenko, Y.; Lu, H. S. M.; Gooding, E. A.; DeGrado, W. F.; Hochstrasser, R. M. Peptide conformational dynamics and vibrational stark effects following photoinitiated disulfide cleavage. *Journal of Physical Chemistry B* 1997 101 8607-8616.
80. Kolano, C.; Helbing, J.; Bucher, G.; Sander, W.; Hamm, P. Intramolecular disulfide bridges as a phototrigger to monitor the dynamics of small cyclic peptides. *Journal of Physical Chemistry B*. 2007, 111, 11297-11302.
81. Prompers, J. J.; Hilbers, C. W.; Pepermans, H. A. M. Tryptophan mediated photoreduction of disulfide bond causes unusual fluorescence behavior of *Fusarium solani* pisi cutinase. *FEBS Letters* 1999 456 409-416.
82. Tucker, M. J.; Couter, J. R.; Chen, J. X.; Atasoylu, O.; Smith, A. B.; Hochstrasser, R. M. Tetrazine phototriggers: probes for peptide dynamics. *Angewandte Chemie-International Edition*. 2010, 49, 3612-3616.
83. Tucker, M. J.; Abdo, M.; Courter, J. R.; Chen, J. X.; Brown, S. P.; Smith, A. B.; Hochstrasser, R. M. Nonequilibrium dynamics of helix reorganization observed by transient 2D IR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*. 2013, 110, 17314-17319.

84. Syage, J. A.; Lambert, W. R.; Felker, P. M.; Zewail, A. H.; Hochstrasser, R. M. Picosecond excitation and *trans-cis* isomerization of stilbene in a supersonic jet – dynamics and spectra. *Chemical Physics Letters*. 1982, 88, 266-270.
85. Lednev, I. K.; Ye, T. Q.; Hester, R. E.; Moore, J. N. Femtosecond time-resolved UV-visible absorption spectroscopy of *trans*-azobenzene in solution. *Journal of Physical Chemistry*. 1996, 100, 13338-13341.
86. Kumita, J. R.; Smart, O. S.; Woolley, G. A. Photo-control of helix content in a short peptide. *Proceedings of the National Academy of Sciences of the United States of America*. 2000, 97, 3803-3808.
87. Kumita, J. R.; Flint, D. G.; Smart, O. S.; Woolley, G. A. Photo-control of peptide helix content by an azobenzene cross-linker: steric interactions with underlying residues are not critical. *Protein Engineering, Design, and Selection*. 2002, 15, 561-569.
88. Chen, E.; Kumita, J. R.; Woolley, G. A.; Kliger, D. S. The kinetics of helix unfolding of an azobenzene cross-linked peptide probed by nanosecond time-resolved optical rotatory dispersion. *Journal of the American Chemical Society*. 2003, 125, 12443-12449.
89. Kumita, J. R.; Flint, D. G.; Woolley, G. A.; Smart, O. S. Achieving photo-control of protein conformation and activity: producing a photo-controlled leucine zipper. *Faraday Discussions*. 2003, 122, 89-103.
90. Ihalainen, J. A.; Bredenbeck, J.; Pfister, R.; Helbing, J.; Chi, L.; van Stokkum, I. H. M.; Woolley, G. A.; Hamm, P. Folding and unfolding of a photoswitchable

peptide from picoseconds to microseconds. Proceedings of the National Academy of Sciences of the United States of America. 2007, 104, 5383-5388.

91. Yoshikawa S, Okeeffe DH, Caughey WS. Investigations of cyanide as an infrared probe of heme protein ligand-binding sites. *Journal of Biological Chemistry* 1985; 260:3518–3528.
92. Reddy KS, Yonetani T, Tsuneshige A, Chance B, Kushkuley B, Stavrov SS, Vanderkooi JM. Infrared spectroscopy of the cyanide complex of Iron(II) myoglobin and comparison with complexes of microperoxidase and hemoglobin. *Biochemistry*. 1996; 35:5562–5570.
93. Getahun Z, Huang CY, Wang T, De Leon B, DeGrado WF, Gai F. Using nitrile-derivatized amino acids as infrared probes of local environment. *Journal of the American Chemical Society* 2003; 125:405–411.
94. Suydam IT, Boxer SG. Vibrational Stark effects calibrate the sensitivity of vibrational probes for electric fields in proteins. *Biochemistry*. 2003; 42:12050–12055.
95. Huang CY, Wang T, Gai F. Temperature dependence of the CN stretching vibration of a nitrile-derivatized phenylalanine in water. *Chemical Physics Letters* 2003; 371:731–738.
96. Tucker MJ, Getahun Z, Nanda V, DeGrado WF, Gai F. A new method for determining the local environment and orientation of individual side chains of membrane-binding peptides. *Journal of the American Chemical Society* 2004; 126:5078–5079.

97. Fafarman AT, Webb LJ, Chuang JI, Boxer SG. Site-specific conversion of cysteine thiols into thiocyanate creates an IR probe for electric fields in proteins. *Journal of the American Chemical Society* 2006; 128:13356–13357.
98. Mukherjee S, Chowdhury P, DeGrado WF, Gai F. Site-specific hydration status of an amphipathic peptide in AOT reverse micelles. *Langmuir*. 2007; 23:11174–11179.
99. Maienschein-Cline MG, Londergan CH. The CN stretching band of aliphatic thiocyanate is sensitive to solvent dynamics and specific solvation. *Journal of Physical Chemistry A*. 2007; 111:10020–10025.
100. Watson MD, Gai XS, Gillies AT, Brewer SH, Fenlon EE. A vibrational probe for local nucleic acid environments: 5-Cyano-2'-deoxyuridine. *Journal of Physical Chemistry B*. 2008; 112:13188–13192.
101. Fang C, Bauman JD, Das K, Remorino A, Arnold E, Hochstrasser RM. Two-dimensional infrared spectra reveal relaxation of the nonnucleoside inhibitor TMC278 complexed with HIV-1 reverse transcriptase. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105:1472–1477.
102. Lindquist BA, Haws RT, Corcelli SA. Optimized quantum mechanics/molecular mechanics strategies for nitrile vibrational probes: Acetonitrile and paratolunitrile in water and tetrahydrofuran. *Journal of Physical Chemistry B*. 2008; 112:13991–14001.

