1-1-2016

Using Molecular Constraints and Unnatural Amino Acids to Manipulate and Interrogate Protein Structure, Dynamics, and Self-Assembly

Beatrice Markiewicz
University of Pennsylvania, bmarki@sas.upenn.edu

Using Molecular Constraints and Unnatural Amino Acids to Manipulate and Interrogate Protein Structure, Dynamics, and Self-Assembly

Abstract
Protein molecules can undergo a wide variety of conformational transitions occurring over a series of time and distance scales, ranging from large-scale structural reorganizations required for folding to more localized and subtle motions required for function. Furthermore, the dynamics and mechanisms of such motions and transitions delicately depend on many factors and, as a result, it is not always easy, or even possible, to use existing experimental techniques to arrive at a molecular level understanding of the conformational event of interest. Therefore, this thesis aims to develop and utilize non-natural chemical modification strategies, namely molecular cross-linkers and unnatural amino acids as site-specific spectroscopic probes, in combination with various spectroscopic methods to examine, in great detail, certain aspects of protein folding and functional dynamics, and to manipulate protein self-assemblies. Specifically, we first demonstrate how strategically placed molecular constraints can be used to manipulate features of the protein folding free energy landscape, thus, allowing direct measurement of key components via temperature-jump kinetic studies, such as folding from a transition-state structure or the effect of internal friction on the folding mechanism. Secondly, we utilize a photolabile non-natural amino acid, Lys(nvoc), to probe the mechanism of protein misfolding in a β-hairpin model and identify an aggregation gatekeeper that tunes the aggregation propensity. We further develop a method where the induced-charge produced by photocleavage of Lys(nvoc) can be used to target and destabilize hydrophobic regions of amyloid fibril assemblies, resulting in complete disassembly, Finally, we highlight new useful properties of a site-specific spectroscopic probe, 5-cyanotryptophan (TrpCN), by demonstrating (1) how the frequency and linewidth of the infrared nitrile stretching vibration is sensitive to multiple hydrogen bonding interactions and solvent polarity, (2) that the fluorescence emission, quantum yield, and lifetime is extremely sensitive to hydration, and serves as a convenient fluorescence probe of protein solvation status, and (3) that the unique characteristics of TrpCN can be used to target the structure, local environment, and mechanism of the tryptophan gate in the M2 membrane proton channel of the influenza A virus.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Chemistry

First Advisor
Feng Gai

Subject Categories
Biophysics | Chemistry | Physical Chemistry

USING MOLECULAR CONTSTRAINTS AND UNNATURAL AMINO ACIDS TO MANIPULATE AND INTERROGATE PROTEIN STRUCTURE, DYNAMICS, AND SELF-ASSEMBLY

Beatrice N. Markiewicz

A DISSERTATION

in

Chemistry

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

_____

Feng Gai
Edmund J. and Louise W. Kahn Endowed Term Professor of Chemistry

Graduate Group Chairperson

_____

Gary A. Molander
Hirschmann-Makineni Professor of Chemistry

Dissertation Committee

Marsha Lester (Chair), Edmund J. Kahn Distinguished Professor
Ivan Dmochowski, Professor of Chemistry
Sergei Vinogradov, Associate Professor of Biochemistry and Biophysics

USING MOLECULAR CONTSTRAINTS AND UNNATURAL AMINO ACIDS TO MANIPULATE AND INTERROGATE PROTEIN STRUCTURE, DYNAMICS, AND SELF-ASSEMBLY

DEDICATION

To my family

ACKNOWLEDGEMENTS

Looking back on the last several years, my time at Penn has been enriching, rewarding, and of course challenging. I consider myself to be extremely fortunate to have had such a strong support system, and there many people to thank.

The person most instrumental in developing and shaping my scientific career is my thesis advisor, Professor Feng Gai. His constant dedication to not only me, but all of his students is really admirable. One of the most memorable lessons he has taught me, which has been motivating me throughout my graduate career, is that success in science may involve a little bit of luck, but it mainly comes from your own perseverance and hard work. His mentorship, enthusiasm toward science, willingness to teach and discuss how to approach research, and words of encouragement, has created an environment where I was able to mature both intellectually and individually. Professor Feng Gai truly devoted an immense amount of time, passion, and support to help me grow as a scientist and as a person, and for that I am extremely grateful and thankful.

I also want to thank my committee members, Professors Marsha Lester, Ivan Dmochowski, and Sergei Vinogradov, for their guidance and encouragement over the years. Their valuable input has allowed me to improve my abilities as a scientist. Also, I greatly appreciate the efforts of Dr. Lester, Dr. Dmochowski, and Dr. Tobias Baumgart in helping me receive external funding. I thank my collaborators Professor Bill DeGrado, Dr. Hyunil Jo, Dr. Thomas Lemmin, Dr. Giacomo Fiorin, Dr. Steven Brown, and Professor Amos B. Smith for their help with protein samples, molecular dynamics simulations, synthesis of compounds, and scientific inquiry.

Also, I sincerely thank the former and current Gai lab members and the UOPL Staff Members: Dr. Robert Culik, Dr. Ileana Pazos, Dr. Wenkai Zhang, Dr. Thomas Measey, Dr. Bei Ding, Dr. Thomas Troxler, Dr. Jian-Xin Chen, Dr. Kwang Im Oh, Dr. Natalia Rubtsova, Rachel Abaskharon, Mary Rose Hilaire, Chun-Wei Lin, Ismail Ahmed, Debopreeti Mukherjee, and Jeffrey Rodgers. None of the work here would have been possible without the support of the group and I thank everyone for contributing to such a collaborative environment. I consider myself so lucky to have worked with such a brilliant group of people. I especially want to thank Robert Culik who has been an exceptional mentor and friend, and who played a major role in my training when I first started. Also, I want to thank Lali, whose vibrant personality made each day in the lab an enjoyable one. I would also like to thank my collaborator Wenkai, who has an amazing excitement toward science, a natural ability to mentor, and imparts humor and wisdom simultaneously, makings for very delightful conversations. I want to specially thank Tom Troxler as well for his collaboration and mentorship. I would also like to take the time to thank all the friends I have made in the department.

Most of all, I need to thank my family: my parents, Krystyna Markiewicz, Jack Wieckowski, and my grandma, Maria Ryczaj, for all their love, understanding, support, and all that they have done to help me succeed in life. To my boyfriend Walter Dorfner, we entered this journey together, experienced the ups and downs of graduate school together, and now can celebrate our achievements together. I thank him for the unwavering love and support.

ABSTRACT

USING MOLECULAR CONTSTRAINTS AND UNNATURAL AMINO ACIDS TO MANIPULATE AND INTERROGATE PROTEIN STRUCTURE, DYNAMICS, AND SELF-ASSEMBLY

Beatrice N. Markiewicz

Feng Gai

Protein molecules can undergo a wide variety of conformational transitions occurring over a series of time and distance scales, ranging from large-scale structural reorganizations required for folding to more localized and subtle motions required for function. Furthermore, the dynamics and mechanisms of such motions and transitions delicately depend on many factors and, as a result, it is not always easy, or even possible, to use existing experimental techniques to arrive at a molecular level understanding of the conformational event of interest. Therefore, this thesis aims to develop and utilize non-natural chemical modification strategies, namely molecular cross-linkers and unnatural amino acids as site-specific spectroscopic probes, in combination with various spectroscopic methods to examine, in great detail, certain aspects of protein folding and functional dynamics, and to manipulate protein self-assemblies. Specifically, we first demonstrate how strategically placed molecular constraints can be used to manipulate features of the protein folding free energy landscape, thus, allowing direct measurement of key components via temperature-jump kinetic studies, such as folding from a transition-state structure or the effect of internal friction on the folding mechanism.

Secondly, we utilize a photolabile non-natural amino acid, Lys(nvoc), to probe the mechanism of protein misfolding in a β-hairpin model and identify an aggregation gatekeeper that tunes the aggregation propensity. We further develop a method where the induced-charge produced by photocleavage of Lys(nvoc) can be used to target and destabilize hydrophobic regions of amyloid fibril assemblies, resulting in complete disassembly, Finally, we highlight new useful properties of a site-specific spectroscopic probe, 5-cyanotryptophan (Trp_{CN}), by demonstrating (1) how the frequency and linewidth of the infrared nitrile stretching vibration is sensitive to multiple hydrogen bonding interactions and solvent polarity, (2) that the fluorescence emission, quantum yield, and lifetime is extremely sensitive to hydration, and serves as a convenient fluorescence probe of protein solvation status, and (3) that the unique characteristics of Trp_{CN} can be used to target the structure, local environment, and mechanism of the tryptophan gate in the M2 membrane proton channel of the influenza A virus.

TABLE OF CONTENTS

List of Tables

List of Figures

1 Introduction

1.1 Protein Folding

Proteins, simple linear polymers composed of unique amino acid sequences, spontaneously fold into compact three-dimensional structures that comprise the machinery necessary to execute biological functions in living organisms. A milestone discovery by Anfinsen,[1] which demonstrated for the first time that a protein can reversibly interchange between a denatured/inactive and folded/active state as a function of solution conditions, led to the onset of *in-vitro*-based protein folding studies. From this, Anfinsen proposed the physical underpinnings that have shaped the field, most notably, the "thermodynamic hypothesis", which suggests that under physiological conditions a protein adopts a native state that minimizes its Gibbs free energy. This led to the belief that the molecular driving forces (i.e., intra-chain or solvent interactions) are solely influenced by the amino acid sequence.[2] Since then, in a multidisciplinary endeavor bridging biologists, chemists, and physicists, one of the most sought after questions in structural biology focuses on unraveling the molecular mechanisms of protein folding and self-assembly.

Deciphering mechanistic details, which include a description of folding pathways, the order in which native structural elements are formed, and the kinetics of those processes, is key to developing predictive protein folding measures.[3] As Levinthal first alluded to in 1969,[4] folding must occur through a bias search to reach a native state on a biologically relevant timescale. Two folding models were later proposed to explain this. The first suggests that folding progresses through a single pathway with well-defined

intermediates.[5] The second hypothesis argues folding pathways are heterogeneous and dictated by a thermodynamic- and kinetic-controlled conformational search on a multi-dimensional funneled free energy landscape.[6,7] Regardless, both views accept the general idea that the nature of a folding reaction is governed by its potential energy surface, which in turn, is dependent on sequence. However, the vast number of degrees of freedom of a polypeptide chain makes it impossible to experimentally assess the energetics of each and every conformational state. Thus, a bottom-up approach, using small single-domain protein scaffolds, has been employed to study folding dynamics in a more simplistic manner. In such cases, the folding/unfolding kinetics can be interpreted in the context of a one-dimensional reaction coordinate containing two wells, describing the native and denatured state separated by a free energy barrier.[8] Consequently, the transition state (TS) structure associated with barrier-crossing became relevant to understanding the folding mechanism and landscape.

The Φ-value analysis method, pioneered by Alan Fersht,[9] has been influential in providing residue-level folding details, particularly by identifying approximate structures of the TS ensemble via measuring the kinetic and thermodynamic response arising from a point mutation.[10] Subsequently, one of the governing principles that arose is that proteins fold by adopting native substructures in the TS (i.e., nucleation-condensation).[11] With the development of nanosecond laser-induced optical-triggering techniques,[12] detection of these early kinetic events, such α-helix and β-sheet secondary structure formation, loop formation, and polypeptide chain collapse, was made possible.[13,14] For example, investigations of helix-coil transitions of Ala-based α-helix peptides showed that

complete folding (i.e., nucleation followed by propagation) occurs within ~1 μs.[15] This rapid timescale suggested that formation of secondary structures greatly precedes formation of tertiary contacts. On the other hand, kinetics of folding of designed β-hairpins used to study β-sheet folding, are on the order of several μs.[16] This slower rate arises because folding must overcome a larger energetic barrier due to longer-range inter-strand interactions compared to single helices. More interestingly, downhill and/or barrierless folding processes have also been observed,[17] where the timescale of folding is instead limited by diffusion on a rough free energy landscape, arising from weak kinetic traps from intra-protein or protein-solvent interactions. In this folding scenario, the absence of a significant energetic bottleneck allows all intermediates to be kinetically accessible.[18] Therefore, a recent effort has been directed toward reengineering proteins for the sake of manipulating the underlying landscape and removing high energy barriers to folding.[19] In addition, modern single-molecule fluorescence techniques have been used to identify regions of a free energy coordinate that manifest internal friction, and have quantitatively shown how it affects the folding behavior of a two-state protein.[20,21] Furthermore, the study of ultra-fast folding proteins is still highly pursued, as current computer power allows molecular dynamics (MD) simulations to access timescales that can now be directly compared with experiment.[22]

Despite advances in detecting protein folding processes, we are still limited to coarse-grained structural representations, as the spectroscopic signals frequently arise from multiple structural elements in the polypeptide chain. Appropriately, improving spatial resolution and/or confining a protein to a well-defined initial conformation are

essential to gaining structural details. In this regard, this thesis aims to enhance site-specificity in a protein (1) to investigate the local electrostatic and hydration environment and (2) to interrogate features of the folding free energy landscape, particularly TS structures and local internal friction.

1.2 Protein Misfolding and Aggregation

Under partially or fully unfolded conditions, protein monomers can encounter non-native interactions, which triggers intermolecular self-assembly, in a polymer-like fashion, to a kinetically trapped state.[23,24] This misfolding pathway leads to formation of highly-ordered fibril structures comprised of repeating β-sheets, called amyloids, abolishing the native protein function. Thus, accumulation of amyloid fibrils *in vivo* are associated with several fatal diseases, such as cystic fibrosis, Alzheimer's, Parkinson's, and Type II diabetes.[25] High resolution structural studies have shown that the fibril structure is stabilized by intermolecular hydrogen bonds of the backbone and hydrophobic interactions, and this generic structure is maintained regardless of sequence.[23] Studies have shown that self-assembly to such a fibril/aggregate structure is initiated from a nucleation event, much like in native folding. Gazit[26] hypothesized that π-π stacking of hydrophobic residues can thermodynamically drive this self-assembly and can also direct the orientation of propagation. As such, several mutational studies of amyloid-related peptide fragments lost aggregation propensity upon removal of crucial hydrophobic sidechains.[27-29] During the course of self-assembly, partially unfolded intermediates convert to oligomeric species, later forming protofibrils, which then coalesce to form a

mature fibril structure.[30] As a result, a large amount of work from the past decade has utilized fluorescence and imaging tools to investigate the kinetics of this process,[31] in an effort to identify and isolate intermediates that could be targeted for inhibition. Also, several triggering methods using switch peptides or light-activated molecules have been used to monitor amyloid self-assembly from a monomeric state.[31] This thesis aims to demonstrate a reverse approach that can be used to disassemble amyloid fibrils via site-specific light-active amino acids by targeting the π-interactions of the peptides in question.

1.3 Thesis Overview

The objective of the work presented in this dissertation is to develop new experimental approaches to investigate protein structure, folding dynamics, and self-assembly. Typical ensemble measurements used to study proteins are hampered by a lack of spatial resolution. As a result, the general strategy presented here uses site-directed chemical modifications of peptides, in several different forms, coupled with time-resolved and steady-state spectroscopies to tease out dynamic and structural details with enhanced site-specificity. Specifically, the extrinsic non-natural modifications employed to the protein structures involve one of the following: (1) molecular cross-linking, (2) a photolabile non-natural amino acid, and (3) a nitrile-derivatized non-natural amino acid spectroscopic probe. Each method offers unique characteristics that can be used to experimentally assess the properties of protein structures.

To begin, Chapter 2 provides a summary of the theoretical basis used to explain the thermodynamics and kinetics of protein folding and unfolding, as well as the theory used to assess the details of the protein folding free energy landscape. This leads into Chapter 3, which presents the details and background of the experimental methods utilized throughout this thesis, including the motivation behind the non-natural chemical modifications used in these studies.

In Chapter 4, we demonstrate how a site-specific sidechain cross-linker could be utilized to constrain and populate a thermodynamically accessible state that mimics the TS structure, and therefore allows for direct measurement of folding from a TS-like configuration. We test the feasibility of this method on a well-studied β-hairpin and use a disulfide sidechain cross-linker to restrain the first backbone-backbone hydrogen bond, in conjunction with laser-induced temperature-jump IR spectroscopy. The measured conformational relaxation kinetics of this variant shows a fast phase relaxing an order of magnitude faster (~500 ns) than the wild-type, indicating that the barrier to folding is abolished.

In Chapter 5, we use an *m*-xylene helix cross-linker to site-specifically assess the effect of internal friction in the most crowded region of a model protein. This leads to a significant decrease in the folding and unfolding rate while demonstrating no effect on protein stability. These results indicate that the *m*-xylene linker acts as a local crowder, thus inducing a frictional drag on the free energy landscape. We estimate that interactions arising from the linker increase the roughness of the landscape by 0.4–1.0 $k_B T$.

In Chapter 6, we investigate the aggregation behaviors of two β-hairpins, Trpzip1 and Trpzip2. Here we utilize a photolabile lysine analogue, Lys(nvoc), and a change in solution conditions, to show that the aggregation mechanism proceeds through native edge-to-edge interactions, and that the Lys8 position serves as an aggregation gatekeeper. Furthermore, we also demonstrate that Lys(nvoc) could be used to site-specifically tune the rate of peptide aggregation.

In Chapter 7, we employ a charge triggering process, where generation of one or several charged sidechains, using a photolabile non-natural amino acid and light, in an otherwise hydrophobic environment is sufficient to cause protein/peptide aggregates, amyloid fibrils, or other types of matrices, like peptide hydrogels, to disassemble. To test this method, we utilize a photolabile lysine analog, Lys(nvoc), which is hydrophobic but yields a charged lysine upon photocleavage of the nvoc moiety, to photo-trigger the dissociation of model amyloid peptides ($hIAPP_{22-27}$ and $A\beta_{16-22}$). Furthermore, we show that it is possible to spatiotemporally control the morphology of a self-assembling peptide hydrogel (FKFEFKE) by substitution of a nonpolar residue (phenylalanine) with Lys(nvoc), and that the placement of Lys(nvoc) can affect the morphology and viscoelastic properties of the gel.

In Chapter 8, we expand the utility of 5-cyanotryptophan (Trp_{CN}) as a site-specific IR probe, and assess the factors that contribute to the IR stretching frequency of the nitrile (C≡N). The frequency and lineshape are highly sensitive to local environment, especially hydration status. By correlating the C≡N stretching frequency to well-known Kamlet-Taft parameters of different solvents (i.e., polarizability (π^*), hydrogen accepting

ability (β), hydrogen bond donating ability (α)), we find that there is a linear correlation when plotted against σ ($\pi^* + \beta - \alpha$). This result indicates that direct hydrogen bonding interactions with C≡N and N-H, as well as the non-specific interactions arising from the polarity of the solvent; all contribute to the final C≡N frequency.

In Chapter 9, we expand the utility of Trp_{CN} as a site-specific fluorescence probe and characterize the photophysical properties. We find that Trp_{CN} has an enhanced sensitivity to hydration, which is reflected in its quantum yield, emission maximum, and fluorescence lifetime measurements. Specifically, when Trp_{CN} transitions from a hydrated to hydrophobic environment, its quantum yield and fluorescence lifetime increases by an order of magnitude. We show that this stark contrast in Trp_{CN} signal can be exceptionally useful as a probe of hydration, and demonstrate its applicability to study binary solvent mixtures, membrane-peptide binding, and hydration sites of a folded protein.

In Chapter 10, we combine the principals that were learned from Chapter 8 and 9, and use Trp_{CN}, combined with steady-state and time-resolved IR and fluorescence measurements, to site-specifically probe the structure and hydration dynamics of the tryptophan (Trp) gate in the transmembrane domain of the M2 proton channel. Our findings suggest that the water density around the Trp gate does show appreciable changes between the open and closed forms. We speculate that the lack of sufficient water molecules near Trp needed to establish a continuous water wire poses an additional energetic bottleneck for proton conduction.

Finally, Chapter 11 will offer a summary of the present findings as well as future directions.

2 Theory

2.1 Thermodynamics of Protein Folding

As Christian Afinsen proposed almost 50 years ago, the "thermodynamic hypothesis" of protein folding generally asserts that the native conformation of a protein is the one with the lowest Gibbs free energy.[1] To further explain how a protein converges to this state efficiently, Wolynes and coworkers[6,7] proposed that the protein folding free energy landscape is funnel-like, with a bias toward a single basin and/or conformation.[5] Since then, this notion, representing a protein's conformational space versus free energy, has served as a quantitative model to describe, via statistical thermodynamics, the microscopic degrees of freedom, or in other words, the density of states (i.e., conformational entropy), at each free energy level.[2] These theories have been elaborated upon in several reviews.[32,33] However, to keep with the relevance of the work discussed in this thesis, obtaining quantitative thermodynamic parameters of a two-state folding protein will be discussed in the context of a thermal unfolding reaction.[34,35] The proteins studied herein, follow a two-state model, where the protein assumes one of two ensembles, folded (F) and unfolded (U),

$$F \underset{k_\mathrm{f}}{\overset{k_\mathrm{u}}{\rightleftharpoons}} U \tag{2.1.1}$$

where k_u represents the unfolding reaction rate constant and k_f is the folding reaction rate constant. Specifically, the free energy of the protein folding reaction ($\Delta G_r°(T)$) at temperature T is described by

$$\Delta G_r°(T) = \Delta H_r°(T) - T\Delta S_r°(T)\,. \tag{2.1.2}$$

Here, $\Delta H_r°(T)$ and $\Delta S_r°(T)$ represent the change in enthalpy and entropy of the reaction, respectively. More importantly, $\Delta G_r°(T)$ is the difference of the free energy between the denatured state ($G_U°$) and the native state ($G_F°$) as follows,

$$\Delta G_r°(T) = G_U° - G_F°. \tag{2.1.3}$$

The temperature dependence of $\Delta H_r°(T)$ and $\Delta S_r°(T)$, can be described by the following two equations,

$$\Delta H_r°(T) = \Delta H_r°(T_m) + \Delta C_p(T - T_m) \tag{2.1.4}$$

$$\Delta S_r°(T) = \Delta S_r°(T_m) + \Delta C_p \ln(\frac{T}{T_m}). \tag{2.1.5}$$

where T_m is a reference temperature, usually the melting temperature (i.e., the point where system is 50% folded and 50% unfolded). The change in heat capacity (ΔC_p) is assumed to stay relatively constant over the limited temperature range.[35] Taking Eq. 2.1.4 and Eq. 2.1.5 and substituting into Eq. 2.1.2, the resulting expression for $\Delta G_r°(T)$ can be described by the following:

$$\Delta G_r°(T) = \Delta H_r° + \Delta C_p(T - T_m) - T[\Delta S_r° + \Delta C_p \ln(\frac{T}{T_m})]. \tag{2.1.6}$$

Furthermore, the transition temperature is defined as $T_m = \Delta H_r°(T_m)/ \Delta S_r°(T_m)$. Commonly, T_m is used to assess the overall thermodynamic stability of the folded structure relative to its thermally denatured state. As a result, a mutational analysis of a protein sequence combined with thermal denaturation experiments can be used to dissect how a particular sidechain interaction contributes to stabilizing the final folded structure.[34]

2.2 Kinetics of Protein Folding

Continuing in the context of a simple two state folding model, as shown in Eq. 2.1.1, the classical kinetics methodology of a unimolecular reaction fluctuating between two states, *F* and *U,* can be used to summarize the kinetic behaviors of the protein models studied herein.[36] The differential rate expressions for this reaction describing the change in concentration of folded [*F*] and unfolded [*U*] protein molecules are,[37,38]

$$\frac{d[F](t)}{dt} = -\frac{d[U](t)}{dt} = k_f[U](t) - k_u[F](t)\,. \tag{2.2.1}$$

When this system is at equilibrium conditions, $\frac{d[F](t)}{dt} = -\frac{d[U](t)}{dt} = 0$, as the concentrations of [*F*] and [*U*] do not change. As a result, we find that [*U*]=[*U*]$_{eq}$ and [*F*]=[*F*]$_{eq}$, where the subscript *eq* represents equilibrium conditions. Also, since k$_f$[*U*]=k_u[*F*], the equilibrium constant, *K*, can be described by[37]

$$K = \frac{[U]}{[F]} = \frac{k_u}{k_f}\,. \tag{2.2.2}$$

Nonetheless, when a protein ensemble is perturbed to a non-equilibrium condition at a time *t*, via relaxation methods such as a temperature perturbation, [*F*] and [*U*] can be re-written to reflect an evolution to a new equilibrium state (denoted by the subscript 2,*eq*)[39]

$$[F](t) = [F]_{2,eq} + \Delta[F](t) \tag{2.2.3}$$

$$[U](t) = [U]_{2,eq} + \Delta[U](t)\,.$$

By substituting Eq. 2.2.3 into Eq. 2.2.1, the initial rate expression describing the time-dependent change in [*F*] becomes,[39]

$$\frac{d\Delta[F](t)}{dt} = k_f([U]_{2,eq} + \Delta[U](t)) - k_u([F]_{2,eq} + \Delta[F](t)). \quad (2.2.4)$$

Keeping in mind that $-\Delta[U]=\Delta[F]$ and $k_f[U]_{2,eq}=k_u[F]_{2,eq}$, Eq. 2.2.4 can be simplified to[39]

$$\frac{d\Delta[F](t)}{dt} = -(k_f + k_u)\Delta[F](t). \quad (2.2.5)$$

Integration of Eq. 2.2.5 yields,

$$\Delta[F](t) = \Delta[F]_0 e^{-(k_f + k_u)t} \quad (2.2.6)$$

where we assume that $\Delta[F]$ at $t = 0$ is equal to $\Delta[F]_0 = [F]_{1,eq} - [F]_{2,eq}$ (i.e., 1,*eq* represents the initial equilibrium state).[39] Here, the observed experimental relaxation rate, k_r, is the sum of the forward and reverse rate constants, and is also equal to the inverse of the relaxation time, τ,

$$k_r = k_f + k_u = \frac{1}{\tau}. \quad (2.2.7)$$

Thus, these relationships demonstrate that a perturbation from equilibrium results in an exponential relaxation rate equal to the folding and unfolding reaction rates. When we quantify the equilibrium constant K, of the protein folding reaction, we can use the observed k_r and Eq. 2.2.7 and Eq. 2.2.2 to specifically determine k_f and k_u.

2.3 Details of the Folding Free Energy Landscape

The mechanism by which a protein folds is embedded in the features of its multi-dimensional free energy landscape.[33] However, a protein chain has innumerable degrees of freedom, making it extremely difficult to experimentally resolve the details of its complex energy surface. For a simple two-state folder, the problem can be reduced to

one-dimension (i.e., unfolded and folded potential wells separated by a barrier). The protein folding reaction dynamics can be correlated with details of the one-dimensional reaction coordinate using Kramer's theory, which establishes a relationship between the folding rate constant k and the folding free energy barrier height ($\Delta G^{\ddagger}$),[19,40]

$$k = \frac{\omega_R \omega_B}{2\pi\gamma} e^{(-\Delta G^{\mp}/RT)} \tag{2.3.1}$$

where the pre-exponential factor terms define the shape and roughness of the underlying free energy landscape. Here, ω_R and ω_B represent frequencies describing the curvatures of the reactant and transition-state potential wells, γ is a frictional coefficient, R is the gas constant, and T is temperature. Specifically, the γ parameter, arises from "friction" on the free energy surface that originates from the dissipation of energy into internal degrees of freedom of the protein (i.e., random sidechain collisions, steric hindrances) or coupling to the solvent bath. To assess these effects experimentally, one approach assumes that γ is directly related with solvent viscosity η_s and protein internal friction α, resulting in a modified Kramer's equation, $k \propto 1/(\eta_s + \alpha)$.[41] Numerous examples have shown that if the folding rate does not follow a linear correlation with solvent viscosity, as is predicted for a barrier-crossing process, then the observed divergence is attributed to friction arising from α alone.[41-44]

In a second approach based on theory developed by Sagnella *et al.*,[45] one can also measure roughness $U^{\ddagger}$ on a two-state free energy landscape in response to a friction-inducing perturbation. For example, they showed previously that $U^{\ddagger}$ can be defined by[45]

$$U^{\mp} = kT \ln \frac{\gamma}{\gamma_{\text{int}}} \tag{2.3.2}$$

where γ_{int} is the intrinsic friction from the solvent and γ is the internal friction within the protein. If we define k_{WT} as the rate constant of a native wild-type protein, and k_P as the rate constant of the same protein subjected to a friction-inducing perturbation (i.e., cross-linker, site-directed mutation, crowding agent), we can express the relationship between the rate constants and internal friction using Kramer's equation,[46]

$$k_{WT} = \frac{A}{\gamma_{int}} e^{(-\Delta G^{\mp}/RT)} \text{ and } k_P = \frac{A}{\gamma_{int} + \gamma} e^{(-\Delta G^{\mp}/RT)} \tag{2.3.3}$$

where $A = \frac{\omega_R \omega_B}{2\pi}$. Assuming that the deviations of its folding/unfolding rate are due entirely to increased internal friction and there is no change in barrier height or A with respect to the wild-type, Eq. 2.3.3 further simplifies to

$$\frac{k_{WT}}{k_P} - 1 = \frac{\gamma}{\gamma_{int}} , \tag{2.3.4}$$

which then finally allows calculation of $U^{\ddagger}$.

3 Methods

3.1 Circular Dichroism Spectroscopy

Far-UV circular dichroism (CD) is not only a well-established technique used to qualitatively assess secondary structures of proteins, but also is a vital tool used to assess the thermodynamics of protein folding.[47,48] Briefly, CD is the differential absorption of left-handed and right-handed circularly polarized light of an optically active medium. This consequential difference in absorption coefficients causes the electric-field vector to adopt an elliptical polarization. As a result, the detected signal of a CD instrument is reported in degrees of ellipticity, which is the angle equal to the tangent of the minor and major axis of the ellipse.[49]

The amide functional group of protein backbone structures serves as the dominant chromophore, and undergoes three electronic transitions, (1) a $n \rightarrow \pi^*$ transition at 220 nm, (2) a $\pi \rightarrow \pi^*$ at 190 nm, and (3) $\pi \rightarrow \pi^*$ at 170 nm. The unique three-dimensional configurations adopted by protein secondary structures (i.e., β-sheet, α-helices, and disordered chains) yield distinguishing CD spectral signatures, which are influenced by factors such as hydrogen bonding and exciton coupling between degenerate states.[49] For example, the common CD signature of α-helices contains a negative band at 222 nm arising from the $n \rightarrow \pi^*$ transition. A second negative band at 208 nm and positive band at 192 nm originates from exciton coupling of the $\pi \rightarrow \pi^*$ transition. Furthermore, β-sheets generate a negative band at 216 nm ($n \rightarrow \pi^*$ transition) and positive band at 195 nm ($\pi \rightarrow \pi^*$ transition). Finally, disordered structures are distinctive, as their signature only shows a single negative feature at 200 nm.

Also important, aromatic sidechains of proteins have sensitive and detectable electronic transitions, such as Trp and Tyr. In the far-UV range (below 260 nm), these transitions usually manifest as exciton couplets, and can serve as indications for tertiary interactions between neighboring strands.[50] In the case of the Trpzip β-hairpin peptides,[51] which are stabilized by two pairs of edge-to-face Trps, a strong couplet is observed at 227 nm and 215 nm when the peptide is folded.

Moreover, as mentioned above, CD spectroscopy is a convenient method to measure the thermodynamics of protein folding via thermal denaturation. This is accomplished by measuring the ellipticity at a single wavelength as a function of temperature to monitor unfolding. To quantitatively assess the thermal stabilities of the resulting curve and obtain thermodynamic parameters, we perform a fit to the following two-state model:[46,47]

$$\theta(T) = \frac{\theta_{\mathrm{U}}(T) + K_{\mathrm{eq}}(T)\theta_{\mathrm{F}}(T)}{1 + K_{\mathrm{eq}}(T)} \tag{3.1.1}$$

$$K_{\mathrm{eq}}(T) = \exp(-\Delta G(T) / \mathrm{R}T) \tag{3.1.2}$$

$$\Delta G(T) = \Delta H_{\mathrm{m}} + \Delta C_{\mathrm{p}}(T - T_{\mathrm{m}}) - T[\Delta S_{\mathrm{m}} + \Delta C_{\mathrm{p}} \ln(\frac{T}{T_{\mathrm{m}}})] \tag{3.1.3}$$

where $\theta(T)$ is the temperature-dependent mean residue ellipticity, $K_{\mathrm{eq}}(T)$ is the equilibrium constant for unfolding, $T_{\mathrm{m}} = \Delta H_{\mathrm{m}}/\Delta S_{\mathrm{m}}$ is the thermal melting temperature, ΔH_{m} and ΔS_{m} are the enthalpy and entropy changes at T_{m}, and ΔC_{p} is the heat capacity change. In addition, the folded and unfolded CD baselines, $\theta_{\mathrm{F}}(T)$ and $\theta_{\mathrm{U}}(T)$, respectively, are assumed to be linear functions of temperature, as follows:

$$\theta_{\mathrm{U}}(T) = c + dT \tag{3.1.4}$$

$$\theta_F(T) = a + bT \tag{3.1.5}$$

where *a*, *b*, *c*, and *d* are treated as constants. For cases where one needs to determine the thermal properties of multiple mutants of single sequence, the subsequent thermal melts are globally fit to the two-state model above, while keeping the folded/unfolded baselines constant.[52]

3.2 Fluorescence Spectroscopy

Fluorescence spectroscopy, measurement of photon emission from an excited singlet state, is one of the most applied and versatile techniques used to investigate protein structure, folding, and dynamics because polypeptide chains contain intrinsic fluorophores, namely the naturally occurring aromatic amino acids (Trp, Tyr, and Phe).[53,54] Of these, Trp is regarded as the most useful for numerous reasons: (1) the molar absorptivity of the ^1L_a and ^1L_b electronic transitions (~280 nm) is large; (2) it has a high quantum yield (QY) in aqueous solution (~0.14); (3) it has a large excited-state dipole moment (~6 D) making its wavelength of emission and fluorescence lifetime exceptionally sensitive to solvent; and (4) Trp is a rare amino acid that is often located in biologically relevant sites that are vital for self-assembly and function. Additionally, the sensitivity of Trp to collisional quenching, from either H_2O or ions,[53] can be used to evaluate solvent accessibility of a particular Trp site. However, larger proteins frequently contain more than one intrinsic fluorophore making interpretation of the spectra quite complicated. Therefore, Trp- and Tyr-based analogues that contain distinctive

photophysical characteristics, such as QY, emission maximum, and fluorescence lifetimes, have been implemented as site-specific fluorescence probes.

Specifically, one popular approach to change the photophysical behavior of an amino acid analogue involves the addition of an atom or functional group to an aromatic ring. Examples of Trp-derived fluorescent probes include 7-azatryptophan,[55] 5-hydroxytrypotphan, 4- and 5-fluorotryptophan,[53] and 5- and 6-cyanotryprotphan.[56,57] Similarly, a widely used Tyr-based non-natural amino acid probe is *p*-cyanophenylalanine, which demonstrates QY dependence on hydrogen bonding strength to the cyano group.[58] Other more structurally diverse small-molecule chromophores have been designed with aromatic moieties such as dansyl amides and coumarins.[53] Particularly, this thesis will emphasize the applicability of 5-cyanotryptophan as a fluorescence probe, as its fluorescence lifetime and quantum yield show a stark contrast in signal upon hydration.

One common method employed to measure the fluorescence lifetime of molecules is Time-Correlated Single Photon Counting (TCSPC). In simple terms, it measures the time difference between initial excitation with a picosecond laser pulse and detection of a single emitted photon, over a series of cycles.[53,59] Specifically, the photon counting rate is limited by the acquisition time of the detection electronics, and as a result, capturing a fluorescence decay process within the typical timescale (picoseconds to nanoseconds), is difficult via multiple photon detection after a single excitation pulse.[53] Therefore, the detection rate is fine-tuned so that only one photon is counted for every ~20-100 excitation pulses, and the accumulated data over numerous cycles are plotted as a

histogram illustrating the photon count versus the time difference (Δt) between excitation and photon detection. The subsequent distribution of Δt reproduces the exponential decay of the fluorescence lifetime. Finally, the fluorescence lifetime parameters are calculated by deconvolution with the instrument response function and fit to a single- or multi-exponential function.[59]

Other advanced fluorescence spectroscopic techniques, which are not applicable to this thesis, but are reviewed in detail elsewhere, involve single-molecule FRET detection to study protein folding and the details of its free energy landscape,[60,61] rapid temperature-jump fluorescence detection for detection of ultrafast protein folding events,[62] and fluorescence correlation spectroscopy for translational diffusion of a protein within a known confocal volume.[54]

3.3 Infrared Spectroscopy

Over the last 60 years, infrared (IR) spectroscopy, particularly in the mid-IR spectral range (i.e., 4000 – 200 cm^{-1}), has evolved into a time-efficient and powerful analytical tool to decipher and quantify global protein structures.[63] One of the distinct advantages of this method is that measurements can be made in biologically relevant media, such as in the presence of membrane mimetics, which is otherwise difficult with other structural techniques. However, due to the size and chemical complexity of proteins, there are an extremely large number of oscillators within a given structure, and the overlapping transitions combined with vibrational couplings between inherent modes, limits us to a crude secondary structure representation. Regardless, the abundance of information that

can be gained from this is invaluable, particularly in cases where IR difference spectroscopy is applied to monitor a protein reaction.[64]

Specifically, the molecular vibrations that contain the most structurally sensitive information is embedded within the amide I absorption band of the protein backbone.[64] This IR transition, which absorbs at a frequency between 1600-1700 cm^{-1}, originates primarily from the C=O stretching vibration and is also weakly coupled to the out-of-phase CN stretching, CCN deformation, and in-phase NH bending.[65] More importantly, the frequency and lineshape of the amide I mode is sensitively dependent on protein secondary structure, giving rise to the following absorption bands; (1) α-helices – narrow band at ~1650-1640 cm^{-1}, (2) parallel β-sheets – narrow band 1620 cm^{-1}, (3) anti-parallel β-sheets – split amide I mode with a high intensity band at 1630 cm^{-1} and a weaker band at 1685 cm^{-1}, and (4) disordered structures – broad, featureless band centered at 1650 cm^{-1}.[66] In particular, transition dipole coupling (TDC),[67] hydrogen bonding,[68] and inhomogeneous band broadening,[66] all contribute to the sensitivity of the amide I absorption. First, TDC occurs when the vibrational modes of neighboring oscillators interact, and the nature of the coupling is determined by distance and orientation of the dipoles relative to each other.[66] This mechanism manifests in two ways, exciton transfer and exciton splitting. The former involves an excitation energy migration to adjacent oscillators, resulting in a delocalization of the excited state over a distance.[69] The latter represents the prominent mechanism of anti-parallel β-sheets, where coupling of in-phase (inter-strand) and out-of-phase (intra-strand) modes produces exciton splitting of the

amide I band,[70] and can be as substantial as ~70 cm^{-1} in well-ordered assemblies. This characteristic band shape can be used a structural marker for protein aggregation.[71]

Secondly, when an oscillator, such as C=O, is engaged in a hydrogen bond, either with a neighboring N-H of the protein backbone or solvent molecule, the redistribution of electron density weakens the frequency of the stretching vibration resulting in a red-shift on the order of ~20 cm^{-1}.[68] As a consequence, the hydrogen bonding pattern within a secondary structure (i.e., $i \rightarrow i+4$ hydrogen bonds of an α-helix or intermolecular hydrogen bonds of β-sheets) also contributes to the overall amide I frequency and lineshape.[64] In addition, solvent-mediated hydrogen bonding plays a critical role in determining the amide I frequency, as the absorption of solvated α-helices can be red-shifted by ~20 cm^{-1}.[72] Band broadening due to conformational heterogeneity also occurs in parallel,[73] thus alluding to the third defining parameter of the amide I mode.

Specifically, inhomogeneous broadening in IR spectroscopy describes a process where absorption of a single vibrational mode yields a large frequency distribution due to a heterogeneous microenvironment.[74] This is most useful in determining the degree of order within the protein sample. For example, a flexible and disordered polypeptide chain will produce a broad amide I band due to a non-uniform interaction with the solvent. On the other hand, highly-ordered β-sheets result in a narrow and sharp amide I lineshape resulting from a uniform structure and environment.[66] The same principals can be applied more rigorously to monitor site-specific vibrational transitions of proteins arising from either particular sidechains, such as the C=O stretching of Glu, Asp, Asp, and Gln,[75] or from extrinsically modified non-natural amino acid probes, which will be discussed in

detail in a later section. In this context, changes in lineshape due to inhomogeneous broadening can distinguish whether a particular site (1) is sequestered in a hydrophobic versus a solvent-exposed environment, (2) contains a bound ligand, and/or (3) is engaged in a protein-protein interactions.[66]

Finally, as mentioned earlier, generating an IR difference spectrum as a function of some perturbation is a convenient means to observe conformational changes in proteins. Most relevant to protein folding studies and the topics of this thesis, measuring the temperature dependence of the amide I transition provides a method to identify the frequencies regimes that reveal the largest changes in signal due to unfolding, which is relevant in determining probing frequencies for temperature dependent time-resolved IR studies of the amide I mode.

3.4 Attenuated Total Reflectance Infrared Spectroscopy

The investigation of protein conformation and orientation in a lipid environment has posed difficulties for structural biologists. A facile technique to study the structure of membrane bound and/or inserted peptides is polarized Attenuated Total Reflectance Fourier transform infrared (ATR-FTIR) spectroscopy. A schematic representation as well as a complete derivation of the theory has been outlined in extensive detail in several reviews.[76-78] Briefly, an incident IR beam is cast through an internal reflection element (IRE) of high refractive index where reflections with the surface generate an evanescent wave at the sample-IRE interface. As the intensity of the evanescent wave drops off exponentially with distance, the light-sample interaction reaches a finite penetration

depth on the order of several hundred nanometers, allowing for surface sensitive detection.[76] Deposition of a protein-membrane bilayer on an IRE provides a uniformly oriented sample with respect to the normal of the surface. Coupled with polarized IR light, the changes in absorption due to electric field orientation allow one to calculate the direction of a specific transition dipole moment.[77] For example, a dipole moment oriented parallel to the incident polarized light will yield a maximum IR response. As a result, the orientational information is embedded within the dichroic ratio (R), which is the ratio between the integrated area of parallel ($A_{\parallel}$) versus perpendicular ($A_{\perp}$) absorption bands:

$$R = \frac{A_{\parallel}}{A_{\perp}} \tag{3.4.1}.$$

This can be used to calculate an experimental order parameter (S_{Exp}), where ε_x, ε_y, and ε_z are the electric field amplitudes of the evanescent wave, and can vary depending on whether a thick film or thin film approximation is employed.[77]

$$S_{\mathrm{Exp}} = \frac{\varepsilon_x^2 - R\varepsilon_y^2 + \varepsilon_z^2}{\varepsilon_x^2 - R\varepsilon_y^2 - 2\varepsilon_z^2} \tag{3.4.2}.$$

More importantly, a well-developed formalism, which in actuality can be applied to any system that is axially symmetric, is utilized to calculate the orientation of a transmembrane α-helix ($S_{\mathrm{Helix\ Angle}}$) with respect to the membrane normal, using the dichroic ratio determined from the amide I absorption,[79]

$$S_{\mathrm{Helix\ Angle}} = \frac{S_{\mathrm{Exp}}}{S_{\mathrm{Dipole}} S_{\mathrm{Membrane}}} \tag{3.4.3}.$$

Within this calculation, it is common for the order parameter of the lipid molecular axis (S_{Membrane}) to be equal to 1.[79] Furthermore, in the case of an α-helix in particular, the

orientation of a transition dipole moment (S_{Dipole}), namely the amide I mode is described by,

$$S_{\text{Dipole}} = \frac{3\cos^2\theta - 1}{2} \tag{3.4.4}$$

where θ is the angle between the dipole and the molecular axis of the helix. Without site-specific isotopic labeling of a backbone unit,[80] or other supporting structural data, the angle of an amide I vibration is impossible to obtain. Typically, a commonly accepted value for the amide I mode is θ = 39-40°.[81] Note, Eq. 3.4.4 can also be implemented to calculate the angle of a transition dipole moment arising from an IR vibration of a site-specific probe, such as a nitrile-derivatized non-natural amino acid.[82] Finally, the helix tilt angle (β) with respect to the membrane axis, can be determined by the following equation:

$$S_{\text{Helix Angle}} = \frac{3\cos^2\beta - 2}{2} \tag{3.4.5}.$$

Taken together, ATR-FTIR allows the study of protein structures that require a membrane environment to fold, and provides an avenue to create a biological relevant, ordered lipid bilayer and quantitatively decipher orientational parameters. .

3.5 Laser-Induced Temperature-Jump Infrared Spectroscopy

The fundamental criterion for monitoring a chemical reaction via laser-induced temperature-jump (*T*-Jump) spectroscopy is that a temperature dependence must be detectable within the initial equilibrium population.[12] Luckily, proteins certainly meet this requirement, as their free energy changes with temperature. Thus, a rapid temperature

perturbation can function as a convenient trigger to shift the equilibrium of the protein ensemble. The resulting relaxation kinetics to a new equilibrium position are governed by both folding and unfolding channels, and in particular, is true for a two-state folding scenario.[13]

The *T*-jump IR spectroscopy technique implemented in the studies of this thesis, contains a pump-probe design,[62,83,84] where first, a rapid *T*-jump of the solution, taking place within picoseconds, is generated by a high-power laser pulse (10 mJ) that excites the OD stretch overtone (absorbing at 1.9 μm) of the D_2O solvent. To achieve the desired excitation wavelength, a gas pressurized Raman cell is used to produce a Stokes-shift from an Nd:YAG fundamental pulse (i.e., 1064 nm → 1900 nm). To probe the temperature-induced transient conformational changes of the proteins, we use a continuous-wave mid-IR laser tuned to the amide I region (1600-1700 cm^{-1}). Here, the magnitude of the *T*-jump is dictated by the following factors: the absorption coefficient, solution density, heat capacity, and the density of energy generated by the 1.9 μm pulse along the propagation axis. Consequently, the weak extinction coefficient of the OD stretch overtone and a narrow pathlength (50 μm) of the sample cell allows us to maintain constant and even heating. Furthermore, the beam size used for the *T*-jump pulse is ~1 mm, and in order to achieve uniform detection of the protein ensemble at a given temperature, the probe pulse diameter is limited to a significantly smaller diameter.[12] Also, the rate of thermal diffusion of the heated volume, ranging from 0.2 – 100 ms, determines the detection time window;[12] as a result, the observable kinetics should occur faster than this limit. Furthermore, D_2O solvent alone has minor temperature dependence

in the amide I region. To evaluate the final temperature of the excited area, the optical density detected from the D_2O buffer and known initial temperature are referenced against a temperature-dependent calibration curve of D_2O measured at the same frequency and pathlength. Typically, the *T*-jumps are on the order of 5-10 ºC. Subsequently, the buffer signal is subtracted from the sample signal to obtain the final relaxation trace at a particular temperature. Overall, one of the main advantages of *T*-jump IR spectroscopy is that it easily yields protein folding relaxation kinetics using a reversible trigger and intrinsic vibrational chromophores.[13] More recently however, monitoring the change in the kinetic response after invoking destabilizing point mutations,[85] or other conformational modifications like cross-linking,[86] within the studied protein, has allowed scientists to elucidate fundamental mechanistic events in a more site-specific manner.

3.6 Molecular Constraints to Direct Folding

The majority of naturally occurring proteins lack a structural element or motif that can be used as a molecular constraint to direct a self-assembly process of interest towards a specific direction, making it difficult to precisely control the outcome. Thus, a great deal of effort has been made in recent years to develop new strategies to extrinsically modify protein structures, aiming to manipulate and/or control protein/peptide self-assemblies. In particular, the strategy of using a chemical cross-linker to confine the molecule of interest to a specific conformation has gained popularity. This method is not only useful in controlling the structural framework of proteins, but also in helping understand the molecular mechanism of the self-assembly process.

Various sidechain-to-sidechain cross-linking strategies have been put forth to reduce the conformational entropy of the unfolded state, thus forcing short peptides to fold into α-helices or β-sheets. Such examples include disulfide bond formation,[87,88] ring-closing metathesis,[89,90] lactam bridge formation,[91,92] hydrocarbon bridges,[93,94] and hydrazone[95] and oxime linkages.[96] In particular, due to the ease of incorporation and natural abundance of cysteines in biological systems, cysteine alkylation[97] has become a popular method for incorporating cross-linkers that stabilize α-helical conformations.[98,99] Besides these sidechain-to-sidechain covalent constraints, attempts have also been made to cross-link two backbone atoms together.[100,101]

The strategy of chemical cross-linking is not only useful for increasing the stability of folded conformations, but also is an effective tool to study the mechanisms of protein folding and, in particular, the structure of the folding transition state ensemble.[102-106] For example, Sosnick and coworkers have exploited the metal binding property of histidine and have used divalent metal ions (e.g., Zn^{2+} and Co^{2+}) to create an unconventional linker in protein systems of interest via engineered bihistidine sites.[107] They showed that it is possible to use this cross-linking approach, in combination with thermodynamic and kinetic measurements, to determine whether a specific native structural element is formed in the transition state ensemble. Similarly, covalent cross-linkers, such as dichloroacetone structures and disulfide bridges, have also been utilized to elucidate the nature of the folding transition state ensemble.[102,103,105,106]

Original Publication

This section has been adapted from Science China Chemistry, Beatrice N. Markiewicz, Robert M. Culik, and Feng Gai, (2014) 57, 1615-1624. DOI: 10.1007/s11426-014-5225-5, with permission from Springer.

3.7 Phototriggers to Initiate Protein Conformational Changes

Due to the ease and convenience of manipulating light, phototriggering is becoming a widely used method to initiate protein conformational changes. In comparison to other initiation methods, phototriggering offers several advantages, since light-triggered reactions, such as isomerization and cleavage, (1) can take place on ultrafast timescales, (2) can lead to significant changes in backbone geometry, and (3) can provide precise conformational control between equilibrium and non-equilibrium states. However, an effective trigger must meet several criteria: (1) it should be selectively initiated by light that has a wavelength higher than 310 nm, as lower-wavelength light could excite protein backbone and sidechain electronic transitions; (2) it should occur on a timescale that is faster than that of the kinetic event of interest; (3) it should have a sufficiently high quantum yield; (4) it should produce relatively inert byproducts; and (5) it should be easily incorporated into the protein system of interest.

Examples of irreversible phototriggers include tetrazines,[108] hydrazines, and dimethoxybenzoins linkers.[109] Additionally, photolabile molecules (i.e., photocages) constitute another type of irreversible phototrigger,[110-112] and have also been used to trigger protein conformational events.[113] The main idea is to use light to remove a moiety that disrupts native structure formation from the protein of interest, thus initiating folding. The advantage of using photocages, in comparison to photoresponsive cross-linkers, is

that they can be localized to a single amino acid sidechain, thus offering greater site-specificity. Examples of photocages that have been used to study protein folding include 4,5-dimethoxy-2-nitrobenzene[114] and 4-(bromomethyl)-6,7-dimethoxycoumarin.[115]

Reversible phototriggers, such as azobenzene and its derivatives[116-121] have become the most popular choice. Similar to stilbene,[122-124] azobenzene undergoes ultrafast isomerization at the excited electronic state,[125,126] which can effectively change the distance between the two ends of the azobenzene linker and modulate protein structure. In protein conformational studies, di-iodoacetamide azobenzene, which can be inserted into proteins via cysteine alkylation,[127,128] is a widely used reversible phototrigger.

Because of their apparent importance, a wide variety of studies have been conducted to understand the mechanisms of amyloid and fibril formation and a recently pursued direction is to use light and phototriggers to manipulate and control the self-assembly process. For example, azobenzene has been introduced into various aggregation-prone systems to serve as a photoswitch between different aggregation or molecular states. One example demonstrated reversible switching between aggregated and non-aggregated states of azobenzene-linked amyloid peptides.[129] In another application, an azobenzene-containing Aβ(1-42) peptide was used to investigate whether turn nucleation is the rate-limiting step in fibril self-assembly.[130] Similar to applications in protein folding studies, light-induced decaging also finds novel use in studying the assembly of amyloid structures. For instance, a coumarin-derived photocage was applied

to trigger an intramolecular acyl migration, thus initiating aggregation of Aβ(1-42) upon illumination.[131]

In addition, others have applied such photoswitching strategies to control the integrity of peptide hydrogels since their molecular packing and architectures are similar to those of amyloid fibrils.[132] Since light is an external stimulus that offers precise spatial and temporal control, biologically compatible light-responsive hydrogels are expected to find great use in applications where such controls are required. In one application, a photoswitch, either an azobenzene[133-135] or a spiropyran,[136] which links hydrogel-forming dipeptides together, was used to control π-π stacking between sidechains and thus the morphology of the hydrogel with light. In another application, photodegradable hydrogels were designed by linking a biaryl-substituted tetrazole to a small peptide, which undergoes a rapid intramolecular photoclick reaction that leads to disruption of the hydrogel matrix. Moreover, Schneider and coworkers[137] have shown that the light-activated release of a nitrobenzyl-cage from a cysteine residue can trigger the self-assembly of the MAX1 amphiphilic β-hairpin peptide, which spontaneously self-assembles into hydrogels in its folded conformation.

Original Publication

This section has been adapted from Science China Chemistry, Beatrice N. Markiewicz, Robert M. Culik, and Feng Gai, (2014) 57, 1615-1624. DOI: 10.1007/s11426-014-5225-5, with permission from Springer.