103. Oh KI, Choi JH, Lee JH, Han JB, Lee H, Cho M. Nitrile and thiocyanate IR probes: molecular dynamics simulation studies. *Journal of Chemical Physics*. 2008; 128:154504.
104. Lindquist BA, Furse KE, Corcelli SA. Nitrile groups as vibrational probes of biomolecular structure and dynamics: an overview. *Physical Chemistry Chemical Physics*. 2009; 11:8119–8132.
105. Taskent-Sezgin H, Chung J, Patsalo V, Miyake-Stoner SJ, Miller AM, Brewer SH, Mehl RA, Green DF, Raleigh DP, Carrico I. Interpretation of p-cyanophenylalanine fluorescence in proteins in terms of solvent exposure and contribution of side-chain quenchers: a combined fluorescence, IR and molecular dynamics study. *Biochemistry*. 2009; 48:9040–9046.
106. Liu J, Strzalka J, Tronin A, Johansson JS, Blasie JK. Mechanism of interaction between the general anesthetic halothane and a model ion channel protein, II: fluorescence and vibrational spectroscopy using a cyanophenylalanine probe. *Biophysical Journal*. 2009; 96:4176–4187.
107. Boxer SG. Stark realities. *Journal of Physical Chemistry B*. 2009; 113:2972–2983.
108. Ghosh A, Remorino A, Tucker MJ, Hochstrasser RM. 2D IR photon echo spectroscopy reveals hydrogen bond dynamics of aromatic nitriles. *Chemical Physics Letters*. 2009; 469:325–330.

109. Waegele MM, Tucker MJ, Gai F. 5-Cyanotryptophan as an infrared probe of local hydration status of proteins. *Chemical Physics Letters*. 2009; 478:249–253.
110. Aschaffenburg DJ, Moog RS. Probing hydrogen bonding environments: solvatochromatic effects on the CN vibration of benzonitrile. *Journal of Physical Chemistry B*. 2009; 113:12736–12743.
111. Ha JH, Lee KK, Park KH, Choi JH, Jeon SJ, Cho M. Integrated and dispersed photon echo studies of nitrile stretching vibration of 4-cyanophenol in methanol. *Journal of Chemical Physics*. 2009; 130:204509.
112. Marek P, Mukherjee S, Zanni MT, Raleigh DP. Residue-specific, real-time characterization of lag-phase species and fibril growth during amyloid formation: a combined fluorescence and IR study of pcyanophenylalanine analogs of islet amyloid polypeptide. *Journal of Molecular Biology*. 2010; 400:878–888.
113. McMahon HA, Alfieri KN, Clark CAA, Londergan CH. Cyanylated cysteine: a covalently attached vibrational probe of protein-lipid contacts. *Journal of Physical Chemistry Letters*. 2010; 1:850–855.
114. Inouye H, Gleason KA, Zhang D, Decatur SM, Kirschner DA. Differential effects of Phe19 and Phe20 on fibril formation by amyloidogenic peptide A beta 16–22 (Ac-KLVFFAE-NH₂). *Proteins: Structure, Function, and Bioinformatics*. 2010; 78:2306–2321.

115. Webb LJ, Boxer SG. Electrostatic fields near the active site of human aldose reductase: 1. New inhibitors and vibrational Stark effect measurements. *Biochemistry*. 2008; 47:1588–1598.
116. Stafford AJ, Ensign DL, Webb LJ. Vibrational Stark effect spectroscopy at the interface of Ras and Rap1A bound to the Ras binding domain of Raf1GDS reveals an electrostatic mechanism for protein-protein interaction. *Journal of Physical Chemistry B*. 2010; 114:15331-15344.
117. Fafarman AT, Sigala PA, Herschlag D, Boxer SG. Decomposition of vibrational shifts of nitriles into electrostatic and hydrogen-bonding effects. *Journal of the American Chemical Society* 2010; 132:12811–12813.
118. Miyake-Stoner SJ, Miller AM, Hammill JT, Peeler JC, Hess KR, Mehl RA, Brewer SH. Probing protein folding using site-specifically encoded unnatural amino acids as FRET donors with tryptophan. *Biochemistry*. 2009; 48:5953–5962.
119. Kondoh A, Yorimitsu H, Oshima K. Nucleophilic aromatic substitution reaction of nitroarenes with alkyl- or arylthio groups in dimethyl sulfoxide by means of cesium carbonate. *Tetrahedron*. 2006; 62:2357–2360.
120. Sano K, Ikegami Y, Uesugi T. Initial intramolecular formation of mercapturic acid. *Biological and Pharmaceutical Bulletin*. 2001; 24:1324–1328.
121. Timmerman P, Barderas R, Desmet J, Altschuh D, Shochat S, Hollestelle MJ, Hoppener JWM, Monasterio A, Casal JJ, Meloan RH. A combinatorial approach for the design of complementarity-determining region-derived

- peptidomimetics with in Vitro anti-tumoral activity. *Journal of Biological Chemistry*. 2009; 284:34126–34134.
122. Andrews SS, Boxer SG. Vibrational stark effects of nitriles I. Methods and experimental results. *Journal of Physical Chemistry A*. 2000; 104:11853–11863.
123. Waegele MM, Gai F. Computational modeling of the nitrile stretching vibration of 5-cyanoindole in water. *Journal of Physical Chemistry Letters*. 2010; 1:781–786.
124. J. W. Neidigh, R. M. Fesinmeyer, N. H. Andersen. Designing a 20-residue protein. *Nature Structural Biology*. 2002, 9, 425–430.
125. C. D. Snow, B. Zagrovic, V. S. Pande. The Trp cage: Folding kinetics and unfolded state topology via molecular dynamics simulations. *Journal of the American Chemical Society*. 2002, 124, 14548–14549.
126. C. Simmerling, B. Strockbine, A. E. Roitberg. All-atom structure prediction and folding simulations of a stable protein. *Journal of the American Chemical Society*. 2002, 124, 11258–11259.
127. S. Chowdhury, M. C. Lee, G. M. Xiong, Y. Duan. Ab initio folding simulation of the Trp-cage mini-protein approaches NMR resolution. *Journal of Molecular Biology*. 2003, 327, 711–717.
128. J. W. Pitera, W. Swope. Understanding folding and design: Replica-exchange simulations of "Trp-cage" fly miniproteins. *Proceedings of the National Academy of Sciences of the United States of America* 2003, 100, 7587–7592.

129. G. V. Nikiforovich, N. H. Andersen, R. M. Fesinmeyer, C. Frieden. Possible locally driven folding pathways of TC5b, a 20-residue protein. *Proteins Structure, Function, and Genetics*. 2003, 52, 292–302.
130. A. Schug, T. Herges, W. Wenzel. Reproducible protein folding with the stochastic tunneling method. *Physical Review Letters*. 2003, 91, 158102.
131. P. Carnevali, G. Toth, G. Toubassi, S. N. Meshkat. Fast protein structure prediction using Monte Carlo simulations with modal moves. *Journal of the American Chemical Society*. 2003, 125, 14244–14245.
132. S. Chowdhury, M. C. Lee, Y. Duan. Characterizing the rate-limiting step of Trp-cage folding by all-atom molecular dynamics simulations. *Journal of Physical Chemistry B* 2004, 108, 13855–13865.
133. P. J. Steinbach. Exploring peptide energy landscapes: A test of force fields and implicit solvent models. *Proteins Structure, Function, and Bioinformatics*. 2004, 57, 665–677.
134. M. Ota, M. Ikeguchi, A. Kidera. Phylogeny of protein-folding trajectories reveals a unique pathway to native structure. *Proceedings of the National Academy of Sciences of the United States of America* 2004, 101, 17658–17663.
135. A. Linhananta, J. Boer, I. MacKay. The equilibrium properties and folding kinetics of an all-atom G(o)over-bar model of the Trp-cage. *Journal of Chemical Physics*. 2005, 122, 114901–114915.

136. A. S. N. Seshasayee. High-Temperature unfolding of a Trp-cage mini-protein: a molecular dynamics simulation study. *Theoretical Biology and Medical Modelling*. 2005, 2, 7–11.
137. F. Ding, S. V. Buldyrev, N. V. Dokholyan. Folding Trp-cage to NMR resolution native structure using a coarse-grained protein model. *Biophysical Journal*. 2005, 88,147–155.
138. A. Irbäck, S. Mohanty. Folding thermodynamics of peptides. *Biophysical Journal*. 2005, 88, 1560–1569.
139. J. L. Alonso, P. Echenique. Relevant distance between two different instances of the same potential energy in protein folding. *Biophysical Chemistry*. 2005, 115, 159–168.
140. J. Chen, W. Im, C. L. Brooks. Balancing solvation and intramolecular interactions: Toward a consistent generalized born force field. *Journal of the American Chemical Society*. 2006, 128, 3728–3736.
141. L. X. Zhan, J. Z. Y. Chen, W. K. Liu. Computational study of the Trp-cage miniprotein based on the ECEPP/3 force field. *Proteins Structure, Function, and Genetics*. 2007, 66, 436–443.
142. D. Paschek, H. Nymeyer, A. E. Garcia. Replica exchange simulation of reversible folding/unfolding of the Trp-cage miniprotein in explicit solvent: On the structure and possible role of internal water. *Journal of Structural Biology*. 2007, 157, 524–533.

143. L. Yang, M. P. Grubb, Y. Q. Gao. Application of the accelerated molecular dynamics simulations to the folding of a small protein. *Journal of Chemical Physics*. 2007, 126, 125102.
144. D. A. C. Beck, G. W. N. White, V. Daggett. Exploring the energy landscape of protein folding using replica-exchange and conventional molecular dynamics simulations. *Journal of Structural Biology*. 2007, 157, 514–523.
145. S. Piana, A. Laio. A bias-exchange approach to protein folding. *Journal of Physical Chemistry B* 2007, 111, 4553–4559.
146. J. Copps, R. F. Murphy, S. Lovas. VCD spectroscopic and molecular dynamics analysis of the Trp-cage miniprotein TC5b. *Biopolymers* 2007, 88, 427–437.
147. A. Kentsis, T. Gindin, M. Mezei, R. Osman. Calculation of the Free Energy and Cooperativity of Protein Folding. *PLoS ONE* 2007, 2, e446.
148. Z. H. Hu, Y. H. Tang, H. F. Wang, X. Zhang, M. Lei. Dynamics and cooperativity of Trp-cage folding. *Archives in Biochemistry and Biophysics*. 2008, 475, 140–147.
149. P. Hudáky, P. Straner, V. Farkas, G. Varadi, G. Toth, A. Perczel. Cooperation between a salt bridge and the hydrophobic core triggers fold stabilization in a Trp-cage miniprotein. *Biochemistry* 2008, 47, 1007–1016.
150. W. X. Xu, Y. G. Mu. Ab initio folding simulation of Trpcage by replica exchange with hybrid Hamiltonian. *Biophysical Chemistry*. 2008, 137, 116–125.

151. D. Paschek, S. Hempel, A. E. Garcia. Computing the stability diagram of the Trp-cage miniprotein. *Proceedings of the National Academy of Science of the United States of America* 2008, 105, 17754–17759.
152. S. Wu, P. I. Zhuravlev, G. A. Papoian. High Resolution Approach to the Native State Ensemble Kinetics and Thermodynamics. *Biophysical Journal*. 2008, 95, 5524–5532.
153. X. Q. Yao, Z. S. She. Key residue-dominated protein folding dynamics. *Biochemical and Biophysical Research Communications*. 2008, 373, 64–68.
154. S. Kannan, M. Zacharias. Folding simulations of Trp-cage mini protein in explicit solvent using biasing potential replica-exchange molecular dynamics simulations. *Proteins Structure, Function, and Genetics*. 2009, 76, 448–460.
155. J. Cerný, J. Vondrasek, P. Hobza. Loss of Dispersion Energy Changes the Stability and Folding/Unfolding Equilibrium of the Trp-Cage Protein. *Journal of Physical Chemistry B* 2009, 113, 5657–5660.
156. Y. Chebaro, X. Dong, R. Laghaei, P. Derreumaux, N. Mousseau. Replica Exchange Molecular Dynamics Simulations of Coarse-grained Proteins in Implicit Solvent. *Journal of Physical Chemistry B* 2009, 113, 267–274.
157. D. Matthes, B. L. de Groot. Secondary Structure Propensities in Peptide Folding Simulations: A Systematic Comparison of Molecular Mechanics Interaction Schemes. *Biophysical Journal*. 2009, 97, 599–608.

158. F. Marinelli, F. Pietrucci, A. Laio, S. Piana. A Kinetic Model of Trp-Cage Folding from Multiple Biased Molecular Dynamics Simulations. *PLoS Computational Biology*. 2009, 5, e1000452.
159. Z. Gattin, S. Riniker, P. J. Hore, K. H. Mok, W. F. van Gunsteren. Temperature and urea induced denaturation of the Trp-Cage mini protein TC5b: A simulation study consistent with experimental observations. *Protein Science*. 2009, 18, 2090–2099.
160. M. Gao, H. Q. Zhu, X. Q. Yao, Z. S. She. Water dynamics clue to key residues in protein folding. *Biochemical and Biophysical Research Communications*. 2010, 392, 95–99.
161. C. Velez-Vega, E. E. Borrero, F. A. Escobedo. Kinetics and mechanism of the unfolding native-to-loop transition of Trp-cage in explicit solvent via optimized forward flux sampling simulations. *Journal of Chemical Physics*. 2010, 133, 105103.
162. N. J. Bruce, R. A. Bryce. Ab Initio Protein Folding Using a Cooperative Swarm of Molecular Dynamics Trajectories. *Journal of Chemical Theory and Computation*. 2010, 6, 1925–1930.
163. M. S. Lee, M. A. Olson. Protein Folding Simulations Combining Self-Guided Langevin Dynamics and Temperature-Based Replica Exchange. *Journal of Chemical Theory and Computation*. 2010, 6, 2477–2487.