3.8 Site-Specific Infrared Reporters of Proteins

Achieving a highly dynamic and spatially resolved picture of biological processes (i.e., protein folding, ligand binding, sidechain structural transitions) in real time has proved to be a tremendously difficult task for scientists. While we have been able to achieve invaluable atomic resolution structures of proteins via X-ray and solid-state NMR techniques,[138,139] these measurements are relatively stationary snapshots and cannot be easily applied to membrane bound proteins. To sufficiently capture fast protein motions, a high temporal resolution is required. Time-resolved infrared (IR) spectroscopic techniques have proven to be extremely advantageous in this regard, due to the fast timescale of molecular vibrations (femtosecond to picoseconds).[140] However, in a protein and/or even a small peptide, the sum of all intrinsic vibrations, emerging from the amide backbone and sidechains, can generate overlapping signals that are spatially unresolved.[66] As a result, over the last two decades a library of site-specific IR probes of proteins have evolved to enhance structural resolution within IR measurements, and have been comprehensively reviewed elsewhere.[80,141-143]

Briefly, a suitable IR marker for protein studies must contain the following prerequisites: (1) it can be easily integrated into biologically relevant sites of a protein, (2) the frequency of the IR vibrational mode is in a spectrally transparent region, (3) it is responsive to changes in local environment (e.g., local electric field, hydration, polarity), (4) it has a large cross-section and extinction coefficient, and most importantly, (5) the extrinsic modification must be non-perturbing to the protein structure and function. Two common approaches are to extrinsically modify the protein at the backbone or sidechain

level.[142] In one case, increasing the reduced mass of an IR mode results in a vibrational frequency red-shift, resulting in an isolated and structurally-specific IR band. An effective and well-known example in the literature utilizes isotope-edited atoms in the protein backbone (i.e., $^{12}C=^{16}O$ modified to $^{13}C=^{18}O$),[80,144] as well as deuterium-modified sidechains (i.e., C-H modified to C-D)[145] to accomplish this task. Additionally, non-natural amino acid spectroscopic IR probes, which contain an IR active extrinsic functional group on the sidechain, have also been used extensively in studies of protein structure and dynamics.[141] Because these chemically customized sidechains are more biosynthetically compatible (both *in vitro* and *in vivo*),[146-148] they can be more widely applicable in biophysical investigations. Examples include azide-,[149] ester carbonyl-,[150] metal carbonyl-,[151] and cyano-modified[152] sidechains. In particular, nitrile-derivatized amino acid IR probes have proven to be extremely useful to address biological questions,[142] and a variety of analogs have been developed thus far, such as *p*-cyanophenylalanine,[153] thiocyanate,[154] and 5-cyanotryptophan.[155]

The nitrile (C≡N) functional group offers several advantages as a site-specific IR marker. First, its vibrational stretching frequency is in a spectrally unconjested region (2100 – 2400 cm^{-1}), it has a high extinction coefficient when bonded to an aromatic sidechain (~250 cm^{-1}), and most of all, its frequency and lineshape have shown a dependence on local electric field[156] as well as hydration status.[155] For example, thiocyanate in conjunction with vibrational Stark spectroscopy has been used to assess the local electrostatics of enzyme active sites[157] and protein-protein interactions.[158] Finally, nitrile-containing aromatic sidechains also possess characteristic fluorescence

properties,[58,155,159] rendering them as dual site-specific IR and fluorescence probes. This point will be expanded on in the context of 5-cyanotryptophan later in this thesis.

4 How Quickly Can a *β*-Hairpin Fold from Its Transition State?

Reprinted with permission from Journal of Physical Chemistry B, Beatrice N. Markiewicz, Lijiang Yang, Robert M. Culik, Yi Qin Gao, and Feng Gai, (2014) 118, 3317-3325. DOI: 10.1021/jp500774q, http://pubs.acs.org/doi/abs/10.1021/jp500774q,[160]

4.1 Introduction

Protein folding occurs spontaneously, as the process lowers the free energy of the system upon formation of the folded state. During folding, however, the entropic loss can transiently outweigh the enthalpic gain, thus resulting in one or more free energy bottlenecks or transition states along a given folding pathway.[2,19,32] Because, for a given protein, identifying the structure of the folding transition state(s) is key to elucidating its folding mechanism, these high energy states have been the subject of many previous studies.[161] Due to the transient nature of these states, however, it is extremely difficult, if not impossible, to directly observe and study them experimentally. As such, only indirect methods[9,162-164] have been used to yield structural information about folding transition states. For example, one such method, Φ-value analysis, uses site-specific sidechain mutations in conjunction with stability and kinetics measurements to infer if a certain sidechain becomes native-like at the transition state.[10,11,165] While these methods have proven invaluable in the study and understanding of how proteins fold, they cannot be used to isolate the folding transition state of interest for further structural and dynamic investigations. Thus, it would be very useful to devise a method that can convert a folding transition state to a thermodynamically stable and accessible state. Herein, we propose, based on the transition state analog (TSA) methodology commonly used in

mechanistic studies of enzymatic reactions,[166,167] that it is possible, at least for small protein systems, to use cross-linking strategies to accomplish this goal.

In enzymatic studies, the TSA represents a stable, nonreactive enzyme-substrate complex that mimics the transition state of the reaction of interest.[168,169] Since the TSA is thermodynamically accessible and also captures the key binding interactions of the transition state complex,[170] it thus allows for a detailed structural characterization of the transition state, which is otherwise difficult to achieve.[171] For protein folding, another advantage of being able to engineer TSAs is that, besides what is mentioned above, it would enable us to measure folding dynamics that are otherwise inaccessible, i.e., the conformational dynamics at the downhill side of the folding free energy barrier. For small and two-state folding proteins, the folding transition state often consists of a relatively small number of key sidechain-sidechain and backbone-backbone hydrogen bond interactions. Thus, a viable strategy that could be used to create a folding TSA is to cross-link some of these key interactions via covalent bonds. To test the feasibility of this idea, herein we apply it to one of the simplest folding systems, the β-hairpin.

Since the folded structure of β-hairpins consists of a series of backbone-backbone hydrogen bonds, some of which are formed in the transition state,[85,172-178] an ideal approach to create a folding TSA would be to covalently cross-link one or multiple of these hydrogen bonds. Converting a backbone-backbone hydrogen bond to a chemical bond without introducing significant structural perturbations, however, is challenging. Thus, instead we seek to use a sidechain disulfide cross-linker to help, albeit in an indirect manner, restrain a particular native backbone-backbone hydrogen bond that is

predicted to form in the transition state. A disulfide bond may introduce strain to these hydrogen bonded sites, however, experiments on a small model system have shown that the favorable enthalpic contribution of the cross-linker could compensate for this potential geometric distortion.[179] While the strategy of cross-linking has been widely used to increase protein stability,[99,104,179-182] trigger protein unfolding,[108,183] and to interrogate, in the context of Φ-value analysis,[102,103,105,106,184] folding mechanisms, to the best of our knowledge it has not been used to create a thermodynamically stable conformation that structurally resembles the folding transition state. As many studies have shown, the major folding pathway of Trpzip β-hairpins[185-197] involves a transition state wherein the turn structure is at least partially formed. Thus, we propose, using Trpzip4 as a testbed, to create a β-hairpin folding TSA by forcing the formation of a backbone-backbone hydrogen bond critical to the stability of the β-turn between Asp6 and Thr11 (Table 4.1). As indicated (Figure 4.1), such a conformational constraint, if effective, is expected to divide the unfolded state ensemble into two structurally distinguishable subpopulations: one with a native or native-like turn (U_A), whereas the other is completely unstructured (U_B). If U_A behaves like a folding TSA, we expect that its folding rate will be significantly increased with respect to the wild-type. On the other hand, the folding kinetics of U_B are expected to be similar to those of the wild-type because both pathways (i.e., $U_B \rightarrow F$ and $U_B \rightarrow U_A \rightarrow F$) involve the formation of the turn, the rate limiting step in Trpzip4 folding.[190] Indeed, consistent with this picture, our results show that the conformational relaxation kinetics of this cross-linked Trpzip4 variant, induced by a rapid temperature-jump (*T*-jump) and measured via time-resolved infrared (IR)

spectroscopy,[198] are biphasic, with one phase having a relaxation rate that is almost identical to that of the wild-type and another relaxing an order of magnitude faster. Further evidence supporting our hypothesis (Figure 4.1) is that *T*-jump measurements on another cross-linked Trpzip variant, whose two ends are connected via a disulfide (Table 4.1), only show single-exponential relaxation kinetics, as are observed for the wild-type, but with a slower relaxation rate, due mainly to a decrease in the unfolding rate.

4.2 Experimental Section

All peptides were synthesized using standard 9-fluorenylmethoxy-carbonyl (FMOC) methods on a PS3 peptide synthesizer (Protein Technologies, Woburn, MA) and purified by reverse-phase high-performance liquid chromatography (HPLC). Amino acids were purchased from Advanced ChemTech (Louisville, KY). MALDI-TOF mass spectrometry was used to characterize the identity of the synthesized peptides. Disulfide formation was accomplished, using a published protocol,[199] via dimethyl sulfoxide (DMSO). Specifically, an appropriate amount of pure peptide solid was dissolved in a 20% DMSO solution in H_2O, and the resulting mixture was stirred for 4 hours at room temperature, allowing the oxidation reaction to complete. Subsequently, a second round of HPLC was carried out to purify the disulfide cross-linked peptide, and the identity was further verified by MALDI-TOF mass spectrometry. All peptide samples were prepared by directly dissolving the lyophilized peptide solid in D_2O and the final pH of the peptide samples was approximately 3. For the uncross-linked peptide samples, a reducing agent, tris(2-carboxyethyl)-phosphine hydrochloride (TCEP), was also added at a concentration of approximately 10 times that of the peptide, to ensure that disulfide bonds were

completely removed. The peptide concentration was determined optically using the absorbance at 280 nm and ε_{280} = 22,125 cm^{-1} M^{-1} and, for all peptide samples, residual trifluoroacetic acid (TFA) from peptide cleavage has been removed via DCl exchange.

Static and Time-Resolved Spectroscopic Measurements

The instruments and conditions used to collect the spectroscopic data, including static circular dichroism (CD) and infrared (IR) and time-resolved IR measurements, are identical to those used previously.[200] For the IR measurements, the peptide concentration was approximately 2 mM, prepared in D_2O (pH 3). The probing frequency in the *T*-jump experiment was 1626 cm^{-1}, and the *T*-jump amplitude was in the range of 8-12 °C.

Molecular Dynamics Simulations Using Integrated Tempering Enhanced Sampling

Molecular dynamics (MD) simulations were performed for wild-type and cross-linked Trpzip4 using the AMBER 11 package. The peptides were modeled with the AMBER FF96[201] all-atom force field and the solvent was modeled with the GB^{OBC}/SA implicit solvent model.[202] The salt concentration is set to 0.2 M, and the default surface tension is 5×10^{-3} kcal/mol/$Å^2$. The SHAKE[203] algorithm with a relative geometric tolerance of 10^{-5} is used to constrain all chemical bonds. Nonbonded cutoffs were not used in the simulations. For both wild-type and cross-linked Trpzip4, ten independent trajectories were carried out for 200 nanoseconds (2 μs in total). In each trajectory, the fully extended structure of the polypeptide was first subjected to 2500 steps of minimization, and then the temperature of the system was established by velocity rearrangement from a Maxwell-Boltzmann distribution at 300 K. Afterwards, the system was maintained at 300

K using the weak-coupling algorithm with a coupling constant of 0.5 ps^{-1}. The integrated tempering enhanced sampling (ITS)[204-206] method was used in the production run of each trajectory. In the present study, 100 temperatures, evenly distributed in the range of 270-470 K, were used in the ITS method to ensure the efficient sampling of the desired energy.

4.3 Results and Discussion

We chose the Trpzip4 β-hairpin (Table 4.1) as our model system because of its small size and the large body of information on its folding mechanism.[16,188,190,207-211] The major folding pathway, as suggested by previous studies,[188,190] begins with turn formation, which is a thermodynamically unfavorable event and hence results in a folding free energy barrier, followed by the sequential creation of backbone-backbone hydrogen bonds further away from the turn in a 'zipping out' manner. Thus, based on this picture and also the NMR structure[212] which indicates that the first interstrand hydrogen bond is formed between the amides of Asp6 and Thr11, we propose to use a disulfide to reduce the number of possible configurations available near these residues with the expectation that this restriction is sufficient to produce an unfolded species that has a native-like turn and, therefore, behaves like a folding TSA. To accomplish this, we first mutated Asp6 and Thr11 to cysteine, and then the resulting mutant was placed under oxidizing conditions to promote disulfide bond formation. In addition, we have studied another cross-linked variant of Trpzip4 with a disulfide formed at the peptide ends that serves as a control. For convenience, the sequences and abbreviations of all the peptides studied are summarized in Table 4.1.

Effect of Cross-Linking on the Thermal Stability of Trpzip4

As shown (Figure 4.2), the far-UV CD spectra of both cross-linked and uncross-linked peptides are in line with that of the wild-type,[190] exhibiting a positive band at 228 nm.[51] Since this CD feature signifies the π-π* exciton-coupling of the paired Trp residues in the folded state,[50,213,214] these results suggest that the mutations and disulfide constraints used in this study do not significantly perturb the native fold of Trpzip4. In addition, the chemical shifts obtained from the 1D ^{1}H-NMR spectrum of TZ4-T-CL agree with previously published data,[212] and are well dispersed, which indicates a well folded secondary structure (Figure 4.3). As expected, CD thermal unfolding measurements indicate, for a given sequence, that cross-linking increases the thermal stability of the β-hairpin, in comparison to that of the uncross-linked peptide (Figure 4.4 and Table 4.2). Nevertheless, what is more interesting is that the T_m (~67 °C) of TZ4-T-CL is almost identical to that (~70 °C) of the wild-type,[188] indicating that the added disulfide constraint at this site does not significantly perturb β-hairpin stability, presumably because the enthalpic stabilization gained from the cross-linking is mostly offset by the conformational entropic loss in the unfolded state.[215] On the other hand, as observed in similar studies,[181,216] cross-linking the two ends of the Trpzip4 peptide results in a hyperstable β-hairpin with a T_m > 100 °C. This result is not surprising considering that the unfolding process, which, according to the zipping out folding mechanism, should 'unzip' the backbone-backbone hydrogen bonds of the hairpin starting from the strands' termini.

Effect of Mutation and Cross-Linking on the Relaxation Kinetics of Trpzip4

The conformational relaxation kinetics of these Trpzip4 variants were measured by a laser-induced *T*-jump IR technique, as was used in the study of the Trpzip4 wild-type peptide. As shown (Figure 4.5), unlike what was observed for the wild-type, the conformational relaxation of TZ4-T-CL, probed at 1626 cm^{-1}, proceeds with two distinct and well-separated kinetic phases, indicating that the disulfide cross-linking indeed, as proposed, changes the folding mechanism of Trpzip4, as the wild-type, in response to a *T*-jump, only shows single-exponential decays. Interestingly, as shown (Figure 4.6), the relaxation rate of the slow phase (τ_2) is almost identical to that of the wild-type in the temperature range between 40-45 °C,[188] whereas the relaxation rate of the fast phase (τ_1) is an order of magnitude faster. To confirm that the difference in the relaxation kinetics of TZ4-T-CL and the wild-type originates from the disulfide cross-linker, we also measured the relaxation kinetics of the uncross-linked variant of this peptide (TZ4-T-UL). As shown (Figure 4.7), the *T*-jump induced relaxation kinetics of TZ4-T-UL can be well described by a single-exponential function with a rate constant that is slower than that of the wild-type. Further analysis of the relaxation rates, based on a two-state model,[46] indicates that this slowing down predominantly arises from a decrease in the folding rate (Figure 4.7). For example, at 40 ^{o}C the folding and unfolding rate constants of TZ4-T-UL are determined to be $(13.3\ \mu s)^{-1}$ and $(66.0\ \mu s)^{-1}$, respectively, compared to $(6.6\ \mu s)^{-1}$ and $(48.4\ \mu s)^{-1}$ of the wild-type. This result is in agreement with the notion that the turn is at least partially, if not completely, formed in the transition state as these cysteine mutations, as shown above (Figure 4.4), destabilize the native fold. Additionally, the

relaxation rate of the slow phase in TZ4-T-CL is at least 2 times faster than that of TZ4-T-UL, thus ruling out the possibility that the double-exponential kinetics observed in TZ4-T-CL could be due to a mixture of oxidized and reduced disulfides.

To further confirm that the double-exponential relaxation behavior observed is unique to TZ4-T-CL, we have also studied a second disulfide Trpzip4 variant, TZ4-E-CL, where the cross-linker is introduced at the termini of the β-hairpin and, hence, is not directly involved with interstrand hydrogen bonding or key sidechain interactions. As shown (Figure 4.8), similar to those observed for the wild-type and TZ4-T-UL, the *T*-jump induced conformational relaxation kinetics, measured only at high temperatures because of the peptide's high stability (Table 4.2), are single-exponential. Thus, these results provide further evidence supporting the notion that the disulfide cross-linker in TZ4-T-CL is unique in that it alters the folding mechanism of Trpzip4.

Evidence Suggesting the Population of a Folding TSA in the Unfolded State

The fact that the conformational relaxation of TZ4-T-CL occurs in a distinctively different manner than that of Trpzip4 wild-type suggests that the cross-linker acts to introduce either an additional folding pathway by forming two distinct unfolded conformational states, as indicated in Figure 4.1, or an on-pathway folding intermediate (i.e., I in a sequential folding mechanism U→I→F). Distinguishing between these two possibilities is not easy, as both could give rise to double-exponential relaxation kinetics with two drastically different rate constants. However, MD simulations provide evidence indicating the presence of two unfolded populations. As a result, we propose that the folding mechanism follows the cartoon shown in Figure 4.1, where there are two

pathways to the folded state, one with a barrier (i.e., from U_B) similar to the wild-type and one (i.e., from U_A) with a much smaller barrier or no barrier at all.

The above results support the notion that the disulfide cross-linker in TZ4-T-CL modifies the mechanism of Trpzip4 folding by creating a partially folded, thermodynamically accessible state, U_A, which folds on an ultrafast timescale (~500 ns at 40 ^oC). In addition, a simple calculation, using the relaxation rate constants of U_A and U_B, suggests that the difference in their folding free energy barrier heights (i.e., $\Delta\Delta G^{\ddagger}$) is about 2.4 k_BT. Considering that U_B has roughly the same relaxation rates as the wild-type and that small proteins typically have a free energy barrier in the range of 2–4 k_BT,[18,19,217-219] these results suggest that the folding of U_A (Figure 4.1) proceeds without encountering any significant free energy barriers. In other words, we believe that U_A behaves like a TSA and its folding rate, approximately $(500\ ns)^{-1}$, reports on the dynamics of a fundamental event in β-hairpin folding, namely, the process taking the system from the transition state to the folded state.

To provide further evidence supporting the proposed folding mechanism, we carried out free energy calculations on TZ4-T-CL. Specifically, we generated the folding free energy landscape of TZ4-T-CL using MD simulations at 313 K as a function of turn residues, which include Asp7-Lys10, and the residues outside the turn region that are also involved in interstrand hydrogen bonding (referred to as β-strand residues). As shown (Figure 4.9), the simulations clearly indicate that two major unfolded populations are present, with one having a folded turn structure (Figure 4.10) and likely corresponding to the proposed U_A state. In addition, there is no apparent barrier between this unfolded state

and the folded state, indicative of a downhill folding pathway. The second unfolded population observed in the simulations corresponds to a fully unfolded structure (Figure 4.10) with none of the turn and β-strand residues being native-like and, thus, is consistent with the proposed U_B state. Unlike the partially unfolded state, folding from this fully unfolded state involves a free energy barrier of ~3.0 k_BT, which agrees well with the value of ~2.4 k_BT calculated from experiments. Similarly, simulations indicate that converting the fully unfolded state to the partially unfolded state also involves a free barrier that amounts to ~3.0 k_BT. For comparison, we also computed the folding free energy landscape of Trpzip4 wild-type. As shown (Figure 4.11), the folding free energy barrier is ~5.0 k_BT, which is in close range to the barrier height observed for U_B in TZ4-T-CL. However, if folding were assumed to begin from an unfolded state similar to U_A, the free energy barrier is significantly higher, showing that the disulfide cross-linker can indeed play a key role in modulating the folding free energy landscape.

In addition, a recent study by Dyer and coworkers[220] showed that a designed mini-β-hairpin, CLN025, a variant of Chignolin, with a preformed turn in the unfolded state, has a folding rate of $(\sim 100\ \text{ns})^{-1}$ at 40 °C. They attribute this ultrafast rate to an early hydrophobic collapsed structure that results in a free energy landscape with a minimal folding barrier. As a consequence, the folding rate is limited only by local rearrangements required to accommodate native hydrogen bond formation. Since Hamm and coworkers have shown that the rate of hydrogen bond formation occurs on a picosecond timescale,[221] then the difference in the folding rates of CLN025 and TZ4-T-CL (from U_A) most likely reflects the difference in the times required to bring the two chains to their

native geometries in these two cases. In other words, the rate of β-hairpin structural evolution, from the transition state, should be limited by the chain diffusion rate or the rate of loop closure and, as a result, the longer folding time of TZ4-T-CL (from U_A) is due to its longer chain length. To provide further support of this notion, we analyzed how the folding rate depends on chain length, using the number of native hydrogen bonds as a proxy. Assuming that the turn is preformed for both CLN025 and TZ4-T-CL (from state U_A), folding then involves propagation of two or four hydrogen bonds, respectively. As shown (Figure 4.12), the folding time shows a power law dependence on the number of native hydrogen bonds to be propagated and, perhaps more interestingly, the value of the exponent (2.3) is almost identical to that (2.4) determined by Makarov and coworkers[222] for end-to-end loop closure time with respect to length for unstructured polymer chains. Not only does this finding provide further evidence indicating that the folding of TZ4-T-CL, when it starts from state U_A, encounters a small, if any, free energy barrier, but it also suggests that the time it takes to form native contacts in a peptide chain in a downhill folding scenario can be estimated by the rate of loop closure. Finally, another line of evidence supporting the aforementioned power law relationship is that the rate of adding an extra β-strand onto a folded three-stranded β-sheet protein, which involves formation of four interstrand hydrogen bonds in a barrierless manner, was also found to be approximately $(500\ ns)^{-1}$ at 40 ^{o}C.[223]

4.4 Conclusions

The transition state is the hallmark of protein folding dynamics. However, due to its transient nature, taking a snapshot of the folding transition state with sufficient structural resolution is inaccessible by current experimental techniques. Thus, it would be helpful to devise a method that could create stable structural analogues of the transition state. Here, we propose that it is possible to utilize a sidechain cross-linker to restrict a particular backbone-backbone hydrogen bond site, thus allowing for the creation of a thermodynamically stable state analogous to the transition state. In a proof-of-principle study, we apply this idea to a small β-hairpin model, Trpzip4, the transition state of which has been shown to involve turn formation. By strategically introducing a disulfide constraint in the turn region that, we believe, would facilitate native turn formation even in the unfolded state, we find that the conformational relaxation kinetics of the disulfide-bond-containing Trpzip4 has two phases, indeed indicative of the presence of an additional state. Further evidence supporting the notion that this cross-linked Trpzip4 has an unfolded state that mimics the folding transition state of the wild-type is that the folding rate of this state, about $(500\ \text{ns})^{-1}$ at 40 ^{o}C, is approximately an order of magnitude faster than the wild-type. In addition, a simple analysis of the folding rate obtained from these results reveals that cross-linking the turn induces a free energy barrier decrease of ~2.4 $k_B T$. Furthermore, MD simulations performed on the cross-linked Trpzip4 variant also corroborate the notion that two distinct unfolded populations are present, one with a preformed turn that folds via a barrierless pathway, and a second fully unfolded state that encounters a folding free energy barrier similar to that of the wild-

type. More interestingly, we find that the time required to propagate a number of native hydrogen bond contacts after the major folding barrier follows a similar length dependence as observed in loop closure kinetics.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978 to FG), the Natural Science Foundation of China (21125311 and 21233002 toYQG), and the Ministry of Science and Technology of China (2012CB917304 toYQG). We also thank Dr. George Furst at the University of Pennsylvania for assistance with NMR measurements.

Peptide	Sequence
Trpzip4	GEWTWDDATKTWTWTE
TZ4-T-CL	GEWTW**C**DATK**C**WTWTE
TZ4-T-UL	GEWTW**C**DATK**C**WTWTE
TZ4-E-CL	**C**EWTWDDATKTWTWT**C**
TZ4-E-UL	**C**EWTWDDATKTWTWT**C**

Table 4.1 Name and sequence of the β-hairpin peptides used in the current study. Underlines indicate disulfide cross-linking.

Peptide	T_m (°C)	ΔH_m (kcal mol^{-1})	ΔS_m (cal K^{-1} mol^{-1})	ΔC_p (cal K^{-1} mol^{-1})
Trpzip4	70.4	20.2	58.8	374
TZ4-T-CL	67.2 ± 5.2	16.7 ± 1.3	49.1 ± 3.8	337 ± 26
TZ4-T-UC	62.0 ± 3.0	20.2 ± 1.0	60.4 ± 2.9	451 ± 22
TZ4-E-CL	>100	-	-	-
TZ4-E-UL	57.6 ± 2.1	16.5 ± 0.6	50.0 ± 1.8	303 ± 11

Table 4.2 Thermodynamic unfolding parameters of β-hairpin peptides obtained from CD measurements. Trpzip4 parameters were taken from previously reported data.[188]

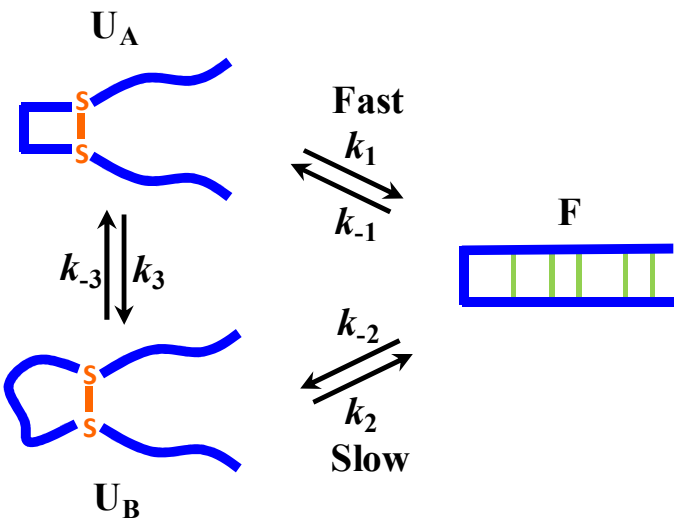


Figure 4.1 Cartoon representation of the proposed folding mechanism wherein two distinguishable unfolded states, U_A and U_B, are populated, due to the disulfide cross-linker in the turn region.

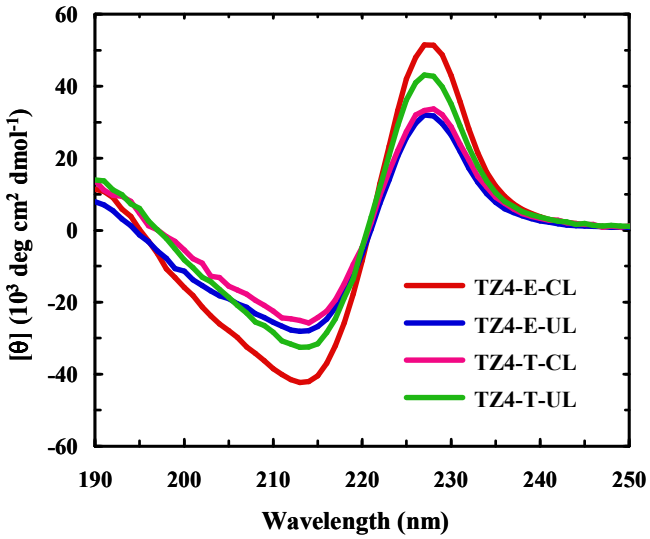


Figure 4.2 CD spectra of the cross-linked and uncross-linked Trpzip4 variants, as indicated.

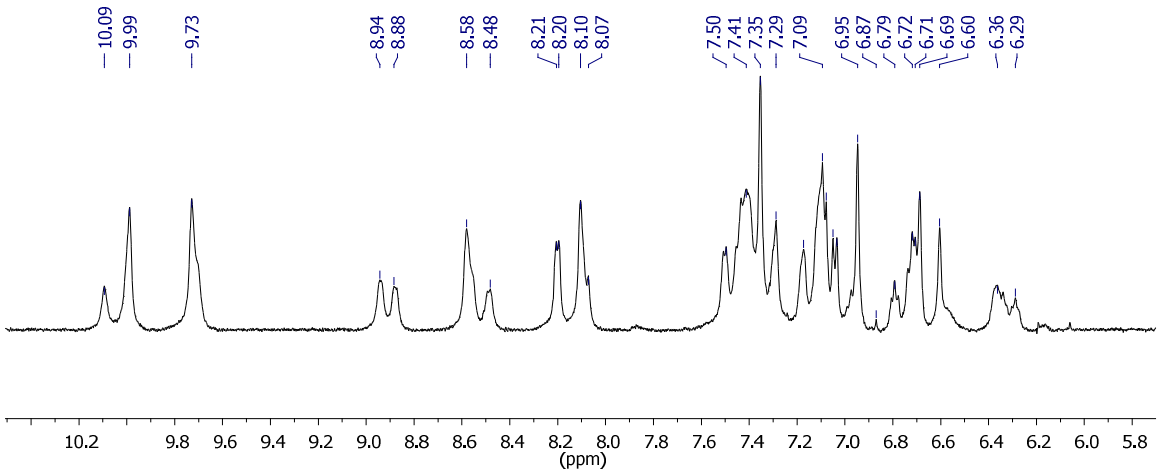


Figure 4.3 ^{1}H-NMR spectrum of TZ4-T-CL in the amide proton region, collected at room temperature on a Bruker AVIII (cryo500) (Bruker, Billerica, MA) NMR spectrometer. The NMR sample was prepared by dissolving lyophilized peptide in H_2O/D_2O (90%/10%) to a final concentration of 1 mM (pH ~3).

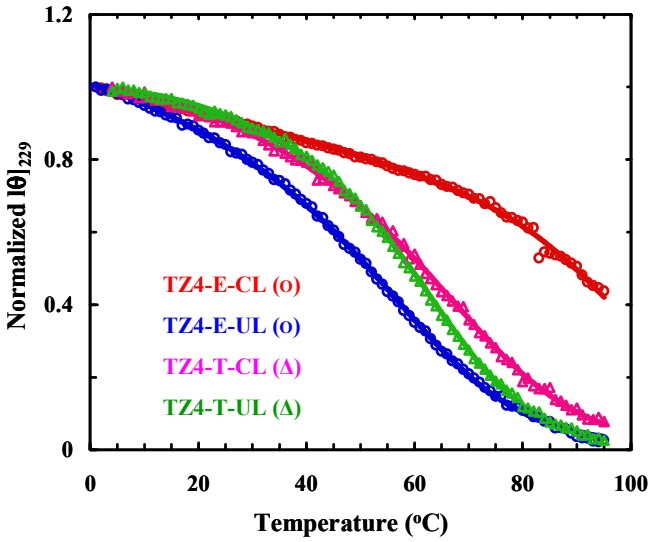


Figure 4.4 Normalized CD thermal unfolding curves of the cross-linked and uncross-linked Trpzip4 variants, as indicated. The solid lines are global fits of these data to a two-state model (Chapter 3 Eq. 3.1.1 – 3.1.5) and the resultant thermodynamics parameters are listed in Table 4.2.

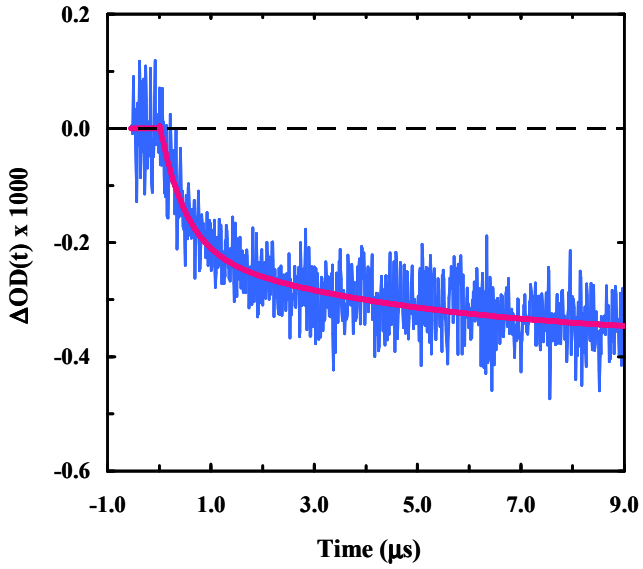


Figure 4.5 Relaxation kinetics of TZ4-T-CL in response to a *T*-jump from 28.9 to 39.7 °C. The smooth line represents the best fit of this curve to the following double exponential function: $\Delta OD(t)=A+B_1 \cdot \exp(-t/\tau_1)+B_2 \cdot \exp(-t/\tau_2)$, with $B_1 = 1.75\times10^{-4}$, $\tau_1 = 0.50 \pm 0.05$ μs and $B_2 = 1.56\times10^{-4}$, $\tau_2 = 4.7 \pm 0.6$ μs.

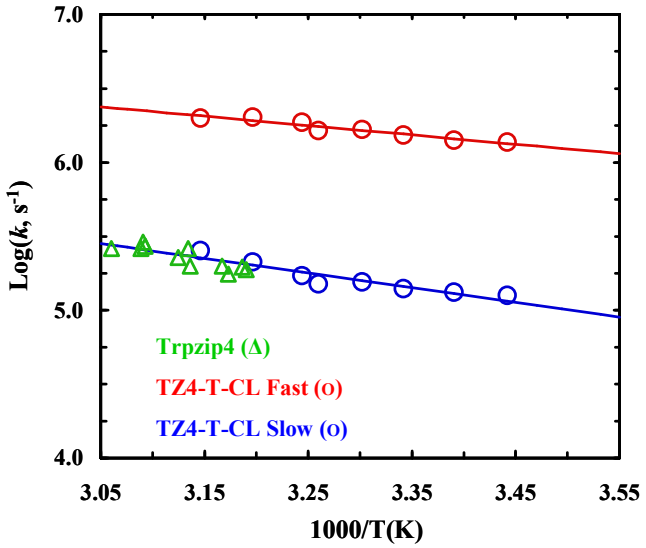


Figure 4.6 Temperature dependence of the fast and slow relaxation rate constants of TZ4-T-CL, as indicated. The green circles represent the relaxation rate constants of the wild-type Trpzip4 near 40 ^{o}C, which are reproduced from ref.188.

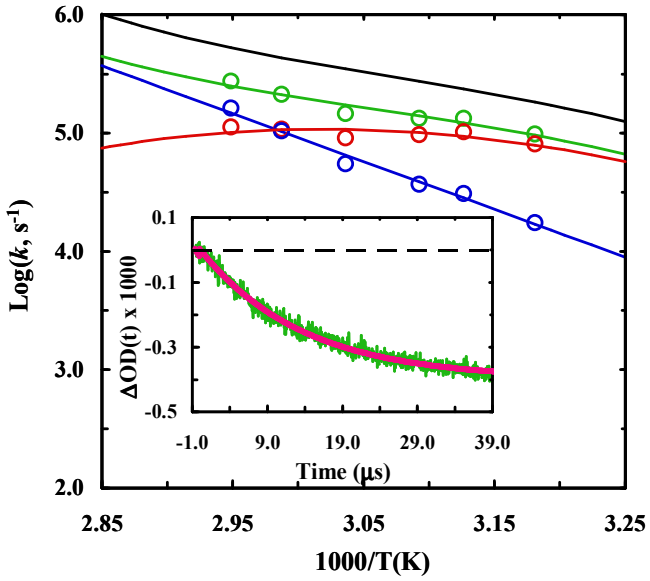


Figure 4.7 Temperature dependence of the relaxation (green), folding (red) and unfolding (blue) rate constants of TZ4-T-UL. The black smooth line represents the relaxation rate constant of the wild-type Trpzip4, reproduced from ref.188. Shown in the inset is a representative relaxation curve (green) of TZ4-T-UL in response to a *T*-jump from 33.2 to 41.2 °C, and the smooth line represents the best fit of this curve to a single-exponential function with a relaxation time constant of 13.6 ± 1.4 μs.

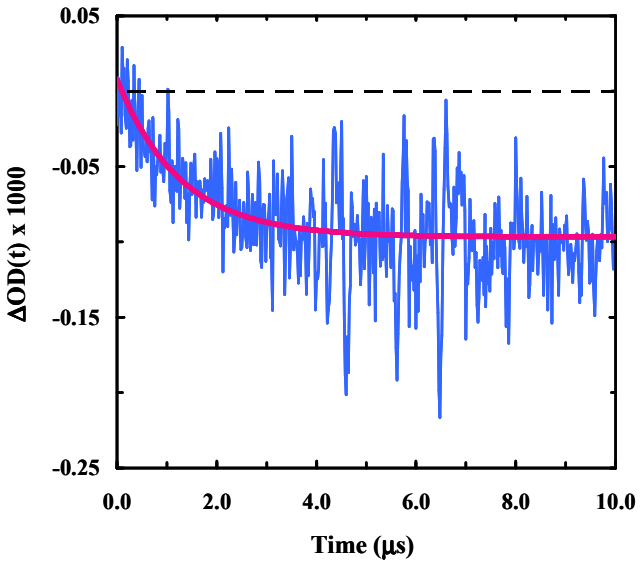


Figure 4.8 Relaxation kinetics of TZ4-E-CL in response to a *T*-jump from 69.7 to 82.0 °C. The smooth line represents the best fit of this curve to a single-exponential function with a time constant of 1.3 ± 0.5 μs.

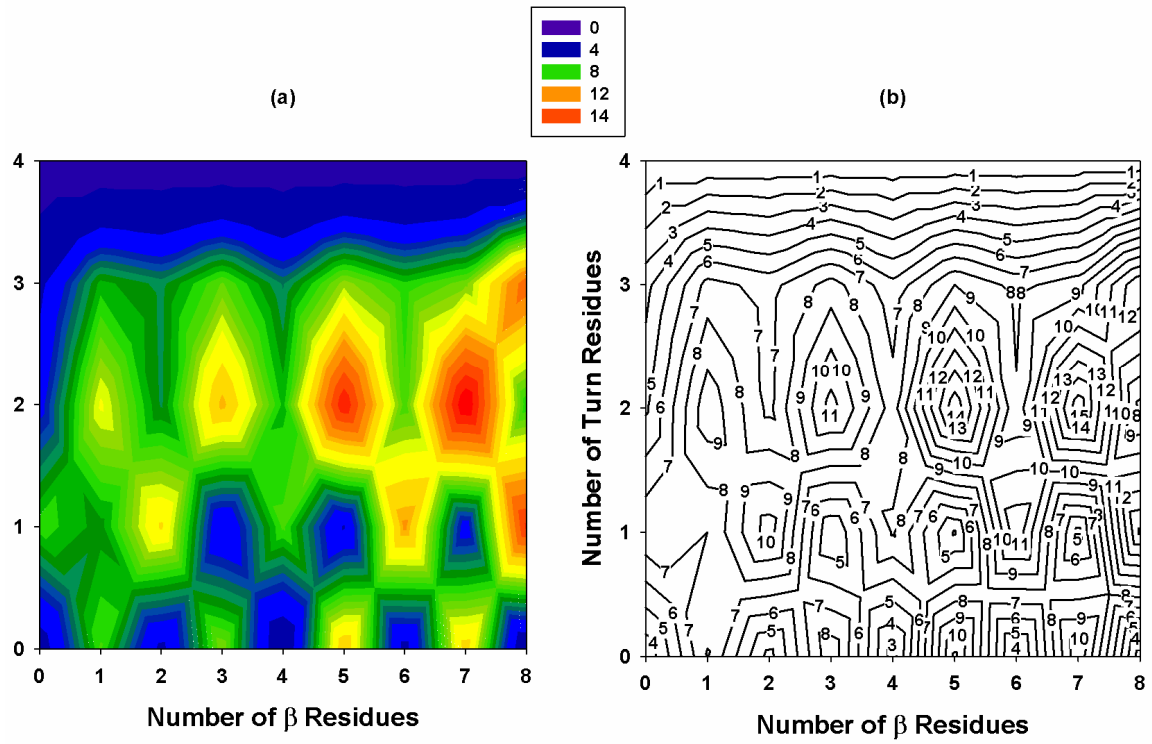


Figure 4.9 Simulated free energy landscape of TZ4-T-CL presented as function of β-strand residues versus the number of turn residues at 313K. The interval between the contour lines is 1 k_BT.

(A) **(B)**

Figure 4.10 (A) A representative structure of the U_A-like unfolded state and (B) the U_B-like unfolded state obtained from MD simulations.

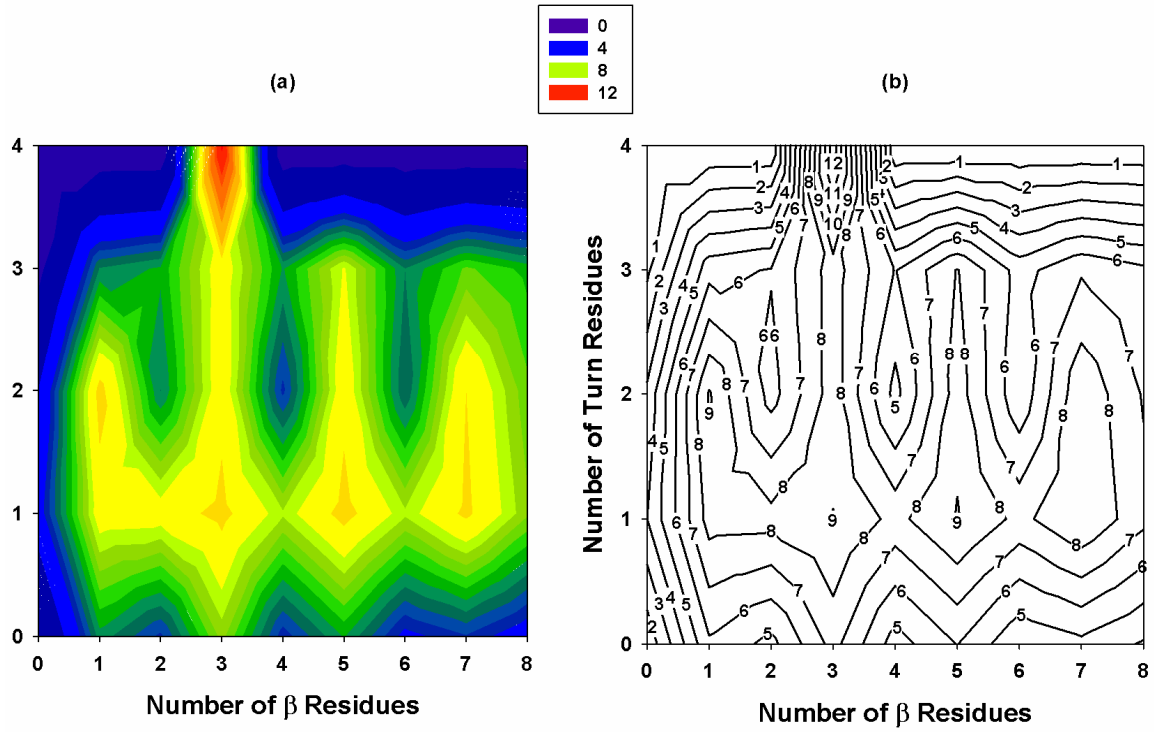


Figure 4.11 Simulated free energy landscape of Trpzip4 wild-type presented as a function β-strand residues versus the number of turn residues at 313K. The interval between the contour lines is 1 $k_B T$.

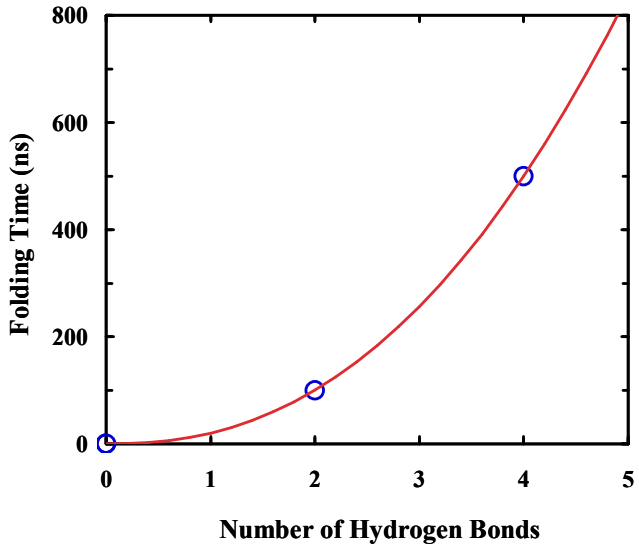


Figure 4.12 Dependence of the folding time (τ) on the number of native hydrogen bonds (n_{H}). The smooth line represents the best fit of these points to the following equation: $\tau = \tau_0 n_{\mathrm{H}}^{\alpha}$, with $\tau_0 = 20$ ns and $\alpha = 2.3$.

5 Assessment of Local Friction in Protein Folding Dynamics Using a Helix Cross-Linker

Reprinted with permission from Journal of Physical Chemistry B, Beatrice N. Markiewicz,‡ Hyunil Jo, ‡ Robert M. Culik, William F. DeGrado, and Feng Gai, (2013) 117, 14688-14696. DOI: 10.1021/jp409334h.[46] (‡ Denotes equal authorship)

5.1 Introduction

Since the native state of proteins is stabilized by many weak forces and consists of well-packed and ordered structural elements, the process of folding is expected to contain a certain degree of energetic and/or steric/topological frustrations.[5,24,32,224-228] For energy landscapes that give rise to a single free-energy bottleneck (i.e., two-state folding), such frustrations, which could arise from various local motions and interactions, typically manifest themselves as an internal frictional force or drag acting on the conformational motion along the folding coordinate.[229-231] Theory and simulations have provided us with useful insights regarding how internal friction affects protein conformational dynamics.[232-243] While the frictional force exerted by the solvent can be experimentally evaluated by altering the bulk viscosity,[232,244-249] quantifying the effect of various sources of internal friction on protein folding is nevertheless more challenging. For example, when two chains (e.g., two α-helices) become sufficiently close during folding, one expects that the further motion of one chain will be affected by the other, due to various local attractive and/or repulsive interactions between them, as well as the steric effect. This type of internal friction, which is prevalent in protein folding and may also play an important role in controlling the dynamics of conformational transitions occurring in or near the native potential well, has not, to the best of our knowledge, been studied

systematically. This is due, at least in part, to the fact that for any naturally occurring proteins it is almost impossible to isolate and independently assess the contribution of a specific structural element to the overall internal frictional effect. Herein, we show that the local frictional effect arising from a nearby chain in proteins could be estimated by using an external structural linker.

The most commonly used experimental method to determine or infer the effect of internal friction on protein folding and conformational dynamics is to measure how the rate of the dynamic event of interest varies with solvent viscosity.[42-44,250] If the rate does not show a linear dependence on the reciprocal of viscosity, as expected for a barrier-crossing process in solution, then the discrepancy is attributed to an additional friction term arising from the protein itself. For example, Eaton and co-workers[42] have utilized this strategy to show that the conformational reconfiguration rate of myoglobin, in response to photodissociation of the CO ligand, is subject to internal friction. Similarly, by assessing polymer chain dynamics in solvents containing denaturant or different viscosities via either FRET or fluorescence quenching measurements,[20,21,43,251,252] Buscaglia *et al.*[251] have shown that even loop formation in simple polypeptides is affected by internal friction, and Schuler and co-workers[20] have shown that it is possible to determine the relative amount of internal friction at different points along the folding coordinate of two α-spectrin domains (i.e., R16 and R17). Specifically, they found that the effect of internal friction is highly localized in the early transition state, suggesting that there are particular interactions stemming from local frustrations that contribute to the rate-limiting step of folding.[20] In addition, by measuring

and analyzing the viscoelastic response of a single protein molecule, several groups have demonstrated that force spectroscopy can also be used to characterize the internal friction associated with the unfolding of several proteins.[253-256] While these previous studies provided significant insights into the effect of internal friction on the folding dynamics of the protein systems studied, they were unable to disentangle the relative contributions of various sources of friction, such as those stemming from a single chain element. Herein, we attempt to use a structural cross-linker to help estimate how local friction, resulting simply from local crowding or the excluded volume effect, affects the dynamics of protein folding. Specifically, we employ the miniprotein trp-cage[257,258] and a helix cross-linker to demonstrate the feasibility of this method.

The use of a cross-linker to constrain the structural integrity of a specific fold and/or to investigate the protein folding mechanism is not new. For instance, disulfide[103,215] and dichloroacetone[105] cross-linkers have been used, in conjunction with Φ-value analysis, to characterize the folding transition state ensemble of proteins of interest. In addition, photoresponsive cross-linkers, such as those based on an azobenzene[259] or a tetrazine[108] moiety, have been used to initiate a targeted folding or unfolding process.[108,183,259-261] In this regard, the study of Hamm and co-workers[183,260,261] on helix–coil transition kinetics is particularly relevant to the present work, as it demonstrates that an azobenzene cross-linker could increase the internal friction along the α-helix folding coordinate, due to interactions between bulky sidechains and the linker. Recently, Jo *et al.*[99] have shown that *m*-xylene is one of the most efficient helix staples for short peptides. Because of the structural rigidity of *m*-

xylene, which will minimize its entropic effect on the folding dynamics of interest, herein we chose to use it as a chain surrogate.

The miniprotein trp-cage was chosen for the following reasons: (1) the study of Qiu and Hagen[41] indicates that, despite its small size, the folding kinetics of this miniprotein are subject to the influence of internal friction, (2) trp-cage contains only one α-helix, spanning residues 1 to 9, which simplifies the choice of location for cross-linker incorporation, and (3) more importantly, a number of experimental[200,262,263] and computational[264-284] studies have shown that this α-helix is formed in the folding transition state, thus making trp-cage an ideal candidate to interrogate the effect of local friction through the incorporation of an appropriate helix cross-linker. Our hypothesis is that when the cross-linker is placed near the hydrophobic core region of the trp-cage its frictional effect will become large enough to be observed in kinetic experiments. In particular, if the cross-linker is further designed to point away from the interior of the protein, we expect that its interaction with the nonhelical part of the trp-cage will be minimized, and as a result, the incorporation of the cross-linker will not induce a significant change in trp-cage stability but will lead to a significant decrease in the folding/unfolding rate, due to an increased frictional force.

5.2 Experimental Section

Peptide Synthesis and Cross-Linking

All peptides were prepared by standard 9-fluorenylmethoxy-carbonyl (FMOC) solid-phase peptide synthesis methods and purified by reverse-phase high-performance liquid

chromatography (HPLC). The detail of the cross-linking method has been described elsewhere.[99] Briefly, a solution of the targeted trp-cage cysteine mutant (10 mg, 4.8 μmol) in 50 mM ammonium bicarbonate buffer (20 mL, pH 8) was mixed with a freshly prepared 0.1 M solution of *m*-dibromoxylene in DMF (100 μL) and stirred for 2 h at room temperature. The reaction was then quenched by addition of 1 M HCl (1 mL), followed by lyophilization. The cross-linked peptide product was further purified and characterized. Specifically, peptides were purified on a preparative PROTO 300 C4 column (250x20 mm, 10 μm) using a 5-min isocratic condition of 4.5% (vol/vol) acetonitrile/water in 0.1% TFA followed by a 20-min linear gradient of 4.5% (vol/vol) acetonitrile/water to 90% (vol/vol) acetonitrile/water in 0.1% TFA with a flow rate of 15 ml/min. ESI-MS (with QTRAP 3200) was used to confirm the peptide molar mass. Residual trifluoroacetic acid (TFA) from the synthesis was removed via three rounds of lyophilization against a 0.1 M DCl solution.

Peptide Cyanylation

Lyophilized peptide was first dissolved in 2 mM tris(2-carboxyethyl)phosphine (TCEP), 4 M guanidine hydrochloride, and 100 mM phosphate buffer (pH 7) with a final peptide concentration of 200 μM. This peptide solution was then mixed with a 45 mM 2-nitro-5-thiocyanatobenzoic acid (NTCB) solution prepared in 100 mM phosphate buffer (pH 7) with a final NTCB (Sigma Aldrich) to peptide concentration ratio of 6:1. The mixture was incubated at room temperature for 30 min, allowing for cysteine cyanylation. The targeted peptide product was purified by HPLC and verified by mass spectrometry.

Static and Time-Resolved Spectroscopic Measurements

Circular dichroism (CD) measurements were carried out on an Aviv 62A DS spectropolarimeter (Aviv Associates, NJ) with a 1 mm cuvette. The peptide concentration was approximately 40–60 μM in 20 mM phosphate buffer solution (pH 7). Infrared spectra at a resolution of 1 cm^{-1} were collected on a Magna-IR 860 spectrometer (Nicolet, WI) using a home-built CaF_2 sample cell with an optical path length of 52 μm.[285] For both static and time-resolved IR measurements, all peptide solutions had a concentration of approximately 4 mM, prepared in 20 mM phosphate D_2O buffer (pH 7). Time-resolved experiments were carried out on a home-built laser-induced temperature jump (*T*-jump) apparatus[83,285] using a quantum cascade laser (Daylight Solutions, CA) as the IR probe. For the *T*-jump kinetics reported, the probing frequency was either 1668 or 1620 cm^{-1}, and the *T*-jump amplitude was in the range of 10–15 °C.

Molecular Dynamics (MD) Simulation

MD simulations were carried out using the Nanoscale Molecular Dynamics (NAMD) program (version 2.7)[286] and the CHARM36 or CHARM22 force field. The peptide of interest was immersed in 1692 TIP3P water molecules in a 40 Å cubic box.[287] For simulation of the cross-linked trp-cage variants, the force field parameters of the xylene cross-linker were generated from those of phenylalanine and cysteine. For each simulation, the temperature was gradually increased from 0 to 368 K with an increment of 20 K every 500 time steps. The temperature was then held constant once it reached its final value, 368 K. After energy minimization of the entire system, a production run of 10 ns was performed. Full electrostatics were calculated every second step using the

particle-mesh Ewald (PME) method. A 2 fs time step was used to integrate the equations of motion, and the structural coordinates were saved every 1 ps for further analysis. During the simulation, the pressure was maintained at 1 atm using the Nosé–Hoover Langevin piston method, and the temperature was controlled by Langevin dynamics. In addition, periodic boundary conditions were used to reduce edge effects, and a cutoff of 12 Å was used for nonbonded interactions.

5.3 Results and Discussion

We chose the 10b variant (sequence: DAYAQWLKDGGPSSGRPPPS) of the trp-cage miniprotein[258] as our model system because its folding kinetics and mechanism have been studied previously.[200,288] To introduce the *m*-xylene cross-linker via cysteine alkylation,[99] we first synthesized two 10b cysteine variants, the first containing double cysteine mutations at positions 4 and 8 and the second having double cysteine mutations at positions 1 and 5. Incorporation of the *m*-xylene moiety into these peptides yielded two cross-linked 10b variants (hereafter referred to as 4–8-CL-Trp-cage and 1–5-CL-Trp-cage, respectively). The reason that we chose these positions is because, as shown (Figure 5.1), the *m*-xylene cross-linker in both cases is expected to point away from the interior of the protein and, thus, should not directly interact with the hydrophobic core and the key Asp9–Arg16 salt bridge in the folded state. MD simulations indeed confirm this point (Figure 5.2). However, since the N-terminal Asp1 residue, which has been shown to be critical to the α-helix stability in trp-cage,[258] is altered in 1–5-CL-Trp-cage, we expect that its stability will decrease. In contrast, we expect that the stability of 4–8-CL-Trp-cage will be similar to that of the wild type. Furthermore, because the *m*-xylene

moiety in 4–8-CL-Trp-cage encloses two residues (i.e., Trp6 and Leu7) that are part of the hydrophobic core, we expect that in this case the cross-linker will lead to an additional frictional force along the folding coordinate, thus slowing down both the folding and unfolding rates. On the other hand, for 1–5-CL-Trp-cage, because the cross-linker is moved away from the most congested region of the protein, we expect that the frictional effect induced by the cross-linker will be significantly reduced.

Effect of the m-Xylene Cross-Linker on Folding Thermodynamics

The folding thermodynamics of 4–8-CL-Trp-cage and 1–5-CL-Trp-cage, as well as their uncross-linked counterparts (hereafter referred to as 4–8-UC-Trp-cage and 1–5-UC-Trp-cage), were examined by CD spectroscopy. For 1–5-UC-Trp-cage, we directly used the mutated sequence for experiments, whereas for 4–8-UC-Trp-cage, to prevent any potential interaction between the free cysteines and Asp9, we further modified the peptide by cyanylating the cysteines as has been described previously.[289] As shown (Figure 5.3), the far-UV CD spectra of 4–8-CL-Trp-cage and 1–5-CL-Trp-cage are consistent with that of the wild type,[200,258,290] indicating that these peptides are folded at low temperatures. Further CD thermal unfolding measurements monitored at 222 nm confirm the role of the cross-linker, which effectively increases the thermal stability of the trp-cage structure in comparison to the uncross-linked variants (Figure 5.4). To better quantify the thermal stabilities of these trp-cage peptides, we globally fit their CD unfolding curves to the two-state model outlined in Chapter 3 Eq. 3.1.1 – 3.1.5. However, in these experiments, b, c, and d were treated as global fitting parameters and a_i was peptide dependent and treated as a local fitting parameter.

As shown (Figure 5.4), this model fits the data satisfactorily. In addition, the resultant thermodynamic quantities for unfolding meet our expectations (Table 5.1). For example, in both cases the T_m of the cross-linked peptide is much higher than that of the uncross-linked variant, validating the role of *m*-xylene as a stabilizing α-helix cross-linker.[99] To verify that the cross-linker does not alter the trp-cage fold, we have performed MD simulations on both cross-linked 10b variants for a length of 10 ns and at an elevated temperature of 368 K. These simulation conditions were chosen to allow for the observation of any significant structural changes in the folded state, arising from the cross-linker. As indicated (Figure 5.2), the overall structures of these cross-linked trp-cage variants are comparable to that of the wild type.

More importantly, the T_m (∼54 °C) of 4–8-CL-Trp-cage is almost identical to that (∼55 °C) of the wild type,[200,258] which is not only consistent with our hypothesis but also indicates that 4–8-CL-Trp-cage is an ideal candidate to characterize the effect of local internal friction on the conformational relaxation dynamics due to the added cross-linker. Specifically, this result suggests that the thermodynamic role of the cross-linker is merely to compensate the loss of stability due to the cysteine mutations and that any static interactions between the linker and nonhelical residues of the protein, if any, are minimal. Thus, changes in the conformational kinetics, especially unfolding kinetics, between the wild type and 4–8-CL-Trp-cage, can be attributed to changes in local friction. In addition, the finding that the 1–5-CL-Trp-cage exhibits a lower T_m (27.8 °C) compared to 4–8-CL-Trp-cage is consistent with the fact that the N-terminus α-helical cap, Asp1, is critical to the thermal stability of the trp-cage fold[258] and also demonstrates that a helix cross-linker

beginning at the N-terminus of a peptide is less effective in stabilizing its α-helical structure. In support of this notion, our MD simulations also show that the α-helix in 1–5-CL-Trp-cage becomes more extended in comparison to that of the wild type (Figure 5.5).

Effect of the Cross-Linker on the Folding and Unfolding Kinetics

To determine the effect of the *m*-xylene cross-linker on the folding and unfolding rates of trp-cage, we measured the conformational relaxation kinetics of all four peptides (i.e., two cross-linked and two uncross-linked variants) using a *T*-jump IR technique.[83] Our previous *T*-jump study[200] on the 10b trp-cage indicated that the 3_{10}-helix is populated only at relatively low temperatures (i.e., below T_m), and its *T*-jump induced relaxation is detectable at a probing frequency of 1668 cm^{-1}. As shown (Figure 5.6), the *T*-jump induced relaxation kinetics of 4–8-CL-Trp-cage, probed at 1668 cm^{-1}, also indicate that the unfolding of the 3_{10}-helix occurs independently from the rest of the molecule. Thus, this result further confirms that the incorporation of the *m*-xylene cross-linker in 4–8-CL-Trp-cage does not change the folding mechanism of the trp-cage fold, and therefore it serves as a good mimic of a nearby chain that could lead to local crowding or friction, due to its close proximity to the most congested region of the protein. Because the 3_{10}-helix unfolds independently,[200,271,291-293] below we only discuss the effect of the cross-linker on the global folding/unfolding kinetics of the cage structure.

As shown (Figure 5.7), the *T*-jump induced relaxation kinetics of 4–8-CL-Trp-cage, probed at both 1668 and 1620 cm^{-1} (Figure 5.8) and at a final temperature of higher than 30 °C, can be fit well by a single-exponential function. However, as shown (Figure 5.9), the conformational relaxation rates obtained in the temperature range of

interest are distinctively slower than that of the wild type (i.e., 10b). In fact, both the folding and unfolding rates of 4–8-CL-Trp-cage become slower than those of 10b (Figure 5.9). More specifically, for example, at 35 °C the folding rate of 4–8-CL-Trp-cage is decreased by a factor of 3.8, whereas its unfolding rate slows down by a factor of 2.5 (Table 5.1). Since the stability of the 4–8-CL-Trp-cage is almost identical to that of 10b, we believe that these rate changes reflect at least partially, if not entirely, an increase in the frictional force along the folding coordinate.[294] This argument is based on the fact that a simple adjustment of the transition state position in the current case cannot account for the different degrees of decrease in the folding and unfolding rates. To further verify this assessment, we have also measured the *T*-jump induced relaxation kinetics of 1–5-CL-Trp-cage wherein the cross-linker is placed at the N-terminus of the peptide. In this case, as discussed above, we expect that the linker-induced frictional effect will become less pronounced or even diminished as the cross-linker is surrounding a less obstructed region of the protein. In other words, if our hypothesis is correct, we expect that 1–5-CL-Trp-cage will exhibit a slower folding rate, due to its lower stability, but a comparable unfolding rate, in comparison to 10b. As shown (Figures 5.10 and 5.11), the *T*-jump induced relaxation process of 1–5-CL-Trp-cage also follows first-order kinetics and is slower than that of the wild type. A further analysis indicates, for example at 35 °C, that its folding rate is decreased by a factor of 5.3, whereas its unfolding rate is, within our experimental uncertainties, identical to that of 10b (Table 5.1). Thus, these results provide further compelling evidence in support of the above claim that the cross-linker in 4–8-CL-Trp-cage induces an additional frictional force along the folding–unfolding

coordinate, due to its close proximity to several sidechains that undergo relatively large-amplitude motions upon folding or unfolding.

It is well-known that mutations can change the unfolding rate. Therefore, to ensure that the slower unfolding rate of 4–8-CL-Trp-cage does not simply stem from the cysteine mutations, we also measured the conformational relaxation kinetics of the two uncross-linked trp-cage variants (i.e., 1–5-UC-Trp-cage and 4–8-UC-Trp-cage). In addition, the results of these measurements will help further substantiate our previous study, which depicts a trp-cage folding mechanism wherein the α-helix is either partially or completely formed in the transition state.[200] As shown (Figures 5.12 and 5.13), both peptides exhibit a slower conformational relaxation rate than the wild type. However, as indicated (Table 5.1), this decrease in the relaxation rate originates almost exclusively from a decrease in the folding rate for these peptides. Thus, taken together, these results not only corroborate our early conclusion[200] that helix formation is the rate limiting step in trp-cage folding but also confirm that the slower unfolding rate of 4–8-CL-Trp-cage arises explicitly from the cross-linker. Furthermore, it is interesting to note that the unfolding rate of 4–8-UC-Trp-cage at 35 °C shows a measurable increase from that of 10b. Because 1–5-UC-Trp-cage does not show any appreciable change in its unfolding rate in comparison to that of 10b, this increase most likely arises from the mutation at position 8. This notion is consistent with our previous study,[295] which showed that the native interactions involving the C-terminal region of the helix are developed at the downhill side of the free energy barrier.

Quantifying the magnitude of internal friction and its effect on protein folding and conformational dynamics has been the subject of a number of theoretical and computational studies,[24,32,45,224,296-299] and in particular, several studies have attempted to make a direct connection between internal friction and the roughness of the underlying energy landscape. For example, Zwanzig showed that the conformational diffusion coefficient of a protein on a barrierless one-dimensional free energy surface is a direct measure of the roughness of this surface.[296] Upon expanding this theoretical model, Thirumalai, Straub, and co-workers[45] further demonstrated that for a barrier-crossing process the roughness ($U^{\dagger}$) could be assessed via the following relationship:

$$U^{\dagger} = \mathrm{k_B} T \ln \frac{\gamma}{\gamma_{\mathrm{int}}} \tag{5.1}$$

where γ_{int} represents the friction exerted by the solvent, whereas γ is the internal friction. Below we apply this equation to provide an estimate of how much the helix cross-linker in 4–8-CL-Trp-cage increases the roughness of the trp-cage folding energy landscape, assuming that the deviations of its folding and unfolding rates from those of 10b are due entirely to increased internal friction and that the value of γ_{int} can be approximated by that of 10b. Using the rate equation of transition state theory and these assumptions, we can easily show,

$$\frac{k_{\mathrm{U(WT)}}}{k_{\mathrm{U(CL)}}} - 1 = \frac{\gamma}{\gamma_{\mathrm{int}}} \tag{5.2}$$

where $k_{U(WT)}$ and $k_{U(CL)}$ represent the unfolding rates of 10b and 4–8-CL-Trp-cage, respectively. Using Equations 5.1 and 5.2 and the rates at 35 °C for 10b and 4–8-CL-Trp-cage (Table 5.1), we estimate the values of $U^{\dagger}$ to be $\sim 0.4 k_B T$ and $\sim 1.0 k_B T$ for unfolding

and folding, respectively. The order of magnitude of these values is reasonable given that the energetic roughness of a much larger protein, the α-spectrin domain R17, was estimated to be $\sim 2.0 k_B T$.[247]

The above analysis suggests that the friction along the free energy surface is asymmetric, which is expected if we take into account the sequence of events in the trp-cage folding mechanism. Given that the unfolded side of the free energy barrier involves sidechain interactions in close proximity leading to helix formation, we would expect more local interactions with the cross-linker, resulting in a larger degree of internal friction. Indeed, Hamm and co-workers have shown that a photoswitchable cross-linker on a monomeric α-helix slows down its folding,[183,260,261] due to the increased friction between the cross-linker and sidechains from a combination of steric and non-native interaction effects. In addition, several simulations[279,300] have suggested the presence of compact molten globule intermediates in the unfolded ensemble, which can act as a kinetic trap and may be stabilized by non-native hydrophobic interactions. Thus, it is also possible that the cross-linker increases the friction of folding by decreasing the rate out of such kinetic traps. Conversely, the folded side of the potential has a greater influence from intrinsic steric effects as the polypeptide chain becomes more compact upon folding and is presumably less affected by the added frictional force of the cross-linker.

In summary, we employed an external linker to help assess the extent to which a local structural element increases the friction along the folding coordinate of trp-cage. While the current study only yielded a global view regarding this effect, future studies employing linkers of different lengths and structures may help provide a more detailed

picture regarding the source of interactions that affect the folding–unfolding kinetics. In addition, we believe that the current method is also a useful addition to the toolbox for mechanistic study of protein folding and conformational dynamics. For example, it is possible to use a specific cross-linker to manipulate the folding free energy barrier and to selectively increase/decrease the flux of a particular folding pathway.

5.4 Conclusions

Because of chain connectivity and also the compactness of the native state, protein folding is subject to a frictional force due to, for example, local steric effects. However, assessing the frictional force arising from an individual structural element is difficult, if not impossible, as conventional kinetics studies do not provide independent information regarding the individual contributions to the overall effect of internal friction. To circumvent this difficulty, herein we use a cross-linker to selectively increase the local mass density, and thus the local friction, of a particular region of the protein of interest and to make the internal friction thus induced detectable via kinetic measurements. We apply this strategy to a well-studied miniprotein, trp-cage, and find that a helix cross-linker (*m*-xylene) appended between residues 4 and 8 on its α-helix can induce a significant decrease in both the folding and unfolding rates, which is consistent with the notion that the cross-linker will result in an increase in local crowding or internal friction. Using a simple theoretical model, we further show that the cross-linker used, which could be thought of as an individual chain segment in proteins, could increase the roughness of the folding energy surface by as much as 0.4–1.0k_BT.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978 to FG and GM-054616 to WFD)

	10b	4-8-CL-Trp-cage	4-8-UC-Trp-cage	1-5-CL-Trp-cage	1-5-UC-Trp-cage
ΔH_m (kJ mol^{-1})	58.0	44.5 ± 4.2	19.6 ± 5.9	38.9 ± 4.3	25.4 ± 5.1
ΔS_m (J K^{-1} mol^{-1})	177	133 ± 12.5	72 ± 21	129 ± 14	89 ± 18
ΔC_p (J K^{-1} mol^{-1})	176	176	176	176	176
T_m (^{o}C)	55.0	54.1 ± 4.9	-0.9 ± 3.0	27.8 ± 3.1	12.0 ± 2.9
τ_f (μs)	1.3	5.0 ± 0.6	6.9 ± 1.4	6.9 ± 1.2	13.8 ± 0.6
τ_u (μs)	5.1	12.9 ± 1.5	2.5 ± 0.6	4.8 ± 0.6	5.8 ± 0.3

Table 5.1 Summary of the unfolding thermodynamic parameters obtained from global fitting of the CD thermal melting data of the four peptides. Also listed are the folding and unfolding times of these peptides determined at 35 ^{o}C. The thermodynamic and kinetic parameters of 10b were obtained from ref.200.

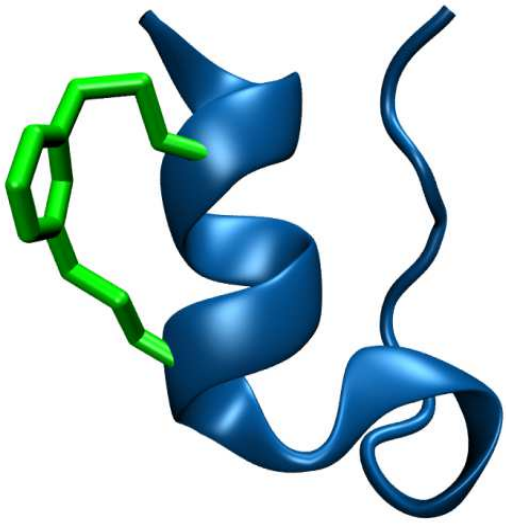
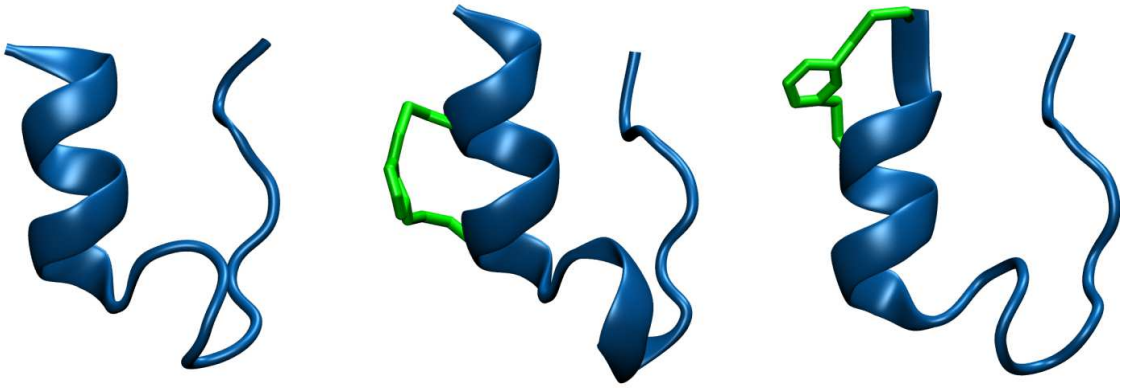


Figure 5.1 Cartoon representation of the 4-8-CL-Trp-cage.

(A)



(B)

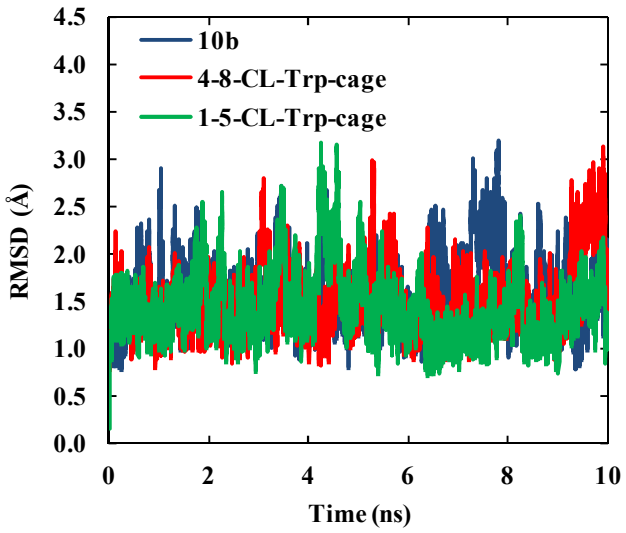


Figure 5.2 (A) MD snapshots of the wild-type 10b trp-cage (left), 4-8-CL-Trp-cage (center), and 1-5-CL-Trp-cage (right), taken at 9.5 ns in the simulation. (B) Backbone RMSD values of these peptides as a function of simulation time.

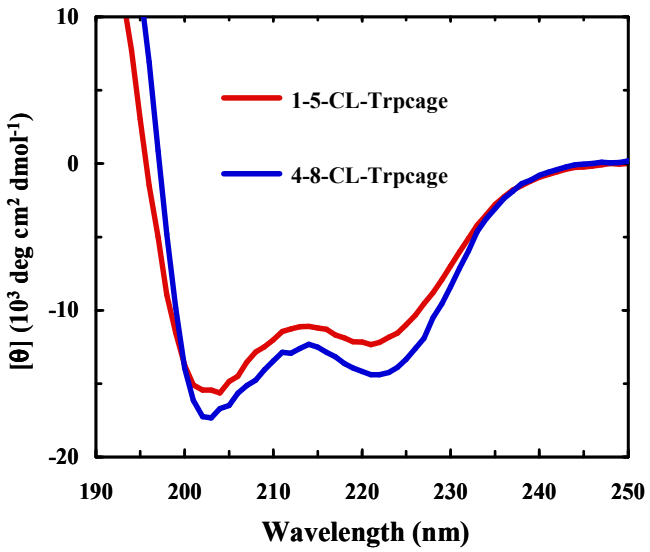


Figure 5.3 Far-UV CD spectra of the two cross-linked trp-cage peptides, as indicated. These data were collected at 1 ºC.

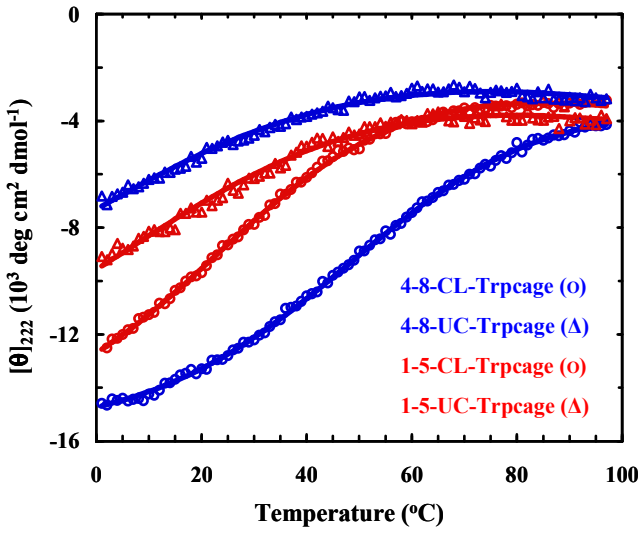


Figure 5.4 CD T-melts of the cross-linked and uncross-linked trp-cage variants, as indicated. The solid lines are global fits of these data to a two-state model discussed in text.

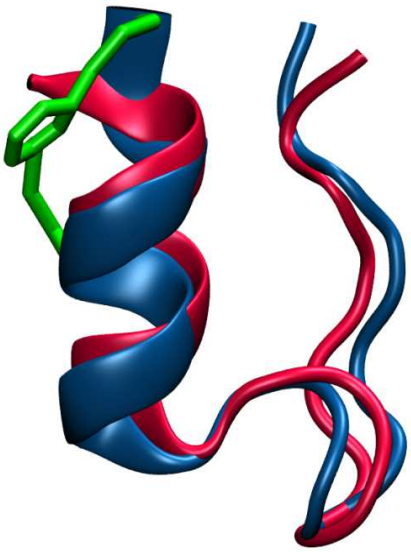


Figure 5.5 Comparison of the MD structures of 1-5-CL-Trp-cage (blue) and wild type 10b trp-cage (red), aligned with respect to the peptide backbone. These snapshots were taken at 9.5 ns. The α-helix length (i.e., the distance between the α-carbons of residues 1 and 9) was determined to be 14.31 Å for 1-5-CL-Trp-cage, which is significantly extended compared to that of 10b trp-cage (11.98 Å). On the other hand, the α-helix in 4-8-CL-Trp-cage has a length of 12.61 Å, which is more similar to that of the wild-type.

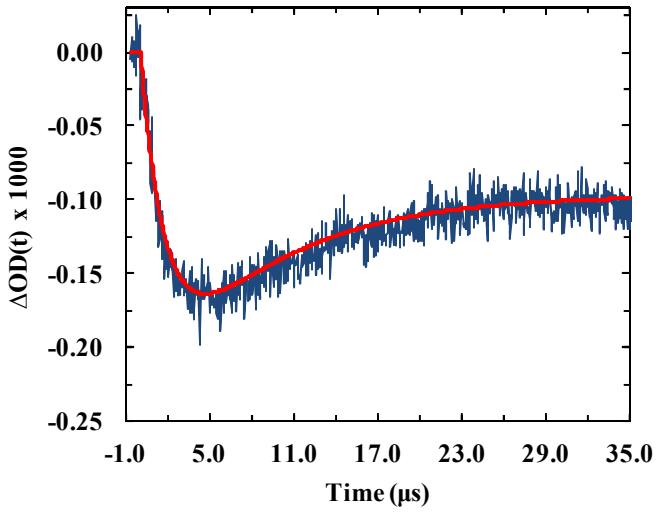


Figure 5.6 Relaxation kinetics of 4-8-CL-Trp-cage in response to a *T*-jump from 3.8 to 12.8 ^{o}C, probed at 1668 cm^{-1}. The smooth line represents the best fit of this curve to a double exponential function with a negative phase (τ = 1.8 μs) and a positive phase (τ = 8.3 μs).

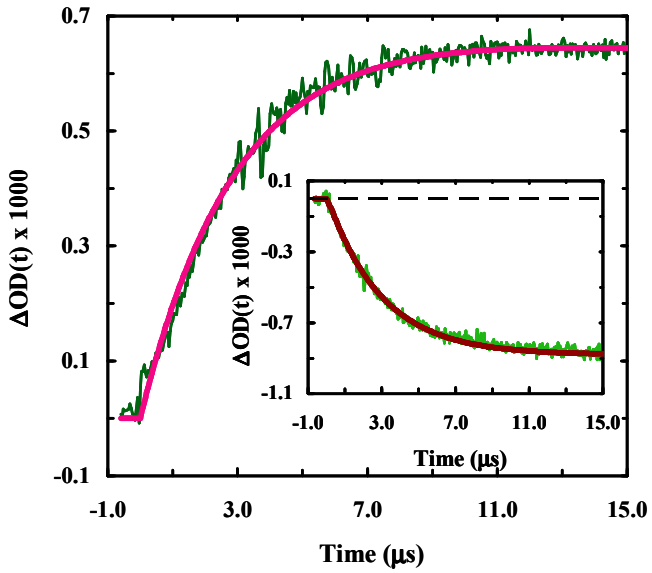


Figure 5.7 Relaxation kinetics of 4-8-CL-Trp-cage in response to a *T*-jump from 29.3 to 42.1 °C, probed at 1668 cm^{-1}. The smooth line represents the best fit of this curve to a single exponential function with a relaxation time constant of 2.7 μs. Shown in the inset are the relaxation kinetics of the same peptide obtained with a probing frequency of 1620 cm^{-1} in response to a *T*-jump of 29.1 to 43.4 °C. Fitting this relaxation curve to a single exponential function yielded a relaxation time constant of 2.9 μs.

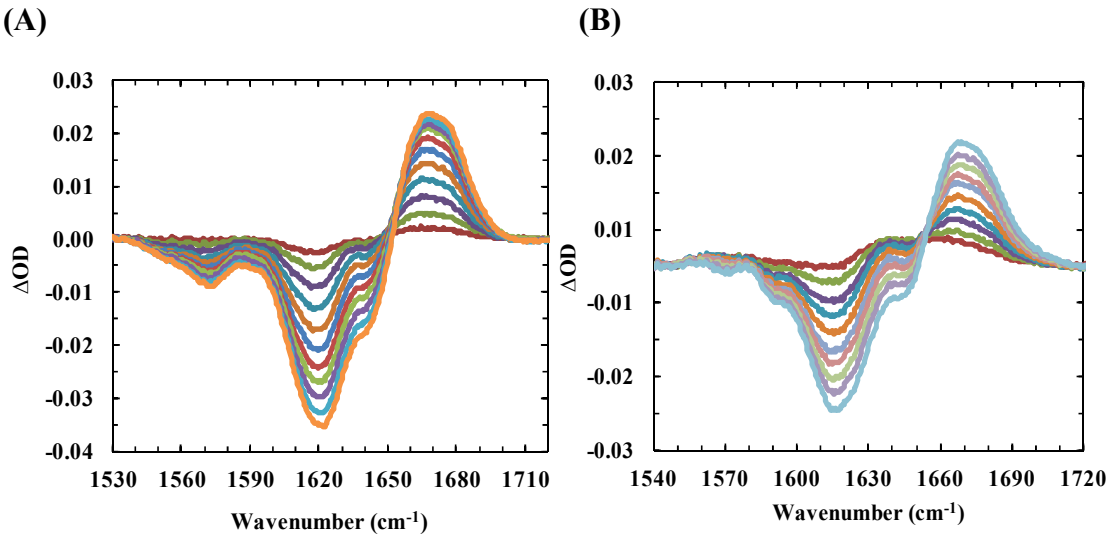


Figure 5.8 Difference FTIR spectra of (a) 1-5-CL-Trp-cage and (b) 4-8-CL-Trp-cage, generated by subtracting the FTIR spectrum at 3.4 °C from those measured at higher temperatures (the highest temperature was 80.2 °C). These FTIR spectra were used to determine the probing frequencies for the *T*-jump experiments.

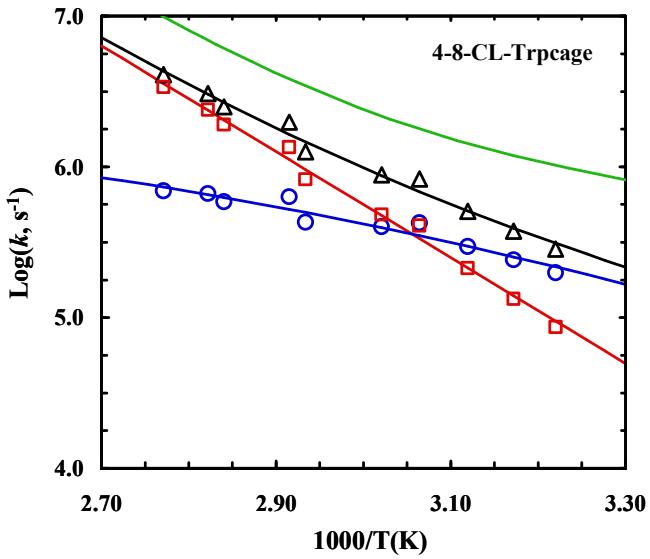


Figure 5.9 Temperature dependence of the relaxation rate constant (triangle), folding rate constant (circle) and unfolding rate constant (square) of 4-8-CL-Trp-cage. The green smooth line represents the relaxation rate constant of the wild type 10b trp-cage.[200]

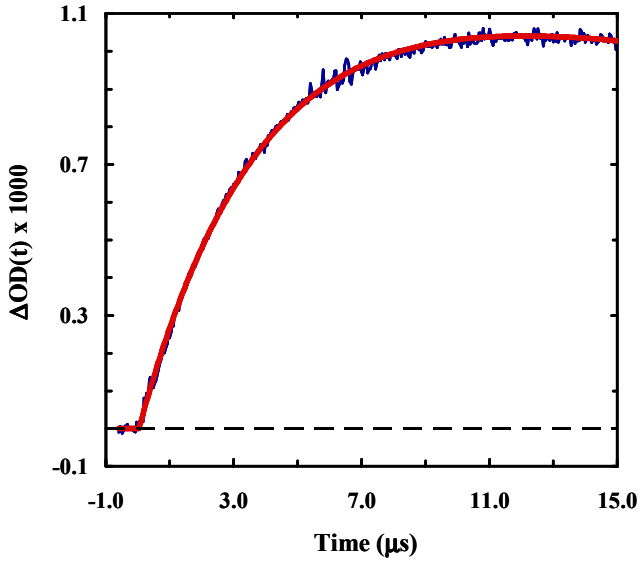


Figure 5.10 Relaxation kinetics of 1-5-CL-Trp-cage in response to a *T*-jump from 20.5 to 36.4 °C, probed at 1668 cm^{-1}. The smooth line represents the best fit of this curve to a single exponential function with a relaxation time constant of 2.8 μs.

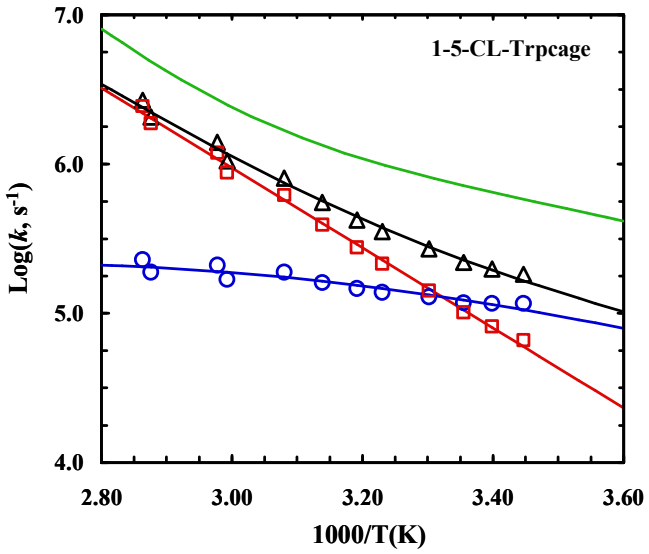


Figure 5.11 Temperature dependence of the relaxation rate constant (triangle), folding rate constant (circle) and unfolding rate constant (square) of 1-5-CL-Trp-cage. The green line represents the relaxation rate constant of the wild type 10b trp-cage.[200]

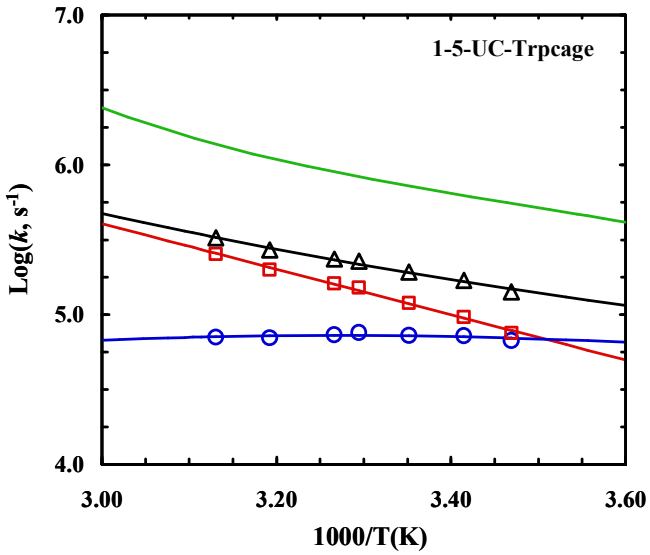


Figure 5.12 Temperature dependence of the relaxation rate constant (triangle), folding rate constant (circle) and unfolding rate constant (square) of 1-5-UC-Trp-cage. The green line represents the relaxation rate constant of the wild type 10b trp-cage.[200]

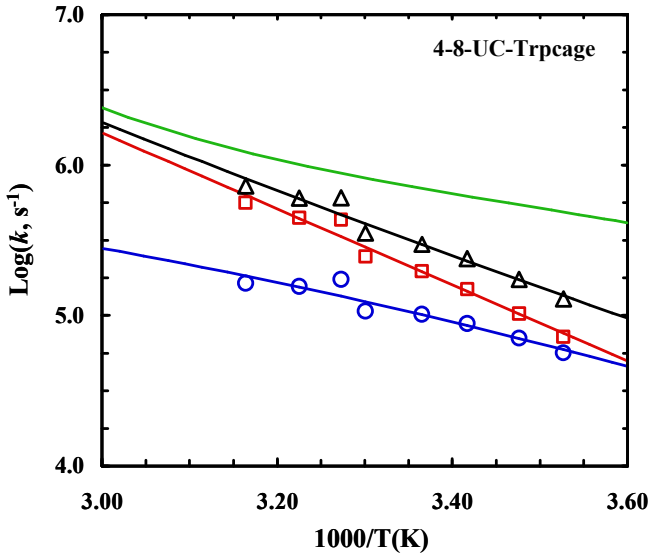


Figure 5.13 Temperature dependence of the relaxation rate constant (triangle), folding rate constant (circle) and unfolding rate constant (square) of 4-8-UC-Trp-cage. The green line represents the relaxation rate constant of the wild type 10b trp-cage.[200]

6 Aggregation Gatekeeper and Controlled Assembly of Trpzip β-Hairpins

6.1 Introduction

Protein and peptide aggregation and amyloid formation are commonly associated with various pathological disorders[302-304] and, thus, have been the subject of many studies. In addition, aggregation poses a major obstacle in *de novo* protein design and also in mechanistic studies of protein folding where relatively high protein or peptide concentrations are required. While it is easily recognized that many factors can come into play in determining the aggregation propensity of a given protein or peptide system, in practice the most commonly used strategy in protein design is to incorporate a certain number of charged residues to prevent or alleviate aggregation.[305-311] For example, Marqusee and Baldwin have shown that the solubility of alanine-based α-helical peptides in aqueous solution can be significantly increased by dispersing either lysine or glutamate residues in the peptide sequence of interest.[312,313] For β-sheet systems, however, the situation can be much more complicated, as the edge strands are often poised for further intermolecular strand–strand association,[308,314-319] and as a result, only a small difference in the peptide sequence could lead to a significant difference in aggregation propensity.[27,320-322] One distinctive example, which is also the focus of this study, is that two designed, closely related β-hairpins, Trpzip1 and Trpzip2,[212] show very different aggregation behaviors. Previous studies[188,212,323-325] indicated that Trpzip2 remains completely monomeric in the concentration range of 5–12 mM at acidic pH, whereas

Trpzip1 has previously shown measurable aggregation at concentrations of >500 μM.[217,326] For this reason, Trpzip2 has been extensively used as a model to study the mechanism of β-hairpin folding,[52,188,189,193,194,213,323,325,327-330] whereas the stronger aggregation propensity of Trpzip1 has made it a less attractive system. As shown (Table 6.1), these two β-hairpins differ only in the order of the two amino acids in the turn region (i.e., NG vs GN). Considering the fact that both peptides adopt a stable β-hairpin conformation in solution at room temperature and their sequences are almost identical, this difference is surprising.

As shown (Figure 6.1), a comparison of the averaged nuclear magnetic resonance (NMR) structures[212] of Trpzip1 and Trpzip2 indicates that the major structural variation between these two β-hairpins is in the relative orientation of the Lys8 sidechain. Specifically, in Trpzip1, the sidechain of Lys8 is oriented orthogonal to the β-hairpin axis so that it points straight outward, and away from the Trp hydrophobic cluster, whereas in Trpzip2, the same sidechain points in an upright parallel direction with respect to the β-hairpin axis. It is well-known that a solvent-exposed Lys sidechain is relatively flexible and can fluctuate among several rotamer conformations. Therefore, to obtain a more quantitative assessment of the difference in the Lys8 orientations of these two peptides, we measured the dihedral angles of Lys8 in an ensemble of 20 NMR structures obtained from the Protein Data Bank (PDB)[212] using built-in functions in Visual Molecular Dynamics (VMD).[331] The results indicate that the greatest angle disparity arises from the difference in χ_2, which describes the angle between the C_α–C_β and C_γ–C_δ planes. For Trpzip1, the Lys8 sidechain always adopts a *trans* configuration along χ_2 (171 ± 11°).

However, for Trpzip2, χ_2 fluctuates between a *trans* and *gauche*(+) configuration. When the averaged structure of all frames of Trpzip2 was evaluated, χ_2 of Lys8 is ~70 ± 10°, indicative of a *gauche*(+) rotamer along the C_β–C_γ bond.[332,333] Thus, we hypothesize that the difference in the aggregation propensities of Trpzip1 and Trpzip2 can be explained by this variation in χ_2. In other words, Lys8 in Trpzip2 is an effective aggregation gatekeeper,[306-308] which prevents edge-to-edge β-hairpin association through unfavorable electrostatic interactions between neighboring Lys8 contacts. To this end, we further point out that our hypothesis is based on the assumption that the β-hairpin unit in the peptide aggregates possesses a native or native-like turn structure, which, in conjunction with the confinement effect induced by peptide association, would place Lys8 in a specific configuration that could disfavor aggregation.

To test this hypothesis, we examined the aggregation kinetics of both peptides under different concentration and pH conditions. In addition, for Trpzip2, we also used a chemical approach to eliminate the positive charge of Lys8 by replacing it with Lys(4,5-dimethoxy-2-nitrobenzyloxycarbonyl).[334,335] The latter is a lysine analogue [hereafter termed Lys(nvoc)] with a photolabile hydrophobic moiety and has been used to control the disassembly of peptide aggregates and hydrogels via illumination.[336,337] Should Lys8 indeed serve as an aggregation gatekeeper of Trpzip2, we expect that this Trpzip2 mutant (hereafter termed Trpzip2-K) will exhibit a significantly stronger aggregation propensity.

6.2 Experimental Section

Materials and Sample Preparation

D_2O (D, 99.96%) and deuterium chloride (D, 99.5%) were purchased from Cambridge Isotope Laboratories (Andover, MA). Fmoc-Lys(4,5-dimethoxy-2-nitrobenzyloxycarbonyl)-OH [Fmoc-Lys(nvoc)-OH] was purchased from Anaspec, Inc. (Fremont, CA), and used without further purification. FMOC-protected amino acids were purchased from Advanced Chem Tech (Louisville, KY). All peptides were synthesized on a PS3 peptide synthesizer (Protein Technologies, Woburn, MA) and purified by reverse-phase high-performance liquid chromatography (HPLC). The identity of each peptide was further verified by matrix-assisted laser desorption ionization (MALDI) mass spectrometry. Residual trifluoroacetic acid (TFA) from peptide synthesis was removed by multiple rounds of lyophilization against a 0.1 M DCl solution. All peptide samples were prepared by directly dissolving the lyophilized peptide solid in D_2O, and the pH of the peptide samples was approximately 3, unless explicitly indicated separately. The peptide concentration was determined optically using the absorbance at 280 nm, with an ε_{280} of 22760 cm^{-1} M^{-1}.

Acylation of Lys12

Lyophilized peptide was first dissolved in 100 mM phosphate buffer (pH 7.4) to a final peptide concentration of 250 μM. This peptide solution was then mixed with a 30 mM *N*-acryloxysuccinimide (Sigma-Aldrich) solution prepared in 100 mM phosphate buffer (pH 7.4) containing 10% dimethyl sulfoxide with a final *N*-acryloxysuccinimide:peptide concentration ratio of 8:1. The reaction mixture was stirred for 9 h at 4 °C. The resulting peptide product was purified by HPLC and verified by MALDI mass spectrometry.

Circular Dichroism (CD) and Fourier Transform Infrared (FTIR) Measurements

CD data were collected on an Aviv 62A DS spectrometer (Aviv Associates) using a 1 mm sample cuvette. FTIR spectra were collected on a Magna-IR 860 spectrometer (Nicolet) at 2 cm^{-1}resolution using a temperature-regulated, 52 μm CaF_2 sample cell.[285]

Photocleavage Experiments

Irradiation of samples was conducted by placing the sample in the optical path of a FluoroLog fluorometer (HORIBA Jobin Yvon), at room temperature. The excitation wavelength was set to 355 nm with a slit width of 0.75 cm. The excitation intensity is approximately 8.8 mW cm^{-1}, estimated on the basis of the measured power of the excitation light and the beam diameter. The nvoc moiety has an extinction coefficient[336] (ε_{350}) of 5485 M^{-1} cm^{-1} and a photochemical yield (Φ_{365}) of 0.023.[111,338,339]

Atomic Force Microscopy (AFM) Measurements

AFM experiments were performed in air at room temperature, using a multimode atomic force microscope (model 5500, Agilent, Santa Clara, CA), equipped with a 90 μm closed loop piezoscanner. Five μL of a sample solution was applied to a freshly cleaved mica surface and allowed to sit for ~10 s, rinsed with 100 μL of Millipore water, and subsequently dried with a slow stream of N_2 gas. Tapping-mode imaging was conducted with a silicon cantilever, where the tip radius was <10 μm and the force constant was 40 N/m (Ted Pella, Redding, CA). Height and deflection images were obtained with a scan rate of 1.6 Hz and a tapping frequency of 285 kHz. Multiple images were obtained for each sample at different locations on the mica substrate to confirm the presence of fibrils.

6.3 Results and Discussion

Aggregation Kinetics of Trpzip1 and Trpzip2

The aggregation kinetics of Trpzip1 and Trpzip2 were examined using FTIR and CD spectroscopy. In particular, the amide I′ band (amide I band in D_2O) of the peptide was used as an IR probe of the aggregation process, as this band has proven to be sensitive to intermolecular β-sheet association. For example, the development of a narrow amide I′ band at approximately 1615 cm^{-1} is indicative of peptide aggregation to form parallel β-sheets,[340] whereas the appearance of a pair of narrow bands, i.e., a strong one at ∼1618 cm^{-1} and a weak one at ∼1685 cm^{-1}, signifies the formation of antiparallel β-sheets.[70,71]

As shown (Figure 6.2), the amide I′ bands of Trpzip1 obtained at different concentrations indicate that its aggregation rate is concentration-dependent, as expected. For example, at ∼2 mM and 25 °C, the peptide sample becomes almost completely aggregated after 24 h, as judged by the full development of the 1616 cm^{-1} band, whereas at ∼350 μM, aggregate formation becomes detectable after just 1 day. In comparison, even at a much higher concentration (i.e., ∼10 mM), Trpzip2 does not show any signs of aggregation under the same conditions (Figure 6.2, inset). As shown (Figure 6.3), further time-dependent measurements indicate that the aggregation kinetics of Trpzip1 at a concentration of 1.2 mM, determined by the growth of the 1616 cm^{-1} band, follow a biexponential function; the magnitude of the signal increases quickly within the first 10 h and then slowly reaches a plateau over ∼150 h. Repeating this measurement at a higher peptide concentration, which results in a faster overall aggregation rate, reproduces this biphasic kinetic pattern (Figure 6.4). Similar biphasic growth kinetics have been observed

in other peptide aggregation studies,[341-343] which were attributed to a separation in the time scales of the fibril nucleation and elongation processes.

Aggregation Mechanism and Gatekeeper

In principle, the initial aggregation step can occur through interactions between two folded β-hairpins or two unfolded peptides. However, a simple calculation, based on the thermal stabilities of Trpzip1 and Trpzip2,[212] indicates that at 25 °C the unfolded concentrations of Trpzip1 (at a total concentration of 350 μM) and Trpzip2 (at a total concentration of 10 mM) are 89 and 900 μM, respectively. Thus, these results strongly suggest that the aggregation of these β-hairpins is not initiated by association of two unfolded peptides; instead, it is triggered by dimerization of two folded β-hairpins. This notion is consistent with that put forward by Richardson and Richardson,[306] who showed that naturally occurring β-sheets can cause aggregation via edge-to-edge β-sheet interactions and that an effective strategy used by nature to defend against this is to place a charged sidechain (or gatekeeper residue) on the hydrophobic face of an edge β-strand to mask the aggregation-prone regions with a solvent favorable interaction. In the current case, we hypothesize that Lys8 in both peptides acts as an aggregation gatekeeper residue; however, it is more effective to prevent Trpzip2 from aggregating. Because Lys8 in Trpzip1 and Trpzip2 appears on the opposite face of the hydrophobic core (i.e., the four Trp residues), we note that the aggregation gatekeeper role of this charged residue is somewhat different from that discussed by Richardson and Richardson, but the overall idea remains the same: a charge is used to disfavor the process of intermolecular association. As proposed in Figure 6.1, in the early stages of the aggregation process,

should two Trpzip2 β-hairpins stack to form a dimer in a parallel fashion, the positively charged Lys8 sidechains can lead to a more unfavorable intermolecular electrostatic interaction, because of their upward-pointing rotamer geometry and thus greater proximity. As a result, the aggregation propensity of Trpzip2 is weaker than that of Trpzip1. To test this hypothesis, we employed two strategies to eliminate the charge of Lys8 and then investigated how this change affects the aggregation propensities of these β-hairpins. In the first case, we used pH to neutralize the charge, whereas in the second case, we replaced Lys8 with a neutral Lys derivative, Lys(nvoc). The added advantage of using Lys(nvoc) is that the nvoc group can be removed via light, converting the mutant back to its parent sequence. In other words, we expect that the Lys(nvoc) modification in Trpzip2 not only will enhance the aggregation propensity of the peptide considerably but also can render the aggregates thus formed photodissociable, a feature that may find important applications in bioengineering.

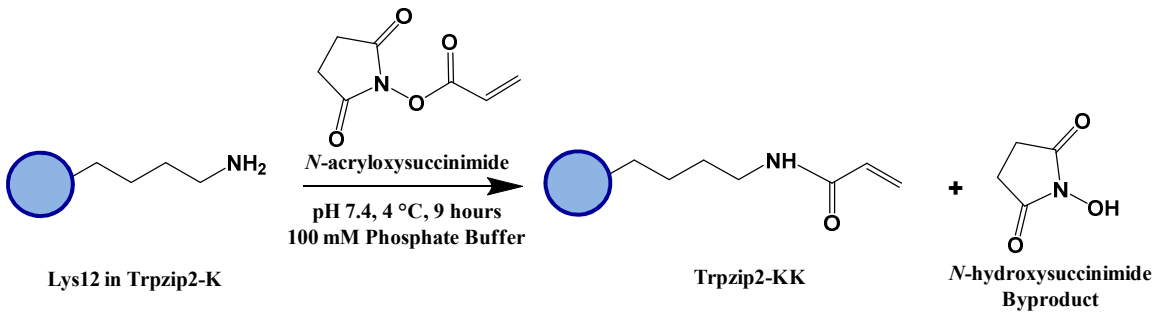
As shown (Figure 6.3, inset), at a concentration of 0.7 mM and pH 13, the aggregation process of Trpzip1, as judged by the intensity of the 1616 cm^{-1} band, is complete within the first 10 h of dissolution, which is faster than the aggregation rate of Trpzip1 at acidic pH. Similarly, the aggregation of Trpzip2 can also be induced by increasing the pH to 13 (Figure 6.5). Taken together, these results provide strong evidence supporting the gatekeeper role of Lys8 mentioned above. However, unlike that of Trpzip1, the aggregation process of Trpzip2 does not seem to be complete even after incubation for 6 days, indicating that there are other factors that also play a role in determining the aggregation kinetics.

It is well-documented that the β-hairpin structure in peptide fibrils tends to have extended and flat β-strands as opposed to the twisted conformations characteristically observed in native monomeric β-sheet proteins.[344-349] Thus, for the initially formed peptide dimer or oligomers consisting of native or native-like β-hairpins to further propagate to produce well-ordered and stacked parallel β-sheets, many native sidechain–sidechain interactions need to be broken to facilitate new intermolecular interactions, such as hydrogen bonding among neighboring monomers. In other words, the native β-hairpin must partially unfold to relax into a flat β-sheet unit, a structure required for the growth of the aggregate nucleus into long fibrils. Indeed, as shown (Figure 6.6), the CD spectrum of the aggregated Trpzip1 lacks the distinctive positive band at 227 nm observed for the folded Trpzip1,[50,214] indicating that the native edge-to-face Trp–Trp packing is disrupted upon the formation of aggregates.[350] This observation is consistent with an aggregation mechanism that requires flattening of the native β-hairpin structure upon incorporation of the peptide into a tightly packed and well-organized fibrillar matrix. While this study does not allow us to describe further structural details, it is reasonable to assume that the Trp residues play an important role in aggregate formation, presumably via non-native hydrophobic stacking. This requirement of native structural change or relaxation would argue that the higher the β-hairpin stability, the more difficult it becomes for the aggregate nucleus to propagate to form mature aggregates or fibrils. Thus, we attribute the slow aggregation growth rate of Trpzip2 at pH 13 to its high stability (T_m = 72 °C).

Because there is another Lys residue (i.e., Lys12) in the peptide sequence, the results obtained at pH 13 may not entirely reflect the effect of Lys8. Therefore, in the second study, we examined the aggregation kinetics of a Trpzip2 mutant wherein Lys8 is replaced with a photolabile lysine analogue, Lys(nvoc).[335-337] As shown (Figure 6.7), the amide I′ band of this mutant (Trpzip2-K) indicates that it aggregates quickly at very low peptide concentrations. Interestingly, the aggregates thus formed adopt an antiparallel β-sheet structure, as judged by the pair of bands centered at 1616 and 1685 cm^{-1}. Because both Trpzip1 and Trpzip2 form parallel β-sheet aggregates, these results not only support the notion that Lys8 is an effective aggregation gatekeeper in Trpzip2 but also indicate that this non-natural lysine residue, which is strongly hydrophobic,[351] can alter the aggregation pathway. In addition, the CD spectrum of the aggregated Trpzip2-K sample shows clearly the presence of a positive band at 228 nm (Figure 6.8), indicating that the native Trp–Trp packing is preserved to a certain extent in the aggregates. This is an interesting finding considering that the aggregates formed by the wild-type peptide do not support native Trp–Trp interactions and hence corroborates the aforementioned notion that Lys(nvoc), because of its higher hydrophobicity, can play a key role in determining the aggregation rate and pathway. This result further substantiates our initial hypothesis. Previously, the vertically pointing Lys8 sidechains served as an aggregation deterrent because of the repulsive electrostatic interaction, but in the mutant case, having Lys(nvoc) in this orientation provides another or possibly stronger avenue for aggregation-prone hydrophobes to associate. Moreover, as shown (Figure 6.9), upon removal of the nvoc group using light,[336,337] which converts Lys(nvoc) to Lys, the

aggregates formed by Trpzip2-K spontaneously disassemble to yield Trpzip2 monomers. Thus, this result provides additional evidence supporting the aggregation gatekeeper role of Lys8 in Trpzip2.