164. R. Day, D. Paschek, A. E. Garcia. Microsecond simulations of the folding/unfolding thermodynamics of the Trp-cage miniprotein. *Proteins Structure, Function, and Genetics*. 2010, 78, 1889–1899.
165. W. Zheng, E. Gallicchio, N. Deng, M. Andrec, R. M. Levy. Kinetic Network Study of the Diversity and Temperature Dependence of Trp-Cage Folding Pathways: Combining Transition Path Theory with Stochastic Simulations. *Journal of Physical Chemistry B* 2011, 115, 1512–1523.
166. L. L. Qiu, S. A. Pabit, A. E. Roitberg, S. J. Hagen. Smaller and faster: The 20-residue Trp-cage protein folds in 4 μ s. *Journal of the American Chemical Society*. 2002, 124, 12952–12953.
167. M. R. Bunagan, X. Yang, J. G. Saven, F. Gai. Ultrafast folding of a computationally designed Trp-cage mutant: Trp²-cage. *Journal of Physical Chemistry B* 2006, 110, 3759–3763.
168. H. Neuweiler, S. Dose, M. Sauer. A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate. *Proceedings of the National Academy of Sciences of the United States of America* 2005, 102, 16650–16655.
169. K. H. Mok, L. T. Kuhn, M. Goez, I. J. Day, J. C. Lin, N. H. Andersen, P. J. Hore. A pre-existing hydrophobic collapse in the unfolded state of an ultrafast folding protein. *Nature* 2007, 447, 106–109.
170. Z. Ahmed, I. A. Beta, A. V. Mikhonin, S. A. Asher. UV-resonance Raman thermal unfolding study of Trp-cage shows that it is not a simple two-state

- miniprotein. *Journal of the American Chemical Society* 2005, 127, 10943–10950.
171. R. H. Zhou. Trp-cage: Folding free energy landscape in explicit water. *Proceedings of the National Academy of Sciences of the United States of America* 2003, 100, 13280–13285.
172. C. Y. Huang, Z. Getahun, Y. J. Zhu, J. W. Klemke, W. F. Degrado, F. Gai. Helix formation via conformation diffusion search. *Proceedings of the National Academy of Sciences of the United States of America* 2002, 99, 2788–2793.
173. C. Y. Huang, Z. Getahun, T. Wang, W. F. Degrado, F. Gai. Time-resolved infrared study of the helix-coil transition using C-13-labeled helical peptides. *Journal of the American Chemical Society*. 2001, 123, 12111–12112.
174. B. Barua, J. C. Lin, V. D. Williams, P. Kummler, J. W. Neidigh, N. H. Andersen. The Trp-cage: optimizing the stability of a globular miniprotein. *Protein Engineering, Design, and Selection*. 2008, 21, 171–185.
175. D. V. Williams, A. Byrne, J. Stewart, N. H. Andersen. Optimal Salt Bridge for Trp-Cage Stabilization. *Biochemistry* 2011, 50, 1143–1152.
176. S. Bagchi, C. Falvo, S. Mukamel, R. M. Hochstrasser. 2D-IR Experiments and Simulations of the Coupling between Amide-I and Ionizable Side Chains in Proteins: Application to the Villin Headpiece. *Journal of Physical Chemistry B* 2009, 113, 11260–11273.
177. D. F. Kennedy, M. Crisma, C. Toniolo, D. Chapman. Studies of Peptides Forming 3_{10} -Helices and Alpha-Helices and Beta-Bend Ribbon Structures in

- Organic Solution and in Model Biomembranes by Fourier-Transform Infrared-Spectroscopy. *Biochemistry* 1991, 30, 6541–6548.
178. R. A. G. D. Silva, S. C. Yasui, J. Kubelka, F. Formaggio, M. Crisma, C. Toniolo, T. A. Keiderling. Discriminating 3_{10} from alpha helices: Vibrational and electronic CD and IR absorption study of related Aib-containing oligopeptides. *Biopolymers* 2002, 65, 229–243.
179. S. Williams, T. P. Causgrove, R. Gilmanishin, K. S. Fang, R. H. Callender, W. H. Woodruff, R. B. Dyer. Fast events in protein folding: Helix melting and formation in a small peptide. *Biochemistry* 1996, 35, 691–697.
180. P. A. Thompson, W. A. Eaton, J. Hofrichter. Laser temperature jump study of the helix reversible arrow coil kinetics of an alanine peptide interpreted with a 'kinetic zipper' model. *Biochemistry* 1997, 36, 9200–9210.
181. C. Y. Huang, J. W. Klemke, Z. Getahun, W. F. DeGrado, F. Gai. Temperature-dependent helix-coil transition of an alanine based peptide. *Journal of the American Chemical Society*. 2001, 123, 9235–9238.
182. R. H. Zhou. Exploring the protein folding free energy landscape: coupling replica exchange method with P3ME/RESPA algorithm. *Journal of Molecular Graphics and Modelling*. 2004, 22, 451–463.
183. J. Juraszek, P. G. Bolhuis. Sampling the multiple folding mechanisms of Trp-cage in explicit solvent. *Proceedings of the National Academy of Sciences of the United States of America* 2006, 103, 15859–15864.

184. Y. Xu, R. Oyola, F. Gai. Infrared study of the stability and folding kinetics of a 15-residue beta-hairpin. *Journal of the American Chemical Society*. 2003, 125, 15388–15394.
185. A. L. Serrano, M. J. Tucker, F. Gai. Direct Assessment of the alpha-Helix Nucleation Time. *Journal of Physical Chemistry B* 2011, 115, 7472–7478.
186. S. Mukherjee, P. Chowdhury, M. R. Bunagan, F. Gai. Folding kinetics of a naturally occurring helical peptide: Implication of the folding speed limit of helical proteins. *Journal of Physical Chemistry B* 2008, 112, 9146–9150.
187. H. Maity, M. Maity, M. M. G. Krishna, L. Mayne, S. W. Englander. Protein folding: The stepwise assembly of foldon units. *Proceedings of the National Academy of Sciences of the United States of America* 2005, 102, 4741–4746.
188. Matouschek, A.; Serrano, L.; Fersht, A. R. The Folding of an Enzyme. 4. Structure of an Intermediate in the Refolding of Barnase Analyzed by a Protein Engineering Procedure. *Journal of Molecular Biology*. 1992, 224, 819-835.
189. Koh, J. T.; Cornish, V. W.; Schultz, P. G. An experimental approach to evaluating the role of backbone interactions in proteins using unnatural amino acid mutagenesis. *Biochemistry* 1997, 36, 11314-11322.
190. Lu, W. Y.; Qasim, M. A.; Laskowski, M.; Kent, S. B. H. Probing intermolecular main chain hydrogen bonding in serine proteinase-protein inhibitor complexes: Chemical synthesis of backbone-engineered turkey ovomucoid third domain. *Biochemistry* 1997, 36, 673-679.

191. Beligere, G. S.; Dawson, P. E. Design, synthesis, and characterization of 4-ester CI2, a model for backbone hydrogen bonding in protein alpha-helices. *Journal of the American Chemical Society*. 2000, 122, 12079-12082.
192. Deechongkit, S.; Nguyen, H.; Powers, E. T.; Dawson, P. E.; Gruebele, M.; Kelly, J. W. Context-dependent contributions of backbone hydrogen bonding to beta-sheet folding energetics. *Nature* 2004, 430, 101-105.
193. Bunagan, M. R.; Gao, J.; Kelly, J. W.; Gai, F. Probing the Folding Transition State Structure of the Villin Headpiece Subdomain via Side Chain and Backbone Mutagenesis. *Journal of the American Chemical Society*. 2009, 131, 7470-7476.
194. Alemán, C. On the ability of modified peptide links to form hydrogen bonds. *Journal of Physical Chemistry A* 2001, 105, 6717-6723.
195. Reiner, A.; Wildemann, D.; Fischer, G.; Kiefhaber, T. Effect of thiopeptide bonds on alpha-helix structure and stability. *Journal of the American Chemical Society*. 2008, 130, 8079.
196. Bachmann, A.; Wildemann, D.; Praetorius, F.; Fischer, G.; Kiefhaber, T. Mapping backbone and side-chain interactions in the transition state of a coupled protein folding and binding reaction. *Proceedings of the National Academy of Sciences of the United States of America*. 2011, 108, 3952.
197. Sanchez, I. E.; Kiefhaber, T. Origin of unusual Phi-values in protein folding: Evidence against specific nucleation sites. *Journal of Molecular Biology*. 2003, 334, 1077.

198. Shalaby, M. A.; Grote, C. W.; Rapoport, H. Thiopeptide synthesis. Alpha-amino thionoacid derivatives of nitrobenzotriazole as thioacylating agents. *Journal of Organic Chemistry*. 1996, 61, 9045.
199. Miwa, J. H.; Margarida, L. A.; Meyer, A. E. Improved acidolytic deprotection conditions for the Fmoc-based solid-phase synthesis of thioxo peptides. *Tetrahedron Letters* 2001, 42, 7189.
200. Caba, J. M.; Rodriguez, I. M.; Manzanares, I.; Giralt, E.; Albericio, F. Solid-phase total synthesis of trunkamide A. *Journal of Organic Chemistry*. 2001, 66, 7568.
201. Batjargal, S.; Wang, Y. J.; Goldberg, J. M.; Wissner, R. F.; Petersson, E. J. Native Chemical Ligation of Thioamide-Containing Peptides: Development and Application to the Synthesis of Labeled alpha-Synuclein for Misfolding Studies. *Journal of the American Chemical Society*. 2012, 134, 9172-9182.
202. Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. Folding dynamics and mechanism of beta-hairpin formation. *Nature* 1997, 390, 196-199.
203. Maness, S. J.; Franzen, S.; Gibbs, A. C.; Causgrove, T. P.; Dyer, R. B. Nanosecond temperature jump relaxation dynamics of cyclic beta-hairpin peptides. *Biophysical Journal*. 2003, 84, 3874.
204. Meier, S.; Guthe, S.; Kiefhaber, T.; Grzesiek, S. Foldon, the natural trimerization domain of T4 fibrin, dissociates into a monomeric A-state form containing a stable beta-hairpin: Atomic details of trimer dissociation and local

- beta-hairpin stability from residual dipolar couplings. *Journal of Molecular Biology*. 2004, 344, 1051.
205. Dyer, R. B.; Maness, S. J.; Peterson, E. S.; Franzen, S.; Fesinmeyer, R. M.; Andersen, N. H. The mechanism of beta-hairpin formation. *Biochemistry* 2004, 43, 11560.
206. Petrovich, M.; Jonsson, A. L.; Ferguson, N.; Daggett, V.; Fersht, A. R. Phi-Analysis at the experimental limits: Mechanism of beta-hairpin formation. *Journal of Molecular Biology*. 2006, 360, 865.
207. Narayanan, R.; Pelakh, L.; Hagen, S. J. Solvent Friction Changes the Folding Pathway of the Tryptophan Zipper TZ2. *Journal of Molecular Biology*. 2009, 390, 538-546.
208. Xu, Y.; Du, D. G.; Oyola, R. Infrared Study of the Stability and Folding Kinetics of a Series of beta-Hairpin Peptides with a Common NPDG Turn. *Journal of Physical Chemistry B* 2011, 115, 15332.
209. Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Tryptophan zippers: Stable, monomeric beta-hairpins. *Proceedings of the National Academy of Sciences of the United States of America*. 2001, 98, 5578-5583.
210. Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. Stability of cyclic beta-hairpins: Asymmetric contributions from side chains of a hydrogen-bonded cross-strand residue pair. *Journal of the American Chemical Society*. 2003, 125, 388-395.

211. Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. Heterogeneous folding of the trpzip hairpin: Full atom simulation and experiment. *Journal of Molecular Biology*. 2004, 336, 241-251.
212. Dempsey, C. E.; Piggot, T. J.; Mason, P. E. Dissecting contributions to the denaturant sensitivities of proteins. *Biochemistry* 2005, 44, 775-781.
213. Guvench, O.; Brooks, C. L. Tryptophan side chain electrostatic interactions determine edge-to-face vs parallel-displaced tryptophan side chain geometries in the designed beta-hairpin "trpzip2". *Journal of the American Chemical Society*. 2005, 127, 4668-4674.
214. 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, 17025-17027.
215. Pitera, J. W.; Haque, I.; Swope, W. C. Absence of reptation in the high-temperature folding of the trpzip2 beta-hairpin peptide. *Journal of Chemical Physics* 2006, 124, 141102.
216. Zhang, J.; Qin, M.; Wang, W. Folding mechanism of beta-hairpins studied by replica exchange molecular simulations. *Proteins Structure, Function, and Bioinformatics* 2006, 62, 672-685.
217. Streicher, W. W.; Makhatadze, G. I. Calorimetric evidence for a two-state unfolding of the beta-hairpin peptide trpzip4. *Journal of the American Chemical Society*. 2006, 128, 30-31.