To further verify the notion that it is Lys8, not Ly12, that plays a key role in mediating the aggregation process, we conducted another photocleavage experiment on an aggregate sample formed by a Tripzip2-K derivative wherein the charged Lys12 sidechain was converted to a neutral species. Specifically, the amine group of Lys12 in Trpzip2-K was allowed to react with a common acylation agent, *N*-hydroxysuccinimide (NHS) ester, to form an amide bond (Scheme 6.1),[352,353] and the resulting peptide is termed Trpzip2-KK. Similar to Trpzip2-K, this peptide readily aggregates at low concentrations, as expected. As shown (Figure 6.10), however, when the native sidechain of Lys8 is recovered via photocleavage of the nvoc group, the aggregates formed by Trpzip2-KK disassemble. Because the resulting peptide contains only one charged Lys sidechain at position 8, this result thus substantiates our hypothesis that Lys12 does not play a significant role in preventing Trpzip hairpins from aggregating.



Scheme 6.1 Reaction scheme of acylation of the primary amine of Lys12 in Trpzip2-K. The resulting peptide is referred to as Trpzip2-KK.

Finally, to rule out the possibility that the weaker aggregation propensity of Trpzip2 is a direct outcome of its higher thermal stability, we tested a second variant in which Trp4 was mutated to Ala (hereafter termed Trpzip2-W4A). As expected (Figure 6.16), this mutation significantly decreases the thermal stability of the β-hairpin (T_m of ~23 °C). However, as indicated (Figure 6.11), Trpzip2-W4A (10 mM) does not show any detectable aggregation even after incubation for 5 days. Thus, this result further corroborates the proposed notion that the difference in the aggregation propensities of Trpzip1 and Trpzip2 stems from the difference in the Lys8 orientations and is not due to their difference in stability. Furthermore, this result suggests that any interactions between Lys8 and Trp4 are not critical in preventing the β-hairpin from aggregating.

It is well-known that a stronger turn-promoting sequence can increase the stability of β-hairpins. However, the effect of turn sequence on the aggregation propensity of β-hairpins has not been systematically examined. On the basis of results obtained from this study, we can begin to think of several possible scenarios. If aggregation proceeds from an unfolded conformation, and the native turn structure is not preserved in the aggregates, the effect of a specific turn sequence on aggregation would be directly correlated with its effect on the β-hairpin stability. On the other hand, if aggregation is initiated by association of folded or partially folded β-hairpins, then the effect of a specific turn sequence on aggregation becomes more subtle. Depending on how it directs the distribution of key charged residues, a turn sequence could prevent or retard β-hairpin aggregation by creating unfavorable intermolecular electrostatic interactions or facilitate aggregate formation by weakening any repulsive interactions. To this end, the

aggregation gatekeeper notion used in this study should not be simply interpreted as an independent action of one amino acid; rather, it should be discussed in the context of the underlying aggregation mechanism and aggregate structures.

Aggregate Stability and Morphology

Because both the strength of hydrophobic interactions and the unfolded population increase with an increase in temperature, many proteins and peptides show a stronger tendency to aggregate at higher temperatures.[354-357] Interestingly, temperature has the opposite effect on Trpzip1 aggregation. As shown (Figure 6.12), Trpzip1 aggregates readily dissociate at higher temperatures. On the other hand, the aggregates formed by Trpzip2-K do not show any detectable heat-induced dissociation (Figure 6.13). These results suggest that the aggregates formed by Trpzip1 are less stable and less rigid than those formed by Trpzip2-K, which is corroborated by AFM measurements. As shown (Figure 6.14), the AFM image of a Trpzip2-K aggregate sample shows a well-defined fibrillar network, with a homogeneous distribution of fibrils approximately 3.8 nm wide, consistent with previously engineered β-hairpin aggregates.[358] In comparison, the AFM image of a Trpzip1 aggregate sample reveals a more heterogeneous morphology, with the presence of variously sized fibrils and amorphous aggregates (Figure 6.15). Thus, taken together, the FTIR and AFM results indicate that the nvoc moiety in Lys(nvoc) not only significantly increases the aggregation rate of the peptide by eliminating the native aggregation gatekeeper in Trpzip2 but also guides the β-hairpins in the fibrils to stack in an antiparallel fashion. In other words, these results suggest that intermolecular Lys(nvoc) interactions can provide a strong driving force for peptide association and thus

a new avenue for fibril formation. We believe this is an important finding as it suggests that it is possible to use Lys(nvoc) to control the fibrillization rate, as well as the fibrillar architecture, of the peptide of interest. In addition, another advantage of using Lys(nvoc) is that it makes the fibrils photoresponsive, a feature that could be exceedingly useful in certain bioengineering applications.[112,137,359,360]

6.4 Conclusions

Protein and peptide aggregation can have dire biological consequences. For example, it may lead to degenerative diseases *in vivo* and dysfunction of protein and peptide therapeutics *in vitro*. Therefore, many studies have been conducted in the past, with the aim of understanding the important factors that control protein and peptide aggregation and devising strategies to prevent it from happening. Herein, we study the aggregation properties of two closely related β-hairpins, Trpzip1 and Trpzip2, seeking to gain further insight into the mechanism of this phenomenon. Despite the minor difference in their turn sequences (i.e., NG vs. GN), these two peptides exhibit totally different aggregation propensities; at acidic pH, Trpzip1 readily aggregates at micromolar concentrations, while under the same conditions, Trpzip2 does not show detectable aggregation even at concentrations of tens of millimolar. On the basis of the difference in their NMR structures and the fact that both peptides form aggregates rich in parallel β-sheets, we propose that (1) aggregation is initiated by association of two folded β-hairpins via edge-to-edge interactions and (2) Lys8 acts as an aggregation gatekeeper in both cases and its higher efficiency in preventing Trpzip2 from aggregating arises from the vertically

pointing sidechain rotamer preference. To test this hypothesis, we utilized two strategies, one by increasing the pH and the other by mutating Lys8 to a non-natural amino acid, Lys(nvoc), to examine how elimination of the positive charge on Lys8 affects the aggregation kinetics. We found that at pH 13 both Trpzip1 and Trpzip2 aggregate faster, which is consistent with the notion that Lys8 behaves as an aggregation gatekeeper. Further evidence in support of this hypothesis is that the Lys(nvoc) Trpzip2 mutant aggregates quickly, even at submillimolar concentrations, to form antiparallel amyloid-like fibrils that can be disassembled via photocleavage of the nvoc group. Moreover, our findings are consistent with an aggregation mechanism in which folded β-hairpins first associate to form a nucleus and the subsequent growth of this nucleus requires partial unfolding of the native structure. Finally, our results indicate that Lys(nvoc), because of its high hydrophobicity, can alter the aggregation mechanism and, hence, can be used to control, in conjunction with light, the morphology and structure of peptide fibrils.

Acknowledgements

We thank Yue Zhang, Ethan Glor, and Prof. Zahra Fakhraai at the University of Pennsylvania for assistance with AFM measurements. We also gratefully acknowledge the financial support from the National Institutes of Health (GM-065978). B.N.M is supported by an NIH Ruth Kirschstein National Research Service Award Predoctoral Fellowship (F31AG046010).

Peptide	Sequence	T_m (°C)
Trpzip1	SWTWE**GN**KWTWK	49.8 ± 0.3[a]
Trpzip2	SWTWE**NG**KWTWK	71.9 ± 0.1[a]
Trpzip2-K	SWTWE**NG**(K*)WTWK	56.8 ± 0.5[b]
Trpzip2-W4A	SWTAE**NG**KWTWK	23.7 ± 2.1[b]
Trpzip2-KK	SWTWE**NG**(K*)WTW(K‡)	-

Table 6.1 Sequence and thermal melting temperatures (T_m) of the peptides studied.

a. From Cochran *et al*.[212]

b. Figure 6.16

K* represents Lysine-dimethoxy-2-nitrobenzyloxycarbonyl

K‡ represents the acylated Lysine (Scheme 6.1)

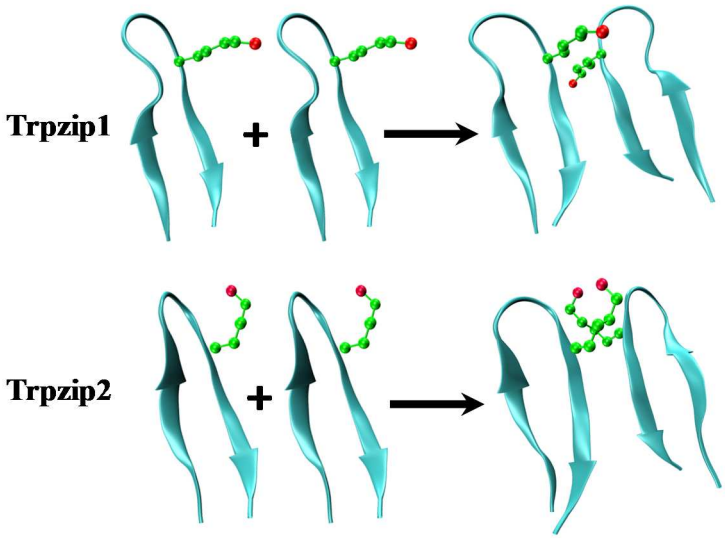


Figure 6.1 NMR structures of Trpzip1 (PDB: 1LE0) and Trpzip2 (PDB: 1LE1), as indicated, and the proposed dimerization scheme, showing the difference in the orientations of the Lys8 sidechains.

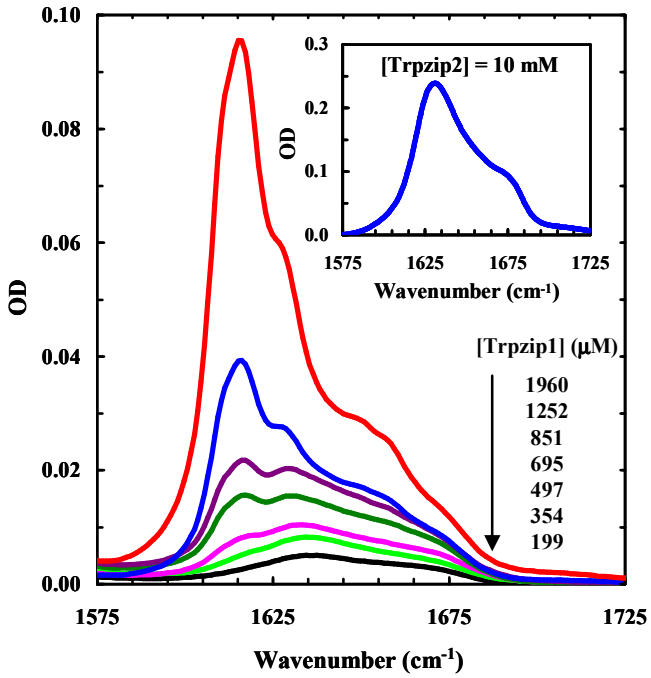


Figure 6.2 Amide I′ spectra of Trpzip1 at different concentrations, as indicated**.** Shown in the **i**nset is the amide I′ band of Trpzip2 at 10 mM. These data were collected after the peptide samples had been incubated for 24 hours at 25 ^{o}C.

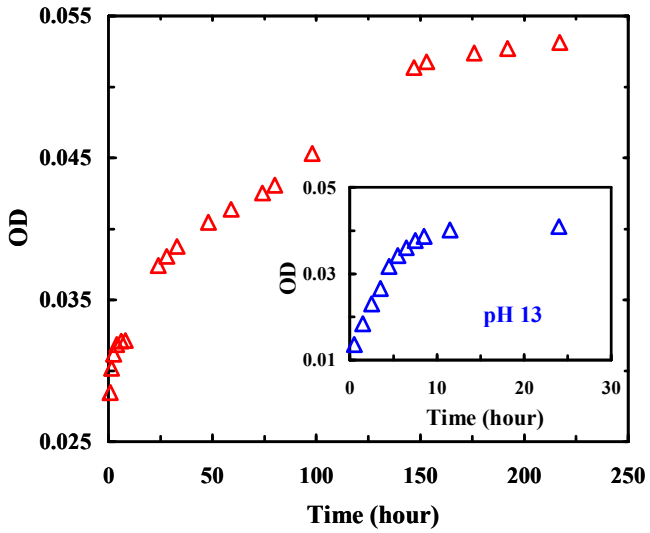


Figure 6.3 Intensity of the 1616 cm^{-1} band of Trpzip1 (1.2 mM, pH 3) as a function of incubation time, showing the aggregation kinetics of this peptide at acidic pH**.** For comparison, the aggregation data of Trpzip1 (0.7 mM) obtained at pH 13 are shown in the inset. The corresponding FTIR spectra are presented in Figure 6.17.

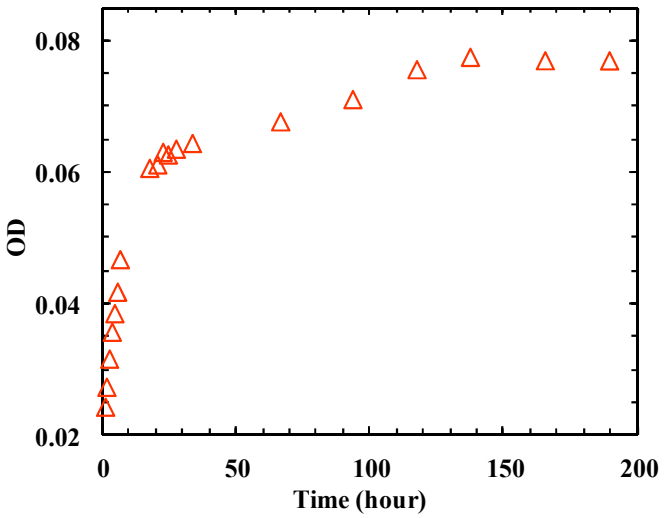


Figure 6.4 Absorbance of Trpzip1 (2 mM, pH 3) at 1616 cm^{-1} versus incubation time, showing its aggregation kinetics.

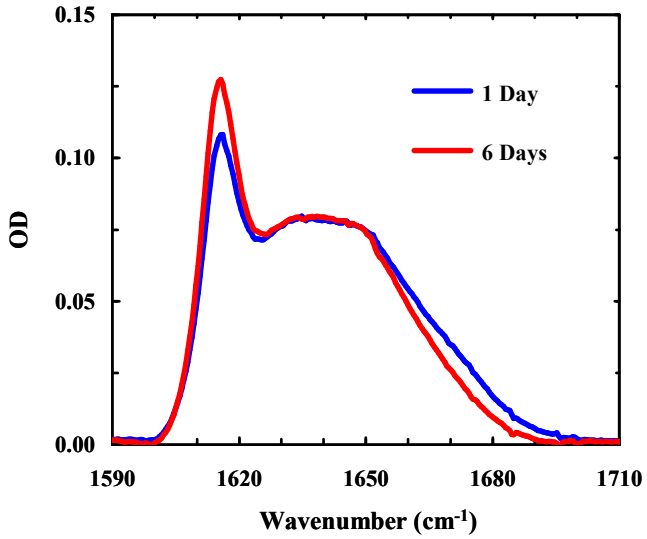


Figure 6.5 Amide I′ spectra of Trpzip2 (~6 mM) at pH 13 measured after the peptide sample had been incubated for 1 day and 6 days, as indicated.

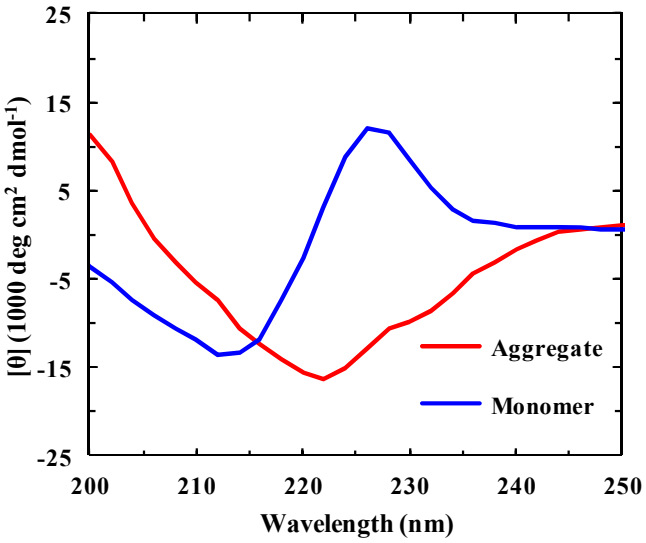


Figure 6.6 CD spectra of Trpzip1 monomer and aggregates, as indicated.

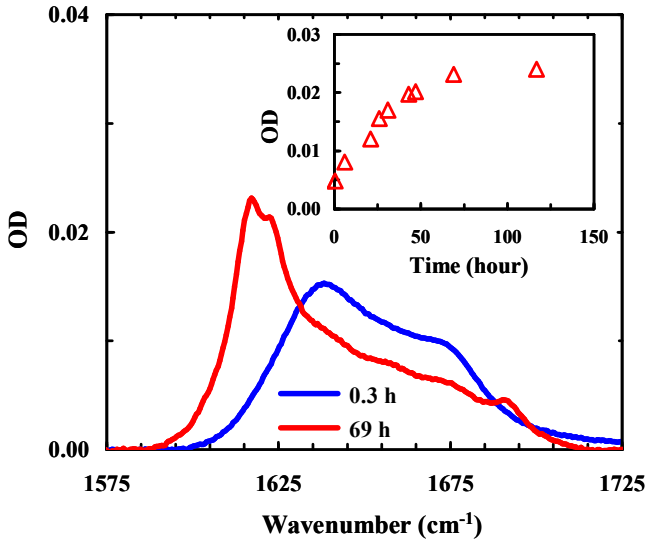


Figure 6.7 Representative amide I′ spectra of Trpzip2-K (0.5 mM, pH 3) obtained after different sample incubation times, as indicated. Shown in the inset is the intensity of the 1616 cm^{-1} band as a function of incubation time. The band intensities were obtained from the FTIR spectra shown in Figure 6.18.

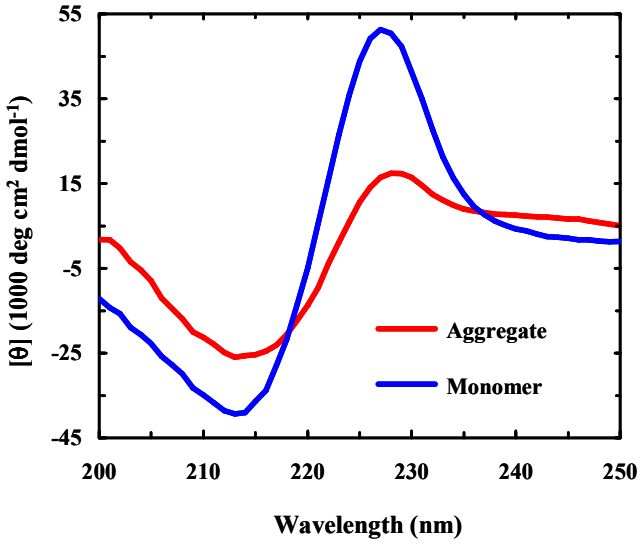


Figure 6.8 CD spectra of monomeric and aggregated Trpzip2-K samples (40 μM, pH 3), as indicated. The aggregated sample was prepared by diluting a more concentrated peptide sample (0.5 mM) that had been incubated for 14 days to allow aggregate formation.

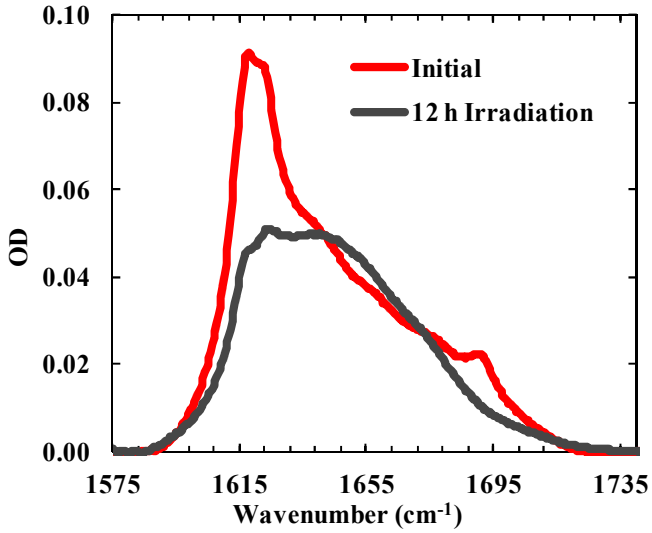


Figure 6.9 FTIR spectrum of an aggregated Trpzip2-K sample (3.7 mM in D_2O) in the amide I′ region (red). Upon irradiation of this sample with 355 nm light for approximately 12 hours, the spectrum (grey) shows significant changes from its initial shape, indicating that light excitation can induce disassembly of the aggregates.

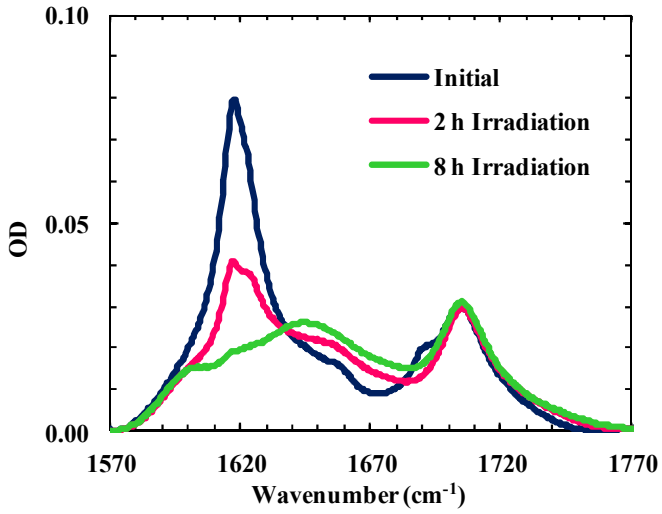


Figure 6.10 Amide I′ bands of an aggregated Trpzip2-KK sample (2.2 mM, D_2O) obtained under different conditions, as indicated. These spectra show that photocleavage of the nvoc group on Lys8 results in aggregate disassembly. The band located near 1700 cm^{-1} arises from the C=O stretching vibration of the Lys12 sidechain acrylamide.

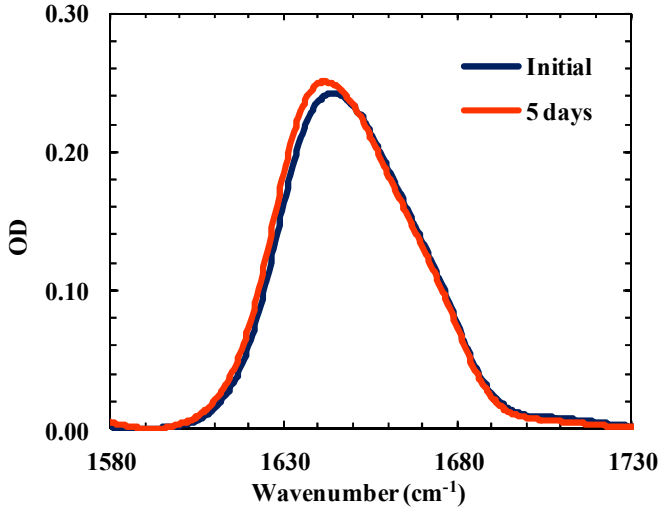


Figure 6.11 Amide I′ bands of Trpzip2-W4A (10 mM, pH 3) obtained at two incubation times, as indicated.

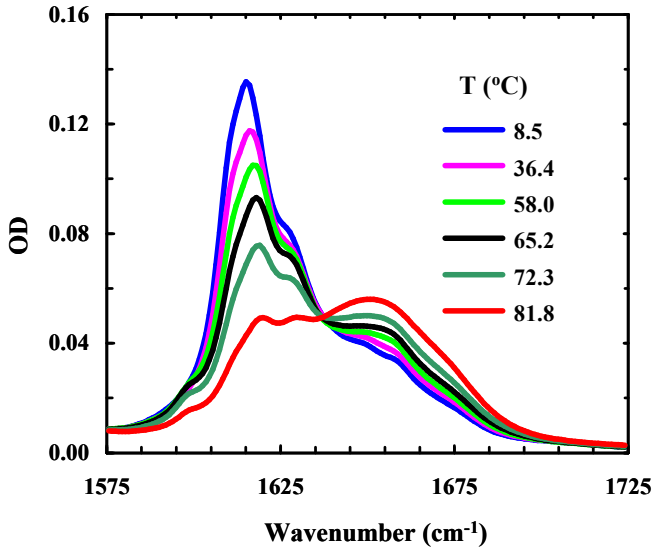


Figure 6.12 Amide I′ band of Trpzip1 (2.4 mM, pH 3) as a function of temperature.

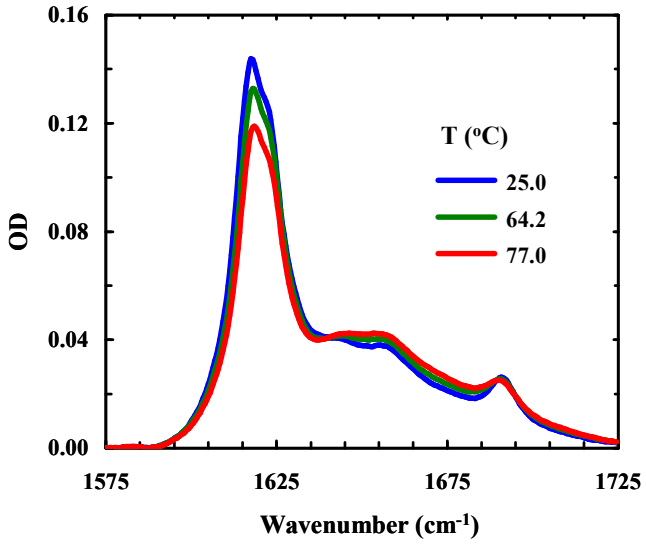


Figure 6.13 Amide I′ band of Trpzip2-K (4 mM, pH 3) as a function of temperature.

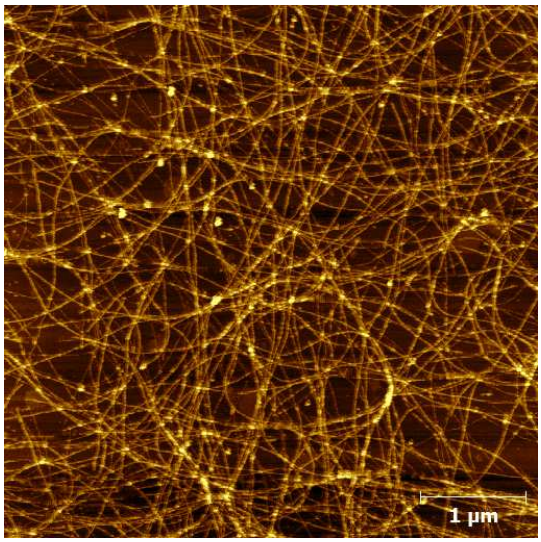


Figure 6.14 Representative AFM image of the peptide fibrils formed by Trpzip2-K after an incubation period of 14 days.

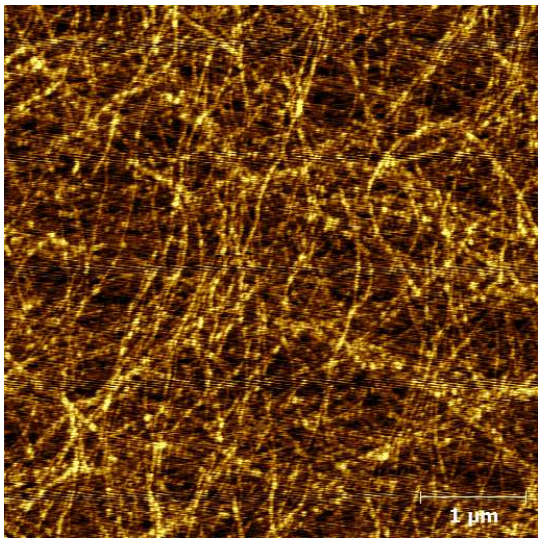


Figure 6.15 Representative AFM image of the peptide fibrils and aggregates formed by Trpzip1 after an incubation period of 14 days.

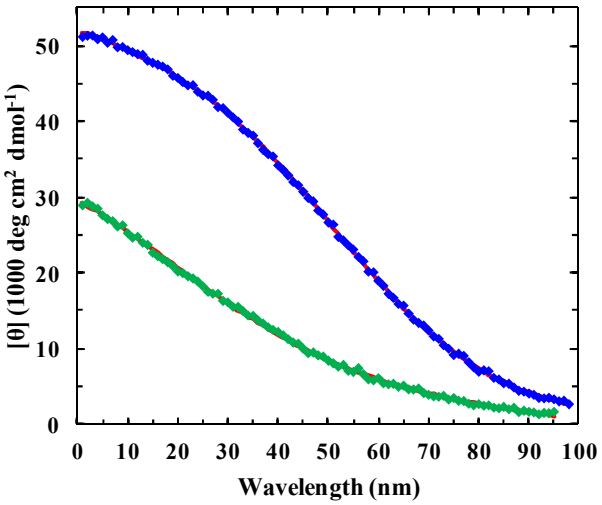


Figure 6.16 CD *T*-melt data of Trpzip2-K (blue) and Trpzip2-W4A (green) monitored at 227 nm. The solid line (red) is the best fit of these data to a two state model.

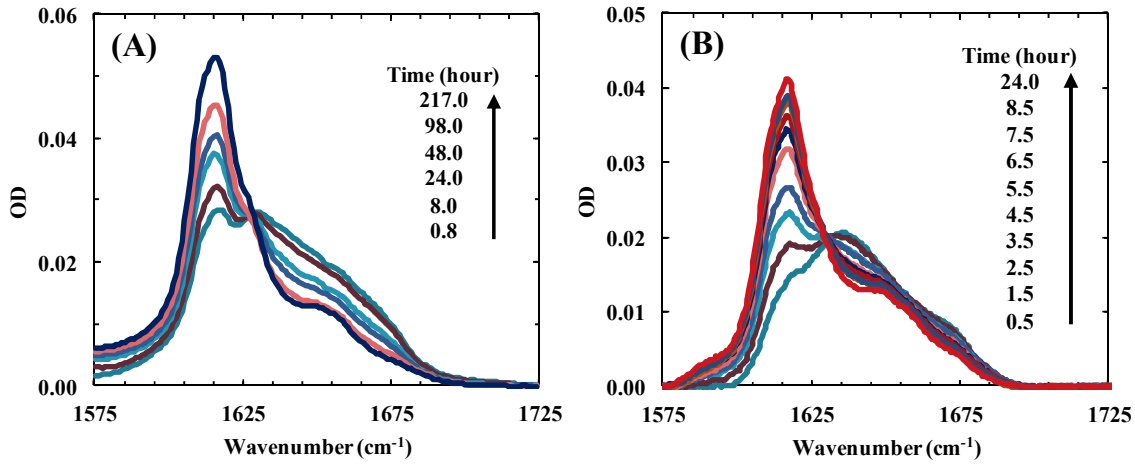


Figure 6.17 FTIR spectra of Trpzip1 in the amide I′ region obtained at different incubation times, as indicated, and solution conditions: (A) 1.2 mM peptide and pH 3 and (B) the 0.7 mM peptide and pH 13.

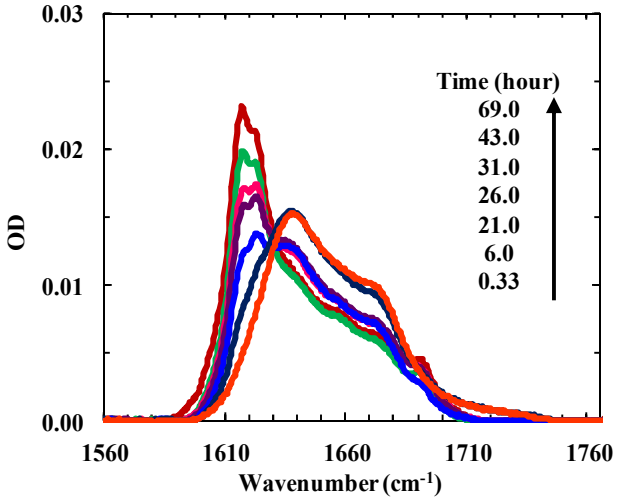


Figure 6.18 FTIR spectra of Trpzip2-K (0.5 mM, pH 3) in the amide I′ region obtained at different incubation times, as indicated.

7 Photo-Initiated Disassembly of Fibrils: Applications to Amyloids and Peptide Hydrogels

Portions of this chapter have been adapted with permission from Chemical Physics Letters, Thomas J. Measey,‡ Beatrice N. Markiewicz,‡, and Feng Gai, (2013) 580, 135-140. DOI: 10.1016/j.cplett.2013.06.055.[337] (‡ Denotes equal authorship)

7.1 Introduction

The phenomenon of protein self-assembly, under either native or nonnative conditions, has gained a great deal of attention given its connection to several pathological conditions, such as Alzheimer's, Parkinson's, and Type II Diabetes.[25,361] The onset of these conditions is correlated with the accumulation of insoluble deposits (or amyloids) that adopt a characteristic fibrillar, cross-β-sheet morphology.[25,132] At the same time, the design of protein/peptide self-assemblies into such structures is also actively pursued to generate biological architectures and scaffolds, i.e., peptide hydrogels, that are posited to yield many new applications, ranging from regenerative medicine to the delivery of therapeutics.[362] Earlier studies support the notion that within these structures, hydrophobic sidechain interactions play a key role in mediating self-assembly and stabilization of β-sheet arrangements, emphasizing the importance between polypeptide sequence and fibril forming propensity.[26,132,316,345,363,364] Therefore, a major motivation in amyloid research is to develop suitable inhibitors using small molecules[365-368] and peptide fragments derived from parent sequences[28,369,370] to interfere with these stabilizing non-covalent interactions. Similarly, the exploration of peptide hydrogels relies on the rational design of sequences that promote favorable hydrophobic contacts to facilitate fibril assembly.[362] However, currently there is a lack of experimental techniques that are

capable of disassembling mature protein/peptide aggregates or fibrils, making the reversal of protein aggregation processes difficult. The key hypothesis of this study is founded on the idea that by introducing a moiety, which can produce a light-activated charge, into the fibril core, we can significantly weaken the hydrophobic interactions crucial to the assembly of aggregates, fibrils, or hydrogel matrices, and may be able to disrupt these higher order structures. To test this hypothesis, we used a well characterized photolabile non-natural amino acid, Lysine-4,5-dimethoxy-2-nitrobenzyloxycarbonyl (Lys(nvoc)), and several disease related and hydrogel forming peptide systems. We believe this method provides a novel means to manipulate the conformations of otherwise irreversible structures, potentially rendering it useful for a broad range of biological applications ranging from therapeutic research to the *de novo* design of bionanomaterials.

The use of light-activated triggers to precisely tune biological processes has been broadly applied, for example in regulating gene expression, neuronal activity, protein activation, and peptide folding/unfolding.[112,371,372] Specifically, "caged" (or light controllable) compounds, usually involving an aromatic moiety, have been site-specifically introduced into amino acid sidechains, mRNA phosphate backbones, and hormones in efforts to activate biological activity upon light exposure.[112,373-375] For protein systems, photocages and photo-linkers have been used to disrupt coiled-coil conformations in short peptides,[376] and to trigger aggregation of Amyloid-β (Aβ)[131] and other β-sheet forming peptides.[129,137]

Herein, we use a lysine analog, Lys(nvoc), which has been shown to produce a native lysine sidechain upon light activation in a biologically safe spectral region (Figure

7.1).[335] Prior studies on larger proteins have suggested that a charged lysine buried within an otherwise hydrophobic core stimulates an energetic process involving both local unfolding and increased hydration, which supports the expected outcome with smaller peptides via our proposed method.[377,378] The Lys(nvoc) cage however, provides light controlled versatility between a hydrophobic (prior to photocleavage) and a hydrophilic (after photocleavage) sidechain. Due to the key role that lysine plays in many biological interactions and processes, Lys(nvoc) has been used to regulate protein functionality,[335,379] and has been successfully integrated into cellular environments.[380-382] Taken together, these examples demonstrate the applicability of Lys(nvoc) in the experiments presented here while also revealing the potential viability of incorporating Lys(nvoc) into living cells, furthering its biological value.

We employ two well studied model amyloid systems, i.e., those associated with Alzheimer's disease (Aβ peptide) and Type II Diabetes (hIAPP peptide), to test the feasibility of the proposed method. In particular, we use a bottom-up approach, working with truncated versions of these peptides identified to be amyloidogenic regions critical to the self-assembly of fibrils.[28] Both Aβ and hIAPP contain small amyloidogenic fragments, or in other words, segments that form fibrils similar in morphology to the full length peptides, namely, $A\beta_{16-22}$ and $hIAPP_{22-27}$. Moreover, these peptide fragments contain a phenylalanine (Phe) residue shown to affect the integrity of higher order structures and kinetics of self-assembly, highlighting the importance of hydrophobic sidechain interactions in amyloid formation.[27,29,383-385] A previous study in the Gai lab confirmed that this Phe19 to Lys(nvoc) mutation in $A\beta_{16-22}$ is conservative in that it

preserves the intrinsic aggregation propensity, but it also showed that the fibrils formed by this mutant can be disassembled by light.[336] Based on this notion, we demonstrate that introducing Lys(nvoc) in the place of Phe residues in $hIAPP_{22-27}$ also conserves the aggregation propensity of this peptide due to the aromatic nature of the photocage, while photocleavage produces a charge at these positions to disassemble fibrils. Later, we show that the same disassembly effect can be accomplished in the full length $A\beta_{1-42}$ peptide. In addition, to expand its biological applicability, we were able to disassemble fibrils formed by the co-aggregation of the photocaged mutant with the native structures of the wild-type peptide.

Moreover, we test the proposed method on peptide hydrogels, which are structurally and characteristically similar to amyloids, and form fibril networks that are extremely hydrated, emphasizing their potential use in biological applications.[362] Pioneering work by Zhang and coworkers led to the discovery of several ionic-complementary peptides, comprised of alternating polar and nonpolar residues, which arrange into layers of hydrophobic and hydrophilic surfaces, thus undergoing optimal electrostatic interactions to facilitate self-assembly into highly ordered fibrillar networks.[362,386] As a result, we test our method on a model hydrogel system (sequence: FKFEFKFE, Figure 7.2),[387] to investigate whether incorporation of Lys(nvoc) into key hydrophobic positions (i.e., Phe to Lys(nvoc) mutations) can trigger the disassembly of the hydrogel architecture in a spatiotemporal manner.

7.2 Experimental Section

Materials

D_2O (D, 99.96 %) and deuterium chloride (D, 99.5 %) were purchased from Cambridge Isotope Laboratories (Andover, MA). Fmoc–Lys(4,5–dimethoxy–2–nitro–benzyloxycarbonyl)–OH (Fmoc–Lys(Nvoc)–OH) was purchased from Anaspec, Inc. (Fremont, CA), and used without further purification. Fmoc-protected amino acids were purchased from Advanced Chem Tech (Louiseville, KY). Rink amide resin (Subs. = 0.28) was purchased from Novabiochem (San Diego, CA). All peptides were synthesized using standard fluorenylmethoxycarbonyl (FMOC) chemistry protocol on a PS3 peptide synthesizer (Protein Technologies, Tuscon, AZ), and purified by reverse-phase HPLC (1100 Series; Agilent Technologies, Santa Clara, CA). The identity of the peptides were verified using MALDI-TOF mass spectrometry (Voyager-DE RP, Applied Biosystems, Foster City, CA).

Sample Preparation

Peptides were first dissolved in a 50/50 mixture of 0.1 M deuterium chloride (DCl) in D_2O/acetonitrile, and the resulting solution was lyophilized overnight, allowing the removal of residual trifluoroacetic acid (TFA), which absorbs in the amide I' region of the peptide IR spectra. The amyloid peptide samples (Aβ and hIAPP) were prepared by directly dissolving the lyophilized peptide in 20 mM phosphate buffer D_2O solution (pH 7.0). The samples were incubated for at least 24 hours at room temperature or until complete fibrillization was achieved. The hydrogel peptide samples used in the subsequent experiments were prepared by directly dissolving the lyophilized peptide in a

D_2O solution with the targeted NaCl concentration (100 mM) and sonicated (20 – 60 minutes) to obtain a homogeneous solution. Samples were left to rest overnight to ensure complete hydrogelation. The pH of all samples was between 2 and 3 before gelation. Similarly, the peptide concentration of all samples was measured prior to inducing aggregation/gelation and prior to the addition of NaCl in the gels. Specifically, the concentration of the caged-modified peptides were determined with the absorbance of the nvoc moiety at 350 nm, using an extinction coefficient of 5485 M^{-1} cm^{-1} for a single mutation and 10,970 M^{-1} cm^{-1} for a double mutation.[336] The concentration of the wild-type hydrogel peptide (sequence: FKFEFKFE) was determined via the absorbance of phenylalanine residues at 257.5 nm.[388] Furthermore, the concentration of the wild-type $hIAPP_{22-27}$ peptide was determined using the backbone absorption at 214 nm and an extinction coefficient of 10,094 M^{-1} cm^{-1}.[389]

FTIR Measurements

FTIR spectra were collected at 25 °C on a Nicolet Magna-IR 860 spectrometer at a resolution of 1 cm^{-1}. The sample was injected into an assembled CaF_2 cell that was divided into two compartments (one for the reference, and one for the sample), using a homemade Teflon spacer. This enabled the collection of the sample and reference spectra under identical conditions. An automated translational stage was used to cycle the sample and reference sides into and out of the IR beam, to correct for the slow instrument drift. For each cycle, 8 single-beam spectra were collected on each side, and the final spectra correspond to the average of 32 such cycles. The pathlength of the assembled cell

was 52 μm determined via the interference fringes of the empty cell. The spectra were baseline-corrected using instrumental software (OMNIC 6.1a).

AFM Measurements

The majority of the AFM experiments were performed in air at room temperature, using a multimode atomic force microscope (Nanoscope IIIa; Digital Instruments, Santa Barbara, CA), equipped with an E-type piezoscanner. The AFM images obtained for two hydrogel peptides (i.e., KFE17 and KFE4) were performed in the TappingMode (Air) setting using a Bruker Dimension Icon AFM (Santa Barbara, CA). 5 μL of sample solution was applied to a freshly-cleaved mica surface for ~5 seconds, rinsed with 300 μL millipore water, and subsequently dried with a slow stream of N_2 gas. Tapping-mode imaging was carried out with a silicon probe (TESP) from Veeco (Camarillo, CA). Height and deflection images were obtained with a scan rate of ~1 Hz, integral gain of 0.4, and a proportional gain of 0.6. Multiple images were obtained for each sample at different locations on the mica substrate, to confirm either the presence or absence of fibrils.

Photocleavage Measurements

Irradiation of samples was carried out by placing the sample in the optical path of a FluoroLog fluorometer (HORIBA Jobin Yvon), at room temperature. The excitation wavelength was set to 350 nm with a band width of 30 nm (i.e., the excitation slit width was set to 0.75 cm). The intensity of the excitation light at the sample was estimated to be about 8.8 mW cm^{-2}. To allow for a direct comparison of FTIR spectra and AFM images of the irradiated sample, both the filled FTIR cell and an aliquot of the stock

solution (for AFM) were placed together in the optical path of the 350 nm light, and irradiated for a given amount of time.

Rheological Measurements

Rheological measurements were performed on a TA Instruments RFS II rheometer at 25 °C using an 8 mm parallel plate and a gap size of 250 μm. Approximately 24 μL of the aged hydrogel sample (10 mM peptide in 100 mM NaCl D_2O solution) was placed on the stage. Three frequency sweeps were performed and averaged for each sample after equilibration on the stage.

7.3 Results and Discussion

7.3.1 Disassembly of Amyloid Fibrils

To initially test the feasibility and robustness of this method, we employed the Lys(nvoc) modifications within short amyloidogenic fragments derived from the full length Aβ and hIAPP peptides. A previous study carried out by our lab has shown that fibrils formed by an $A\beta_{16\text{-}22}$ (sequence: ^{16}KLVFFAE^{22}) mutant, which contains Lys(nvoc) at the native Phe19 position (referred to hereafter as Aβ-F19C), can be sufficiently disassembled via photocleavage at 350 nm. This mutation was initially designed based on evidence demonstrating that Phe19 in both $A\beta_{16\text{-}22}$ and the full length Aβ peptide is critical to fibril formation.[29] To verify whether this method can be universally applied to other model systems, we chose the $hIAPP_{22\text{-}27}$ fragment (sequence: ^{22}NFGAIL^{27}), which is the shortest hIAPP peptide fragment that exhibits similar aggregation propensity and toxicity as the full length hIAPP chain.[383] Because the self-assembly of $hIAPP_{22\text{-}27}$ heavily

depends on Phe23,[27] we mutated this position to Lys(nvoc) (hereafter referred to as hI-F23C). As shown in Figure 7.3a, the FTIR spectra of hI-F23C demonstrates that after an incubation period of ~24 hours, this peptide forms antiparallel β-sheet aggregates as indicated by the high intensity band at ~1620 cm^{-1} and weak band at ~1680 cm^{-1}.[70] More importantly, the corresponding AFM image in Figure 7.3b reveals that hI-F32C self-assembles into aggregates that form relatively small homogeneously sized fibrils. After a 2.5 hour irradiation period at 350 nm to remove the nvoc moiety, a decrease in intensity of the ~1620 cm^{-1} and ~1680 cm^{-1} transition and concurrent increase in the disordered band intensity at 1650 cm^{-1}, signifies the disappearance of β-sheet secondary structure. Likewise, the disassembly of the fibrils was confirmed by the corresponding AFM image (Figure 7.3b).

Based on the two model peptides tested thus far, Aβ-F19C previously[336] and hI-F23C herein, we have shown that their Phe positions serve as an aggregation "hot spot", or in other words, a point of intermolecular contact along the self-assembly process. It would be ideal to be able to incorporate the photocaged mutants into these positions in the native structures to render the disassembly of fibrils formed by the wild-type sequences. To test this capability which could potentially expand its biological usefulness, we attempted to co-aggregate both hI-F23C and Aβ-F19C with their wild-type counterparts (i.e., $hIAPP_{22-27}$ and $A\beta_{16-22}$) so that the Lys(nvoc) moiety can sequester within the highly hydrophobic region and promote charge-induced disassembly. Two different approaches were used. The first involves co-aggregation of the monomers of the photocaged mutant and the wild-type sequence, which was accomplished by initially

dissolving a mixture of the two peptides and subsequently allowing aggregation between the two structures. For example, as shown in Figure 7.4, when monomers of hI-F23C were mixed and incubated with monomers of wild-type $hIAPP_{22-27}$ in a 1:1 ratio, the initial FTIR spectral signature implies that it self-assembles into well formed β-sheets. Upon irradiation of the sample at time intervals of 1, 3, and 12 hours, the previously formed aggregate structures have been abolished via Lys(nvoc) cleavage. This evidence suggests that within this model system, the Lys(nvoc) sidechain of the hI-F23C mutant is amenable to integration with hydrophobic regions of the native structures, and also, that the conversion to a charged Lys via photocleavage serves as an avenue for destabilization of the high order structures.

Furthermore, the second experimental approach to demonstrate this idea was designed based on the notion that amyloid fibril conformations continue to be structurally dynamic even after equilibrium is reached. Previously, Carulla and coworkers performed hydrogen exchange measurements of fibrils that suggested there is a 'molecular recycling' mechanism, where monomers within solution occasionally replace monomers in fibrils.[390] More recently, a 2D-IR study on $hIAPP_{1-37}$ revealed that there are significant structural changes that occur even after the fibrils appear to be in equilibrium.[391] Computational studies have also shown that monomers in solution can undergo a 'dock and lock' mechanism to integrate into the underlying structure of preformed fibrils.[392-394] Based on these proposed mechanisms, we tested whether monomers of the photocaged variants can also potentially exchange with monomers in mature fibrils formed by the wild-type. Wild-type $Aβ_{16-22}$ was incubated for 2 weeks to grow mature fibril structures.

Afterwards, monomeric Aβ-F19C was added to the solution at an equal molar ratio. The solution was left to equilibrate for approximately 1 month and FTIR spectra were taken before and after irradiation with 350 nm light. As shown in Figure 7.5, the spectrum before irradiation shows signatures (at ~1620 cm^{-1} and 1690 cm^{-1}) typically observed in antiparallel β-sheet aggregates. After 2 hours and 12 hours of light irradiation, there is a visible decrease in both the bands at 1620 cm^{-1} and 1690 cm^{-1}, and an increase at 1650 cm^{-1}, which suggests a decrease in β-sheet structure with a concomitant increase in disordered conformations. Evidence confirming that Aβ-F19C incorporates into the mature fibrils lies in the change of the FTIR weak β-sheet band. Previously, it has been shown that the high frequency band is shifted to ~1705 cm^{-1} in the FTIR spectrum of Aβ-F19C aggregates/fibrils.[336] The high frequency band in the co-aggregated sample remains at 1690 cm^{-1} with no increase in intensity at ~1705 cm^{-1}, suggesting that Aβ-F19C adapts to the anti-parallel β-sheet structures of the $A\beta_{16-22}$ wild-type. Taken together, these results suggest that in both co-aggregation scenarios presented herein, the monomers of the Lys(nvoc) peptide variants are capable of interacting with the native fibril structures, and are efficient to induce light-activated disassembly.

Founded on the ideas presented above, experiments were conducted on the biologically relevant full length $A\beta_{1-42}$ peptide (sequence: DAEFRHDSGY-EVHHQKLVFF-AEDVGSNKGA-IIGLMVGGVV-IA). As was shown previously, the Phe19 position in the short model peptide played a key role in the aggregation/disaggregation process. Therefore, a single Lys(nvoc) mutation was incorporated at the Phe19 position in $A\beta_{1-42}$ (hereafter referred to as $A\beta_{42}$-F19C). A

sample of $A\beta_{42}$-F19C was incubated for 2 weeks and shows characteristic parallel β-sheet structure confirmed by a high intensity band at 1620 cm^{-1} (Figure 7.6). After a 30 minute irradiation period, there is a complete disappearance of the β-sheet structure. This reiterates again that not only is Phe19 especially important for maintaining the highly ordered fibril structure, but the light-induced transition to a charged sidechain can destabilize fibrils formed by even the longer sequences.

Overall, the studies presented here provide a proof of principal concept that outlines the fundamental chemical interactions required to reverse/disassemble highly ordered protein structures from within the core of the fibril. First, we broadly show that when Lys(nvoc) replaces a hydrophobic residue critical to the aggregation of a sequence, it maintains its intrinsic self-assembly pathway, while photocleavage of the nvoc moiety results in disaggregation. This has been effective both within small amyloid sequences as well as the large full length $A\beta_{1-42}$ peptide. More interestingly, the Lys(nvoc) containing peptides can be incorporated and/or co-aggregated with the native sequences to also facilitate disassembly. Note, we do not expect this method to evolve into a therapeutic route, however, the goal of these investigations is to provoke new studies that can implement analogous chemical interactions within these structures using *in vivo* approaches.

7.3.2 Photo-induced Disassembly of Peptide Hydrogels

Much effort has been devoted to the generation of 'smart' hydrogels, in which their structures are responsive to an external stimulus, such as a change in pH,[395] temperature,[396] salt concentration,[397] or light.[136,137,398] Controlling such structures via light

offers the advantage of spatial and temporal tunability. To construct a photoresponsive peptide hydrogel, a phototrigger needs to be incorporated into the peptide of interest. In addition, this phototrigger must meet the following requirements to be effective: (1) it does not alter the peptide's ability to form hydrogels, (2) upon absorbing a photon it induces disassembly of the β-sheet structures underlying the hydrogel's fibrillar networks, and (3) it should be biocompatible. Based on our work on the light triggered disassembly of peptide fibrils[336] and the study of Lauffenburger and coworkers[399] on peptide hydrogels, herein we employed Lys(Nvoc) as the phototrigger and the KFE peptide as a model system (Figure 7.2 and Table 7.1). Lauffenburger and coworkers[399,400] have shown that at peptide concentrations between 1-10 mM and salt concentrations >5 mM, the KFE peptide readily self-assembles into a macroscopic hydrogel with a β-sheet-rich fibril network. More importantly, the amphiphilic nature of the KFE peptide (refer to Figure 7.2) promotes formation of β-sheets with two layers, namely a hydrophilic layer where negatively and positively charged sidechains of adjacent peptides pack together via ionic interactions, and a hydrophobic layer that is stabilized by non-covalent interactions between adjacent non-polar sidechains. This pattern of alternating hydrophobic and electrostatic interactions provides an avenue for us to use light and Lys(Nvoc), which results in a charged lysine upon excitation with a near-UV photon, to modulate these stabilizing forces and, hence, disrupt the structural integrity of the target peptide hydrogel. As listed in Table 7.1, three variants were synthesized that contained Phe to Lys(nvoc) mutations. The chosen sequences were designed to test (1) whether incorporation of two Lys(nvoc) mutations into the

hydrophobic layer would allow self-assembly of similar β-sheet hydrogel architectures as the wild-type, (2) if photocleavage of the aromatic nvoc moiety results in a fast-acting degradation process, (3) the effect of the Lys(nvoc) positions (i.e., the termini versus the central residues) on the viscoelastic and/or structural properties, and (4) the effect of sequence length on the hydrogel structure.