218. Wang, J. P.; Chen, J. X.; Hochstrasser, R. M. Local structure of beta-hairpin isotopomers by FTIR, 2D IR, and ab initio theory. *Journal of Physical Chemistry B* 2006, 110, 7545-7555.
219. Hauser, K.; Krejtschi, C.; Huang, R.; Wu, L.; Keiderling, T. A. Site-specific relaxation kinetics of a tryptophan zipper hairpin peptide using temperature-jump IR spectroscopy and isotopic labeling. *Journal of the American Chemical Society*. 2008, 130, 2984-2992.
220. Settanni, G.; Fersht, A. R. High temperature unfolding simulations of the TRPZ1 peptide. *Biophysical Journal*. 2008, 94, 4444-4453.
221. Takekiyo, T.; Wu, L.; Yoshimura, Y.; Shimizu, A.; Keiderling, T. A. Relationship between Hydrophobic Interactions and Secondary Structure Stability for Trpzip beta-Hairpin Peptides. *Biochemistry* 2009, 48, 1543-1552.
222. Wu, L.; McElheny, D.; Huang, R.; Keiderling, T. A. Role of Tryptophan-Tryptophan Interactions in Trpzip beta-Hairpin Formation, Structure, and Stability. *Biochemistry* 2009, 48, 10362-10371.
223. Cheng, Z. H.; Campbell, R. E. An engineered tryptophan zipper-type peptide as a molecular recognition scaffold. *Journal of Peptide Science*. 2009, 15, 523-532.
224. Huang, R.; Wu, L.; McElheny, D.; Bour, P.; Roy, A.; Keiderling, T. A. Cross-Strand Coupling and Site-Specific Unfolding Thermodynamics of a Trpzip beta-Hairpin Peptide Using C-13 Isotopic Labeling and IR Spectroscopy. *Journal of Physical Chemistry B* 2009, 113, 5661-5674.

225. Hills, R. D.; Lu, L. Y.; Voth, G. A. Multiscale Coarse-Graining of the Protein Energy Landscape. *PLOS Computational Biology*. 2010, 6.
226. Roy, S.; Jansen, T. L. C.; Knoester, J. Structural classification of the amide I sites of a beta-hairpin with isotope label 2DIR spectroscopy. *Physical Chemistry Chemical Physics*. 2010, 12, 9347-9357.
227. Smith, A. W.; Lessing, J.; Ganim, Z.; Peng, C. S.; Tokmakoff, A.; Roy, S.; Jansen, T. L. C.; Knoester, J. Melting of a beta-Hairpin Peptide Using Isotope-Edited 2D IR Spectroscopy and Simulations. *Journal of Physical Chemistry B* 2010, 114, 10913-10924.
228. Hwang, S.; Hilty, C. Folding of a Tryptophan Zipper Peptide Investigated on the Basis of the Nuclear Overhauser Effect and Thermal Denaturation. *Journal of Physical Chemistry B* 2011, 115, 15355-15361.
229. Dempsey, C. E.; Mason, P. E.; Jungwirth, P. Complex Ion Effects on Polypeptide Conformational Stability: Chloride and Sulfate Salts of Guanidinium and Tetrapropylammonium. *Journal of the American Chemical Society*. 2011, 133, 7300-7303.
230. Huang, J. J. T.; Larsen, R. W.; Chan, S. I. The interplay of turn formation and hydrophobic interactions on the early kinetic events in protein folding. *Chemical Communications*. 2012, 48, 487-497.
231. Grishina, I. B.; Woody, R. W. Contributions of Tryptophan Side-Chains to the Circular-Dichroism of Globular-Proteins - Exciton Couplets and Coupled Oscillators. *Faraday Discussions*. 1994, 99, 245-262.

232. Du, D. G.; Gai, F. Understanding the folding mechanism of an alpha-helical hairpin. *Biochemistry* 2006, 45, 13131-13139.
233. Zhu, Y. J.; Fu, X. R.; Wang, T.; Tamura, A.; Takada, S.; Saven, J. G.; Gai, F. Guiding the search for a protein's maximum rate of folding. *Chemical Physics*. 2004, 307, 99-109.
234. Venkatraman, J.; Gowda, G. A. N.; Balaram, P. Design and construction of an open multi stranded beta-sheet polypeptide stabilized by a disulfide bridge. *Journal of the American Chemical Society*. 2002, 124, 4987-4994.
235. Kiehna, S. E.; Waters, M. L. Sequence dependence of beta-hairpin structure: Comparison of a salt bridge and an aromatic interaction. *Protein Science*. 2003, 12, 2657-2667.
236. Serrano, A. L.; Waagele, M. M.; Gai, F. Spectroscopic studies of protein folding: Linear and nonlinear methods. *Protein Science*. 2012, 21, 157-170.
237. Culik, R. M.; Serrano, A. L.; Bunagan, M. R.; Gai, F. Achieving Secondary Structural Resolution in Kinetic Measurements of Protein Folding: A Case Study of the Folding Mechanism of Trp-cage. *Angewandte Chemie-International Edition* 2011, 50, 10884-10887.
238. Snow, C. D.; Qiu, L.; Du, D.; Gai, F.; Hagen, S. J.; Pande, V. S. Trp zipper folding kinetics by molecular dynamics and temperature-jump spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America* 2004, 101, 4077-4082.

239. Du, D.; Zhu, Y.; Huang, C. Y.; Gai, F. Understanding the key factors that control the rate of beta-hairpin folding. *Proceedings of the National Academy of Sciences of the United States of America*. 2004, 101, 15915-15920.
240. Du, D.; Tucker, M. J.; Gai, F. Understanding the mechanism of beta-hairpin folding via phi-value analysis. *Biochemistry* 2006, 45, 2668-2678.
241. Best, R. B.; Mittal, J. Microscopic events in beta-hairpin folding from alternative unfolded ensembles. *Proceedings of the National Academy of Sciences of the United States of America*. 2011, 108, 11087-11092.
242. Xu, Y.; Purkayastha, P.; Gai, F. Nanosecond folding dynamics of a three-stranded beta-sheet. *Journal of the American Chemical Society*. 2006, 128, 15836-15842.
243. S. Costantini, G. Colonna, A.M. Facchiano. Amino acid propensities for secondary structures are influenced by the protein structural class. *Biochemical and Biophysical Research Communications*. 342 (2006) 441-451.
244. D. Bang, A.V. Gribenko, V. Tereshko, A.A. Kossiakoff, S.B. Kent, G.I. Makhatadze. Dissecting the energetics of protein alpha-helix C-cap termination through chemical protein synthesis. *Nature Chemical Biology*. 2 (2006) 139-143.
245. W.E. Stites, J. Pranata. Empirical Evaluation of the Influence of Side-Chains on the Conformational Entropy of the Polypeptide Backbone. *Proteins* 22 (1995) 132-140.