Similar to the wild-type KFE peptide, we found that in the presence of ~100 mM NaCl and at peptide concentrations ranging from 5-10 mM, the KFE35 and KFE17 peptide, both containing two Lys(nvoc) mutations but in different positions, forms a self-supporting macroscopic hydrogel within a few hours, as shown in Figure 7.7. As these hydrogel matrices were subjected to irradiation at 350 nm, a transition to a non-viscous solution was observed (Figure 7.7), confirming the fact that peptides KFE35 and KFE17 can undergo a light-activated structural transformation. As for the shorter fragment KFE4, we did not observe significant changes in the solution viscosity to start.

Furthermore, we employed FTIR spectroscopic measurements to assess the secondary structure of the hydrogels, and monitor the light-induced structural transitions. In agreement with previous literature,[364] the amide I' band of the KFE peptide hydrogel is characteristic of a predominant population of antiparallel β-sheet structures, as evidenced by two sharp bands centered at 1620 and 1685 cm^{-1} (Figure 7.8a). Similarly, the initial FTIR spectra (Figure 7.8b) of the KFE35 hydrogel before irradiation also displays sharp bands in the same region, with a slight enhancement in the high frequency IR transition. While a mechanistic interpretation is difficult without further investigation, we believe the augmentation arises from exciton couplings involving the carbonyl group of

Lys(nvoc) and the peptide amide groups. Likewise, the KFE17 hydrogel structure also adopts an antiparallel β-sheet alignment similar to the wild-type (Figure 7.8c). On the contrary, KFE4 yielded IR spectra (Figure 7.8d) that were vastly different from the other three structures. The appearance of the band at 1625 cm^{-1} indicates that there is in fact β-sheet content present, however, the emergence of several high frequency transitions most likely reveals that there is a heterogeneous distribution of conformations. More importantly, as indicated in the red FTIR traces (Figure 7.8), irradiation of the hydrogel samples result in the complete disassembly of the underlying β-sheet structures within 30-90 minutes, as evidenced by the formation of a broad amide I' band centered at ~1650 cm^{-1}. As a control, the wild-type KFE peptide (i.e., the peptide without the photolabile modification) was irradiated for an equal amount of time, and no structural change was observed. Overall, these results show that the Lys(nvoc) modified peptides are capable of organizing into β-sheets that can be controllably degraded via light.

To verify the presence of fibrils and explore the morphology of the self-assembled hydrogels, we performed AFM measurements on all three photoactive peptides. A combined AFM and molecular dynamics study published by Marini *et al.*[400] on the structural characterization of KFE determined that the initial architectures are comprised of β-sheet helical ribbons which then transition into fibers. According to the AFM images in Figure 7.9, the KFE35 and KFE17 peptide hydrogels also organize into a relatively homogenous fibrillar network, as seen previously for KFE.[400] Moreover, a closer observation of the images revealed several helical ribbons buried underneath the mature fibrils. This was more obvious in the KFE17 image, but it can also be observed for

KFE35 even though the image window is much larger. Specifically, since the β-sheet helical ribbons were characterized as intermediates leading to fibril formation in the original study of KFE,[400] this suggests that even with two Lys(nvoc) mutations the peptides are able to adopt a similar morphology and maintain a similar self-assembly pathway as the wild-type. Unfortunately, the same cannot be said for the KFE4 peptide. The AFM images in Figure 7.9c indicate that there is a mixture of conformations composed of spherical aggregates and fibers, which is consistent with its FTIR spectra showing several high frequency β-sheet bands. Furthermore, even the fibrils thus formed show a dramatically different profile, mainly in that it contains a pitch that is not evident in the other longer peptides. As expected, the inability of KFE4 to transform into a viscous gel-like solution stems from the absence of a well-ordered and homogenous β-sheet network. Our motivation to design this shorter sequence was an attempt at a minimalist forming gel that has the added advantage of photodegradability. Unfortunately, shortening the sequence in this case diminishes the number of intermolecular residue contacts, thus resulting in non-uniform structures.

In addition to specific structural requirements, a hydrogel material must possess distinctive mechanical and viscoelastic properties. For oscillatory shear rheometric measurements of hydrogels, the frequency dependence of the storage modulus (G') and the loss modulus (G") can be used to report on the gel stiffness. Specifically, gels that are rigid and self-supporting yield G' values that are at least 10-fold greater than G",[364,401] and this is a crucial prerequisite for biomedical applications. The frequency sweeps shown in Figure 7.10 do in fact illustrate this trend, where G' is significantly higher than

G". The rheological measurements of wild-type KFE are consistent with previously reported values,[364] and more importantly, KFE35 and KFE17 display similar viscoelastic properties as wild-type KFE. Using the magnitude of G' to judge the strength of the gels, the order of strongest to weakest is KFE17 > KFE35 > KFE > KFE4. This data suggests that positioning of the Lys(nvoc) mutations (i.e., KFE17 versus KFE35) can result in different structural frameworks and therefore varying viscoelastic properties. One reason Lys(nvoc) containing gels have a greater rigidity may be due to the enhanced surface area of hydrophobic interactions induced by the large aromatic moiety. This would in turn create a more stable intermolecular network. If this notion is true, this is most effective when two Lys(nvoc) residues are positioned near the termini of the peptide. While there are slight differences in gel stability among the eight residue peptides (i.e., KFE, KFE35, KFE17), overall the results support the fact that incorporation of Lys(Nvoc) into the amino acid sequence of the KFE peptide does not significantly alter the propensity of the peptide to self-assemble and subsequently transition to a gel state. The major outlier in this case is the KFE4 hydrogel solution which shows little or no gel rigidity, and again, this is consistent with the FTIR spectra and AFM images presented above.

Taken together, the results presented above demonstrate the validity of using light to modulate the structural integrity of a model peptide hydrogel. Our working hypothesis is that generation of a charge, via a photochemical reaction, in a strategically important site in the peptide sequence is sufficient to destabilize the peptide hydrogel's underlying nanofibrillar scaffold. Specifically, we chose a well studied and biocompatible nonnatural amino acid, Lys(Nvoc), as the phototrigger, which, upon photocleavage, is converted to a

charged lysine. Using secondary structure analysis, microscopic imaging, and rheometric analysis, we confirm that the incorporation of Lys(nvoc) into the KFE sequence does not alter the hydrogel-forming propensity of the peptide and also the mechanical properties of the hydrogel thus formed; however, it provides a convenient mechanism of using light to trigger the degradation of peptide hydrogels. More important, the current phototriggering method offers easy control over the spatial and temporal selectivity, an attribute that is invaluable to many applications, as in controlled drug and/or protein release systems, injectable scaffolds for tissue engineering and regenerative medicine, and 3D cell culture.[402]

7.4 Conclusions

Taken together, the compilation of studies presented herein demonstrates a new method that allows the disassembly of various highly ordered structural scaffolds arising from the self-assembly of peptides. The photolabile non-natural amino acid, Lys(nvoc), used in this work, allows tunability of the chemical environment via light in the form of a transition from a hydrophobic to charged hydrophilic sidechain. Thus, incorporation of Lys(nvoc) into aggregation-prone peptides allows generation of a charge via light activation in regions that are otherwise hydrophobic and essential for stabilizing fibril/aggregate environments, finally resulting in their disaggregation. Here, we demonstrated this method on two types of systems, amyloid fibrils and peptide hydrogels. First, we show that when Phe23 in the wild-type $hIAPP_{22-27}$ sequence is replaced with Lys(nvoc) (i.e., hI-F23C), the resulting mutant is not only capable of forming fibrils, but

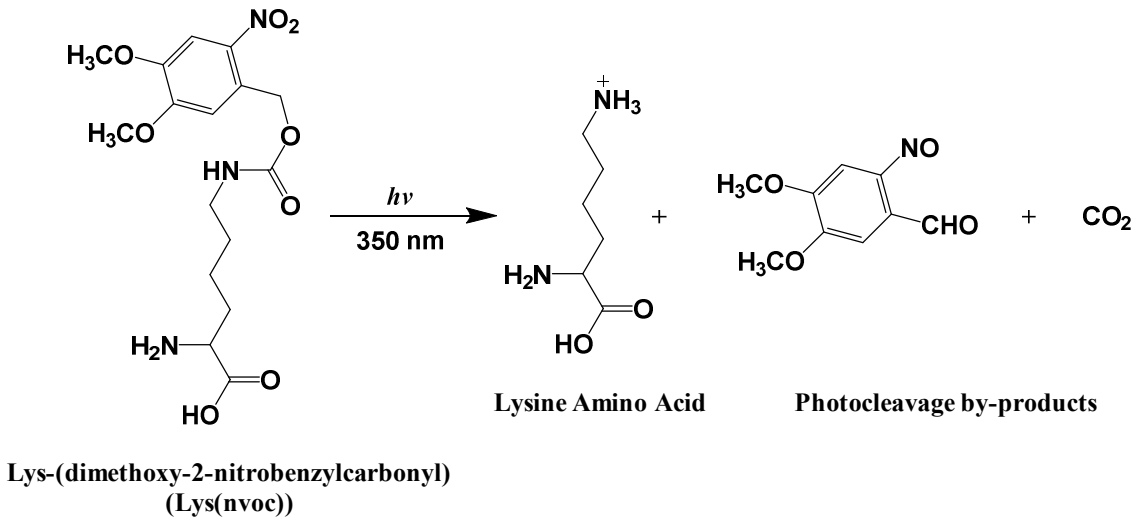
the fibrils thus formed can be disassembled via illumination. The same result was observed when Phe19 was replaced with Lys(nvoc) in the full length $A\beta_{1-42}$ sequence. In addition, to disaggregate fibril structures formed by the native peptide, we demonstrated that after co-aggregation of hI-F23C and the wild-type $hIAPP_{22-27}$ monomers, the peptides form an amalgamate fibril that can be disassembled upon photocleavage of Lys(nvoc). Taking this one step further, we determined that monomers of Aβ-F19C can be added to mature wild-type $A\beta_{16-22}$ fibrils and can also be disassembled to some degree. The second model system we investigated was a well-studied peptide hydrogel sequence (i.e., KFE) composed of alternating hydrophobic and charged residues and as a result, self-assembles into individual non-covalent and ionic interacting β-sheet layers. We illustrate two examples where two Phe residues in the hydrophobic layer were site-specifically modified to Lys(nvoc), (1) Phe1 and Phe7 and (2) Phe3 and Phe5. The resulting designed peptides favorably formed stiff hydrogel scaffolds composed of a dense network of β-sheet fibrils and with viscoelastic properties comparable to its wild-type. More interestingly, our results confirm that the current Lys(nvoc) photocleavage method indeed proves a convenient means to control hydrogel degradation in a site-specific manner. Overall, while here we have mainly focused on the applications toward amyloid related peptides and biomaterial applications, the findings presented above can be applicable to other biological and/or medical applications where a high degree of spatial and temporal control over the structural framework is necessary.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978), the NIH Ruth Kirschstein National Research Service Award Predoctoral Fellowship (F31AG046010), and the National Science Foundation-supported Nano/Bio Interface Center (NBIC) at the University of Pennsylvania. We also thank Anne van Oosten, Prof. Paul Janmey, and the Laboratory for Research on the Structure of Matter (LRSM) at the University of Pennsylvania for assistance with rheological measurements, and Dr. Timothy Wade and the Drexel University Department of Chemistry for assistance with the AFM measurements.

Abbreviation	Sequence
KFE	FKFEFKFE
KFE35	FK**X**E**X**KFE
KFE17	**X**KFEFK**X**E
KFE4	FK**X**E

Table 7.1 Summary of peptide hydrogel sequences. X denotes the Lys(nvoc) non-natural amino acid mutation.



Lys-(dimethoxy-2-nitrobenzylcarbonyl) (Lys(nvoc))

Figure 7.1 Photochemical reaction scheme of Lys-(dimethoxy-2-nitrobenzylcarbonyl) (Lys(nvoc)). Irradiation with UV light results in restoration of a native lysine amino acid and also photocleavage by-products.

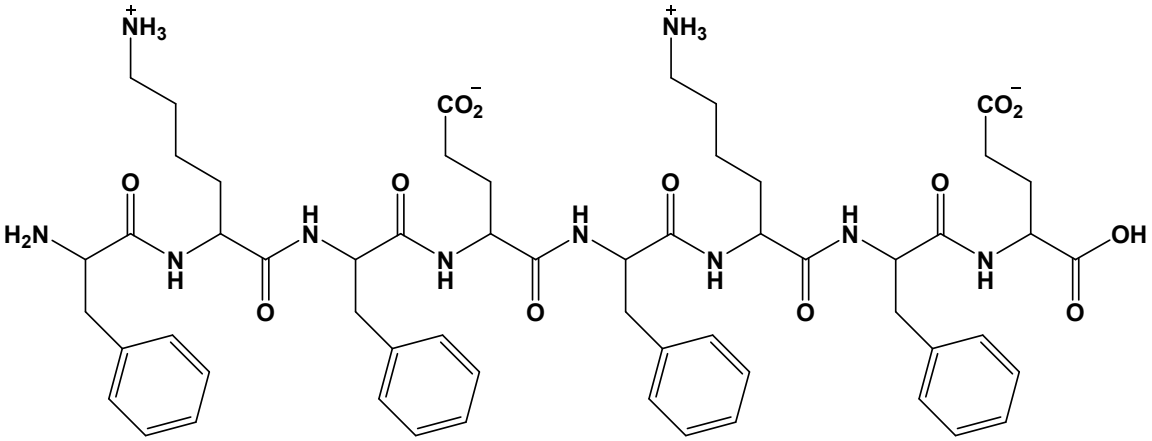


Figure 7.2 Chemical structure of the model peptide hydrogel (KFE, Table 7.1) used in the present studies (sequence: FKFEFKFE).

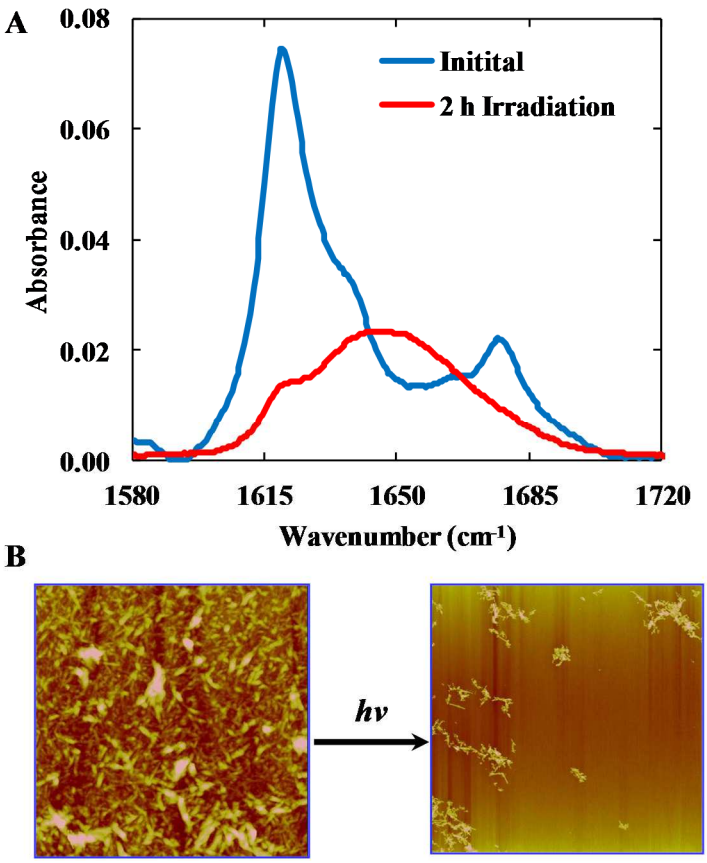


Figure 7.3 (A) FTIR spectra of hI-F23C (2 mM, pH 7) before and after irradiation with UV light at 350 nm, as indicated. (B) AFM images of the hI-F23C fibrils before and after irradiation. The image spans a dimension of 10 x 10 μm.

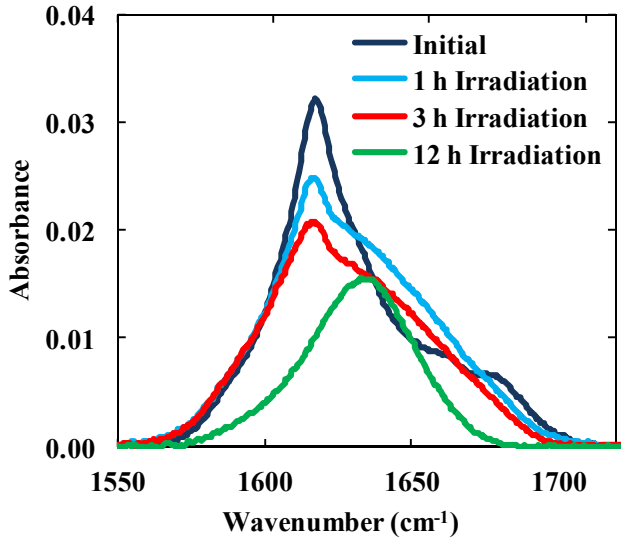


Figure 7.4 FTIR spectra of hI-F23C co-aggregated with wild-type $hIAPP_{22-27}$ in a 1:1 ratio (2 mM, pH 7) before and after periods of UV irradiation at 350 nm, as indicated.

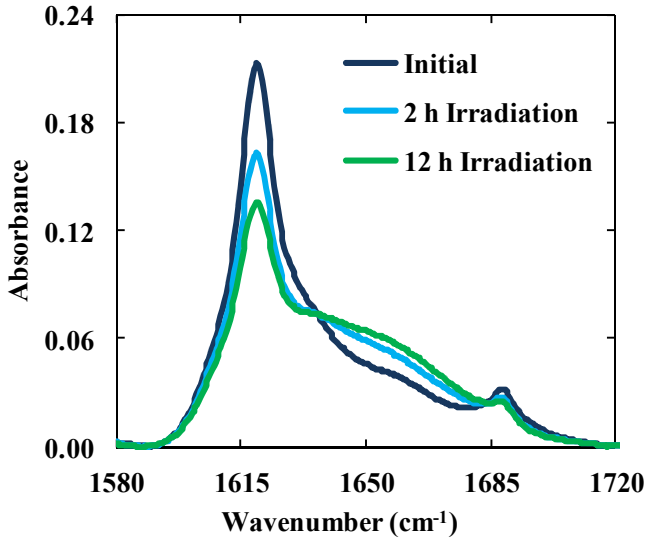


Figure 7.5 FTIR spectra of Aβ-F19C and $A\beta_{16-22}$ in a 1:1 ratio (4 mM, pH 7) with and without light irradiation, as indicated. Aβ-F19C monomers were added to an $A\beta_{16-22}$ solution that was previously allowed to form aggregates for two weeks. The initial spectrum was taken after 1 month for equilibration of this mixture.

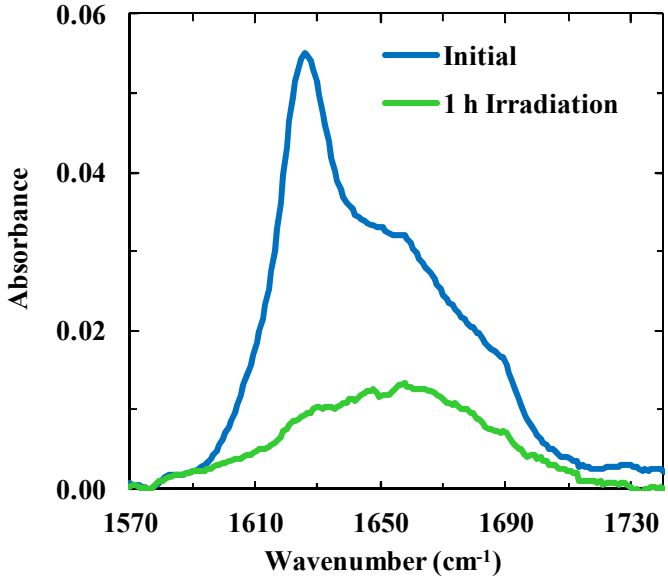


Figure 7.6 FTIR spectra of Aβ$_{42}$-F19C (2 mM, pH 7) before and after an irradiation with UV light, as indicated. Aβ$_{42}$-F19C was incubated for 2 weeks prior to the collection of FTIR spectra.

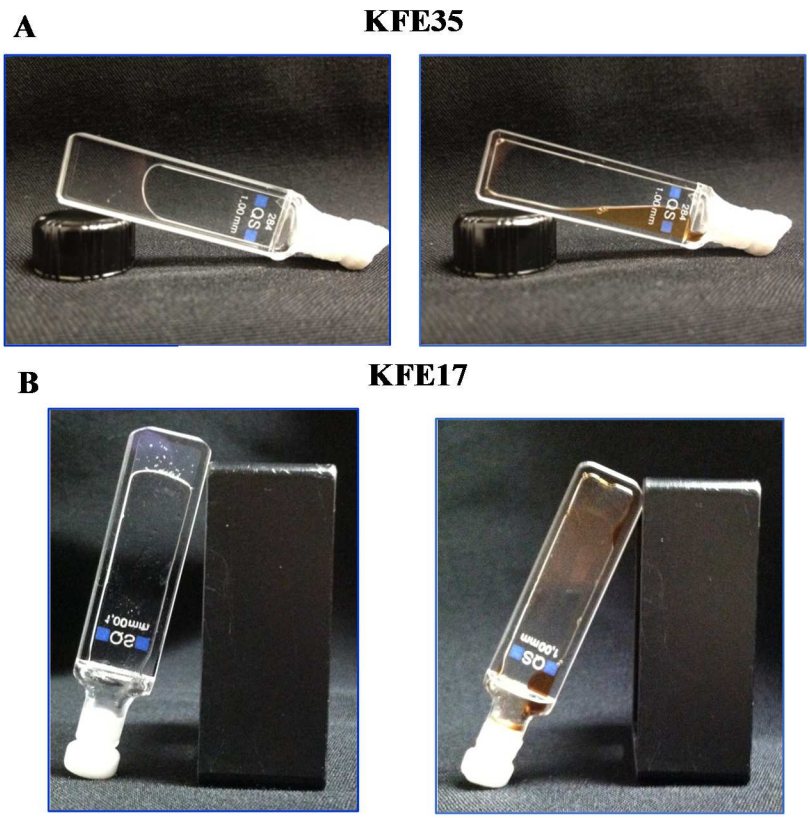


Figure 7.7 Picture of a 10 mM sample of (A) KFE35 and (B) KFE17. (Right) The samples formed a self-supporting gel, showing that the solution is viscous and does not flow. (Left) After ~30 mins of irradiation with UV light at 350 nm, the gel degrades and the sample transitions into a non-viscous solution.

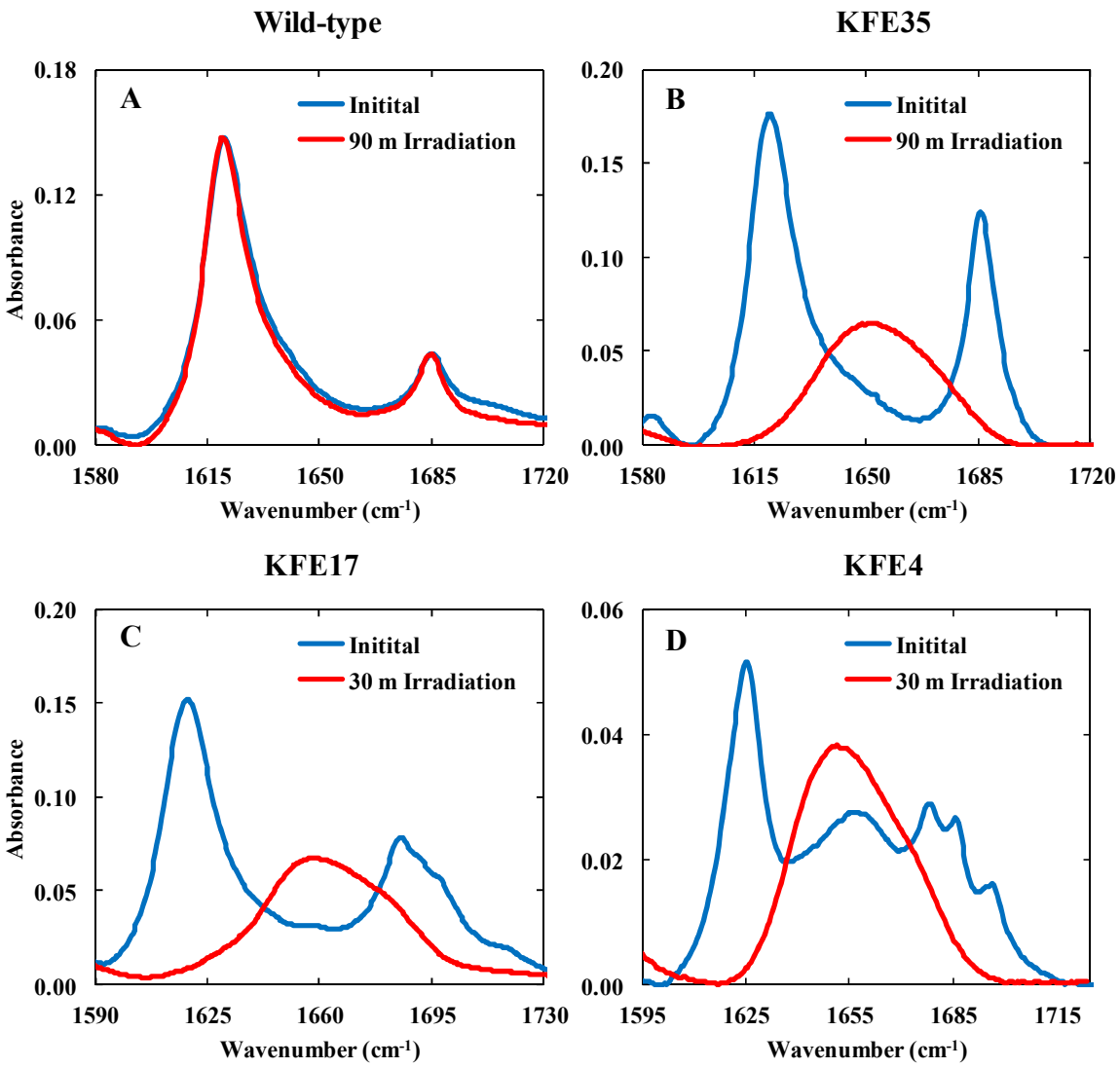


Figure 7.8 (A) FTIR spectra of the KFE hydrogel in the amide I' region before and after irradiation with 350 nm light for 90 min, showing that light is ineffective in this case to change the peptide and hence the hydrogel structure. FTIR spectra of the (B) KFE35 (C) KFE17 and (D) KFE4 hydrogel in the amide I' region before and after irradiation with 350 nm light for 90 min, showing that the underlying peptide secondary structure undergoes a β-sheet to disordered conformational transition.

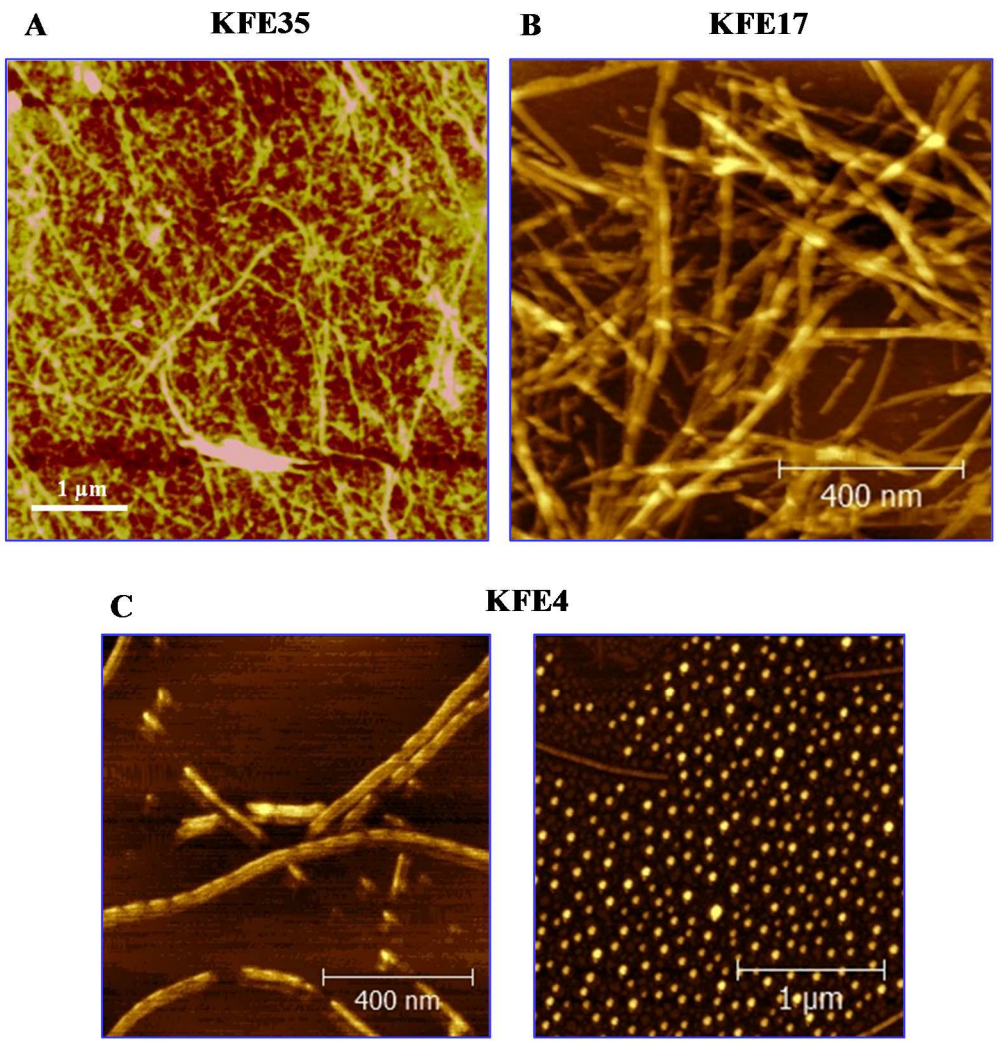


Figure 7.9 AFM images of (A) KFE35 (B) KFE15 and (C) KFE4.

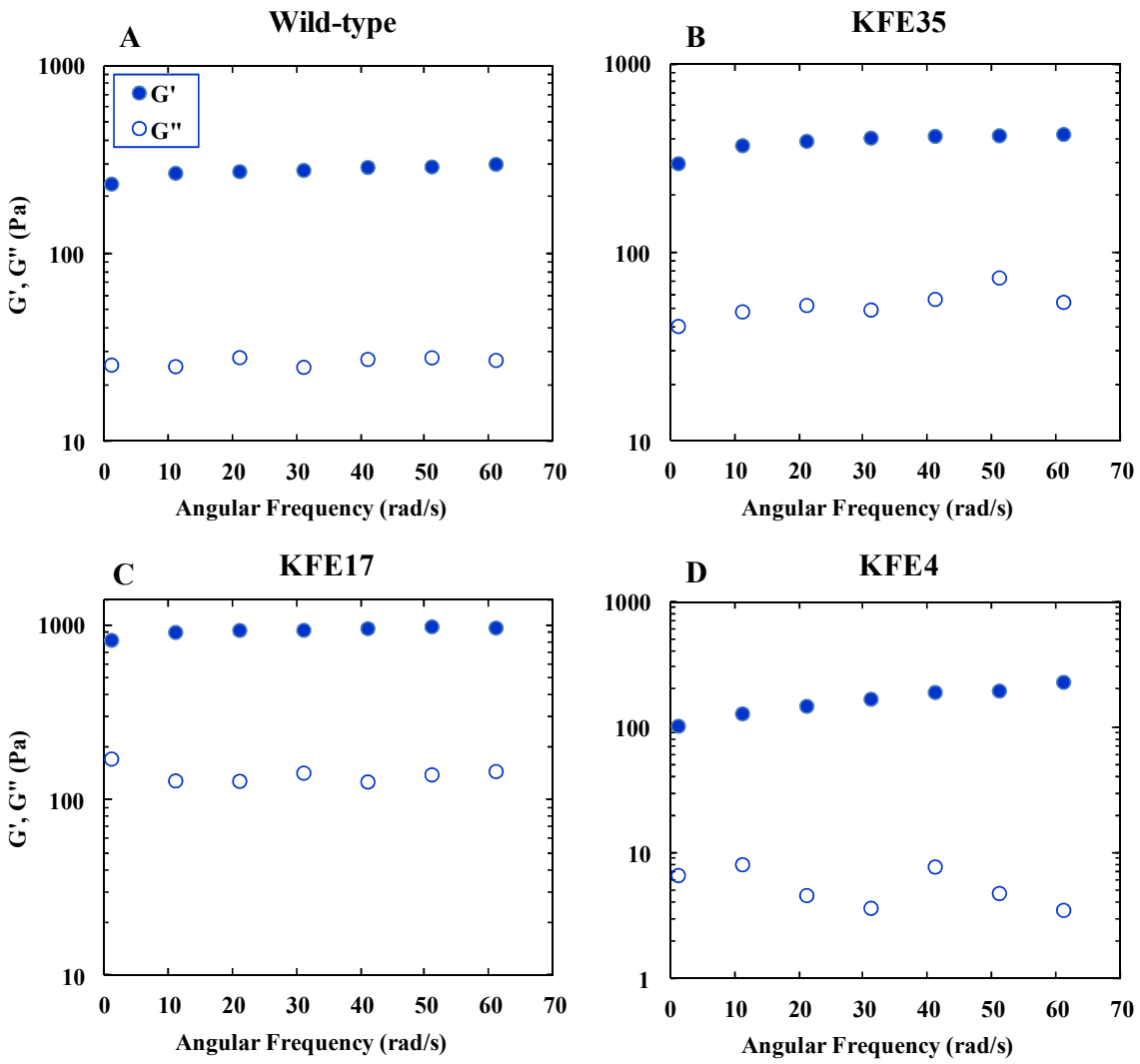


Figure 7.10 Rheological measurements of (A) KFE, (B) KFE35, (C) KFE17, and (D) KFE4 hydrogels. Frequency sweeps were conducted from an angular frequency of 1 to 60 rad/s. The elastic moduli (G') and viscous moduli (G") were measured for each sample.

8 C≡N Stretching Vibration of 5-Cyanotryptophan as an Infrared Probe of Protein Local Environment: What Determines Its Frequency?

Reproduced by permission of the PCCP Owner Societies. Physical Chemistry Chemical Physics, Wenkai Zhang,‡ Beatrice N. Markiewicz,‡ Rosalie S. Doerksen, Amos B. Smith III, and Feng Gai, (2016) 18, 7027-7034. DOI: 10.1039/C5CP04413H.[403] (‡ Denotes equal authorship)

8.1 Introduction

Vibrational spectroscopy is a powerful tool for assessing the structure and conformational dynamics of proteins. While many intrinsic vibrational modes, such as the amide I mode of the protein backbone, have been used for this purpose, they often lack the ability to reveal site-specific information due to spectral overlapping and/or vibrational coupling. To overcome this limitation, the past decade has seen an increased effort towards the development of unnatural amino acid-based infrared (IR) probes that can be used to site-specifically interrogate various structural and environmental properties of proteins, such as the local electrostatic field, hydrogen-bonding (H-bonding) interactions, and degree of hydration.[141,142,156,404,405] One of those unnatural amino acids is 5-cyanotryptophan (Trp_{CN}). Waegele *et al*.[294] found that the C≡N stretching vibrational mode of Trp_{CN}, especially its bandwidth, is highly dependent on the percentage of water in water and tetrahydrofuran (THF) mixtures and, thus, suggested that this unnatural amino acid could be used as an IR probe of the local hydration status of proteins. By performing a combined QM/MM study on 5-cyanoindole, they further showed that, besides the H-bonding interactions between water and the C≡N group, the interactions between water and the aromatic indole ring, especially the H-bonding interactions between water and the

indole amine (N-H) group, also influenced the C≡N stretching frequency.[406] More specifically, they found that the H-bonding interactions occurring at the C≡N and N-H groups had different effects on the C≡N stretching frequency, with one (C≡N) shifting the frequency toward higher wavenumbers and the other (N-H) toward lower wavenumbers. This finding is particularly interesting as it suggests that the indole N-H group of Trp_{CN} could be used to sense local H-bonding dynamics through measurement of the C≡N stretching vibration, especially under conditions where the nitrile group is not directly involved in hydrogen bond (HB) formation. To further verify this notion, herein we carried out static and ultrafast IR studies on 3-methyl-5-cyanoindole (3M5CI), which is the sidechain of Trp_{CN} (Figure 8.1), in different solvents.

Tryptophan (Trp) plays an important role in defining the folding, structure and function of many proteins; it is frequently found at or near sites that are responsible for protein-protein interaction,[407] ligand binding,[408,409] protein-DNA interaction,[410,411] and enzyme catalysis.[412,413] For example, the M2 proton channel of the influenza A virus uses a Trp tetrad to gate and control proton conduction across the viral membrane in an asymmetric manner after endocytosis, which enables uncoating and release of the viral RNA into the host cell for viral replication.[414] In addition, it is well recognized that Trp plays a key role in anchoring membrane proteins and peptides in lipid bilayers as it is preferentially located at the water-membrane interface.[415] For these reasons, Trp, which fluoresces upon excitation with ultraviolet light, has become one of the most frequently utilized fluorophores in the study of the structure-dynamics-function relationship of proteins via fluorescence spectroscopy. On the other hand, the study of the role of Trp in

protein structure and dynamics using IR spectroscopy is scarce. This is because none of the intrinsic IR active vibrational modes of Trp is particularly ideal for being used as a site-specific IR probe of proteins, due to spectral congestion, low extinction coefficients, or insensitivity to local environment.[75,416] Since IR spectroscopy is capable of offering a higher temporal and sometimes structural resolution than fluorescence spectroscopy, it would be quite useful to confer distinct IR utility to Trp by adding a relatively non-perturbing exogenous moiety to the indole ring that displays a strong, localized, and environmentally sensitive absorption band located in a non-congested region of the IR spectrum of proteins. While the C≡N stretching vibration of Trp_{CN} seems to meet these requirements, further experimental study is required to delineate the factors that determine its frequency and to lay a quantitative foundation to use this vibrational mode to investigate various biophysical problems, such as the changes in local hydration environment and H-bonding dynamics. Our results showed that the frequency of the C≡N stretching mode of 3M5CI is not simply dictated by the immediate environment of the nitrile group, but instead, it depends on the microscopic surroundings of the entire molecule. In other words, specific and/or non-specific interactions with the C≡N group, the aromatic ring, and the pyrrole N-H group combined determine the position and width of this vibrational band. In addition, we found that when the nitrile group is buried in an aprotic environment, the C≡N stretching frequency exhibits a simple and linear dependence on the polarizability and H-bonding ability of the solvent. We believe that this finding is particularly encouraging and useful in applying Trp_{CN} to probe local HB dynamics in cases where the benzene ring is situated in a dehydrated environment.

Further time-resolved measurements indicated that an ultrafast process at the vibrationally excited state produces a long-lived ground state bleach signal ($\tau \approx 12.3$ ps), making it possible to use Trp_{CN} to probe dynamic events occurring on the timescale of tens of ps.

8.2 Experimental Section

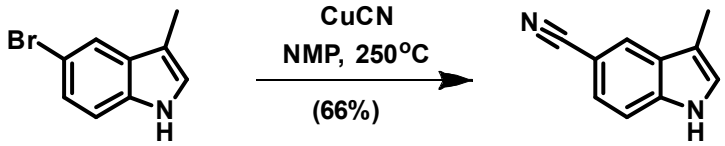
Materials and Sample Preparation

The details of the 3M5CI synthesis are given in the next section. N-methyl-5-cyanoindole (NM5CI) was purchased from Thermo Fischer Scientific (Loughborough, UK) and used as received. The following solvents (spectroscopic grade) were purchased from Acros Organics: methanol (MeOH), 2-propanol, dichloromethane (DCM), dimethyl sulfoxide (DMSO), dimethylformamide (DMF), pyridine, acetophenone, cyclopentanone, tetrahydrofuran (THF), 1,4-dioxane, toluene, tetrachloromethane, and 2,2,2-trifluoroethanol (TFE); and hexafluorobenzene was purchased from Oakwood Products. We checked the water content in two representative solvents, DMSO and THF, using the OH stretching band at ~3500 cm^{-1}. As indicated (Figure 8.2), in both cases the amount of water is negligible. Samples were freshly prepared before use by directly dissolving either 3M5CI or NM5CI in the desired solvent and the final concentration of the solute was approximately 33 mM for static and 100 mM for time-resolved measurements.

Synthesis of 3-Methyl-5-Cyanoindole (3M5CI)

3M5CI was prepared from 3-methyl-5-bromoindole using Scheme 8.1. Specifically, a conical heavy-walled microwave tube was charged with 3-methyl-5-bromoindole (500

mg, 2.39 mmol, 1 equiv), CuCN (320 mg, 3.59 mmol, 1.5 equiv), and anhydrous N-methyl-2-pyrrolidone (NMP) (5 mL). Vigorous shaking of the microwave tube was employed to completely dissolve the reactants. The solution turned green during dissolution. The microwave tube was crimped and sealed and then inserted into a microwave (Biotage Initiator). The contents were heated to 250 °C for 1.5 hours at high absorption level. The reaction solution was allowed to cool, then diluted with Et_2O (30 mL) and then partitioned with water (15 mL) and brine (3 × 15 mL); the organic layer was dried over Na_2SO_4. The oil was loaded onto a silica-gel column and purified by flash chromatography using hexanes/ethyl acetate gradient to yield 246 mg (66%) of a white, amorphous solid. HRMS (ES) found *m/z* 157.0767 [$(M+H)^+$; calcd. for $C_{10}H_9N_2$: 157.0766]; ^{1}H NMR (500 MHz, Chloroform-*d*) δ 8.27 (s, 1H), 7.93 (dt, *J* = 1.6, 0.8 Hz, 1H), 7.41 (d, *J* = 1.4 Hz, 1H), 7.40 (d, *J* = 0.9 Hz, 1H), 7.09 (dd, *J* = 2.3, 1.1 Hz, 1H), 2.34 (d, *J* = 1.2 Hz, 3H). IR (KBr, cm^{-1}) 2218.7 (s).



Scheme 8.1 Reaction scheme of 3-methyl-5-bromoindole with CuCN and NMP to form 3M5CI.

Static and Time-Resolved IR Measurements

All static IR measurements were carried out on a Thermo Nicolet 6700 FTIR spectrometer at a resolution of 1 cm^{-1}. A solvent background has been subtracted for each spectrum shown. 2D IR spectra were obtained on a heterodyne-detected photon-echo setup with a boxcar geometry that has been described in detail elsewhere.[417] IR pump-

probe data were obtained using a transient absorption spectrometer derived from the 2D IR setup. Briefly, the local oscillator and one of the three pump beams were blocked; one of the remaining two beams was used as the pump, and the other was attenuated and directed to the monochromator to act as the probe. The polarization of the probe was set at the magic angle with respect to that of the pump. For both the static and time-resolved measurements, the sample solution was placed between two 2 mm CaF_2 windows separated by either a 25 μm (for time-resolved experiment) or 50 μm (for static measurement) spacer.

Analysis of Decay Kinetics of the Excited State of the C≡N Stretching Vibration

According to the kinetic scheme in the main text (Scheme 8.2), the $(C{\equiv}N)_{v=1}$ state (A), initially prepared by the IR excitation pulse, has two decay channels as discussed below: one to $(C{\equiv}N)_{v=0}$ (C), and the other to $(\text{Dark State})_{v=1}$ (B) which then converts to $(C{\equiv}N)_{v=0}$. This kinetic model, as indicated below,

$$\begin{array}{ccccc} \text{A} & \xrightarrow{k_{11}} & \text{B} & \xrightarrow{k_{20}} & \text{C} \\ & & \text{A} & \xrightarrow{k_{10}} & \text{C} \end{array}$$

gives rise to the following rate equations:

$$\frac{d[\text{A}]}{dt} = -k_{11}[\text{A}] - k_{10}[\text{A}]$$
$$\frac{d[\text{B}]}{dt} = k_{11}[\text{A}] - k_{20}[\text{B}]$$
$$\frac{d[\text{C}]}{dt} = k_{20}[\text{B}] + k_{10}[\text{A}]$$

Solving these differential equations leads to the following population kinetics:

$$[\mathrm{A}] = [\mathrm{A}]_0 e^{-(k_{11}+k_{10})t}$$

$$[\mathrm{B}] = [\mathrm{A}]_0 \frac{k_{11}}{(k_{11}+k_{10})-k_{20}}(e^{-k_{20}t} - e^{-(k_{11}+k_{10})t}) \quad (8.1)$$

$$[\mathrm{C}] = [\mathrm{A}]_0 - [\mathrm{A}] - [\mathrm{B}] = [\mathrm{A}]_0 \{1 - \frac{k_{11}e^{-k_{20}t} + (k_{10}-k_{20})e^{-(k_{11}+k_{10})t}}{(k_{11}+k_{10})-k_{20}}\}$$

The ground state bleach recovery signal is given by $[\mathrm{C}] - [\mathrm{A}]_0$, i.e.,

$$\mathrm{GSB}(t) = -[A]_0 \frac{k_{11}e^{-k_{20}t} + (k_{10}-k_{20})e^{-(k_{11}+k_{10})t}}{(k_{11}+k_{10})-k_{20}}, \quad (8.2)$$

which indicates that the percentage of the slow component is

$$\frac{A_2}{A_1 + A_2} = \frac{k_{11}}{k_{11} + (k_{10} - k_{20})} \quad (8.3)$$

Adding the stimulated emission contribution to Equation 8.2 gives rise to the decay kinetics of the negative-going pump-probe signal, S(t):

$$\mathrm{S}(t) = -[A]_0 \frac{k_{11}e^{-k_{20}t} + (2k_{10} - 2k_{20} + k_{11})e^{-(k_{11}+k_{10})t}}{(k_{11}+k_{10})-k_{20}} \quad (8.4)$$

It is east to show, based on Equation 8.4, that the ratio of the two exponential components in the negative-going pump-probe decay kinetics is $k_{11}/(2k_{10} - 2k_{20} + k_{11})$.

8.3 Results and Discussion

C≡N Vibrational Bands of 3M5CI in Different Solvents

It is well known that protic solvents, such as water and simple alcohols, can form HBs with the C≡N group of aryl and alkyl nitriles.[418] For 3M5CI, such solvents are expected to also interact with the indole N-H group via H-bonding interactions. Thus, it is impossible to distinguish the effects of these two types of H-bonding interactions based

on measurements of the C≡N stretching vibrational band of 3M5CI in such protic solvents alone. Therefore, we have chosen mostly solvents that can only form HBs with a solute through their H-bonding accepting groups, based on the Kamlet-Taft solvent parameters (Table 8.1).[419,420] As expected (Figure 8.3), the C≡N band of 3M5CI shows a clear dependence on solvent. Furthermore, the band shape in each case can be described satisfactorily by a Voigt profile (Figure 8.4) and the corresponding spectral parameters are listed in Table 8.1. A cursory inspection of the results suggests, as predicted by the computational study of Waegele *et al*.,[406] that H-bonding interactions between solvent and the indole N-H group can have a measurable influence on the stretching frequency of the C≡N group located on the other side of the aromatic ring. This can be seen by comparing the spectra obtained in TFE, DMSO and MeOH. In TFE, which is a strong HB donor according to its Kamlet-Taft parameters ($\beta = 0$, $\alpha = 1.51$), the C≡N stretching band is centered at 2230.7 cm^{-1}, whereas in DMSO, which is a strong HB acceptor ($\beta = 0.76$, $\alpha = 0$), the C≡N stretching band is shifted to 2216.1 cm^{-1}. On the other hand, in MeOH ($\beta = 0.62$, $\alpha = 0.93$), which is capable of forming HBs with both the C≡N and N-H groups, the C≡N stretching band is located between those obtained in TFE and DMSO and centered at 2223.6 cm^{-1}.

It is noticeable that the C≡N bands of 3M5CI obtained in protic solvents, especially in TFE, are much broader than those obtained in aprotic solvents (Table 8.1), indicating that in protic solvents the inhomogeneity sensed by the C≡N stretching vibration becomes larger. Previously, Cho and coworkers[421] have shown that the C≡N stretching frequency of acetonitrile is dependent on the HB configuration, leading to

inhomogeneous broadening of the band. As shown (Figure 8.5), the bandwidth of the C≡N stretching vibration of acetonitrile in TFE is approximately 14 cm^{-1}, which is broader than that (~8 cm^{-1}) in THF.[422] This is consistent with the study of Cho and coworkers.[421] In comparison, however, the bandwidth of this vibration in 3M5CI shows a more significant increase upon changing the solvent from THF to TFE (i.e., 8 to 24 cm^{-1}). Taken together, these results indicate that in the case of 3M5CI, besides an inhomogeneous distribution of HB configurations at the C≡N site, other solvent interactions, likely H-bonding interactions with the indole ring also contribute significantly toward the broadening of the bandwidth. This is consistent with a recent study of McLain and coworkers,[423] which showed that in a methanol-water mixture, instead of being solvated by the methyl groups of methanol, the indole ring is in fact forming HBs with the –OH groups of water and methanol. Further evidence supporting this picture is, as shown (Figure 8.6), that for 3M5CI the dependence of the C≡N bandwidth on the overall solvent property, as judged by $\tau = \pi^* + \beta + \alpha$, are different for protic and aprotic solvents. It is worth noting that the bandwidth of the C≡N stretching vibration of 3M5CI in hexafluorobenzene was found to be concentration dependent (Figure 8.7), which indicates that 3M5CI may oligomerize in this case and, as a result, the measured frequency may not be entirely representative of the monomeric form. Therefore, we excluded this solvent in the following analysis.

This difference could be exploited to probe preferential interactions between the indole ring and a specific solvent component. To illustrate this point, we measured the C≡N stretching band of 3M5CI in a binary solvent composed of DMSO and TFE. We

chose this mixture because it is well known that at a relatively high molar fraction DMSO will replace the protic solvent molecules, such as water, that are initially solvating the indole ring.[120] As shown (Figure 8.8), the results obtained at different volume ratios of TFE and DMSO are indeed consistent with this picture. It is clear that at a 50:50 ratio the C≡N stretching band is already similar to that obtained in pure DMSO, which indicates, as expected, exclusion of TFE molecules from the indole ring.

Quantifying the C≡N Stretching Frequency of 3M5CI in Aprotic Solvents

While the simple comparison discussed above revealed a picture that is consistent with the notion that HB formations at both the nitrile and N-H sites affect the C≡N stretching frequency, a more quantitative assessment of the FTIR data is needed in order to extract the exact contribution of the H-bonding interactions at the N-H site. To do so, we first considered solvents that can only form HBs with the N-H group but not the nitrile moiety. In other words, these solvents all have a Kamlet-Taft α parameter of zero (Table 8.1). A previous study by Moog and coworkers[424] indicated that the C≡N stretching frequency of benzonitrile shows a linear dependence on the Kamlet-Taft π^* parameter of solvents with $\alpha = 0$. Since the π^* parameter is a relative scale measuring the solvent's polarizability, this result indicates that the C≡N stretching frequency of benzonitrile, which lacks any HB donating groups, is determined only by the local electrostatic field exerted by the surrounding solvent molecules when the C≡N group is not involved in any direct H-bonding interactions.[424] As indicated (Figure 8.9a), however, the C≡N stretching frequency of 3M5CI does not show a strong linear correlation with the π^* parameter as observed for benzonitrile. Similarly, a strong linear correlation is not observed when

other solvent parameters (i.e., β and ε) are used alone (Figure 8.9b and 8.9c). Thus, these simple analyses indicate that in this case other factors are also at play. Since the most pronounced difference between benzonitrile and 3M5CI, in the context of the current study, is that the latter can interact with the solvent through its N-H group, one needs to consider the effect arising not only from the solvent's polarizability (π^*) but also its HB accepting ability (β). Indeed, we found that the C≡N stretching frequency of 3M5CI obtained in solvents with $\alpha = 0$ exhibits a strong linear dependence on $\gamma = \pi^* + \beta$ (Figure 8.10). This finding provides a strong corroboration for the aforementioned notion that HB formations involving the indole N-H group of Trp_{CN} can be sensed by the C≡N stretching vibration. In addition, and perhaps more importantly, this linear relationship makes it possible to use the C≡N stretching frequency to characterize changes in the H-bonding interactions between the N-H group of a Trp_{CN} residue and a neighboring solvent molecule or sidechain in proteins, especially under conditions where the 5-cyano position is immersed in an aprotic environment.

To further validate the conclusions reached above, we measured the C≡N stretching modes of N-methyl-5-cyanoindole (NM5CI, Figure 8.1) in DMSO since NM5CI is incapable of forming HBs with DMSO due to the added methyl group. As shown (Figure 8.11), the C≡N band of NM5CI in DMSO is centered at 2217.1 cm^{-1}, which is blue-shifted from that of 3M5CI. In addition, its bandwidth (8.4 cm^{-1}) is narrower than that (10.6 cm^{-1}) of 3M5CI. This blue-shift and band narrowing is once again consistent with the idea that H-bonding interactions with the indole N-H group is an important determinant of the C≡N stretching frequency.

Finally, we attempted to derive a simple relationship, using the Kamlet-Taft parameters alone, to describe the C≡N stretching frequencies obtained in all solvents. As discussed above, direct H-bonding interactions with the nitrile group shifts its stretching frequency to higher wavenumbers. Thus, considering the fact that the α parameter is a measure of the HB donating ability of the solvent and the results presented in Figure 8.10 are for solvents with $\alpha = 0$, we hypothesized that the simplest surrogate variable that can reasonably capture the overall effect of a solvent on the C≡N stretching frequency of 3M5CI is $\sigma = \pi^* + \beta - \alpha$. As shown (Figure 8.12), this parameter (σ), to our surprise, proves to be an excellent scale to quantify the C≡N stretching frequencies measured in all solvents, including water. The significant linear correlation exhibited between ω_0 and σ provides a simple and quantitative way to interpret the C≡N stretching frequency of 3M5CI or Trp_{CN}. Despite this success, however, it is worth noting that this linear relationship is obtained based on empirical solvent parameters. Hence, it is impossible to use it to directly yield a microscopic interpretation of the environment of a specific Trp_{CN} residue in proteins. Nonetheless, this relationship validates the notion that, besides the direct interaction with the nitrile group, interactions with other parts of the indole ring can also be sensed by the C≡N stretching vibration, making Trp_{CN} a more versatile IR probe in this regard.

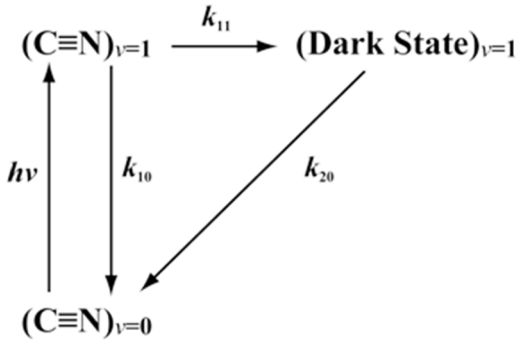
Time-Resolved IR Measurements

While the bandwidth of a linear IR spectrum is informative about the degree of inhomogeneous broadening, the underlying dynamics can only be assessed by nonlinear spectroscopic techniques, such as 2D IR spectroscopy.[425] To further test the feasibility of

using Trp_{CN} to probe such dynamics, we carried out 2D IR measurements on the C≡N stretching mode of 3M5CI in a 50:50 DMSO:TFE mixture. The reason that we chose this particular system as the testbed is based on the following considerations: (1) our results indicated, as discussed above (Figure 8.8), that under this condition the solute experiences a DMSO-like environment, indicating preferential accumulation of DMSO molecules near the indole ring; (2) a previous simulation study by Bagchi and coworkers[120] on Trp solvated by DMSO-water mixtures indicated that when the fraction of DMSO is larger than 15%, preferential interactions between the indole ring and DMSO occur wherein the solvent molecules form a distinct network or cluster surrounding the solute enhanced by favorable solvent-solute hydrophobic interactions; (3) we hypothesized that such a scenario also happens in DMSO-TFE mixtures, which can be tested by measuring the spectral diffusion dynamics[425] of the C≡N stretching vibrations via 2D IR spectroscopy as cluster formation has been shown to result in a prolonged spectral diffusion time.[426,427] As shown (Figure 8.13), 2D IR spectra of the C≡N band of 3M5CI obtained at 4 different waiting times (*T*) clearly indicate spectral diffusion dynamics, as manifested by changes in the contour of the 2D peak corresponding to the 0-1 transition. However, even at the longest waiting time of the experiment, 20 ps, this 2D peak still shows an appreciable tilt towards the diagonal direction, signifying the slowness of the spectral diffusion process. Since the spectral diffusion dynamics of the C≡N stretching vibration in simple, pure liquids typically occur on a few ps timescale,[426,428] this 2D IR result is therefore consistent with our hypothesis

that DMSO can form clusters surrounding the indole ring, making its microscopic environment fluctuate at a slower time scale.

Interestingly, the 2D IR spectra also revealed the presence of a second 1-2 transition. Since this spectral signature is not detectable at T = 0, it most likely corresponds to a dark state (i.e., the corresponding 0-1 transition is forbidden). A similar phenomenon was also observed for the C≡N stretching vibration of cyanophenol in methanol by Cho and coworkers,[429] which was attributed to a combination band arising from the combination of two vibrational modes of the parent molecule. To better characterize the impact of this dark state, we conducted IR pump-probe measurements under the magic angle polarization condition on 3M5CI in DMSO. As shown (Figure 8.14), the time-resolved spectra clearly reveal the existence of three distinguishable spectral features: one negative band centered at ~2216 cm^{-1}, which corresponds to contributions from the ground state bleach (GSB) and stimulated emission (SE) signals, and two positive bands, centered at ~2190 and ~2205 cm^{-1}, respectively. Because the 2190 cm^{-1} band is more intense at earlier delay times and the 2205 cm^{-1} band grows in with time, the simplest model capable of explaining the decay kinetics of these features is a competing relaxation process from the excited state, $(C\equiv N)_{v=1}$, to the ground state, $(C\equiv N)_{v=0}$, (through two channels: one goes through the v = 1 state of the dark state and the other takes the initially prepared excited state population directly to the ground state, as indicated in the following kinetics scheme:



Scheme 8.2 Proposed kinetic model describing a competing relaxation process from the first excited state, $(C\equiv N)_{v=1}$, to the ground state, $(C\equiv N)_{v=0}$, wherein a dark state provides an additional decay channel.

where k_{ij} is the rate constant of the corresponding kinetic step. To determine the key rate constants, we analyzed the transient absorption kinetics at two representative probing frequencies, ω_A = 2190 cm^{-1} and ω_B = 2223 cm^{-1}. Because the transient signal at ω_A contains contribution mostly from the excited state absorption from the v = 1 state of the C≡N stretching vibration, the relaxation kinetics at this frequency should reveal $k_1 = k_{10} + k_{11}$. As shown (Figure 8.15), the signal at ω_A can be satisfactorily described by a single-exponential function with a time constant of 1.3 ± 0.1 ps, indicating that $k_1 = (1.3 \pm 0.1\ ps)^{-1}$. On the other hand, the transient signal at ω_B contains contributions from both channels and, thus, should decay in a double-exponential manner. Indeed, as shown (Figure 8.15), the transient kinetics at this frequency can be fit to a double-exponential function with τ_1 = 1.2 ± 0.1 ps and τ_2 = 12.3 ± 1.6 ps. Based on the rate equations of the above kinetic scheme (see Experimental Section above), it is easy to show that $\tau_1 = (k_1)^{-1}$ and $\tau_2 = (k_{20})^{-1}$ = 12.3 ± 1.6 ps. It is clear that the τ_1 value is consistent with the k_1 value

determined above from the ω_A data. To further determine k_{11}, we took advantage of the fact the ratio (R) between the population going through the two channels is $k_{11}/(2k_{10} - 2k_{20} + k_{11})$. It is straightforward to show (methods section above) that $R = A_2/A_1 = 0.18$, where A_1 and A_2 are the amplitudes of the τ_1 and τ_2 components obtained from the double-exponential fit of the ω_B kinetics. Using this ratio and the above determined value of k_1, we calculated k_{10} and k_{11} to be $(1.8\ ps)^{-1}$ and $(4.7\ ps)^{-1}$, respectively. A previous study has shown that the vibrational lifetime of the C≡N stretching vibration of benzonitrile in DMSO is about 4 ps,[430] which is longer than that (1.8 ps) of 3M5CI. However, due to the second decay channel, a significant population (~31%) is transferred to another excited state, leaving a significant portion of the GSB signal of 3M5CI long lived. This prolonged GSB recovery time (12.3 ps) could be useful in the study of protein dynamic events occurring on the timescale of tens of ps. However, we note that the solute concentration used in the current 2D IR experiments is much higher than that of any typical protein solutions and, hence, the feasibility of this statement requires further test. In addition, future work is needed in order to provide a more comprehensive understanding of the factors that determine the onset and percentage of the aforementioned dark state, especially in aqueous solutions or under biological conditions.

Applying Trp_{CN} to measure local H-bonding dynamics in proteins

In many regards, the indole ring of Trp makes it a unique amino acid: it has a large hydrophobic surface area, a permanent dipole moment, a large aromaticity, and a HB donating group (i.e., the pyrrole N-H). As such, it can interact with the surroundings via different forces. For example, it can form π-hydrogen bonds[431-433] with backbone and/or

sidechain H-bond donating groups and also hydrophobic clusters with other aromatic amino acids. In addition, it tends to immerse itself in an environment where different interactions can occur; for instance, the benzene ring buries in a dehydrated and hydrophobic environment whereas the pyrrole N-H undergoes HB formation. One such example is the aforementioned M2 proton channel, wherein the N-H ends of the key Trp41 residues, according to a recent high-resolution crystal structure,[434,435] point toward the aqueous pore, leaving the benzene rings facing the hydrocarbons of the lipid. Another example is the gramicidin A proton channel,[436,437] where H-bonding interactions involving the N-H groups of several Trp residues are believed to be crucial for stabilizing the functional channel conformation. A third example is transmembrane peptides and proteins, where Trp residues are often located at the water-membrane interface with their N-H groups H-bonded to either water or a lipid headgroup.[438,439] The results obtained in the current study, we believe, provide a foundation to use Trp_{CN} to study such H-bonding interactions using various linear and nonlinear IR spectroscopic methods. Finally, our findings suggest that the C≡N stretching vibration of Trp_{CN} could be also used to study cation-π interactions involving Trp, which are prevalent in many biological systems.

8.4 Conclusions

Trp residues are frequently found at locations that are crucial for structure, interaction and functions of proteins. However, assessment of the microscopic environment of a specific Trp sidechain in proteins using linear and/or nonlinear IR spectroscopy has been hampered by the fact that none of its intrinsic IR active vibrational modes are ideally

suited for this purpose. While a Trp analog, Trp_{CN}, has been suggested to be useful in this regard because the bandwidth of its C≡N stretching vibration is sensitive to hydration, an experimental delineation of the factors that affect this vibration is lacking. Herein, we studied the C≡N stretching vibration of 3M5CI, which is the sidechain of Trp_{CN}, in a series of solvents, aiming to provide a better understanding and characterization of these factors. Our results revealed that the C≡N stretching frequency of 3M5CI depends not only on solvent interactions with the nitrile group, but also on interactions with the indole ring, making Trp_{CN} a versatile IR probe of proteins. Specifically, we found that a single solvent parameter, $\sigma = \pi^* + \beta - \alpha$, where π^*, β, and α are the Kamlet-Taft parameters characterizing the polarizability, HB accepting ability, and HB donating ability of the solvent, respectively, is sufficient to describe the C≡N stretching frequencies obtained in all solvents via a simple linear function. This relationship thus confirms the possibility of using the C≡N stretching vibration of Trp_{CN} to sense the dynamics of H-bonding interactions with the pyrrole N-H group, especially under conditions, as often seen in membrane proteins, where the N-H group is involved in HB formation but the benzene ring is buried in a dehydrated or hydrophobic environment. Linear and 2D IR measurements on 3M5CI in a binary solvent consisting of DMSO and TFE provide further evidence that its C≡N stretching vibration is a sensitive probe of the microscopic environment of the molecule. The results, which are consistent with the literature, show preferential accumulation of DMSO molecules around the indole ring when its mole fraction reaches about 0.5. In addition, the 2D IR data indicated the presence of a second excited state species, leading to a long-lived ground state bleach signal. Kinetic

measurements using IR pump-probe spectroscopy allowed us to further determine the formation time constant (~5.0 ps) of this additional excited state as well as the decay time constant (~12.3 ps) of the long-lived ground state bleach signal.

Acknowledgements

The ultrafast IR experiments were performed on instruments that were developed under an NIH Research Resource Grant (P41-GM104605). B.N.M is supported by an NIH Ruth Kirschstein National Research Service Award Predoctoral Fellowship (F31AG046010).

Solvent	ω_0, cm^{-1}	FWHM, cm^{-1}	π^*	β	α	ε
Water	2224.0[a]	18.0[a]	1.09	0.47[b]	1.51	80.1
MeOH	2223.6	14.2	0.60	0.62	0.93	33.0
2-propanol	2223.6	14.7	0.48	0.95	0.76	20.2
DCM	2221.7	10.1	0.82[b]	0.10[b]	0.13[b]	8.9
DMSO	2216.1	10.6	1.00	0.76	0.00	47.2
DMF	2217.2	9.7	0.88	0.69	0.00	38.2
Pyridine	2217.9	10.3	0.87	0.64	0.00	13.3
Acetophenone	2218.5	10.6	0.90	0.49	0.00	17.4
Cyclopentanone	2218.7	8.9	0.76	0.52	0.00	13.6
THF	2219.8	8.0	0.58	0.55	0.00	7.5
1,4-dioxane	2221.1	8.7	0.55	0.37	0.00	2.2
Toluene	2222.4	8.7	0.54	0.11	0.00	2.4
Tetrachloromethane	2224.5	11.5	0.28	0.00	0.00	2.2
Hexafluorobenzene	2226.9	10.7	0.33[b]	0.02[b]	0.00[b]	2.0
TFE	2230.7	24.1	0.73	0.00	1.51	27.7

Table 8.1 The center frequency (ω_0) and full-width at half maximum (FWHM) of the C≡N stretching band of 3M5CI in different solvents. Also listed for each solvent are its Kamlet-Taft parameters, π^* (polarizibility), β (hydrogen bond acceptor), α (hydrogen bond donor), as well as its dielectric constant (ε).[440] Unless otherwise indicated, all Kamlet-Taft parameters are from ref.419.

[a] IR data in water were from ref.155 and measured for 5-cyanoindole in a 95/5 (v/v) water/MeOH mixture.

[b] Taken from ref.420

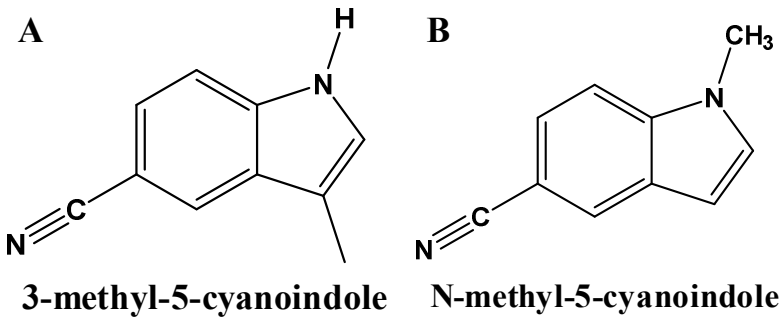


Figure 8.1 Structures of (A) 3-methyl-5-cyanoindole (3M5CI) and (B) 1-methyl-1H-indole-5-carbonitrile (NM5CI).

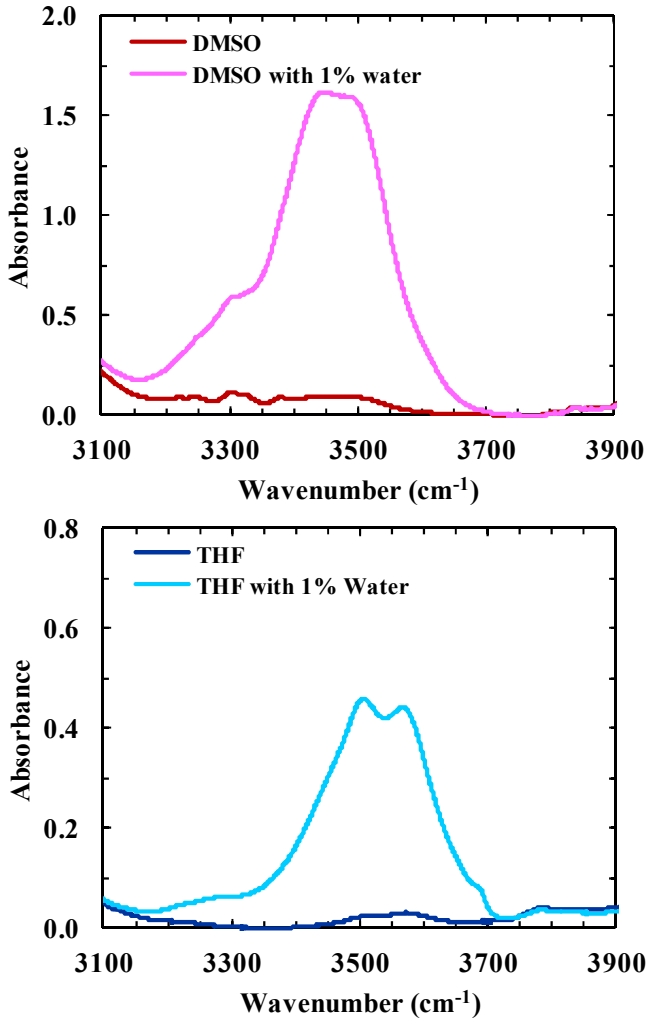


Figure 8.2 FTIR spectra of DMSO (Top) and THF (Bottom) with and without addition of 1% water, as indicated, in the OH stretching band region of water. These spectra indicate that the water content in the original solvents is negligible.

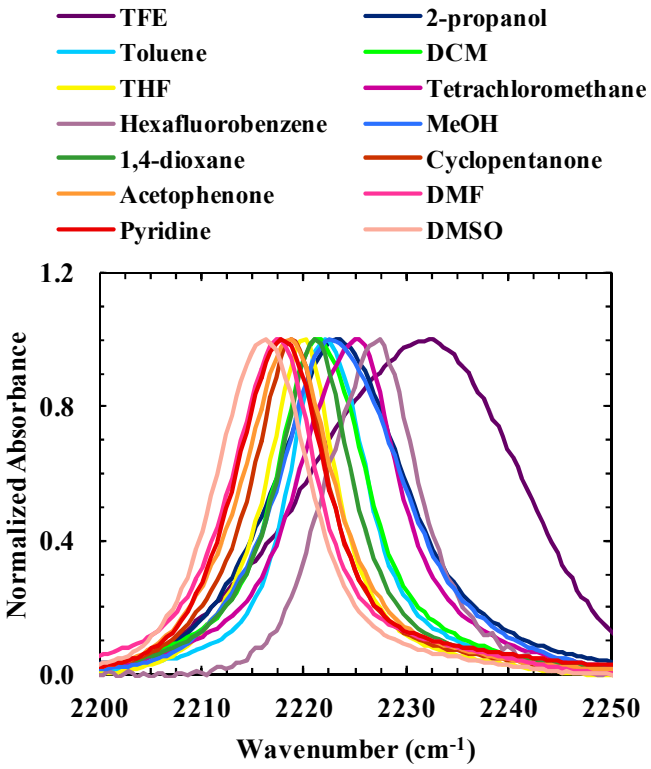


Figure 8.3 C≡N stretching bands of 3M5CI in different solvents, as indicated.

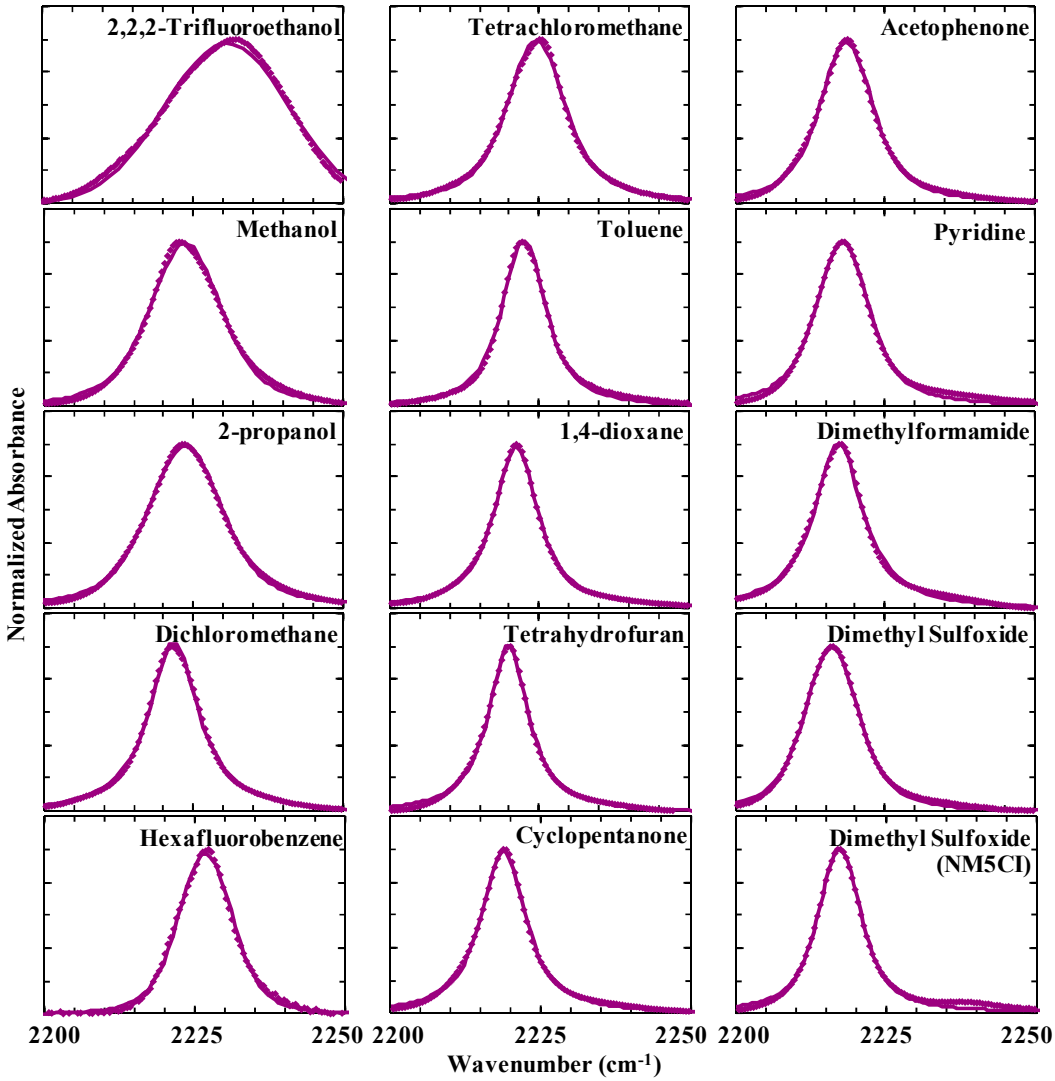


Figure 8.4 C≡N stretching bands of 3M5CI and NM5CI in different solvents, as indicated. The solid line in each case is the fit of the data to a Voigt profile and the resulting peak frequency and FWHM are given in Table 8.1.

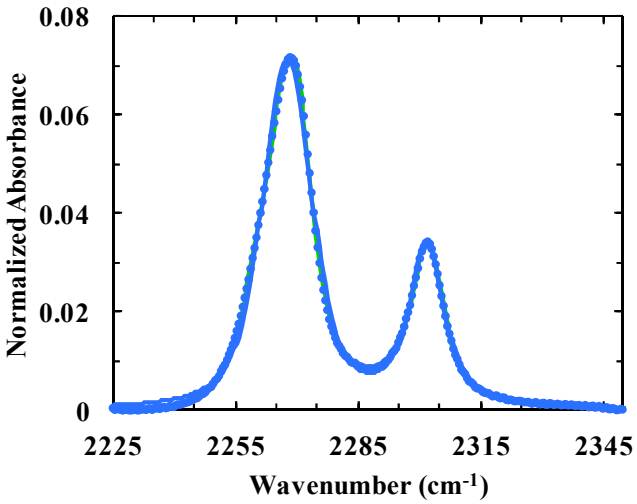


Figure 8.5 FTIR spectrum of acetonitrile in 2,2,2-trifluoroethanol, where the solid line is a fit of the spectrum to a function composed of two Voigt profiles. The peak frequency and FWHM of the C≡N stretching band are 2267.7 cm^{-1} and 14.3 cm^{-1}, respectively. The second, lower intensity band located at 2301.9 cm^{-1} arises from a Fermi resonance.

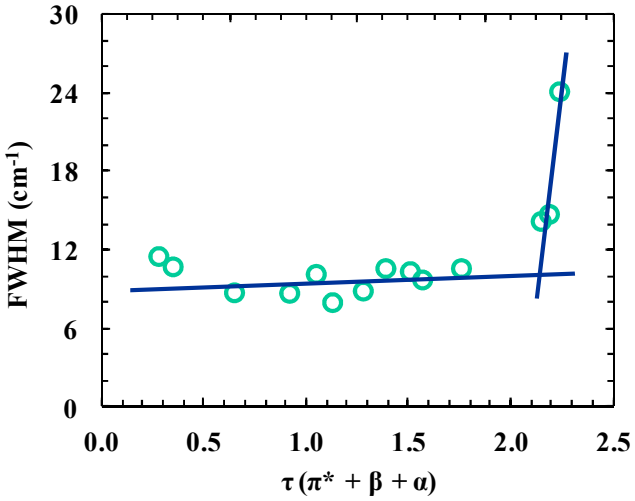


Figure 8.6 Plot of FWHM versus $\tau = \pi^* + \beta + \alpha$.

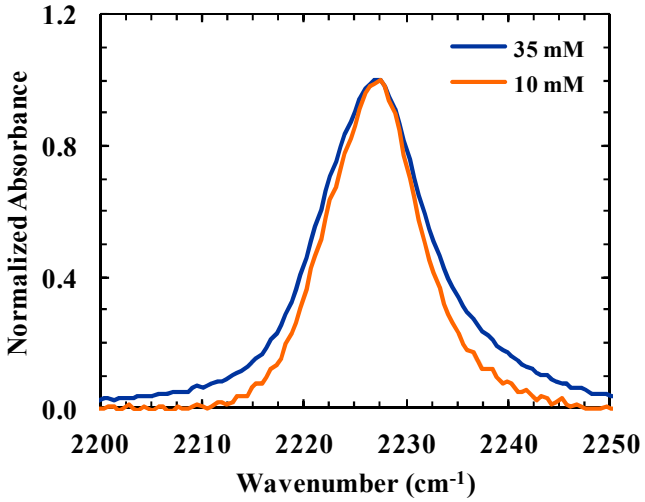


Figure 8.7 C≡N stretching bands of 3M5CI in hexafluorobenzene at 35 mM and 10 mM, as indicated.

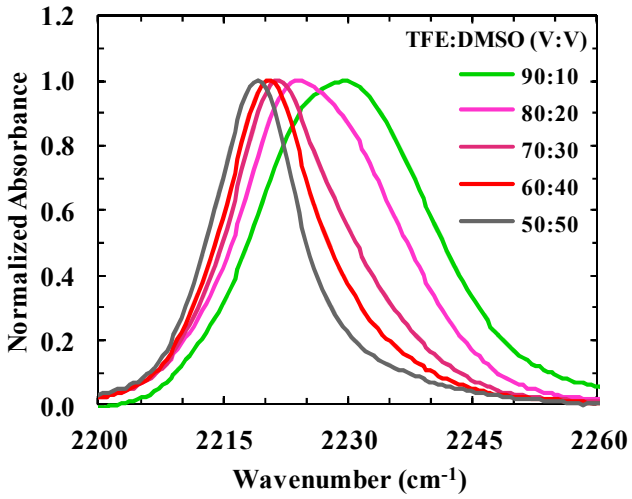


Figure 8.8 C≡N stretching bands of 3M5CI in TFE and DMSO mixtures with different volume ratios, as indicated.

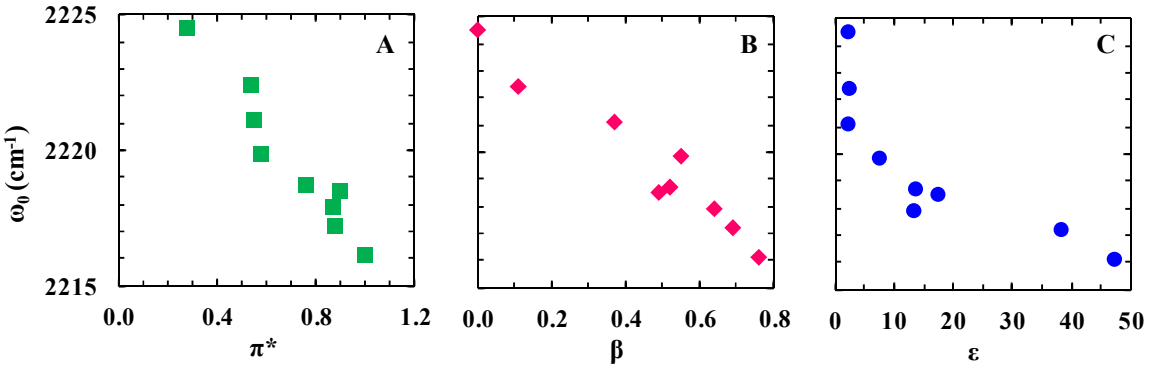


Figure 8.9 Center frequency (ω_0) of the C≡N stretching band of 3M5CI versus π^* (A), β (B), and ε (C). Only frequencies obtained in solvents with $\alpha = 0$ were used.

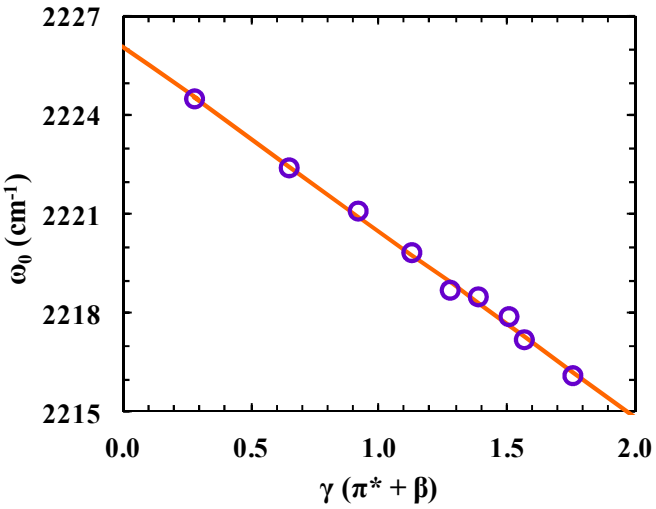


Figure 8.10 Center frequency (ω_0) of the C≡N stretching band of 3M5CI versus the solvent γ parameter, where $\gamma = \pi^* + \beta$. The solid line represents the linear regression of this data, yielding a slope of -5.6 ± 0.2 cm^{-1} and an intercept of 2226.1 ± 0.2 cm^{-1}.

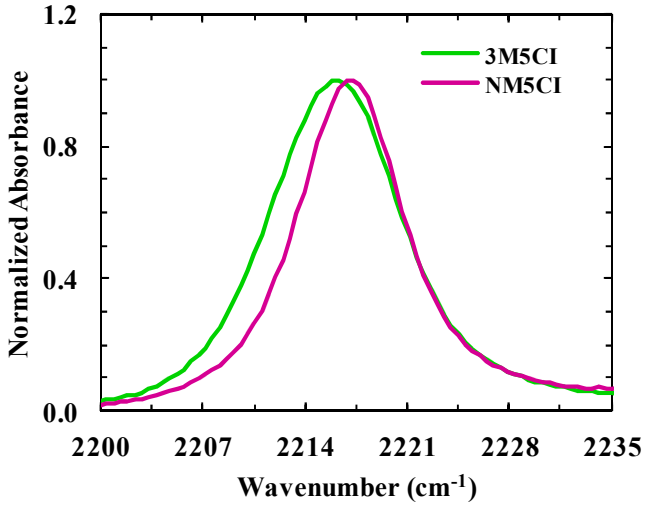


Figure 8.11 Comparison of the C≡N stretching bands of NM5CI and 3M5CI, as indicated, in DMSO.

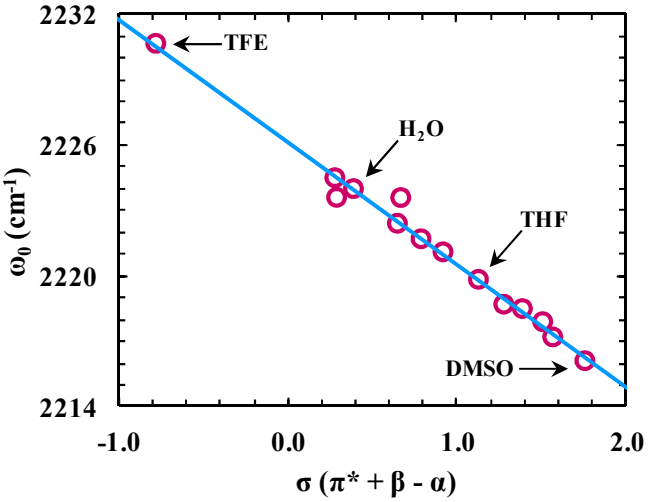


Figure 8.12 Center frequency (ω_0) of the C≡N stretching band of 3M5CI versus the solvent σ parameter, where $\sigma = \pi^* + \beta - \alpha$. The solid line represents the best fit of this data to a line with a slope of -5.6 ± 0.2 cm^{-1} and an intercept of 2226.2 ± 0.2 cm^{-1}.

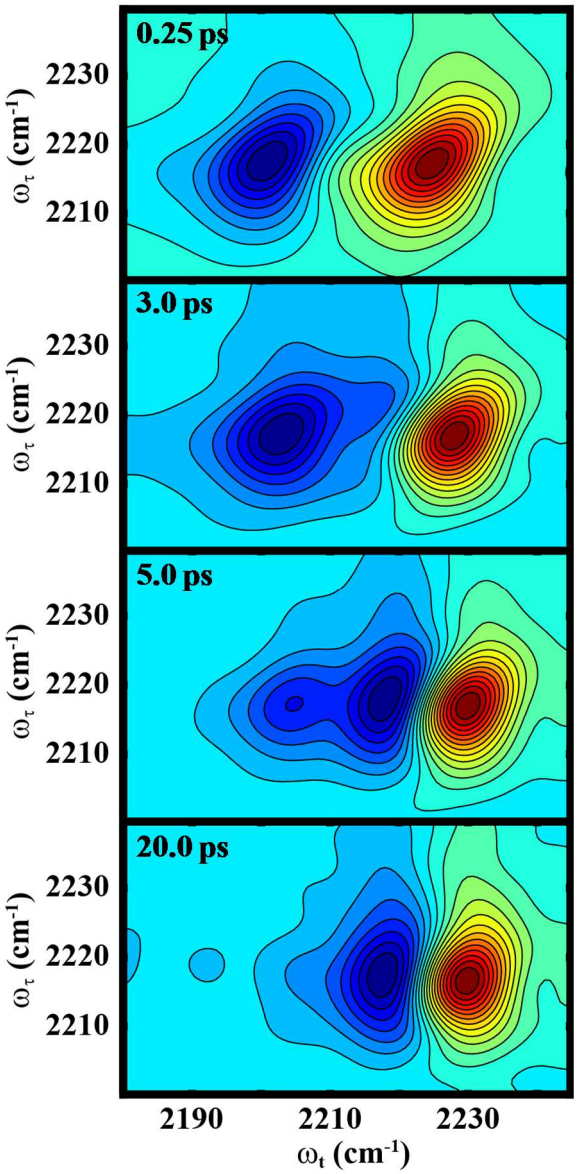


Figure 8.13 2D IR spectra of 3M5CI in a 50:50 TFE:DMSO mixture, at various waiting times (*T*), as indicated.

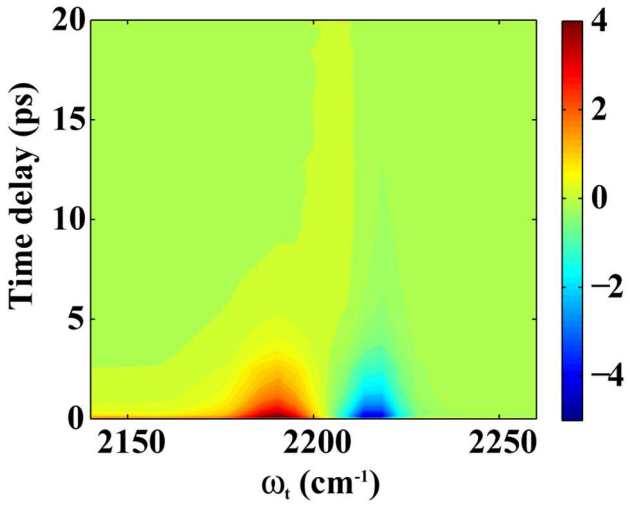


Figure 8.14 Time-resolved absorption spectra of 3M5CI in DMSO.

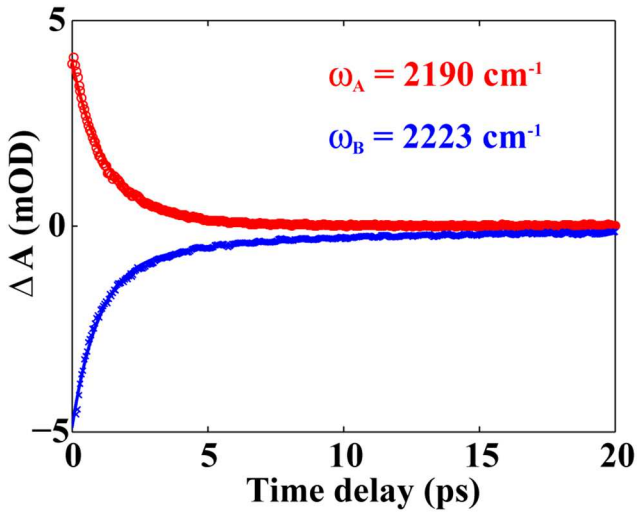


Figure 8.15 Transient absorption kinetics obtained at 2190 cm^{-1} (red circle) and 2223 cm^{-1} (blue cross). The transient data at 2190 cm^{-1} can be fit to a single exponential function (red line) with a time constant of 1.3 ± 0.1 ps. On the other hand, the kinetics at 2223 cm^{-1} can be best described by a double-exponential function (blue line) with the following time constants (amplitude): 1.2 ± 0.1 ps (3.8 ± 0.1 mOD) and 12.3 ± 1.6 ps (0.7 ± 0.1 mOD).

9 Utility of 5-Cyanotryptophan Fluorescence as a Sensitive Probe of Protein Hydration

Reprinted with permission from Journal of Physical Chemistry B, Beatrice N. Markiewicz,‡ Debopreeti Mukherjee,‡ Thomas Troxler, and Feng Gai, (2016) 120, 936-944. DOI: 10.1021/acs.jpcb.5b12233.[56] (‡ Denotes equal authorship)

9.1 Introduction

Among the naturally occurring fluorescent amino acids, tryptophan (Trp) is the most widely used fluorescent probe of protein structure, function, and dynamics.[53,54] This is because 1) its fluorescence properties, such as the emission wavelength, Stokes shift and lifetime, depend on local environment,[54,441,442] 2) its fluorescence can be quenched by various amino acid sidechains[443-445] as well as other molecules or ions,[446-450] and 3) it has a relatively high fluorescence quantum yield (QY) (i.e., ~0.14 in water)[451] and large molar extinction coefficient for transitions to the ^1L_a and ^1L_b excited states, which combined, allows for measurements using dilute protein solutions. While Trp is an exceedingly useful and convenient fluorescence reporter of proteins, it affords, like any other spectroscopic probes, certain limitations and/or disadvantages. For example, in practice it is often difficult to quantitatively assess and interpret Trp fluorescence results, as many different mechanisms can contribute to the excited-state decay process of the indole fluorophore in a protein environment.[452-455] Therefore, much effort has been made to expand and/or improve the utility of Trp fluorescence by exploring the feasibility of using various Trp-based non-natural amino acids, such as 7-azatryptophan,[456] 5-hydroxytryptophan,[457] 4- or 5-fluorotryptophan,[458,459] 6- or 7-cyanotryptophan,[57] and β-(1-Azulenyl)-L-alanine,[460] that possess different photophysical properties from Trp (e.g.,

quantum yield, emission wavelength, and fluorescence decay kinetics). Herein, we continue this effort by showing that the fluorescence emission of a nitrile-derivatized Trp analog, 5-cyanotryptophan (Trp_{CN}), can be used as a sensitive probe of the local hydration status of proteins.

From a practical point of view, for many biological applications a good fluorescence probe should exhibit a significant change in its fluorescence spectrum (i.e., intensity, wavelength, or both) when exposed to different environments. This requirement is especially important for examining processes, such as protein folding and interactions, whereby the local surrounding of the fluorophore undergoes a drastic change, for instance, from a hydrated to a dehydrated environment. While Trp fluorescence has been extensively used in these types of applications, the total fluorescence intensity change is rarely more than two-fold. This is due in part to the fact, as shown (Figure 9.1), that the fluorescence QY of indole is not sensitively dependent on hydration. Of course, one can find special cases where the change is more drastic due to involvement of other quenching mechanisms, such as those arising from specific sidechains.[444] In comparison, the fluorescence QY of 5-cyanoindole (5CI), the sidechain of Trp_{CN}, exhibits a much stronger sensitivity to interactions with H_2O as shown below. This is supported by the fact that upon changing the solvent from water to 1,4-dioxane, a solvent commonly used to mimic the hydrophobic interior of proteins, the total fluorescence intensity of 5CI is increased by a factor of approximately 23, whereas that of indole only increases by a factor of ~1.3 (Figure 9.1). Furthermore, a previous study by Jennings *et al.*[461] demonstrated that the maximum of the fluorescence spectrum of 5CI is shifted from ~315

nm in isopentane to ~391 nm in H_2O. Thus, taken together, the large Stokes shift and a more significant QY change of 5CI in response to hydration suggests that Trp_{CN} could be a more sensitive protein hydration probe than Trp.

To verify this notion, we carried out steady-state and time-resolved fluorescence measurements on 5CI, Trp_{CN}, and a model tripeptide Gly-Trp_{CN}-Gly (hereafter referred to as $GW_{CN}G$) in different solvents. We found, as expected, that exposure to water results in significant quenching of the Trp_{CN} fluorescence, due to a substantial increase in its excited-state nonradiative decay rate. On the other hand, the fluorescence QY and lifetime of Trp_{CN}, either in the free amino acid form or in a peptide environment, were found to increase, on average, by more than an order of magnitude in aprotic solvents compared to those in H_2O. Combined, we believe that these results support the idea that Trp_{CN} can be used as a sensitive fluorescence probe of the local hydration status of proteins. Further evidence confirming this utility of Trp_{CN} comes from several applications wherein we demonstrated that this nonnatural amino acid can be used to 1) detect the preferential accumulation of dimethyl sulfoxide (DMSO) molecules around aromatic sidechains in a disordered peptide, 2) probe the binding of an antimicrobial peptide to lipid membranes, and 3) differentiate the microenvironments of two Trp_{CN} residues in a folded protein.

9.2 Experimental Section

Materials and Sample Preparation

5-cyanoindole (5CI) at 99% purity was purchased from Acros Organics (Morris Plains, New Jersey), Fmoc-5-cyano-L-tryptophan with a purity of >99% was purchased from RSP amino acids (Shirley, MA), and all other amino acids were purchased from Advanced ChemTech (Louisville, KY). The following solvents (spectroscopic grade) were purchased from Acros Organics: methanol (MeOH), ethanol (EtOH), dimethyl sulfoxide (DMSO), acetonitrile (ACN), 1,4-dioxane, tetrahydrofuran (THF, without the BHT preservative), and 2,2,2-trifluoroethanol (TFE). D_2O was purchased from Cambridge Isotope Laboratories, Inc. (Tewksbury, MA). Dodecylphosphocholine (DPC) was purchased from Avanti Polar Lipids Inc. (Alabaster, Alabama). All materials and solvents were used as received. Deprotection of the Fmoc-5-cyanotryptophan to produce the free Trp_{CN} amino acid (Trp_{CN}-NH_2) and synthesis of all peptides (Gly-Trp_{CN}-Gly-NH_2, TZ2W_{CN}-NH_2, TC2W_{CN}-NH_2, and MPXW_{CN}-NH_2) were achieved by using standard 9-fluorenylmethoxy-carbonyl (Fmoc) solid-state methods on a CEM (Matthews, NC) Liberty Blue automated microwave peptide synthesizer. Peptide purification was done by reverse-phase HPLC (Agilent Technologies 1260 Infinity) with a C18 preparative column (Vydac). The Trp_{CN} amino acid and all peptides were constructed on Rink amide resin and thus contained an amidated C-terminus. The mass of every peptide was verified by either liquid-chromatography mass spectrometry (LC-MS) or matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) where appropriate. Samples were freshly prepared before use by directly dissolving the desired compound in the desired solvent, and the final concentration of the solute was approximately 45 μM for static and time-resolved measurements. The UV-Vis spectra (in the region of 250 –

300 nm) of a 5-cyanoindole solution and an indole solution in methanol of equal concentration (determined by weight) are almost identical. Thus, for each sample we used its absorbance and the molar extinction coefficients of Trp at 280 nm ($\varepsilon = 5500\ M^{-1}\ cm^{-1}$) to estimate the solute concentration. Membrane-bound $MPXW_{CN}$ peptide was prepared by solubilizing $MPXW_{CN}$ with DPC in TFE at a 1:70 peptide to lipid ratio. The organic solvent was removed via a nitrogen stream and the left-over film was lyophilized for at least 4 hours to ensure complete solvent removal. Subsequently, the resultant dry film was redissolved in H_2O. The final concentration of the peptide was ~40 μM.

Static and Time-Resolved Fluorescence Measurements

Static fluorescence measurements were obtained with a Jobin Yvon Horiba Fluorolog 3.10 spectrofluorometer at room temperature in a 1 cm quartz cuvette with a 1.0 nm resolution and an integration time of 1.0 nm/s. For all the static measurements, an excitation wavelength of 280 nm was used. Time-resolved fluorescence measurements were collected on a time-correlated single photon counting (TCSPC) system with a 0.4 cm quartz cuvette at 25 °C. The details of the TCSPC system have been described elsewhere.[462] Briefly, a home-built femtosecond Ti:Sapphire oscillator operating at 800 nm and 85 MHz repetition rate was used to generate a 270 nm excitation pulse train in a home-built collinear third harmonic generator. Repetition rate was reduced to 21 MHz by using an electro-optical pulse picking system (Conoptics Inc.). Emission was collected at magic angle polarization condition and in a 90 degree geometry relative to excitation, selected by a short-wavelength bandpass filter (Semrock FF01-357/44) around 360 nm and a long-pass filter (Semrock FF01-300/LP) with a 300 nm cutoff to better suppress

scattered excitation light, and detected with a MCP-PMT detector (Hamamatsu R2809U) and a TCSPC PC-board (Becker and Hickl SPC-730). Fluorescence decays were deconvoluted with the instrument response function (IRF) and fit either to a single-or multi-exponential function in order to minimize χ^2 below an acceptable value (i.e., 1.2) using FLUOFIT (Picoquant GmbH). *N*-acetyl-L-tryptophanamide (NATA) was used as a standard and control, yielding a single exponential decay of 3.0 ns (see Table 9.1) in accordance with the literature.[452] The optical density of the samples at the excitation wavelength was equal to or below 0.2 for both static and time-resolved measurements. Furthermore, the quantum yield of a given sample (QY_S) was determined using the quantum yield of NATA (QY_R) in H_2O at pH 7.0 as a reference and the following equation,

$$QY_S = QY_R \frac{I_S}{I_R} \frac{A_R}{A_S} \frac{n_S^2}{n_R^2} \tag{9.1}$$

where *I* and *A* represent the integrated area of the fluorescence spectrum and the optical density of the sample at 280 nm, respectively, and *n* is the refractive index of the solvent used. The subscripts S and R represent sample and reference, respectively. In addition, the value of QY_R was taken as 0.14,[2] and 1.333 was used for n_R.

9.3 Results and Discussion

To assess the feasibility of using Trp_{CN} fluorescence as a reporter of protein local hydration status, we first systematically examined the steady-state fluorescence properties and fluorescence decay kinetics of 5CI, Trp_{CN}, and $GW_{CN}G$ in a series of solvents with different polarities and hydrogen bonding abilities (Table 9.1). Then, the utility of this

fluorescence probe was tested in three different applications, including detection of a specific solute-solvent interaction, peptide-membrane association, and the hydration status of two Trp_{CN} residues located in different environments within a folded mini-protein.

Steady-State Fluorescence Measurements

As shown (Figure 9.2), both the intensity and maximum wavelength (λ_{max}) of the fluorescence spectrum of 5CI exhibit a strong dependence on solvent. Qualitatively, the λ_{max} of 5CI increases with increasing solvent polarity (Table 9.1), which is in good agreement with the trend previously observed by Jennings *et al.*[461] Furthermore, the λ_{max} of 5CI is red-shifted compared to the fluorescence spectrum of indole, the fluorophore of Trp, obtained in the same solvent. For example, in H_2O the fluorescence spectrum of indole is peaked at ~352 nm, whereas that of 5CI has an emission maximum at ~387 nm. The greater Stokes shift of 5CI can be attributed to the larger change in its dipole moment upon photoexcitation.[463] In addition, and perhaps more importantly, in solvents that can form strong hydrogen bonds (H-bonds), such as H_2O and trifluoroethanol (TFE), the fluorescence QY of 5CI exhibits a significant decrease (Figure 9.2 and Table 9.1). For example, in 1,4-dioxane the fluorescence QY of 5CI is determined to be ~0.13, which is decreased to ~0.005 in H_2O. In comparison, while indole is a brighter fluorophore, its fluorescence QY has less of a dependence on these solvents (i.e., 0.45 in 1,4-dioxane versus 0.28 in H_2O).[464] Thus, these results indicate that, when used to probe a hydration/dehydration event, 5CI would be able to produce a fluorescence signal with a higher contrast than that of indole.

As shown (Figure 9.2), the fluorescence spectra of the free amino acid Trp_{CN} are similar to those of 5CI obtained in the same solvents, although the λ_{max} value in each case is further red-shifted (Table 9.1). This is consistent with the trend already observed for indole and Trp.[55] What is more important, however, is that the strong fluorescence quenching effect of H_2O toward the 5CI fluorophore is maintained (Table 9.1). This result, which is consistent with a previous study,[155] indicates that Trp_{CN} fluorescence could be used to sense the local hydration status of proteins. Results obtained with $GW_{CN}G$ (Figure 9.2 and Table 9.1), also corroborate this notion. For example, the fluorescence QY of Trp_{CN} in $GW_{CN}G$ is increased by more than an order of magnitude upon changing the solvent from H_2O to MeOH. Since MeOH is also capable of forming hydrogen bonds, this finding thus manifests the specific and high sensitivity of Trp_{CN} fluorescence towards H_2O.

Time-Resolved Measurements

To further understand the photophysics of Trp_{CN}, we set out to measure the fluorescence decay kinetics of 5CI, Trp_{CN}, and $GW_{CN}G$ in those aforementioned solvents. As indicated (Figure 9.3 and Table 9.1), the fluorescence decay kinetics of 5CI in all solvents could be fit reasonably well to a single-exponential function (i.e., $\chi^2 < 1.2$), with the exception of those obtained in H_2O and TFE, which required at least a double-exponential function to yield a satisfactory fitting (i.e., $\chi^2 < 1.2$). Consistent with the steady-state measurements, the (intensity weighted) fluorescence lifetimes (0.1 – 0.3 ns) of 5CI in H_2O and TFE are significantly shorter (by at least an order of magnitude) than those in other solvents, which are in the range of 3.0 – 7.1 ns. This finding is, to some extent, surprising, as for

indole the fluorescence lifetime in TFE (0.45 ns) is drastically shorter than in H_2O (4.5 ns).[465] Barkley and coworkers[465,466] have shown that both TFE and H_2O can quench the fluorescence of indole via an excited-state proton-transfer process. While we cannot completely rule out the possibility that a solvent-induced proton-transfer event is responsible for the observed fast excited-state decay kinetics of 5CI in H_2O and TFE, the fact that, unlike indole,[465] the fluorescence QY and decay kinetics of 5CI do not show any measurable difference between H_2O and D_2O (Figure 9.4 and Table 9.1) strongly argues against this scenario. Since both H_2O and TFE can form H-bonds with the nitrile group of 5CI,[403] it is possible that their strong quenching effect arises solely from such H-bonding interactions. However, this possibility can also be ruled out as in both MeOH and EtOH, which are able to form such H-bonds, the fluorescence lifetime of 5CI becomes much longer (Table 9.1). Similarly, the H-bonding interactions between the pyrrole N-H group of 5CI and a solvent molecule, such as H_2O, MeOH, EtOH, and DMSO, are also unlikely to serve as a major nonradiative decay channel as the fluorescence lifetimes of 5CI in those solvents differ significantly (Table 9.1). It is known that both H_2O and TFE can interact with the indole moiety via another type of H-bonding interactions, i.e. those formed between the –OH group of the solvent and the π-electron cloud of the aromatic ring.[423,467,468] In fact, this type of interactions has been suggested to play an important role in the excited-state decay kinetics of indole in TFE.[469] Thus, we tentatively attribute the sub-nanosecond fluorescence decay kinetics of 5CI in H_2O and TFE to such H-bond formations, which increase the nonradiative decay rate of the fluorophore. If this assessment is indeed valid, the much slower fluorescence decay

lifetime of 5CI obtained in other protic solvents (i.e., MeOH and EtOH) indicates that the indole ring is solvated mainly by the methyl groups rather than the –OH group, which is consistent with previous studies.[470,471] In addition, it is noticed that, unlike those obtained in other solvents, the fluorescence decay kinetics of 5CI in H_2O and TFE cannot be satisfactorily described by a single-exponential function (Figure 9.3). There are two possible interpretations. First, 5CI has a relatively low solubility in both H_2O and TFE; hence the non-single-exponential decay kinetics may reflect the heterogeneity of the solute. The second, and perhaps a more probable interpretation is that this deviation manifests the heterogeneity in the solvent-solute interactions, especially those –OH···π-electron H-bonding interactions.