246. B. Anil, B.B. Song, Y.F. Tang, D.P. Raleigh. Exploiting the right side of the ramachandran plot: Substitution of glycines by D-alanine can significantly increase protein stability. *Journal of the American Chemical Society*. 126 (2004) 13194-13195.
247. F.I. Valiyaveetil, M. Sekedat, R. MacKinnon, T.W. Muir. Glycine as a D-amino acid surrogate in the K⁺-selectivity filter. *Proceedings of the National Academy of Sciences of the United States of America* 101 (2004) 17045-17049.
248. D.V. Williams, B. Barua, N.H. Andersen. Hyperstable miniproteins: additive effects of D- and L-Ala mutations. *Organic and Biomolecular Chemistry* 6 (2008) 4287-4289.
249. M.D. Struthers, R.P. Cheng, B. Imperiali. Design of a monomeric 23-residue polypeptide with defined tertiary structure. *Science* 271 (1996) 342-345.
250. A. Rodriguez-Granillo, S. Annavarapu, L. Zhang, R.L. Koder, V. Nanda. Computational Design of Thermostabilizing D-Amino Acid Substitutions. *Journal of the American Chemical Society*. 133 (2011) 18750-18759.
251. J.M. Stewart, J.C. Lin, N.H. Andersen. Lysine and arginine residues do not increase the helicity of alanine-rich peptide helices. *Chemical Communications*. (2008) 4765-4767.
252. R.B. Best, J. Mittal. Balance between alpha and beta Structures in Ab Initio Protein Folding. *Journal of Physical Chemistry B* 114 (2010) 8790-8798.
253. Luo, P.; Baldwin, R. L. Mechanism of helix induction by trifluoroethanol: A framework for extrapolating the helix-forming properties of peptides from

trifluoroethanol/water mixtures back to water. *Biochemistry* 1997, 36, 8413-8421.

254. Kentsis, A.; Sosnick, T. R. Trifluoroethanol promotes helix formation by destabilizing backbone exposure: Desolvation rather than native hydrogen bonding defines the kinetic pathway of dimeric coiled coil folding. *Biochemistry* 1998, 37, 14613-14622.
255. Walgers, R.; Lee, T. C.; Cammers-Goodwin, A. An indirect chaotropic mechanism for the stabilization of helix conformation of peptides in aqueous trifluoroethanol and hexafluoro-2-propanol. *Journal of the American Chemical Society*. 1998, 120, 5073-5079.
256. Gast, K.; Zirwer, D.; Müller-Frohne, M.; Damaschun, G. Trifluoroethanol-induced conformational transitions of proteins: Insights gained from the differences between alpha-lactalbumin and ribonuclease A. *Protein Science*. 1999, 8, 625-634.
257. Hong, D. P.; Hoshino, M.; Kuboi, R.; Goto, Y. Clustering of fluoroine-substituted alcohols as a factor responsible for their marked effects on proteins and peptides. *Journal of the American Chemical Society*. 1999, 121, 8427-8433.
258. Yiu, C. P. B.; Mateu, M. G.; Fersht, A. R. Protein folding transition states: Elicitation of Hammond effects by 2,2,2-trifluoroethanol. *Chembiochem* 2000, 1, 49-55.

259. Hamada, D.; Chiti, F.; Gujjarro, J. I.; Kataoka, M.; Taddei, N.; Dobson, C. M. Evidence concerning rate-limiting steps in protein folding from the effects of trifluoroethanol. *Nature Structural Biology*. 2000, 7, 58-61.
260. Roccatano, D.; Colombo, G.; Fioroni, M.; Mark, A. E. Mechanism by which 2,2,2-trifluoroethanol/water mixtures stabilize secondary-structure formation in peptides: A molecular dynamics study. *Proceedings of the National Academy of Sciences of the United States of America*. 2002, 99, 12179-12184.
261. Chatterjee, C.; Gerig, J. T. Interactions of trifluoroethanol with the Trp-cage peptide. *Biopolymers* 2007, 87, 115-123.
262. Jalili, S.; Akhavan, M. Molecular dynamics simulation study of association in trifluoroethanol/water mixtures. *Journal of Computational Chemistry*. 2010, 31, 286-294.
263. Shao, Q.; Fan, Y.; Yang, L.; Gao, Y. Q. From protein denaturant to protectant: Comparative molecular dynamics study of alcohol/protein interactions. *Journal of Chemical Physics*. 2012, 136, 115101.
264. Anderson, V. L.; Webb, W. W. A desolvation model for trifluoroethanol-induced aggregation of enhanced green fluorescent protein. *Biophysical Journal*. 2012, 102, 897-906.
265. Gerig, J. T. Toward a molecular dynamics force field for simulations of 40% trifluoroethanol-water. *Journal of Physical Chemistry B* 2014 118 1471-1480.

266. Mukherjee, S.; Waagele, M. M.; Chowdhury, P.; Guo, L.; Gai, F. Effect of macromolecular crowding on protein folding dynamics at the secondary structure level. *Journal of Molecular Biology*. 2009, 393, 227-236.
267. Hua, Q. X.; Jia, W. H.; Bullock, B. P.; Habener, J. F.; Weiss, M. A. Transcriptional activator-coactivator recognition: Nascent folding of a kinase-inducible transactivation domain predicts its structure on coactivator binding. *Biochemistry* 1998, 37, 5858-5866.
268. Cho, S. S.; Reddy, G.; Straub, J. E.; Thirumalai, D. Entropic stabilization of proteins by TMAO. *Journal of Physical Chemistry B* 2011, 115, 13401-13407.
269. Lu, H. S. M.; Volk, M.; Kholodenko, Y.; Gooding, E.; Hochstrasser, R. M.; DeGrado, W. F. Aminothietyrosine disulfide, an optical trigger for initiation of protein folding. *Journal of the American Chemical Society*. 1997, 119, 7173-7180.
270. Natansohn, A.; Rochon, P. Photoinduced motions in azo-containing polymers. *Chemical Reviews*. 2002, 102, 4139-4175.
271. Boudebous, H.; Košmrlj, B.; Šket, B.; Wirz, J. Primary photoreactions of the 3',5'-dimethoxybenzoin cage and determination of the release rate in polar media. *Journal of Physical Chemistry A* 2007, 111, 2811-2813.
272. Hoppmann, C.; Barucker, C.; Lorenz, D.; Multhaup, G.; Beyermann, M. Light-controlled toxicity of engineered amyloid beta-peptides. *Chembiochem*. 2012, 13, 2657-2660.