Finally, unlike that of indole,[55] the fluorescence lifetime of 5CI in DMSO is significantly lengthened (i.e., to 7.1 ns) in comparison to those in other solvents, suggesting that Trp_{CN} fluorescence could be used to probe preferential interactions[120] between DMSO and Trp sidechains in a protein environment (see below).

As indicated in Figure 9.3, the fluorescence decay kinetics of Trp_{CN}, in either the free amino acid form or the $GW_{CN}G$ peptide, support the notion that it can be used as a protein hydration probe. While the average fluorescence lifetime of Trp_{CN} obtained in H_2O is slightly increased compared to that of 5CI, its excited-state decay kinetics are still dominated by a fast component (i.e., 0.4-0.5 ns), which is significantly separated from those obtained in other solvents, except TFE (Table 9.1).

Unlike 5CI, Trp_{CN} exhibits non-single-exponential fluorescence decay kinetics in all the solvents studied. Since Trp shows the same behavior and its double- or multi-

exponential fluorescence decay kinetics have been attributed to different sidechain rotamers,[455] we believe that the fluorescence decay kinetics of Trp_{CN} can also be explained by such a rotamer model. Many previous studies have shown that the average fluorescence decay rate of Trp is faster than that of indole under the same solvent conditions, indicating that an additional, nonradiative decay pathway exists due to the presence of a backbone. The current consensus is that this new decay process arises from electron transfer from the indole sidechain to the carbonyl group of adjacent peptide bonds.[454] As shown (Table 9.1), unlike the trend observed for indole and Trp,[55] the average fluorescence lifetime of Trp_{CN} is longer than that of 5CI in the same solvent. More specifically, in most solvents (except H_2O and TFE) the fluorescence decay kinetics of Trp_{CN} and $GW_{CN}G$ consist of two components, with one that is relatively independent of solvent and decays within 2-3 ns, whereas the other, which is slower (τ_F > 5.5 ns) and dominant in most cases, shows a clear dependence on solvent environment. Taken together, these results suggest that the electron-transfer mechanism invoked to explain the fluorescence decay kinetics of Trp is not applicable to Trp_{CN}. Instead, we believe that the fast component arises from a more specific interaction facilitated by differences in geometry between the backbone and a particular sidechain rotamer. While elucidation of the nature of this interaction requires further studies, we hypothesize that it corresponds to a H-bonding-like interaction between the aromatic ring and a backbone amide N-H group. This hypothesis is consistent with the aforementioned assumption that H_2O quenches the fluorescence of 5CI via H-bonding interactions with the π–electron cloud of the ring. In addition, this model is self-consistent, as for rotamers that cannot

engage in such additional interactions due to distance constraints, their fluorescence lifetimes would be longer and depend mostly on the solvent, as observed.

Probing Preferential Interactions with DMSO

To further demonstrate the utility of Trp_{CN} fluorescence as a local solvation reporter, we first applied it to probe the preferential interaction between DMSO and Trp residues within a peptide. The binary mixture of water and DMSO has been extensively studied and used in various applications owing to its non-ideal nature and hence unusual properties.[472] For example, it has been used as a cryoprotectant, an enzyme activator, and a denaturant.[473,474] In particular, several studies[475-477] have indicated that in a binary mixture of water and DMSO preferential solvation of hydrophobic residues, such as Phe and Trp, by DMSO can occur. All of these behaviors can be attributed to the amphiphilic nature of DMSO which allows this solvent to engage in both H-bonding (e.g., with water) and hydrophobic interactions (via methyl groups).[478] As demonstrated above, the fluorescence QY of Trp_{CN} in DMSO is approximately 24 times greater than its QY in H_2O, making it ideally suited to probe any preferential accumulation of DMSO molecules around its indole ring. Specifically, we measured the fluorescence spectra of a peptide that contains two Trp_{CN} residues (sequence: S-Trp_{CN}-TAENGKAT-Trp_{CN}-K), in a series of DMSO-H_2O mixtures. This peptide (hereafter referred to as $2W_{CN}P$) is largely unstructured in aqueous solution according to its CD spectrum (Figure 9.5), thus allowing the Trp_{CN} sidechains to be solvent accessible.

As shown (Figures 9.6 and Figure 9.7), the intensity of the Trp_{CN} fluorescence spectrum of $2W_{CN}P$ depends strongly on the mole fraction of DMSO (χ_{DMSO}) in the

binary mixture, especially in the range of 0.0-0.4. A more quantitative analysis indicates that the relative fluorescence quantum yield of Trp_{CN} in $2W_{CN}P$, as measured by the integrated area (I) of the fluorescence spectrum, exhibits a transition (Figure 9.6) that is similar to that observed in substrate binding kinetics of enzymes. Indeed, this transition (i.e., I vs. χ_{DMSO}) can be satisfactorily described by a modified Hill equation shown below,

$$I = I_{min} + \frac{(I_{max} - I_{min})(a\chi_{DMSO} + b)}{\left(\frac{k}{\chi_{DMSO}}\right)^n + 1} \tag{9.2}$$

where I_{min}, I_{max}, k, a, b, and n (the Hill coefficient) are constants. The linear term, $(a\chi_{DMSO} + b)$, is introduced to account for the slight downward trend of the signal after saturation. As shown in Figure 9.6, the best fit yields a k value of 0.18, indicating that the fluorescence signal reaches a maximum at a χ_{DMSO} value of 0.36. Interestingly, a previous dielectric relaxation study[478] indicated that the maximum H-bonding interactions occurring between H_2O and DMSO take place at χ_{DMSO} = 0.33, leading to formation of H_2O-DMSO-H_2O complexes. Thus, the above result suggests that the fluorescence of Trp_{CN} is able to 'sense' this intrinsic property of the binary mixture. In addition, and perhaps more interestingly, the Hill coefficient obtained from the fitting is n = 3.2, suggesting that the DMSO binding interaction with the indole ring is highly cooperative. Following the common interpretation of the value of the Hill coefficient,[479] the above n value suggests that one Trp_{CN} sidechain can have a maximum of three DMSO binding sites. Hence, from this simple example it becomes evident that Trp_{CN} is not only a useful

fluorescence probe of proteins or peptides, but it can also be used to reveal important physical properties of binary solvent systems that contain water.

Probing Peptide-Membrane Interaction

In the second application, we demonstrated that Trp_{CN} fluorescence can be used to probe peptide-membrane interactions. To do so, we measured the static and time-resolved fluorescence properties of a Trp_{CN} mutant of an antimicrobial peptide, mastoparan X, in the presence and absence of a model membrane. Like the parent,[480] this mutant (sequence: IN-Trp_{CN}-KGIAAMAKKLL), hereafter referred to as $MPXW_{CN}$, is relatively unstructured in aqueous solution, and folds into an α-helical conformation upon binding to DPC micelles (Figure 9.8). As shown (Figure 9.9a and Table 9.1), in the absence of DPC micelles the Trp_{CN} fluorescence spectrum of $MPXW_{CN}$ is peaked at 391 nm and has a low intensity, which is consistent with an unfolded peptide wherein the Trp_{CN} sidechain is mostly exposed to H_2O. In the presence of DPC micelles, however, the Trp_{CN} fluorescence spectrum of $MPXW_{CN}$ not only is blue-shifted (to 372 nm) but also becomes much more intense, which is characteristic of a Trp_{CN} buried in a hydrophobic environment and, hence, consistent with binding of the peptide to the DPC membranes. Further fluorescence lifetime measurements also corroborate this picture. As shown (Figure 9.9b and Table 9.1), the Trp_{CN} fluorescence decay of $MPXW_{CN}$ in H_2O consists of three components, with a sub-nanosecond (i.e., 0.7 ns) component being dominant (~82%). It is clear that this fast decay component corresponds to an ensemble of $MPXW_{CN}$ conformations wherein the Trp_{CN} residue is well hydrated, whereas those slower and minor decay components must arise from conformations wherein the Trp_{CN}

sidechain is less exposed to solvent due, for example, to sidechain-sidechain and/or sidechain-backbone interactions. As expected, in the presence of DPC micelles, the Trp_{CN} fluorescence decay of $MPXW_{CN}$, now dominated by a 7.2 ns component, is similar to those measured for the $GW_{CN}G$ peptide in aprotic solvents (Table 9.1). Taken together, the above results confirm that both the fluorescence intensity and lifetime of a Trp_{CN} residue can be used to probe various biological binding interactions, as long as its hydration status undergoes a change in response to the binding event in question.

Probing Local Hydration Environment

In the third application, we demonstrated that the high sensitivity of the fluorescence lifetime of Trp_{CN} to H_2O makes it especially useful to differentiate between hydrated and dehydrated environments in a protein. Specifically, we carried out fluorescence lifetime measurements on a double Trp_{CN} mutant of a miniprotein, Trp^2-cage.[481] As suggested by its name, this miniprotein, which was computationally designed based on the miniprotein Trp-cage,[258] contains two Trp residues, with one (at position 12) being solvent exposed and the other (at position 6) being buried in a hydrophobic cage.[258] Upon replacing these two Trp residues with Trp_{CN}, we expected that the fluorescence decay kinetics of the resultant double mutant (sequence: NLYIQ-$^6Trp_{CN}$-LKDGG-$^{12}Trp_{CN}$-SSGRPPPS), hereafter referred to as $TC2W_{CN}$, will reflect this difference. As shown (Figure 9.10 and Table 9.1), the fluorescence decay kinetics of $TC2W_{CN}$ in H_2O consist of three exponential components with time constants of 0.4, 1.8, 11.0 ns, respectively. The first and very fast component is similar to the major fluorescence decay component of $GW_{CN}G$ in H_2O, indicating contribution from, and hence detection of, the solvent-

exposed Trp_{CN} residue, whereas the slowest component is closer to that of $GW_{CN}G$ in DMSO, representative of the buried and solvent inaccessible Trp_{CN} residue in the protein. Because both the solvent-exposed and buried Trp_{CN} residues can contribute to the 1.8 ns component in the fluorescence decay kinetics of $TC2W_{CN}$, it is not straightforward to use the relative amplitudes of those three exponentials to directly determine the relative populations of these two differently solvated sidechains. However, if we only consider the other two exponentials (i.e., the 0.4 and 11.0 ns components), which exclusively report the solvent-exposed and buried Trp_{CN} species, and also assume that the radiative rate constant does not change, we can estimate the relative population of the solvent-exposed Trp_{CN} sidechain to be ~58%. As shown (Figure 9.11), the CD spectrum of $TC2W_{CN}$ indicates that it is folded. If its stability is assumed to be comparable to that of the wild-type, which has ~10% unfolded population at room temperature,[481] the solvent-exposed Trp_{CN} should amount to 55%. Thus, the estimate obtained above using the percentages of the 0.4 and 11.0 ns components is reasonable. Taken together, these results, especially the existence of two kinetic components that differ by almost 28 times in their time constants, clearly demonstrate that the fluorescence decay kinetics of Trp_{CN} are sensitive to the presence of H_2O and can be used to differentiate between differently hydrated environments in a protein.

9.4 Conclusions

Any spectroscopic study of protein folding, conformational transition and interactions requires a specific probe whose spectroscopic signature would undergo a change in response to a variation in its environment. In practice, an ideal spectroscopic probe is one

that can produce a large difference or contrast between signals measured before and after the event of interest. Herein, we demonstrated that the nonnatural amino acid, 5-cyanotryptophan, could be used as a sensitive fluorescence probe of proteins. This is because a series of static and time-resolved fluorescence measurements revealed that the fluorescence quantum yield and decay kinetics of 5-cyanotryptophan are sensitively dependent on its hydration status. For example, when fully hydrated its fluorescence quantum yield is approximately 0.01, whereas in a dehydrated environment its quantum yield is increased by at least an order of magnitude. Validation of the potential utility of this nonnatural amino acid as a sensitive local protein hydration reporter was demonstrated in three applications, wherein we used it to probe to probe complex formation in a binary mixture, peptide-membrane interactions, and the hydration environments of two tryptophan residues in a miniprotein. Given that 5-cyanotryptophan has also been shown to be a useful site-specific infrared (IR) probe of proteins,[142] we believe that the present work will further expand its utility as a novel spectroscopic probe to study the structure-dynamics-function relationship of proteins.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (P41GM104605). B.N.M is supported by an NIH Ruth Kirschstein National Research Service Award Predoctoral Fellowship (F31AG046010).

	Solvent	λ_{max} (nm)	# exp	%	τ_F (ns)	τ_{ave} (ns)	QY	χ^2
5CI	H_2O	387	2	91	0.1	0.3	0.005	1.19
				9	0.3			
	D_2O	385	2	88	0.1	0.4	0.002	0.99
				11	0.3			
	TFE	373	2	97	<0.1	0.1	0.007	1.05
				3	0.3			
	MeOH	366	1	100	2.9	2.9	0.06	1.04
	EtOH	366	1	100	4.3	4.3	0.09	1.14
	ACN	358	1	100	4.4	4.4	0.10	1.17
	1,4-dioxane	347	1	100	4.2	4.2	0.13	1.04
	THF	345	1	100	4.4	4.4	0.11	1.15
	DMSO	367	1	100	7.1	7.1	0.15	1.18
Trp_{CN}	H_2O	391	2	91	0.4	1.4	0.01	0.97
				9	1.6			
	TFE	386	3	53	<0.1	1.1	0.02	1.10
				44	0.7			
				3	2.8			
	MeOH	384	2	10	1.8	5.4	0.09	1.09
				90	5.6			
	EtOH	386	2	18	2.3	6.5	0.17	1.20
				82	6.8			
	ACN	381	2	67	3.1	5.5	0.13	1.07
				33	7.5			
	1,4-dioxane	361	2	25	3.4	5.6	0.11	1.11
				75	6.0			
	THF	368	2	18	2.5	5.6	0.12	1.07
				82	5.9			
	DMSO	391	2	21	2.1	16.1	0.24	1.06
				79	16.5			
$GW_{CN}G$	H_2O	394	2	96	0.5	0.8	0.01	0.93
				4	2.1			
	MeOH	383	2	24	0.8	5.7	0.13	1.16
				76	6.0			
	THF	371	2	17	3.0	5.3	0.14	1.04
				83	5.6			
	DMSO	386	2	26	2.0	13.9	0.24	1.13
				74	14.5			
$MPXW_{CN}$	H_2O	391	3	82	0.7	1.6	0.02	0.94
				17	2.0			
				1	6.1			
$MPXW_{CN}$	DPC	372	3	24	2.4	6.7	0.18	1.08
				76	7.2			
$TC2W_{CN}$	H_2O	387	3	40	0.4	9.2	0.06	1.18
				31	1.8			
				29	11.0			
NATA*	H_2O/pH 7	358	1	100	3.0	3.0	0.14	1.03

Table 9.1 Summary of results obtained from all static and time-resolved fluorescence measurements. QY was determined using the value of NATA as standard. τ_{ave} corresponds to the intensity weighted average fluorescence lifetime. *Values adopted from ref.53.

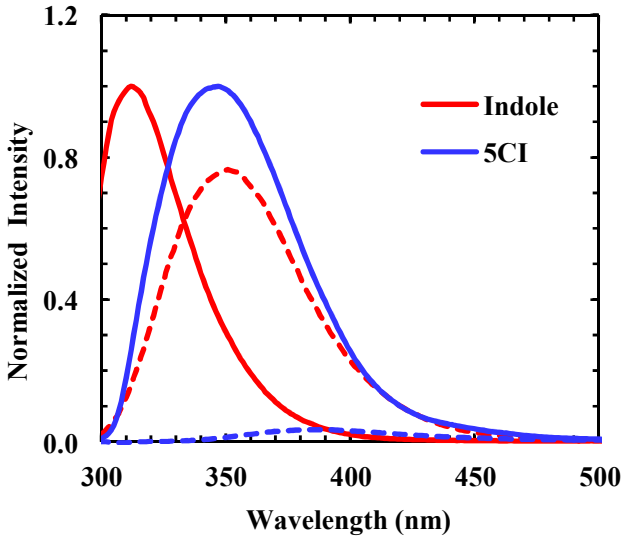


Figure 9.1 Normalized fluorescence spectra of indole and 5CI obtained in 1,4-dioxane (solid lines) and H_2O (dashed lines), as indicated.

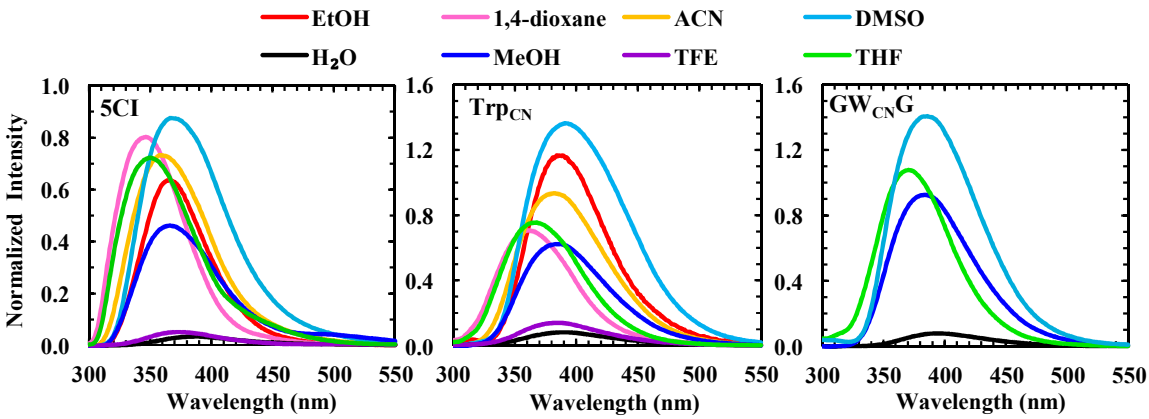


Figure 9.2 Fluorescence spectra of 5CI, Trp_{CN}, and $GW_{CN}G$ in different solvents, normalized against the fluorescence spectrum of NATA obtained in H_2O. In each case, the normalization factor was calculated based on the integrated areas of the corresponding fluorescence spectra.

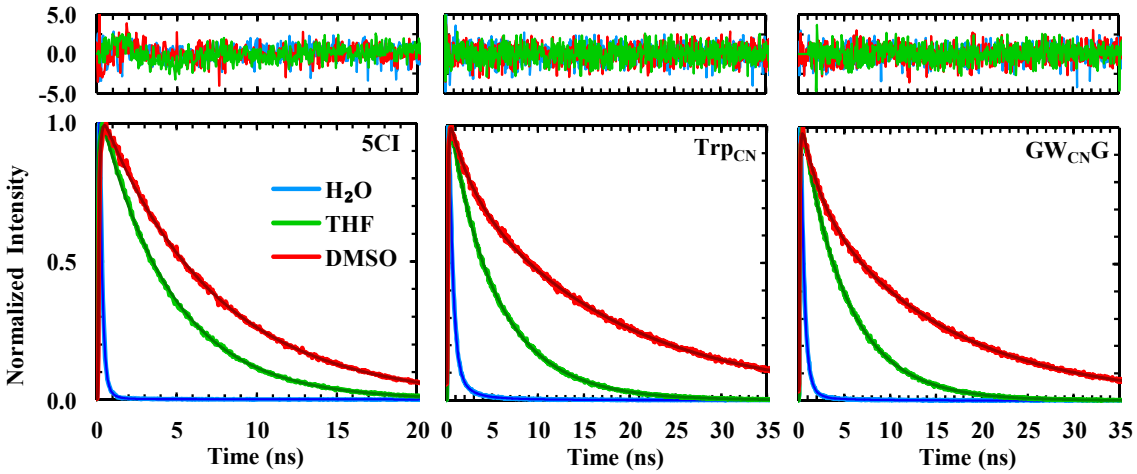


Figure 9.3 Fluorescence decay kinetics of 5CI, Trp_{CN}, and $GW_{CN}G$ in three representative solvents, as indicated. In each case, the smooth line corresponds to a fit of the kinetics to either a single- or double-exponential function and the resulting fitting parameters are listed in Table 9.1. Shown in the top panel are the residuals of the fits of the unnormalized data.

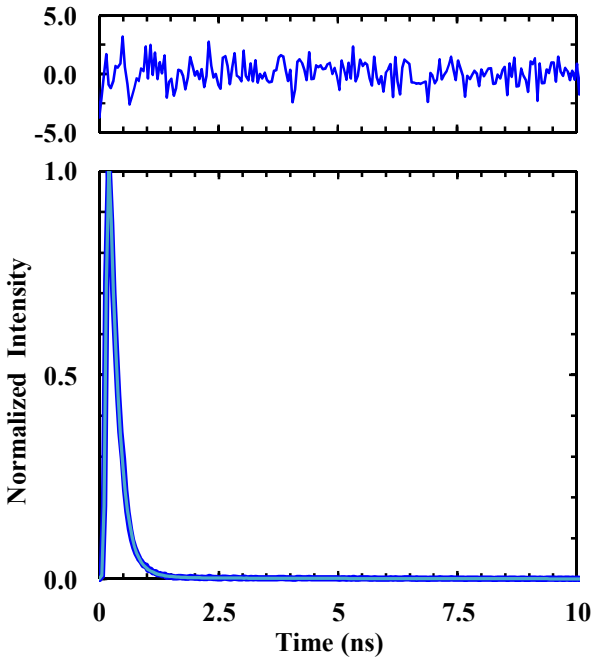


Figure 9.4 Fluorescence decay kinetics of 5CI in D_2O. The smooth line is a fit to a triple-exponential function. The residuals of the fits are shown in the top panel and the resulting fitting parameters are listed in Table 9.1.

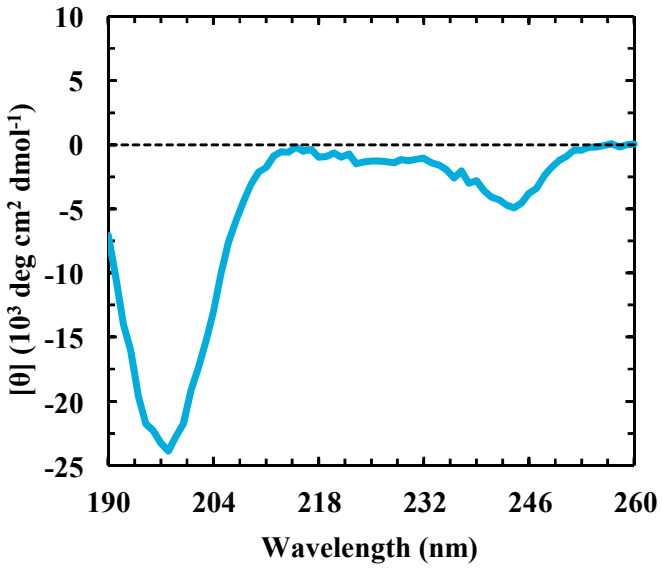


Figure 9.5 CD spectrum of $2W_{CN}P$ in H_2O at 1.0 °C.

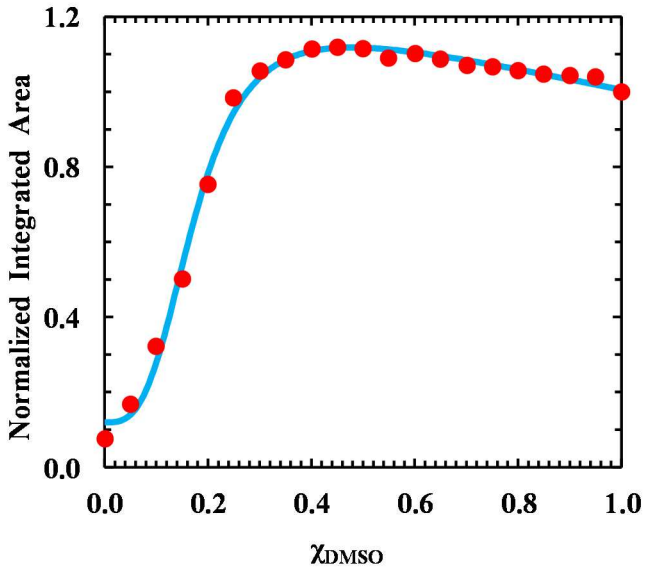


Figure 9.6 Integrated area of the Trp_{CN} fluorescence spectrum of $2W_{CN}P$ versus the mole fraction of DMSO (χ_{DMSO}) in the H_2O-DMSO binary solvent. These data have been normalized such that the value obtained in pure DMSO is 1.0. The smooth line is the best fit of the data to a modified Hill equation (i.e., Equation 9.2) with the following parameters: $I_{min} = 0.11 \pm 0.02$, $I_{max} = 1.10 \pm 0.01$, $k = 0.18 \pm 0.01$, $n = 3.20 \pm 0.29$, $a = -0.30 \pm .01$, and $b = 1.20 \pm 0.01$.

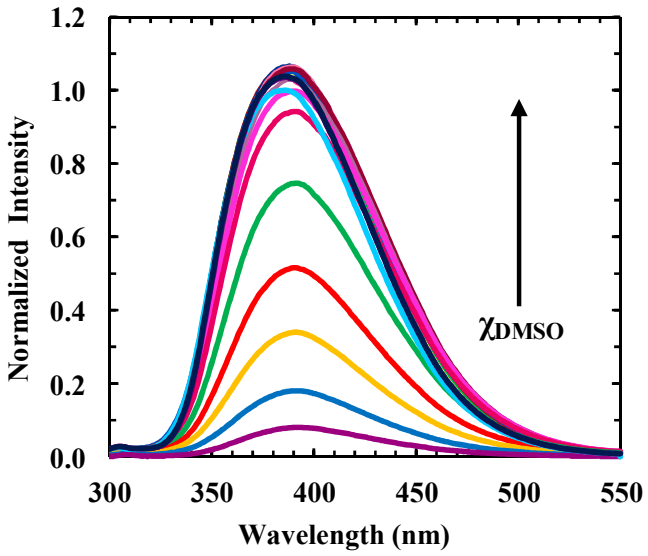


Figure 9.7 Fluorescence spectra of $2W_{CN}P$ in DMSO-H_2O mixtures with different molar fractions of DMSO (χ_{DMSO}).

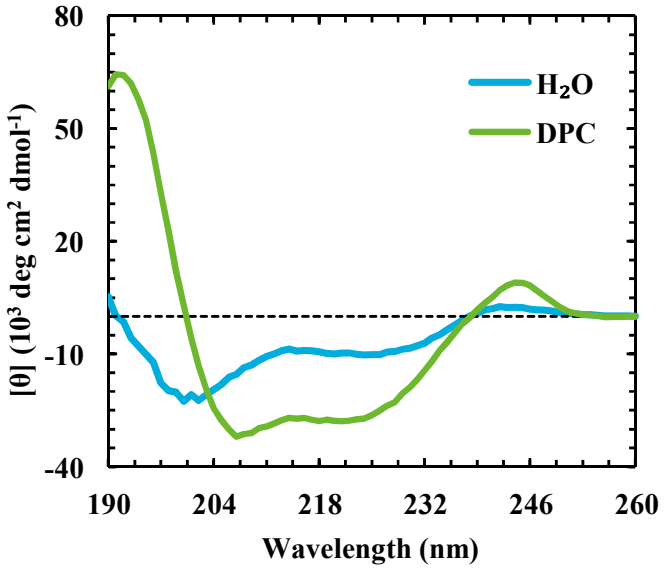


Figure 9.8 CD spectra of $MPXW_{CN}$ in H_2O and in DPC micelles at 25.0 °C, as indicated.

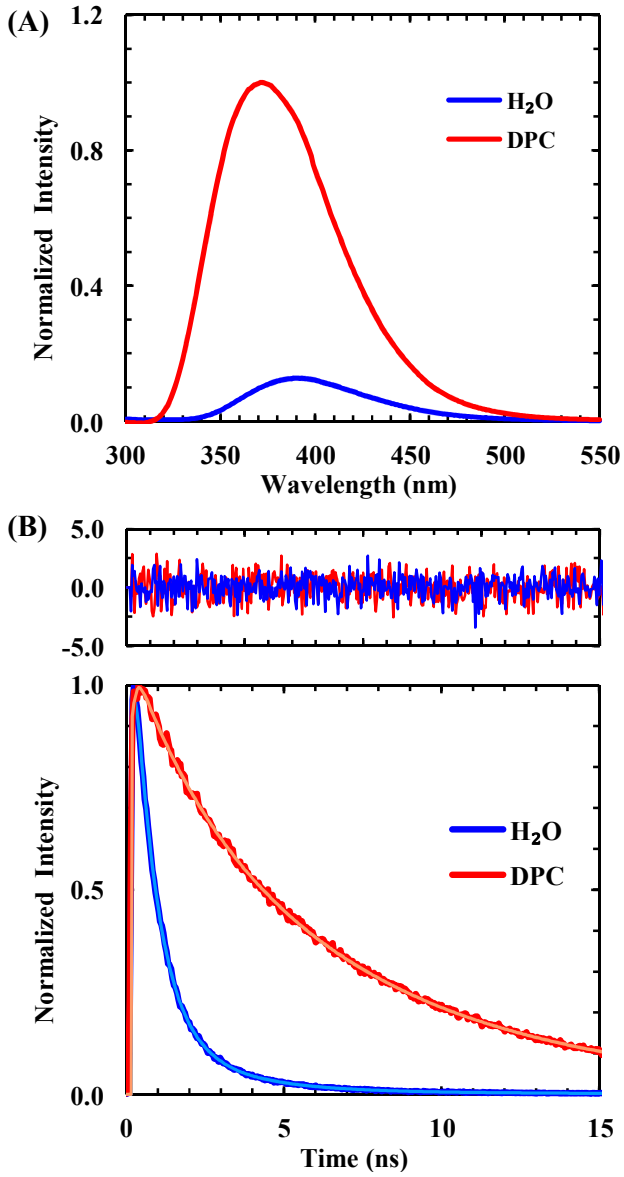


Figure 9.9 (A) Normalized fluorescence spectra of $MPXW_{CN}$ obtained in H_2O and DPC micelles, as indicated. The lipid to peptide ratio was 1:70 with a final peptide concentration of 40 μM. (B) Normalized Trp_{CN} fluorescence decay kinetics of $MPXW_{CN}$ in H_2O and DPC micelles, as indicated. In each case, the smooth line is the best fit of the data to a triple-exponential function and the resulting fitting parameters are listed in Table 9.1. Shown in the top panel are the residuals of the fits of the unnormalized data.

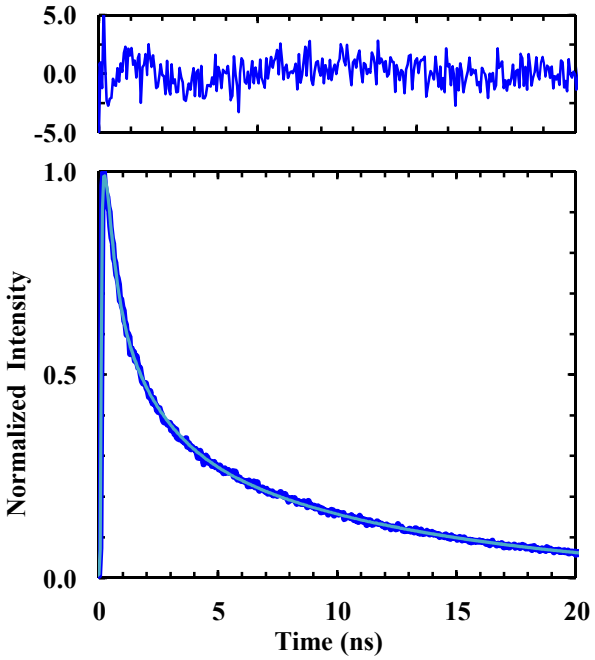


Figure 9.10 Normalized Trp_{CN} fluorescence decay kinetics of $TC2W_{CN}$ in H_2O. The smooth line is the best fit of the data to a triple-exponential function and the resulting fitting parameters are listed in Table 9.1. Shown in the top panel are the residuals of the fit of the unnormalized data.

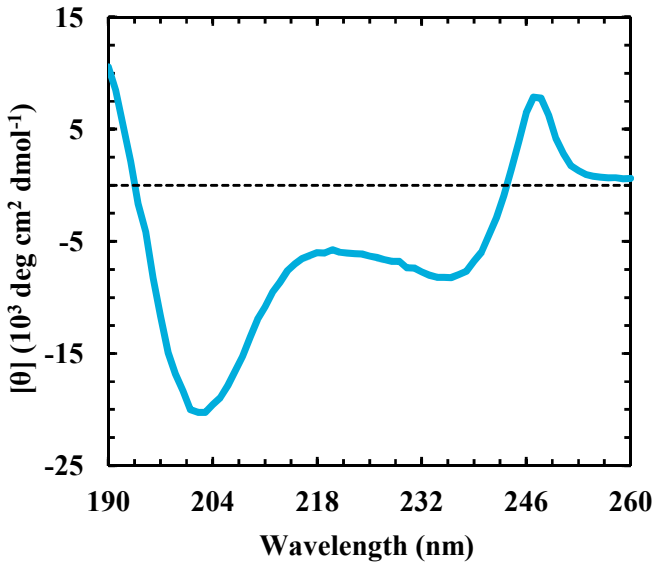


Figure 9.11 Far UV CD spectrum of TC2W$_{CN}$ in H_2O at 1.0 °C. The strong CD couplet at 247 and 236 nm, arising from the exciton coupling between the two Trp$_{CN}$ sidechains, indicates that the peptide is folded. It is worth noting that the Trp$_{CN}$–Trp$_{CN}$ exciton CD band is different from that of Trp–Trp and hence, can be used as a stand-alone structure reporter. We are currently exploring this potential application and will report the findings in a future publication.

10 Infrared and Fluorescence Measurements Provide Insight into the Low Proton Conductance of Influenza A M2 Protein

The text within this chapter is unpublished material by Beatrice N. Markiewicz, Wenkai Zhang, Thomas Lemmin, Hyunil Jo, William F. DeGrado, and Feng Gai. (2016). *Currently Manuscript is in Preparation for Submission.*

10.1 Introduction

The M2 protein of the influenza A virus forms an α-helical homotetrameric channel in the viral envelope, allowing selective proton conduction across the viral membrane upon endosomal acidification on virus entry.[482,483] Because it is vital to the viral life-cycle, the M2 proton channel has been the subject of extensive studies and also targeted for development of anti-influenza drugs.[484-488] It is well recognized that, lining the channel pore, His37 and Trp41 (shown in Figure 10.1), are the two most important residues that are responsible for channel opening and unidirectional or asymmetric proton conduction from the exterior into the interior of the virus.[489] In particular, the role of His37 is to select protons for conduction via its imidazole ring and to trigger the opening of the channel at low pH through protonation of approximately three His37 residues, whereas the Trp41 tetrad functions as a gate to allow asymmetric proton conductance only when the pH is low on the outside of the virus.[490-492] In addition, recent studies[434,484,493,494] indicate that Asp44 plays an essential role in stabilizing the Trp41 gate in the closed state via hydrogen bonding through a structured water network and perhaps more importantly, this stabilization helps the channel achieve its asymmetric proton conduction property.

Over the past decade, several models have been proposed to explain the mechanism of proton transport through the M2 channel. An early model suggests that in

the open state the channel hosts a water wire with which protons conduct via the Grotthuss mechanism.[495-497] A second model, the "shuttle mechanism", proposes that His37 is directly involved in proton conduction by cycling through a protonation-deprotonation process.[498-502] Similarly, a third model, involving the formation of a His37 dimer via hydrogen bonding in the closed state, suggests that protonation of the third His37 disrupts this dimer and promotes a His37-Trp41 cation-π interaction, which occasionally is disrupted by conformational motions of the helix backbone and/or Trp41 causing exposure of His37 to C-terminal water and hence proton release.[489] Finally, a transporter model,[434,503] which involves two distinct conformational ensembles with only one capable of proton conduction, proposes that acidification of the virus interior is achieved by oscillating between these two conformations, triggered by protonation and deprotonation of His37.

One particular property of the M2 channel that makes it unique is that it has a low proton conductance rate. Both whole-cell patch clamp and liposome experiments have put the proton conducting rate in the range of $10^1 – 10^4$ protons per second,[504-506] which is one to three orders of magnitude slower than expected for transmission of a proton through a pore of similar dimension at the same pH.[507] In particular, existing evidence indicates that, when scaled for changes in the permeant ion concentration, the proton conduction rate is in fact attenuated as the pH approaches 5.0.[507] This observation suggests that additional kinetic barriers arise for proton conduction at low pH due to structural fluctuations and/or protonation/deprotonation of His37 necessary for conduction.[504,508] For example, the recent electrophysiological study of DiFrancesco *et*

al.,[509] suggested a recycling step, where after proton release the channel resets its conformation by closing the Trp41 gate and opening the Val27 gate to prepare for proton intake from the exterior. On the other hand, others have proposed that a rate-limiting step arises from the necessity of tautomerization or ring reorientation of the imidazoliums in the low pH, conducting state.[498,502] This type of pre-equilibration is consistent with electrophysiological data showing that His37 equilibrates with protons from the outside of the virus prior to a rate-determining step, which allows protons to diffuse away from the His37 on the cytoplasmic side of the channel.[504] Indeed, Hong and coworkers[510] have determined that the His37-water proton exchange rate ($\sim 10^5$ s^{-1}) is significantly higher than the time-averaged unitary proton flux of M2, indicating that not every proton that is exchanged with His37 exits the channel. Williams *et al.*[435] proposed that such futile exchanges are caused by the His37-Trp41 cation-π interaction, which periodically forms and breaks at low pH. Yet, another possibility is that the channel resides in a predominantly closed state even at low pH and a conducting state is only transiently populated to allow proton release. In a theoretical treatment of M2 proton conductance Zhou found that in order to use a proton transfer model that takes into account the primary and secondary gating of Trp41 and Val27 as well as backbone dynamics to quantitatively describe the proton conductance data of M2, the C-terminal open conformation (i.e., the conformation that allows transfer of protons from His37 to the interior of the virus) needs to be a minor population (<10%).[511] To provide further insight into the proton conducting mechanism of M2, herein we approach the problem from the perspective of channel water.

It is well known that water plays a pivotal role in maintaining a viable pathway for proton conduction in biological systems.[512] While water is an integral part of the M2 channel, as revealed by a high-resolution crystal structure of the transmembrane (TM) domain of the M2 proton channel,[434] only a few previous studies[353,434,513-518] have focused on elucidating the structure and dynamics of the channel water. The current understanding is that the region above His37 is a water-filled pore; thus protons can diffuse up to His37 through this water pore via a Grotthuss mechanism. However, the structure and dynamics of the water in this region have been shown to depend on pH (and the binding of pore-blocking drugs): at high pH (>7), the channel water is more structured and less mobile in comparison to bulk water, whereas at low pH (e.g., 5.0), the water density near Gly34 is increased[353] and the water cluster exhibits bulk-like dynamics.[519] While these previous studies provide invaluable insight into the hydration environment of the region immediately above the key His37 tetrad, less is known about the structure and dynamics of the water clusters below the His37 gate, especially those surrounding the Trp41 tetrad. Recently, Voth and coworkers[517] have calculated the potential mean force of proton conduction across the transmembrane (TM) domain of M2 and identified two energy barriers, with one (~10 kcal/mol) associated with the proton passing through the His37-Trp41 section of the channel. Based on these results, they concluded that deprotonation of the His37 tetrad is the rate-determining step; however, the intrinsic proton release rate is modulated by the structure and dynamics of water below the His37 gate and/or conformational fluctuations of the channel. In addition to the role of water, previous resonance Raman[520] and electrophysiological measurements point

to a tight coupling between His37 and Trp41 at low pH for the most common variants of M2 that contain Asp at position 44.[494] However, other spectroscopic and crystallographic studies suggest that the His37/Trp41 interaction might be largely broken at low pH.[435,521] Thus, considering these previous studies, it would be quite useful to design an experiment that would allow one to directly assess the dynamics of water surrounding the Trp41 tetrad as well as the His37/Trp41 interactions under different conditions.

In order to use infrared (IR) spectroscopy to interrogate the water environment and local electrostatics directly surrounding His37 and Trp41, a site-specific IR probe is required. For this reason, we employed a nitrile-derivatized Trp, 5-cyanotryptophan (Trp_{CN}) in the current study. Previously studies have shown that the frequency and linewidth of the C≡N stretching vibration of Trp_{CN} is highly sensitive to local solvation environment, due to the fact that even solvent interactions that do not directly involve the nitrile moiety, such as H-bonding with the indole amine (N-H) group, can be propagated through the aromatic ring to affect the C≡N stretching frequency.[155,403,522] In addition, the well defined direction of the C≡N stretching vibrational transition dipole moment can be exploited to determine the structural transition of Trp41 in response to channel closing or opening, akin to the strategy of Hong and coworkers[435] that used 5-^{19}F-Trp and NMR spectroscopy to probe the pH-dependent rotameric states of Trp41. An added advantage of using Trp_{CN} is that its fluorescence emission, quantum yield and lifetime, exhibit a strong dependence on hydration, and hence can be used as a complementary method to assess the solvation status of the Trp41 region of the channel.[56] Specifically, we employed the TM domain of M2 (sequence: ^{22}SSDPLVVAASIIGILH^{37}LILW^{41}ILDRL^{46})

as our model system and incorporated the IR probe via Trp41/Trp_{CN} mutation. We chose this M2 construct (hereafter referred to as M2TM) because it has been shown that: 1) it fully recapitulates the tetrameric assembly of the full length M2, 2) there is little, if any, difference in its rate of conduction or amantadine inhibition versus the full-length protein when both are expressed in cellular membranes[523]; 3) the conductance measured for this peptide, when reconstituted in an appropriate lipid composition, is in close agreement with the magnitude of flux expected from whole-cell channel recordings[523]; 4) it is easily synthesized in high yield and purity and has been used in many structural studies.[434,494,521,524-527] Despite this convenience, it is worth noting that this truncated version of M2 does not capture other biological functions of the full length protein, such as facilitating budding of the virus in an ESCRT-independent manner and interaction with the M1 protein. Therefore, the current study does not shed light on these aspects of M2's function. Furthermore, a recent ssNMR study[528] on the M2(21-97) sequence containing the cytoplasmic domain suggests that the pK_a of the His37 residues are slightly increased.

Our results indicate that the water is well-structured around the His37 and Trp41 residues at both high and low pH states and does not show evidence of a significant increase in local hydration upon channel acidification. Furthermore, we find that the Trp41 sidechain undergoes a cation-π interaction with His37 and adopts the *t*90 rotamer that rotates approximately 30° along the χ_2 dihedral angle at low pH. Taken together, these results suggest that there is a hydration bottleneck within the M2TM channel,

particularly in the vicinity of Trp41, which may cause a significant barrier for the proton to diffuse across this region of the channel.

10.2 Experimental Section

Materials and Peptide Synthesis

The peptide (M2TM-W_{CN}) used in this work corresponds to the transmembrane domain of the M2 proton channel with a single mutation at the 41 position, i.e., Trp41/Trp_{CN}, which has the following sequence: ^{22}SSDPLVVAASIIGILHLIL-Trp_{CN}-ILDRL^{46}. All peptides were synthesized using standard 9-fluorenylmethoxy-carbonyl (Fmoc) methods on a PS3 peptide synthesizer (Protein Technologies, Woburn, MA) and cleaved from resin with trifluoroacetic acid (TFA). Fmoc-5-cyanotryptophan was purchased from RSP amino acids (Shirley, MA), and all other amino acids were purchased from Advanced ChemTech (Louisville, KY). 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC), 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoglycerol (POPG), cholesterol, and dodecylphosphocholine (DPC) were purchased from Avanti Polar Lipids Inc. (Alabaster, Alabama). The crude peptides were purified by reverse-phase HPLC (Agilent Technologies 1260 Infinity) with a C18 preparative column (Vydac). For purification of the M2TM peptides, a linear gradient of buffer A (Millipore water) and buffer B (6:3:1 2-propanol:acetonitrile:water), both containing 0.1% TFA, was used, starting from 30% buffer B. For purification of the tripeptide Gly-Trp_{CN}-Gly, buffer B was replaced with acetonitrile and 0.1% TFA. The mass of every peptide was verified by either liquid-

chromatography mass spectrometry (LC-MS) or matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) where appropriate.

LUV Preparation

Large unilamellar vesicles (LUVs) with a 100 nm diameter containing the M2TM-W_{CN} peptide were prepared by first cosolubilizing the peptide with a ternary lipid mixture consisting of POPC, POPG, and cholesterol at 4:1:2 ratios in a mixture of chloroform and ethanol using the desired peptide to lipid ratio. The organic solvent was removed using a stream of nitrogen gas and the resulting film was lyophilized for at least 2 hours to remove any remaining solvent. The film was rehydrated with 1 mL of either a pH 7.4 buffer (50 mM phosphate and 100 mM NaCl) or a pH 5.0 buffer (50 mM cacodylate and 100 mM NaCl) and the resultant solution was then subjected to 8 freeze-thaw cycles, which include slow vortexing, freezing, and thawing. Further extrusion of this lipid-peptide solution using an extruder (Avanti Polar Lipids Inc.) equipped with a 100 nm membrane yielded the desired LUV sample. All LUV solutions were stored at 4 °C and used within 5 days of preparation.

Proton Flux Assay

A detailed description of the liposome flux assay can be found elsewhere.[523] Briefly, the LUVs used in the proton flux measurements were prepared using a similar procedure described above. Specifically, a peptide-lipid film composed of 25 nmol peptide (M2TM or M2TM-W_{CN}) and 25 μmol lipid (the same as above) was rehydrated with 995 μL of pH 7.4 potassium phosphate buffer (15 mM K_xPO_4 and 50 mM K_2SO_4). Immediately upon rehydration, 5 μL of a pH indicator dye, pyranine (100 mM), was added to the

mixture followed by freeze-thaw cycles and extrusion. The pyranine dye on the outside of the liposomes was removed by allowing this LUV solution to pass through a PD-10 column (GE Healthcare Life Sciences). To initiate the proton flux assay, 20 μL of the newly prepared LUV solution was added to 2.5 mL of another buffer, that contains pH 7.4 sodium phosphate buffer (15 mM Na_xPO_4, 50 mM Na_2SO_4), 37.5 μL of *p*-xylene-bis-pyridinium bromide (DPX) (1 M), which is a quencher of pyranine fluorescence, and 4 μL of valinomycin (18 μM). The fluorescence kinetic traces were collected on a Cary Eclipse Fluorescence Spectrophotometer under constant stirring, using a time window of 100 s and a time step of 0.25 s. The deprotonated pyranine fluorescence was excited at 460 nm and monitored at 515 nm, using a spectral width of 5 nm. At the end of each trial, the pyranine fluorescence intensity at 515 nm was collected again (integrated for 20 s) by exciting its pH-independent absorption isosbestic point at 417 nm. Dividing the fluorescence kinetics obtained with 460 nm excitation by this fluorescence signal helps remove any effects arising from differences in the sample concentrations and excitation intensities. Five trials were conducted and then averaged, yielding the reported kinetics.

The normalized fluorescence signal was further converted to intraliposomal proton concentration, $[H^+]$, using a calibration curve (Figure 10.2) determined by measuring the fluorescence of free pyranine dye in a series of potassium phosphate buffers with known pH values. Furthermore, the size and concentration of the LUVs, which were needed to determine the total intraliposomal volume, were measured using a fluorescence correlation spectroscopic setup described in detail elsewhere.[529] Additionally, since the M2TM construct can be randomly oriented, either the N-terminus

pointing toward the interior ($N_{in}C_{out}$) or exterior ($N_{out}C_{in}$) of the liposome, we assumed there was a 1:1 ratio of the two conformations where only the $N_{out}C_{in}$ is conducting, and as a result, divided the totally number of tetramers by a factor of 2.

Circular Dichroism, Fluorescence, and FTIR Measurements

Circular dichroism (CD) spectra were measured on an Aviv 62A DS spectropolarimeter (Aviv Associates, NJ) with a 1 mm cuvette. Fluorescence spectra were collected on a Jobin Yvon Horiba Fluorolog 3.10 spectrofluorometer using a 1 cm quartz cuvette. FTIR spectra were collected on a Thermo Nicolet 6700 FTIR spectrometer at a resolution of 1 cm^{-1} using a home-built CaF_2 sample cell with an optical pathlength of 25 μm. For all the FTIR spectra shown, a background arising from solvent absorption has been subtracted. All the static spectroscopic measurements were conducted at room temperature.

Attenuated Total Reflectance (ATR)-FTIR Spectroscopy

Polarized ATR-FTIR spectra were recorded on a Nicolet 6700 FT-IR spectrometer (Thermo Fisher Scientific Inc., Madison, WI) with a Harrick's Horizon multiple-reflection attachment equipped with a Ge crystal (50 mm × 10 mm × 2 mm SPT45). The instrument was purged with nitrogen gas during data acquisition. Each spectrum corresponds to an average of 128 scans with a resolution of 2 cm^{-1}. A solution of M2TM-W_{CN} peptide in TFE was first mixed with a stock solution of POPC/POPG/cholesterol (4:1:2) in chloroform and ethanol in a 1:25 peptide to lipid ratio. A multilayer lipid film was prepared by introducing 100 μL of the mixture onto the surface of the Ge crystal, which was previously plasma cleaned, and left to dry. After a dry film was obtained, 100 μL of either a pH 7.4 buffer (10 mM phosphate buffer and 10 mM NaCl) or a pH 5.0

buffer (10 mM cacodylate buffer and 10 mM NaCl) was uniformly added to the surface and allowed to dry under nitrogen. Then, the crystal was placed back in a hydration chamber and incubated for at least 24 hours before measurement.

The C≡N orientation and helix tilt angles, all with respect to the membrane normal, were calculated from the dichroic ratio (*R*) of the respective C≡N and amide I bands following the published theoretical protocols.[77,82] In the current study, *R* was calculated using the integrated areas of the bands measured with parallel and perpendicular polarizations.

Trp41 Rotamer Analysis

The Trp41 rotamer analysis using the C≡N orientation angle of Trp_{CN} as a constraint was similar to that used by Hong and coworkers.[435] First, four backbone structures, two for each pH condition (i.e., low and high), were chosen from the available M2TM structures in the Protein Data Bank (PDB) based on the similarity of their helix tilt angles to the experimentally determined values. Specifically, 3LBW and 2KQT were chosen for the high pH condition, and 2KAD and 3C9J were chosen for the low pH condition. Second, for each backbone structure, the Trp41 dihedral angles (χ_1 and χ_2) were systematically varied in a 5° increment using the VMD program,[331] and for each (χ_1 and χ_2) combination the hypothetical C≡N orientation angle was determined with respect to the Z-axis of the channel. Third, for each case the RMSD value between the experimentally determined and the simulated C≡N angles was calculated. Finally, plotting the dependence of the RMSD values on χ_1 and χ_2 yielded the 2D contour plot presented in the text. Out of the possible 7 rotamers that a Trp sidechain can sample, we rejected those that have a low

RMSD value but encounter steric clashes with either the helix backbone or His37 residues.

2D IR Measurements

The samples used in the 2D IR measurements were prepared by first solubilizing the M2TM-W_{CN} peptide and DPC in trifluoroethanol (TFE) at a 1:35 peptide to lipid ratio. A stream of nitrogen gas was used to remove the organic solvent and the resultant film was lyophilized for at least 4 hours to eliminate the residual solvent. Then, this peptide-detergent film was re-dissolved in either a pH 7.4 buffer (50 mM phosphate and 100 mM NaCl) or a pH 5.0 buffer (50 mM cacodylate and 100 mM NaCl). The final concentration of the peptide was approximately 15 mM. The primary reason to use DPC micelles, instead of the lipid membrane used in the static spectroscopic measurements, in the 2D IR experiments was that they significantly decrease the noise arising from the scattered light. DPC micelles have been used in many previous 2D IR studies of similar nature.[518,519] As shown (Figure 10.3), the C≡N stretching band of M2TM-W_{CN} in DPC micelles was almost identical to that obtained in membranes composed of POPC/POPG/cholesterol (4:1:2), suggesting that the use of DPC micelles in the current case has not significantly changed the assembly and function of the M2TM-W_{CN} peptide.

The details of the 2D IR setup have been described elsewhere.[417] In short, a train of 800 nm, 65 fs, and 300 μJ pulses were generated by a Libra Ti:sapphire regenerative amplifier (Coherent, Santa Clara, CA) laser system at 1 kHz, which were used to pump a home-built optical parametric amplifier with different frequency generation capability to generate mid-IR pulses with frequency centered at 2220 cm^{-1}. Then, these mid-IR pulses

were split into three groups and focused on to the sample in a boxcar geometry. The photon echo signal was then collected and overlapped with a local oscillator pulse for heterodyne detection. A grating in the spectrometer was used to disperse the heterodyned signal onto a 64 element mercury−cadmium−telluride (MCT) array detector (Infrared Associates, Stuart, FL). The data were collected at different waiting times (*T*) and the resulting 2D IR spectra were obtained from three Fourier transforms of the raw data. In addition, a tapered cosine window function was applied to all 2D IR time domain data before applying the Fourier transform.