273. Ali, A. M.; Woolley, G. A. The effect of azobenzene cross-linker position on the degree of helical peptide photo-control. *Organic and Biomolecular Chemistry*. 2013, 11, 5325-5331.
274. Beharry, A. A.; Woolley, G. A. Azobenzene photoswitches for biomolecules. *Chemical Society Reviews*. 2011, 40, 4422-4437.
275. Kumita, J. R.; Flint, D. G.; Smart, O. S.; Woolley, G. A. Photo-control of peptide helix content by an azobenzene cross-linker: steric interactions with underlying residues are not critical. *Protein Engineering, Design, and Selection*. 2002, 15, 561-569.
276. Chen, E.; Kumita, J. R.; Woolley, G. A.; Kliger, D. S. The kinetics of helix unfolding of an azobenzene cross-linked peptide probed by nanosecond time-resolved optical rotatory dispersion. *Journal of the American Chemical Society*. 2003, 125, 12443-12449.
277. Kumita, J. R.; Flint, D. G.; Woolley, G. A.; Smart, O. S. Achieving photo-control of protein conformation and activity: producing a photo-controlled leucine zipper. *Faraday Discussions*. 2003, 122, 89-103.
278. Bredenbeck, J.; Helbing, J.; Sieg, A.; Schrader, T.; Zinth, W.; Renner, C.; Behrendt, R.; Moroder, L.; Wachtveitl, J.; Hamm, P. Picosecond conformational transition and equilibration of a cyclic peptide. *Proceedings of the National Academy of Sciences of the United States of America*. 2003, 100, 6452-6457.
279. Bredenbeck, J.; Helbing, J.; Kumita, J. R.; Woolley, G. A.; Hamm, P. Alpha-helix formation in a photoswitchable peptide tracked from picoseconds to

- microseconds by time-resolved IR spectroscopy. *Proceedings of the National Academy of Sciences of the United States of America*. 2005, 102, 2379-2384.
280. Schrader, T. E.; Schreier, W. J.; Cordes, T.; Koller, F. O.; Babitzki, G.; Denschlag, R.; Renner, C.; Löweneck, M.; Dong, S. L.; Moroder, L.; Tavan, P.; Zinth, W. Light-triggered beta-hairpin folding and unfolding. *Proceedings of the National Academy of Sciences of the United States of America*. 2007, 104, 15729-15734.
281. Pfister, R.; Ihalainen, J.; Hamm, P.; Kolano, C. Synthesis, characterization, and applicability of three isotope labeled azobenzene photoswitches. *Organic and Biomolecular Chemistry*. 2008, 6, 3508-3517.
282. Ihalainen, J. A.; Paoli, B.; Muff, S.; Backus, E. H. G.; Bredenbeck, J.; Woolley, G. A.; Caflisch, A.; Hamm, P. Alpha-helix folding in the presence of structural constraints. *Proceedings of the National Academy of Sciences of the United States of America*. 2008, 105, 9588-9593.
283. Buchli, B.; Waldauer, S. A.; Walser, R.; Donten, M. L.; Pfister, R.; Blöchliger, N.; Steiner, S.; Caflisch, A.; Zerbe, O.; Hamm, P. Kinetic response of a photoperturbed allosteric protein. *Proceedings of the National Academy of Sciences of the United States of America*. 2013, 110, 11725-11730.
284. Rajan, R.; Balaram, P. A model for the interaction of trifluoroethanol with peptides and proteins. *International Journal of Peptide and Protein Research*. 1996, 48, 328-336.

285. Diaz, M.; Fioroni, M.; Burger, K.; Berger, S. Evidence of complete hydrophobic coating of bombesin by trifluoroethanol in aqueous solution: An NMR spectroscopic and molecular dynamics study. *Chemistry – A European Journal*. 2002, 8, 1663-1669.
286. Cammers-Goodwin, A.; Allen, T. J.; Oslick, S. L.; McClure, K. F.; Lee, J. H.; Kemp, D. S. Mechanism of Stabilization of Helical Conformations of Polypeptides by Water Containing Trifluoroethanol. *Journal of the American Chemical Society*. 1996, 118, 3082-3090.
287. Buck, M.; Schwalbe, H.; Dobson, C. M. Characterization of Conformational Preferences in a Partly Folded Protein by Heteronuclear NMR Spectroscopy: Assignment and Secondary Structure Analysis of Hen Egg-White Lysozyme in Trifluoroethanol. *Biochemistry*. 1995, 34, 13219-13232.
288. Gente, G.; La Mesa, C. Water—Trifluoroethanol Mixtures: Some Physicochemical Properties. *Journal of Solution Chemistry*. 2000, 29, 1159-1172.
289. Radhakrishnan, I.; Pérez-Alvarado, G. C.; Parker, D.; Dyson, H. J.; Montminy, M. R.; Wright, P. E. Solution Structure of the KIX Domain of CBP Bound to the Transactivation Domain of CREB: A Model for Activator:Coactivator Interactions. *Cell*. 1997, 91, 741-752.
290. Furuki, T.; Shimizu, T.; Chakrabortee, S.; Yamakawa, K.; Hatanaka, R.; Takahashi, T.; Kikawada, T.; Okuda, T.; Mihara, H.; Tunnacliffe, A.; Sakurai,

- M. Effects of Group 3 LEA protein model peptides on desiccation-induced protein aggregation. *Biochimica et Biophysica Acta*. 2012, 1824, 891-897.
291. Klimov, D.; Thirumalai, D. Viscosity Dependence of the Folding Rates of Proteins. *Physical Review Letters*. 1997, 79, 317-320.
292. Zagrovic, B.; Pande, V. Solvent viscosity dependence of the folding rate of a small protein: distributed computing study. *Journal of Computational Chemistry*. 2003, 24, 1432-1436.
293. Jas, G. S.; Eaton, W. A.; Hofrichter, J. Effect of Viscosity on the Kinetics of α -Helix and β -Hairpin Formation. *Journal of Physical Chemistry B*. 2001, 105, 261-272.
294. Fioroni, M.; Diaz, M. D.; Burger, K.; Berger, S. Solvation phenomena of a tetrapeptide in water/trifluoroethanol and water/ethanol mixtures: a diffusion NMR, intermolecular NOE, and molecular dynamics study. *Journal of the American Chemical Society*. 2002, 124, 7737-7744.