Time-Resolved Fluorescence Measurements

Time-resolved fluorescence measurements were collected on a time-correlated single photon counting (TCSPC) system with a 0.4 cm quartz cuvette at 25 °C. The details of the TCSPC system have been described elsewhere.[56,462] Briefly, a home-built femtosecond Ti:Sapphire oscillator was used to generate a 270 nm excitation pulse. Emission was collected at magic angle polarization condition, selected by a short-wavelength bandpass filter (Semrock FF01-357/44) around 360 nm and a long-pass filter (Semrock FF01-300/LP) with a 300 nm cutoff to better suppress scattered excitation light, and detected with a MCP-PMT detector (Hamamatsu R2809U) and a TCSPC PC-board (Becker and Hickl SPC-730). Fluorescence decays were deconvoluted with the instrument response function (IRF) and fit to a multi-exponential function to minimize χ^2 below an acceptable value (i.e., 1.2) using FLUOFIT (Picoquant GmbH).

10.3 Results and Discussion

As indicated above, the site-specific spectroscopic probe, Trp_{CN}, was introduced into M2TM via Trp41/Trp_{CN} mutation (this mutant is hereafter referred to as M2TM-W_{CN}). As shown below, this mutation does not impede the folding and proton conduction activity of M2TM in any significant manner, but does allow us to monitor the local electrostatic and hydration environment of Trp41 through the C≡N stretching vibration via linear and nonlinear IR techniques.

CD and Proton Conductance Measurements of M2TM-W_{CN}

As shown (Figure 10.4), the far-UV CD spectra of M2TM-W_{CN} in vesicles consisting of a ternary mixture of POPC, POPG, and cholesterol at both pH 5.0 and 7.4 (also referred to in the following text as the low pH and high pH state, respectively) show characteristic features (i.e., minima at 208 and 222 nm) of α-helical secondary structures and are similar to those of the wild-type.[530] Interestingly, unlike that of M2TM, the CD spectrum of M2TM-W_{CN} in a membrane environment contains an additional weak band located at 245 nm. The fact that this band disappears when M2TM-W_{CN} is solubilized in a mixture of water and isopropanol (Figure 10.5), wherein the peptide is expected to remain monomeric, suggests that it arises from a M2TM assembly that promotes electronic exciton coupling between nearby Trp_{CN} residues. In addition, the intensity of this CD band depends on pH (Figure 10.4), indicating that the relative orientation and separation distance between these Trp_{CN} sidechains are dependent on pH, as seen for the Trp residues in the M2TM proton channel. Thus, taken together, these CD results suggest that

the chosen Trp41/Trp_{CN} mutation does not disrupt the tetrameric assembly of M2TM in lipid membranes.

To ensure that the Trp_{CN} mutation does not impede the proton conduction function of the tetramer, we further performed proton flux measurements on M2TM and M2TM-W_{CN} channels reconstituted in large unilamellar vesicles (LUVs) following a protocol previously published by Ma *et al.*[523] Briefly, a M2TM or M2TM-W_{CN} vesicle solution was first prepared in potassium phosphate buffer (pH 7.4), and then was diluted by a sodium phosphate buffer solution (pH 7.4) that also contained a potassium ion selective carrier, valinomycin.[505] This mixing caused an electrochemical gradient between the interior and exterior of the liposome, which, when a proton-selective channel is embedded in the lipid bilayer, enabled transfer of potassium ions from the inside to the outside of the liposome through valinomycin and, in the meantime, a proton flow in the opposite direction. This proton flux was detected and quantified by the fluorescence time-profile of a pH-sensitive dye, pyranine, which was preloaded inside the vesicles (see details in Supporting Information). As indicated (Figure 10.6), the cumulative proton flux per M2TM tetramer at this pH was determined to be 5.7 ± 0.4 H^+/s, which is almost identical to that reported previously,[523] and 4.2 ± 0.6 H^+/s for M2TM-W_{CN}. Thus, these results provide direct evidence that the proton conduction activity of the M2TM channel tolerates the Trp41 to Trp_{CN} mutation. In addition, the proton conduction rates obtained here are within an order of magnitude of previously reported values for the M2 channel.[485,506,523,531]

Structural Transition of Trp41

Hong and coworkers have investigated the structure and structural transition of the Trp41 gate in M2TM at acidic pH using ssNMR.[435,491] By replacing Trp41 with 5-^{19}F-Trp in M2TM, they were able to use ^{19}F spin diffusion NMR measurements to characterize the distance between His37 and Trp41 as well as the conformational dynamics of the indole sidechain. Their results indicated that the Trp41 sidechain adopts a *t*90 rotamer at both high and low pH; however, the two dihedral angles of Trp41, χ_1 and χ_2, are increased by ~20^o upon channel activation. Furthermore, they found that at low pH, Trp41 undergoes Gaussian fluctuations in both χ_1 and χ_2, suggesting that the cation-π interaction between His37 and Trp41 is sporadic, and may serve as the trigger to release a proton into the C-terminal region of the channel.

To provide further validation of these previous findings, herein we employed polarized attenuated total reflectance IR (ATR-IR) spectroscopy to assess the orientation of the Trp41 sidechain. This assessment is based on the notion that the C≡N stretching vibration is a local vibrational mode and its transition dipole moment is along the molecular axis of the C≡N moiety, which lies in the same plane of the indole ring. Therefore, the orientation angle of Trp_{CN} in a well-oriented M2TM-W_{CN} tetramer, such as when embedded in a surface-supported lipid bilayer, can be determined by measuring the dependence of the absorbance the C≡N stretching vibration on the polarization of the IR light.[82] As shown (Figure 10.7), the dichroic ratio of the C≡N stretching vibration, defined as the ratio between the integrated areas of the absorption spectra obtained with parallel and perpendicular polarizations, is changed from 1.55 ± 0.24 at pH 7.4 to 1.95 ±

0.18 at pH 5.0. This indicates that Trp_{CN} undergoes a distinct conformational change in response to a protonation state change of the channel. Based on these dichroic ratios, we further calculated the average angle of the C≡N axis with respect to the membrane normal[77,81] or z-axis (θ_{CN}) to be 60.2 ± 4.7° for pH 7.4 and 53.1 ± 3.7° for pH 5 (Table 10.1). Similarly, we used the amide I band of the peptide and polarized ATR-IR measurements to estimate the helix tilt angle of M2TM-W_{CN} in supported lipid bilayers at different pH values. As shown (Figure 10.7a and 10.7b), in both cases the absorbance of the amide I band measured with polarization parallel to the membrane normal is significantly higher than that measured with IR light whose polarization is perpendicular to membrane normal, indicating, as expected, that the helix is oriented perpendicular to the membrane plane. Further calculations using the experimentally determined dichroic ratios indicate that the average helix tilt angle (β) increases from 24.2 ± 9.0° to 39.7 ± 7.4° with respect to the z-axis upon changing the pH from 7.4 to 5.0 (Table 10.1). This pH-induced global structural change is consistent with several other previously published M2TM structures.[521,532,533] In addition, the angle determined herein at high pH is in agreement with the ATR-IR study of Manor *et al.*,[514] which reported a helix tilt angle of 27° at pH 7.0 obtained via a protein model constructed from backbone carbonyl angles calculated from the dichroic ratios of backbone isotope labels. However, their data suggested that the conformational change upon lowering the pH manifests through a helix rotation from 161° to 290° from pH 7.0 to 4.0 instead of an increase in β.

Because θ_{CN} depends on the torsion angles (i.e., χ_1 and χ_2) of the Trp_{CN} sidechain and also on the helix tilt angles of individual helices, the rotameric state of Trp41 cannot

be determined in an absolute sense using the θ_{CN} alone. However, similar to the strategy used by Hong and coworkers,[435] this information can be used as constraints to exclude rotameric states that do not meet the angle criteria obtained from experiment, especially when the backbone conformation is known. In the current case, we began by identifying PDB structures of M2TM that could be used as a structural template for the subsequent rotameric analysis of Trp41. Based on the similarity of their helix tilt angle with the β angle determined from experiment, two PDB structures[434,486] obtained at neutral pH and pH 6.5, 2KQT and 3LBW, were chosen for the high pH case. Similarly, a pH 5.3 crystal structure (PDB: 3C9J)[521] and a ssNMR structure (PDB: 2KAD)[526] obtained at pH 7 were chosen for the low pH state, because the helix tilt angle in both is between 35-38° and their C-terminal pore is relatively more expanded compared to that of the structure models of the high pH state. For each PDB structure, we then allowed the Trp41 sidechain to sample all the possible rotameric conformers by varying its χ_1 and χ_2 dihedral in 5° increments and the value of θ_{CN} was calculated for each and every possible combination. As shown (Figure 10.8) and expected, there are a large number of combinations of χ_1 and χ_2 that can satisfy the experimentally determined C≡N angle constraints. However, out of all the possible rotameric states that tryptophan can adopt, the majority of them can be ruled out due to steric clashes either with the protein backbone or with the His37 sidechain. As indicated (Figure 10.8a), the two allowable rotamers are *t*90 and *t*-105. To further determine which of these two rotamers is the most probable, one needs additional information. As discussed in detail below, the C≡N

stretching frequency allowed us to further conclude that Trp41 adopts the *t*90 rotamer in both the high and low pH states.

To further refine the possible χ_1 and χ_2 values in the t90 rotamer regime, we only considered combinations of χ_1 and χ_2 that satisfy the θ_{CN} angle within ± 7.0° of the reported value to account for the experimental error within the ATR measurement. As indicated (Figure 10.9), this analysis allowed us to determine the minimum RMSD values within this range and hence better resolved rotamer angles (Table 10.2). Interestingly, the (χ_1, χ_2) determined from the current analysis for the low pH state are comparable to those obtained from the corresponding PDB structures, while the (χ_1, χ_2) determined for the closed (or high pH) state deviate significantly from their PDB counterparts. The most notable difference is a ~30° decrease in χ_2, from 79° to 51°, indicating that the IR results support a tighter packing of the Trp41 sidechains in the closed state (Figure 10.9). Similarly, our χ_1 angle is similar to that reported by Williams *et al.*,[435] although our χ_2 angle is significantly different from theirs. In particular, their χ_2 angle suggests that the benzene ring is pointed downward and hence closer to the C-terminus, whereas our result suggests that the benzene ring is more buried within the pore. In addition, the dihedral angles determined from the IR measurements indicate that upon lowering the pH, χ_2 undergoes a significant increase (i.e., from 51° at pH 7 to 89° at pH 5). This increase shifts the benzene ring of Trp41, specifically atoms Cη2, Cζ3, and Cε3, away from the center and into the lipid region, only exposing the N-H side of the pyrrole to the inner pore. Accompanied by a larger helix tilt angle, the His37 imidazole rings come into closer contact with Trp41, potentially allowing an optimal proximity for a cation-π

interaction between adjacent helices as was observed in UV resonance Raman[520] and ssNMR[435] experiments. Therefore, these structural data on Trp41 provide further validation that the mutation strategy used, i.e., the Trp41 to Trp_{CN} mutation does not disrupt the conformational transitions that are critical for the proton conducting function of the M2TM channel.

Local Electrostatic and Hydration Environment of the Trp41 Gate

It has been shown that the C≡N stretching vibrational band of Trp_{CN} is a sensitive IR reporter of the hydration status of its indole sidechain.[155] For example, in water its C≡N band is centered at 2224 cm^{-1} with a broad bandwidth of 18 cm^{-1}, whereas in tetrahydrofuran (THF) the band is shifted to 2220 cm^{-1} and narrowed to 8 cm^{-1}.[155] As shown (Figure 10.10 and Table 10.3), the C≡N band of M2TM-W_{CN} in a hydrated lipid bilayer composed of POPC, POPG, and cholesterol is centered at 2219.8 cm^{-1} at pH 7.4, which becomes slightly red-shifted (~1 cm^{-1}) and narrower when the pH is decreased to 5.0. In comparison, the C≡N stretching vibration of a fully hydrated tripeptide, Gly-Trp_{CN}-Gly, has a peak frequency of 2225.5 cm^{-1} and also a broader bandwidth. Taken together, these results suggest that the Trp_{CN} sidechain in the M2TM-W_{CN} channel is mostly dehydrated and situated in a THF-like environment even when the channel is in its low pH state. This is a significant finding as previous studies have shown that more mobile water will be accumulated in the Gly34 region under more acidic conditions.[353,503,517,519,525] In addition, the observed frequency shift toward lower wavenumbers upon changing the pH from 7.4 to 5.0 is consistent with the notion that the His37 tetrad is transitioned from a +2 state to a +3 state. This is because the associated

increase in electric field at the nitrile site is predicted to decrease its stretching frequency, due to the vibrational Stark effect. Boxer and coworkers[534] have shown that the C≡N stretching vibrational frequency of 5-cyanoindole decreases with increasing electric field with a Stark tuning rate of 0.86 $cm^{-1}/(MV\ cm^{-1})$. Thus, a simple calculation of the electric field change (ΔE) experienced by the C≡N bond allowed us to estimate the frequency shift of the C≡N stretching vibration of Trp_{CN}. Assuming that a single point charge is added to the center of the His37 tetrad and using the PDB structure 3C9J and the Trp41 rotamer obtained above for distance and angle determination, ΔE was calculated to be 4.9 MV cm^{-1} when vacuum permittivity was used, which gives rise to a frequency shift of -4.2 cm^{-1}. Comparing to the 1 cm^{-1} red-shift observed in the experiment, this value suggests that the effective dielectric constant in this region of the channel is on the order of 4. This finding thus supports the notion that the Trp41 tetrad is immersed in a THF-like environment, which has a dielectric constant of 7.6. Moreover, as discussed above, this red-shift would argue against the idea that the Trp41 sidechain adopts the *t*-105 rotamer, as this conformation would point the C5 position, where nitrile is located in Trp_{CN}, directly toward the center of the channel, resulting in a blue-shift[155,535] due to direct H-bonding between the C≡N group and the few water molecules in this region (see below). Finally, the fact that at pH 5.0 the C≡N stretching band becomes narrower is consistent with a stronger cation-π interaction between His37 and Trp41 in the low pH state, which forces Trp41 to sample a more homogenous environment.

Within the M2TM channel, the Trp41 cluster forms the narrowest region of the channel and is sequestered in a hydrophobic environment with the exception of the

confined waters, especially in the closed state.[536] Therefore, the characteristics of the C≡N stretching vibration obtained at pH 7.4, indicative of a more ordered and less hydrated environment for the Trp_{CN} sidechain, is not surprising. At pH 5.0, which leads to the C-terminal region to expand, it is expected, as observed in simulations,[496,503,517] that more water will come to the vicinity of Trp41, leading to an increase in the degree of hydration of the indole ring. However, the IR results suggest otherwise. One concern was that the M2TM-W_{CN} channel may be still closed at pH 5.0 due to the Trp_{CN} mutation. Thus, we also measured the FTIR spectrum of M2TM-W_{CN} in the aforementioned lipid environment at pH 3.0, where all four His37 sidechains are protonated. The result (Figure 10.11) shows that the C≡N stretching frequency is identical to that measured at pH 5.0 and, therefore, we rule out this possibility. Another concern was that the C≡N stretching vibration may lack the necessary sensitivity to reveal this change.

Accordingly, to provide further evidence to support the conclusions reached through IR experiments, we carried out both steady-state and time-resolved fluorescence measurements. Recently, we have shown that the fluorescence properties of Trp_{CN} are highly sensitive to the degree of hydration.[461] For example, when Trp_{CN} transitions from a hydrophobic solvent (i.e., 1,4-dioxane) to H_2O, the fluorescence is severely quenched and results in a decrease in quantum yield by an order of magnitude. Likewise, the emission maximum red-shifts from 361 nm to 391 nm. Additionally, the fluorescence lifetime decay kinetics of Trp_{CN} in H_2O are subnanosecond whereas in 1,4-dioxane the decay kinetics are ~6.0 ns. As shown (Figure 10.12), the Trp_{CN} fluorescence spectra of M2TM-W_{CN} in vesicles composed of POPC, POPG, and cholesterol in a 4:2:1 ratio at both pH

7.4 and 5.0 closely resemble that of Gly-Trp_{CN}-Gly obtained in THF,[56] indicating that the Trp41 sidechain is situated in a relatively hydrophobic environment in both high and low pH states of the channel. Consistent with our results, a previous study on the wild-type M2TM in DHPC/POPC bicelles showed that the native Trp41 fluorescence spectrum is peaked at about 334 nm at both pH 8.0 and 4.5.[537] Interestingly, a closer inspection indicates that the Trp_{CN} fluorescence obtained at pH 5.0 is approximately 1.5 times more intensive than that obtained at pH 7.4 and also shows a small but measurable blue-shift (i.e., from 371 to 368 nm). These behaviors are different from those observed with native Trp fluorescence, which was shown to decrease its intensity when the pH was lowered.[538] Nevertheless, these changes in the fluorescence properties are consistent with the idea that the aforementioned cation-π interactions are strengthened at pH 5.0.[539] Also, since a change in hydration would be more drastically reflected in the fluorescence intensity signal of Trp_{CN},[56] these results corroborate the conclusions we have reached above based on the IR measurements. Moreover, our findings are in line with a previous UV resonance Raman study,[520] which showed that the W17 Raman band of M2, which is a marker of the H-bonding status of the indole N-H, did not show any appreciable change upon lowering the pH from 7.4 to 5.4 and that the relative intensity of the W7 doublet, which measures the overall hydrophobicity of the indole ring environment, also remained unchanged.

To further confirm this outcome, we performed fluorescence lifetime measurements on M2TM-W_{CN} in vesicles at both pH 5.0 and 7.4. As shown in Figure 10.13, the intensity weighted average lifetime (τ_{ave}) for the two pH conditions are 6.9 ns

(pH 5.0) and 6.7 ns (pH 7.4), and do not indicate significant changes in kinetics due to varying hydration environments. As a control, we show that when the Trp_{CN} in M2TM-W_{CN} is more exposed to solvent, as is the case when the peptide was fully solubilized in trifluoroethanol (TFE) without lipid, the average fluorescence lifetime is significantly faster, τ_{ave} = 2.0 ns. From this, we emphasize that if distinct hydration changes were occurring locally near Trp41 in M2TM-W_{CN}, the observed fluorescence photophysics would reflect a measurable change in lifetime due to differences in nonradiative decay pathways. The fact that the kinetics remains similar in both pH scenarios corroborates are results presented above.

Hydration Dynamics of the Trp41 Gate

The motions of water molecules surrounding a vibrator can induce fluctuations in its vibrational frequency, causing its frequency-frequency correlation function (FFCF) to decay on a timescale that characterizes the dynamics of these motions. Two-dimensional IR (2D IR) spectroscopy is capable of assessing such dynamic events and has been widely used to study water dynamics.[143,540] Thus, to further validate the aforementioned findings obtained via static spectroscopic methods, we carried out 2D IR experiments on the C≡N stretching vibration of the tripeptide Gly-Trp_{CN}-Gly in water and the M2TM-W_{CN} channel in a membrane environment. As shown (Figure 10.14), at the same waiting time (*T*) the 2D IR spectrum of the tripeptide is more circular than that of the M2TM-W_{CN} channel, indicating that the spectral evolution or diffusion rates of these systems are different. To quantify these spectral diffusion dynamics, we used the center line slope (CLS) method developed by Fayer and coworkers[541] to determine their respective rate

constants. Specifically, the center line represents the maxima of the υ=0 → υ=1 peak of a series of cuts through the 2D IR spectrum. A plot of the inverse CLS as a function of waiting time (i.e., *T*) provides a measure of the underlying FFCF, and is described here by the following exponential function:

$$\text{CLS}\,(T) = A \cdot e^{-T/\tau} + B, \qquad (10.1)$$

where *A*, *B*, and τ are the amplitude, offset, and spectral diffusion decay time constant. As shown (Figure 10.15 and Table 10.3), the spectral diffusion dynamics of the C≡N vibrational mode in the tripeptide occurs on a timescale of 1.6 ± 0.3 ps, which is characteristic of the H-bonding dynamics of bulk water.[426,428] Thus, this result is consistent with the expectation that the Trp_{CN} sidechain in this case is fully hydrated. On the other hand, as indicated (Figure 10.15), the CLS of M2TM-W_{CN} decays on a slower timescale and with a time constant of approximately 3.0 ± 0.8 ps for both pH 7.4 and 5.0. Taken together, these results indicate that the water in the immediate vicinity of Trp41 in the M2TM channel does not exhibit bulk-water-like dynamics. The slower dynamics suggest that the water molecules in this critical region of the channel form a more structured and hence more strongly hydrogen-bonded network. This is entirely consistent with the notion put forward by Acharya *et al.*[6] and Schnell and Chou[15] that a strongly hydrogen bonded bridging water cluster resides between Trp41 and Asp44 when the channel is closed (see below). However, our finding that this water cluster remains intact even in the low pH condition has strong implication on the rate of proton conduction.

Implication on Proton Conductance

As an integral part of the M2 proton channel, water must be considered in order to achieve a quantitative understanding of the underlying proton conducting mechanism and kinetics. As indicated by a recent high-resolution crystal structure of M2TM,[434] which represents an intermediate state (pH 6.5) between the open and closed states, three well-ordered channel water clusters exist. As shown (Figure 10.1), the first cluster is situated above His37 and consists of six water molecules where four are hydrogen bonded (H-bonded) to the His37 tetrad. The second cluster, consisting of two water molecules with each H-bonded to a His37, is located between His37 and Trp41. Perhaps more importantly, the third cluster, which consists of five well-ordered waters that form H-bonds between the indole N-H of Trp41 and the carboxylate oxygen of Asp44, is located near the C-terminal region of the channel and serves to stabilize the Trp41 gate. It should, however, be appreciated that these clusters were observed in the structure determined at cryogenic temperatures. MD simulations indicate that a similar distribution is also present near room temperature, although the waters have more dynamic character on the nano- to microsecond time scale. Thus, the crystallographically observed water clusters reflect the approximate positions of more dynamic waters at various positions in the channel.[434,542] The water distribution pattern clearly indicates that in order for a proton to diffuse from the first cluster to the third, or vice versa, it has to overcome two barriers imposed by His37 and Trp41, respectively. Thus, this picture justifies, from the perspective of water, why when the channel is closed protons are unable to diffuse across the channel beginning from either direction.

How does this water distribution pattern change upon viral acidification to facilitate proton conduction? Several recent experimental and simulation studies[353,489,503,517,519,525] showed that both the density and mobility of the first water cluster are significantly increased under acidic conditions, indicating that in the low pH state water can readily flow into this region of the channel, hence allowing protons to access His37 from the N-terminal end of the channel. On the other hand, our results indicate that the degree of hydration surrounding the Trp41 tetrad does not show any appreciable change upon decreasing the pH from 7.4 to 5.0, suggesting that the second and third water clusters present at high pH are preserved at low pH. In other words, our results argue against the idea that, even under acidic conditions where the channel has a wider C-terminal pore, more water will gain access to the Trp41 gate region to establish a continuous water wire below His37. This notion is consistent with other experimental findings: (1) it is consistent with the fact that the M2 channel conducts protons in an asymmetric manner, as the lack of a continuous water chain would prevent protons to gain easy access to His37 from the C-terminal side of the channel; (2) it is consistent with the study of Pinto and coworkers,[490] which showed that replacing Trp41 with a smaller residue, such as Ala, Cys, or Phe, results in an increase in both the forward and backward proton flux. This can be rationalized by considering that a smaller sidechain in this region will allow more water to be accumulated and, as a result, either decreasing or eliminating the proton conduction barrier caused by Trp41 in the wild type; (3) it is consistent with the study of Chizhmakov *et al*. and Ma *et al.*,[493,494] which demonstrated that in the absence of Asp44, the rate of proton conductance increases at high pH and also at pH 5.0,

and was attributed to the weakening of the H-bonding network of the water cluster below Trp41.[494] This once again points to the fact that the water network surrounding Trp41 plays an important role in controlling the forward and backward proton flows; and (4) it is consistent with the MD simulation of Voth and coworkers,[517] which showed that there is a significant energy barrier arising from deprotonation of His37 and ion passage through Trp41, and also with the MD simulation of Acharya *et al*.,[434] which showed that when the channel is transitioned from the +2 to +3 state the water cluster between His37 and Trp41 only accumulates one additional water molecule. Finally, and perhaps more importantly, the current study provides new insight into the low proton conductance of the M2 proton channel; it suggests, besides the previously well-recognized barrier imposed by His37,[502,508,510] that a discontinued water wire in the Trp41 region produces a second barrier for proton conduction. A phenomenological kinetic scheme summarizing these points is give below:

$$
\begin{aligned}
(H_3O^+)_A + (His_{37})_A &\rightleftharpoons (His_{37}{}^+)_A + (H_2O)_A \\
(His_{37}{}^+)_A &\rightleftharpoons (His_{37}{}^+)_B \\
(His_{37}H^+)_B + (H_2O)_B &\rightleftharpoons (His_{37})_B + (H_3O^+)_B \\
(H_3O^+)_B &\xrightarrow{k_{BC}} (H_3O^+)_C
\end{aligned}
\quad ,
$$

where in the first step the third His37 uptakes a proton from water cluster A ($(H_2O)_A$), leading to formation of the +3 state of the channel, which is followed by a tautomerization and/or ring reorientation step (i.e., conformational change from $(His_{37}{}^+)_A$ to $(His_{37}{}^+)_B$) required to orient His37 in a position primed to release the proton to water cluster B ($(H_2O)_B$), in the third step. Previously, the proton exchange rate between water

and the imidazole of His37, assessed by a ssNMR study,[510] is at least one to two orders of magnitude faster than the overall proton flux of M2 measured with whole-cell patch-clamp and liposome assays[504-506] and is synchronized with ring reorientations that produce a barrier of ~60 kJ/mol.[502] In the last step, the rate of which is not diffusion-limited ($k_{BC} < k_D$), the proton is transferred from water cluster B to C ($(H_2O)_C$). While the current study does not provide any direct information about the structural nature of this step, it likely suggests that an additional conformational change, either at the sidechain (e.g., Trp41) or the backbone level,[435,489,502,509,517] or a transiently occupied proton diffusive state[511] is required to successfully relay a proton from His37 to the water at the C-terminal end of the channel.

10.4 Conclusions

The tryptophan residue (Trp41) in the transmembrane domain of the M2 protein of influenza A virus serves as a gate to allow asymmetric proton conductance into virion interior under external acid conditions. Herein, we used an analog of tryptophan, 5-cyanotrptophan (Trp_{CN}), and linear and nonlinear infrared spectroscopic methods to interrogate the structure and hydration dynamics of Trp41 of a model M2 proton channel (M2TM) in a membrane environment. Specifically, we mutated the Trp41 residue in the transmembrane domain of the M2 protein (M2TM) with Trp_{CN} and used the C≡N stretching vibration of the resulting peptide (M2TM-W_{CN}) as a site-specific infrared probe. The proton conductivity of the M2TM-W_{CN} channel was found to be comparable to the wild-type M2TM channel, indicating that this mutation does not impede, in any

significant manner, the native function. Consistent with many previous studies on M2TM, polarized ATR-FTIR measurements showed that both the backbone and Trp_{CN} sidechain of the M2TM-W_{CN} channel undergo conformational changes in response to a pH change from 7.4 to 5.0. Furthermore, a rotamer analysis based on the ATR-FTIR results suggested that while Trp41 adopts the *t*90 rotamer conformation, in agreement with a previous study, the calculated dihedral angles place the indole rings in a more compact arrangement when the channel is at high pH. In addition, it was found that upon channel opening, the Trp41 sidechain rotates by +30° along the χ_2 axis, loosening the initially tightly bound Trp41 cluster. As expected, the C≡N stretching frequency and fluorescence measurements indicated that the Trp41 tetrad is sequestered in a THF-like environment at pH 7.4, and the spectral diffusion dynamics of this vibrational mode indicated that the surrounding water is not bulk-like. What is not expected, however, is that these characteristics are preserved at pH 5.0, or in the open state of the channel. Taken together, these findings suggest that (1) there are a small number of water molecules in the Trp41 region of the channel, as observed in a high-resolution crystal structure of M2TM; (2) the water molecules in this critical region of the channel are less mobile compared to bulk water; and (3) unlike the cavity above His37, where more water molecules will accumulate upon channel acidification, the water density surrounding Trp41 does not show any significant increase when the pH is decreased from 7.4 to 5.0. We believe that this invariability of water density between the high and low pH states in this gating region is closely linked to the low rate of proton conduction in the M2 channel, as it suggests that a continuous water wire passing through this region has a low

probability to be established. In other words, besides the energetic barrier inflicted by the His37 reorientation and protonation-deprotonation steps, the inward proton flux encounters a second barrier imposed by the Trp41 tetrad.

Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (P41-GM104605). B.N.M. is an NIH Ruth Kirschstein Predoctoral Fellow (AG-046010).

Parameter	pH 5.0	pH 7.4
R_{CN}	1.95 ± 0.18	1.55 ± 0.24
S_{CN}	0.04 ± 0.07	-0.13 ± 0.11
R_{Helix}	2.31 ± 0.26	2.94 ± 0.37
S_{Helix}	0.16 ± 0.07	0.30 ± 0.07
β (°)	39.7 ± 7.4	24.2 ± 9.0
θ_{CN} (°)	53.1 ± 3.7	60.2 ± 4.7

Table 10.1 Dichroic ratio (R), order parameter (S), helix tilt angle (β), and C≡N orientation angle (θ) of M2TM-W_{CN} obtained based on the ATR-FTIR measurements at pH 5.0 and 7.4.

pH	PDB Structure	$(\chi_1, \chi_2)_{PDB}$	$(\chi_1, \chi_2)_{CN}$
Low	3C9J	(-169°, 84°)	(-177°, 72°)
	2KAD	(-176°, 85°)	(-171°, 89°)
High	3LBW	(178°, 79°)	(-158°, 51°)
	2KQT	(-175°, 78°)	(-164°, 51°)

Table 10.2 Dihedral angles of Trp41, $(\chi_1, \chi_2)_{CN}$, determined from the rotamer analysis detailed in the text. Also listed are the dihedral angles of Trp41, $(\chi_1, \chi_2)_{PDB}$, determined from the corresponding PDB structure.

Peptide	ω_0, cm^{-1}	FWHM, cm^{-1}	*A*	τ, ps	*B*
M2TM-W_{CN} pH 5.0	2218.8 ± 0.5	11.9 ± 1.0	0.48 ± 0.04	2.92 ± 0.52	0.11 ± 0.04
M2TM-W_{CN} pH 7.4	2219.8 ± 0.5	13.7 ± 1.0	0.60 ± 0.08	2.84 ± 0.82	0.15 ± 0.09
Gly-Trp_{CN}-Gly	2225.5 ± 0.5	14.9 ± 1.0	0.42 ± 0.05	1.63 ± 0.33	0.04 ± 0.05

Table 10.3 Center frequency (ω_0) and width (FWHM) of the C≡N stretching vibration determined from linear IR experiments and parameters obtained from an exponential fit of the CLS from 2D IR measurements.

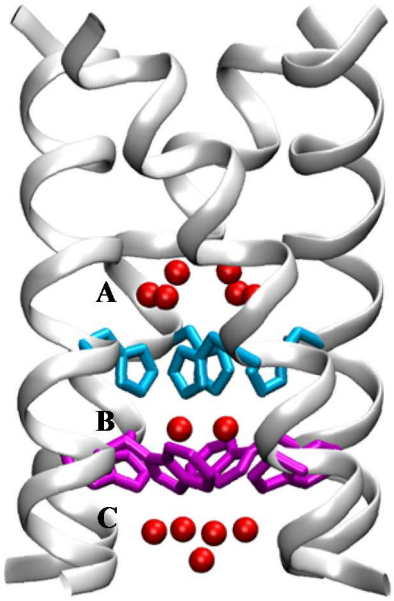


Figure 10.1 X-ray crystallographic structure of the transmembrane domain of the M2 proton channel (Protein Data Bank ID: 3LBW), showing the two key sidechains, His37 (blue) and Trp41 (magenta), as well as three channel water clusters (red) labeled as A-C.

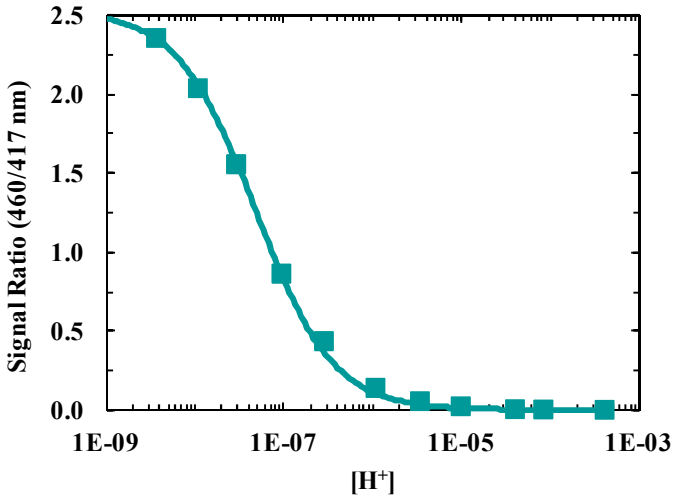


Figure 10.2 Pyranine fluorescence-pH calibration curve. The signal ratio is defined as the ratio between the fluorescence intensities obtained with excitation wavelengths of 460 nm and 417 nm, respectively. The smooth line is a fit of the data to the following equation: $S([H^+]) = A/(1+[H^+]/B)$, where $A = 2.53 \pm 0.02$ and $B = 4.79 \pm 0.18 \times 10^{-8}$ M.

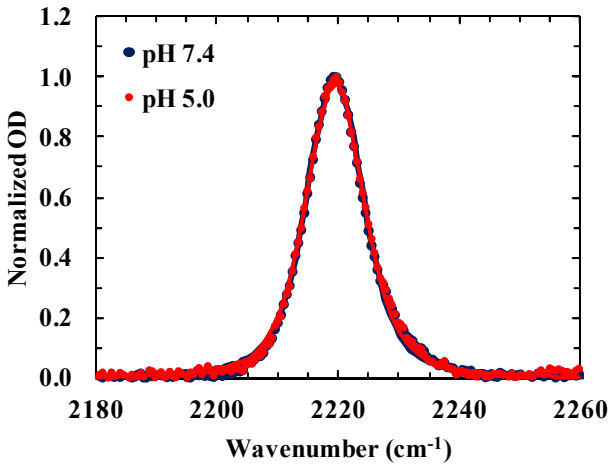


Figure 10.3 Comparison of the C≡N stretching vibrational bands of M2TM-W_{CN} in DPC micelles at pH 5.0 and 7.4. The smooth lines are fits to a Voigt profile.

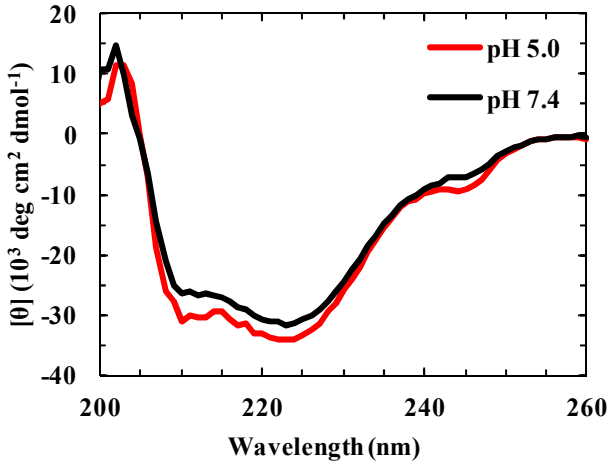


Figure 10.4 CD spectra of M2TM-W_{CN} variant in 100 nm vesicles composed of POPC:POPG:cholesterol (4:1:2) at pH 5.0 and 7.4, as indicated. The peptide concentration was 30 μM and the peptide to lipid ratio was 1:100.

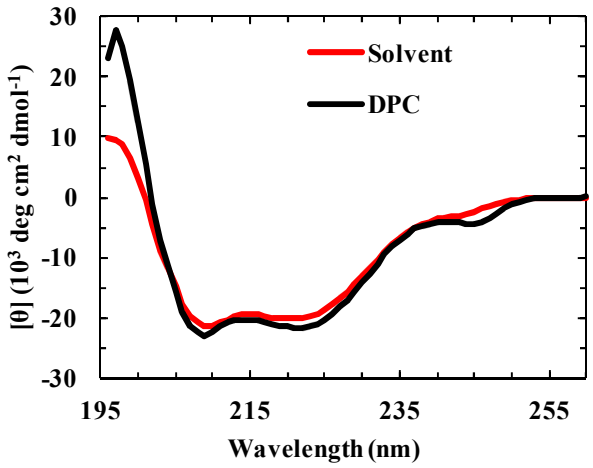


Figure 10.5 CD spectra of M2TM-W_{CN} in DPC micelles (1:35 peptide to lipid ratio) at pH 7.4 and in a solvent consisting of 2-propanol and H_2O (40:60), as indicated. The final peptide concentration was approximately 70 μM.

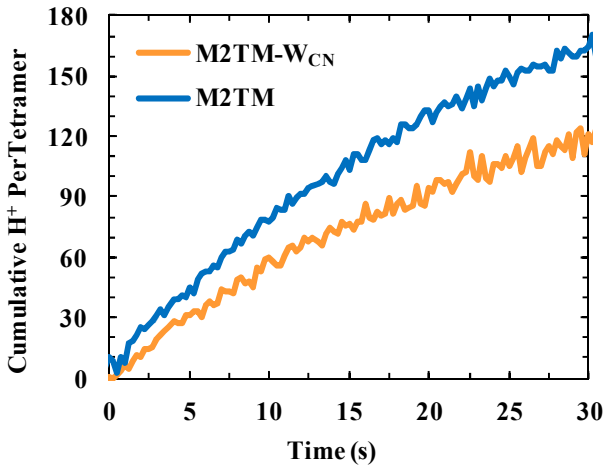


Figure 10.6 Comparison of the proton flux of M2TM-W_{CN} and wild type M2TM, determined from the proteoliposome assay described above. The cumulative proton flux per tetramer was determined to be 5.7 ± 0.4 H^+/s for the wild-type M2TM and 4.2 ± 0.6 H^+/s for M2TM-W_{CN}.

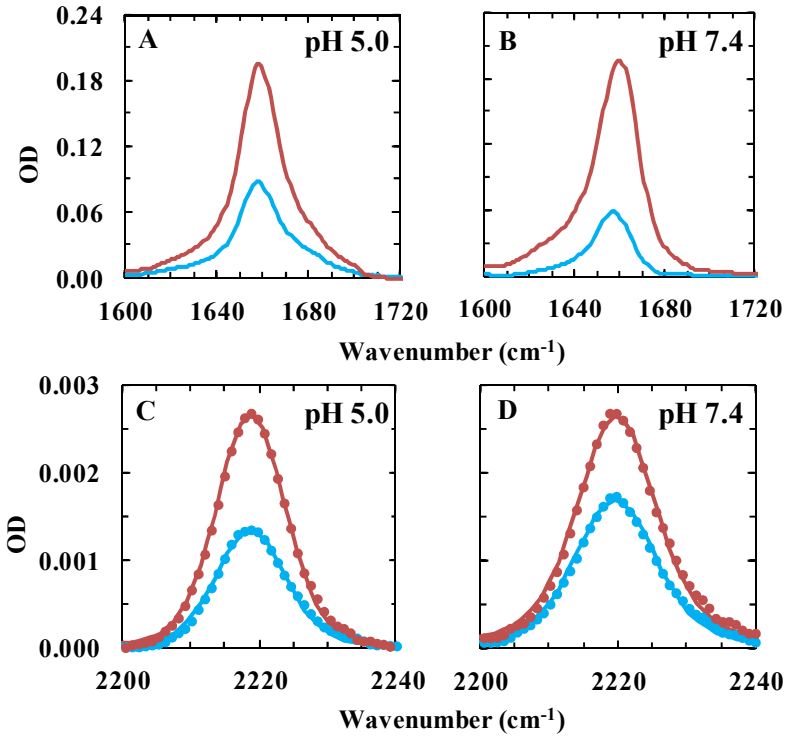


Figure 10.7 ATR-FTIR spectra of M2TM-W_{CN} in a hydrated lipid bilayer composed of POPC:POPG:cholesterol (4:1:2) at pH 5.0 and 7.4, obtained with parallel (red) and perpendicular (blue) polarizations. For comparison, the spectrum measured at pH 7.4 was scaled by a factor of 1.2. (A) and (B): amide I mode of the peptide. (C) and (D): C≡N stretching vibrational band of Trp_{CN}, where the solid line is the best fit of the corresponding spectrum to a Voigt profile and the resultant orientational parameters were given in Table 10.1.

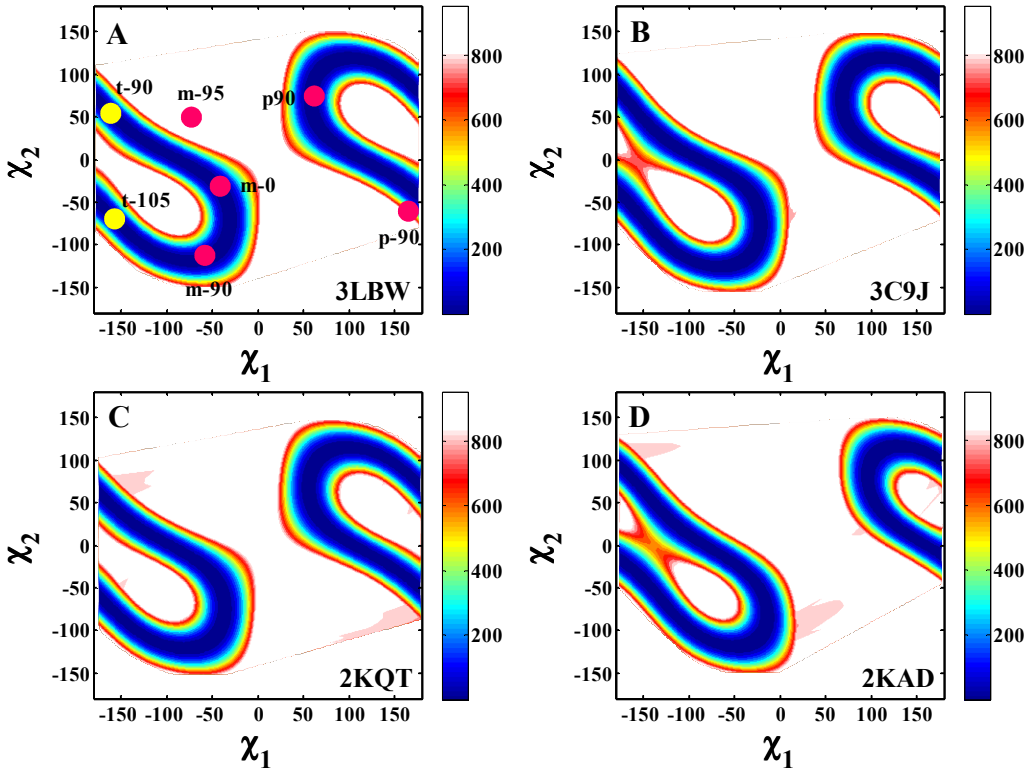


Figure 10.8 2D contour plot of the RMSD between the measured and calculated C≡N angles (θ_{CN}) as a function of the torsion angles (χ_1, χ_2) of Trp41. Four backbone structures, 3LBW (A) and 2KQT (B) for high pH and 3C9J (C) and 2KAD (D) for low pH, were chosen because their helix tilt angles are comparable to those obtained from the ATR-FTIR measurements. The circles depicted in (A) represent all the possible rotameric states of Trp41. Pink circles are the rotameric states that were ruled out based on steric clashes. Yellow circles represent the two rotameric states that were allowed based on the measured θ_{CN}.

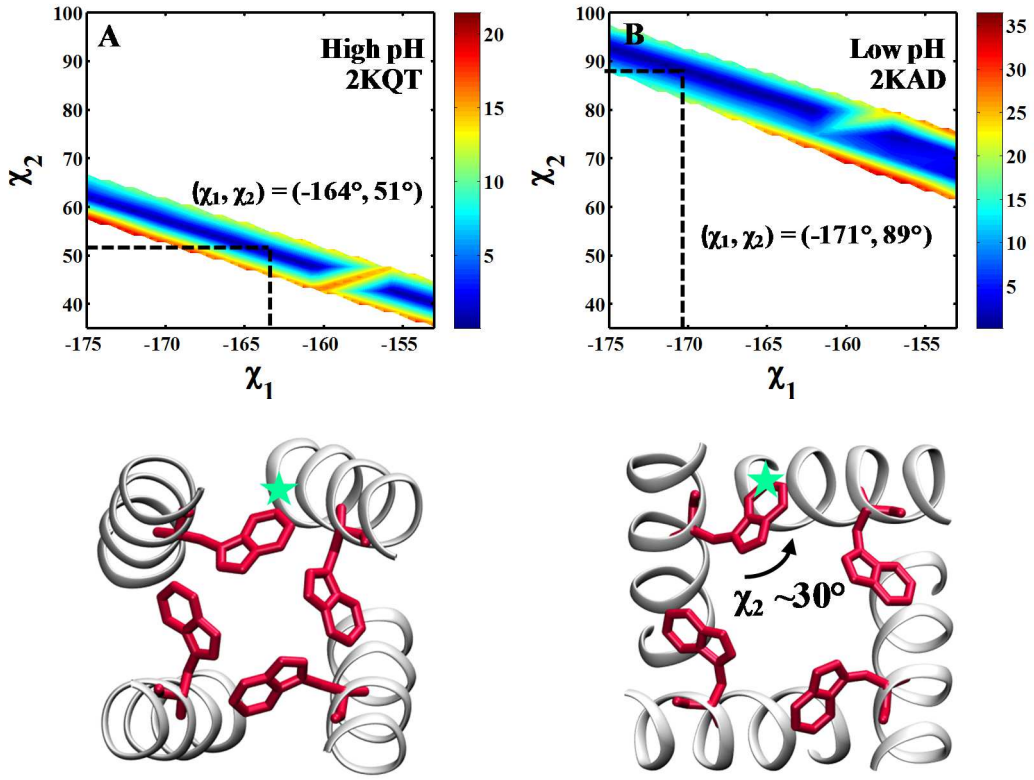


Figure 10.9 Contour plot of the RMSD between the measured θ_{CN} from the ATR-FTIR spectra (Table 10.1) and the calculated θ_{CN} in the *t*90 rotamer range, where only values within ± 7.0° of the experimental value were plotted, as a function of Trp41 (χ_1, χ_2) torsion angles for high pH (A) and low pH (B) cases. The minimum values are depicted in Table 10.2. The PDB structure represent a bottom-up view of the Trp residues; the star is the position of the 5-cyano group.

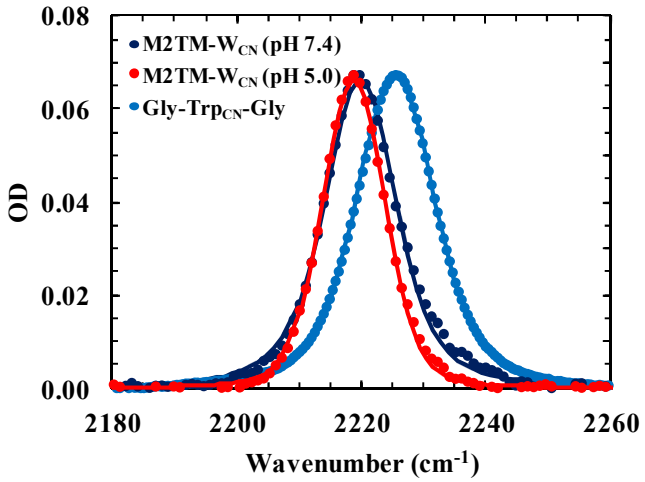


Figure 10.10 Comparison of the C≡N stretching vibrational bands of M2TM-W_{CN} in a lipid bilayer at pH 5.0 and 7.4 with that of Gly-Trp$_{CN}$-Gly in H_2O. The M2TM-W_{CN} spectra are identical to those used in Figure 10.7. For comparison, the spectra of M2TM-W_{CN} was scaled by a factor of 25.The smooth lines are fits to a Voigt profile and the resultant parameters are listed in Table 10.3.

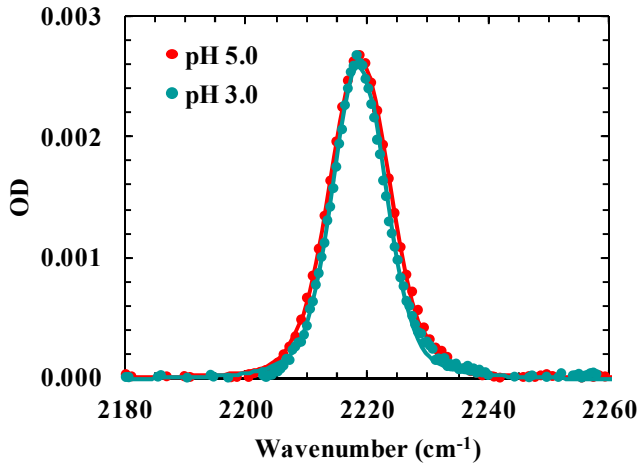


Figure 10.11 Comparison of the C≡N stretching vibrational bands of M2TM-W_{CN} in the aforementioned lipid bilayers at pH 5.0 and 3.0. These spectra were collected using the ATR-FTIR setup with parallel polarization. The smooth lines are fits to a Voigt profile.

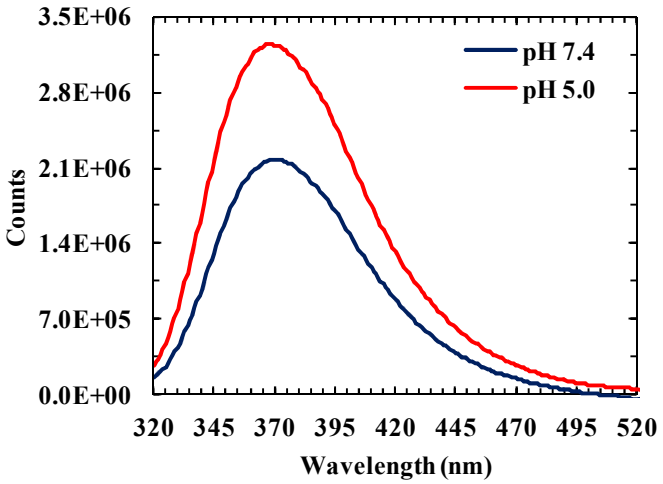


Figure 10.12 Trp_{CN} fluorescence spectra of M2TM-W_{CN} in 100 nm vesicles composed of POPC:POPG:cholesterol (4:1:2) at pH 5.0 and 7.4, as indicated. The peptide concentration was 30 μM and the peptide to lipid ration was 1:100.

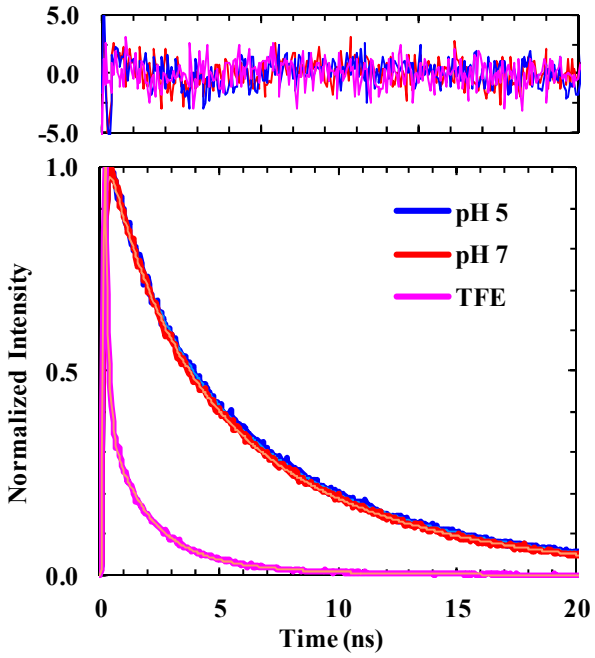


Figure 10.13 Fluorescence kinetics of M2TM-W_{CN} in vesicles composed of POPC, POPG, and cholesterol at pH 5, pH 7, and in TFE solution without lipid, as indicated. The smooth line corresponds to a fit of the kinetics to a multi-exponential function and the intensity weight average lifetimes were determined (τ_{ave}). For pH 5: $A_1 = 0.62$, $\tau_1 = 7.8$ ns, $A_2 = 0.38$, $\tau_2 = 2.5$ ns, and $\tau_{ave} = 6.9$ ns. For pH 7: $A_1 = 0.62$, $\tau_1 = 7.5$ ns, $A_2 = 0.38$, $\tau_2 = 2.4$ ns, and $\tau_{ave} = 6.7$ ns. For TFE solution: $A_1 = 0.84$, $\tau_1 = 0.1$ ns, $A_2 = 0.12$, $\tau_2 = 15$ ns, $A_3 = 0.03$, $\tau_3 = 4.0$ ns and $\tau_{ave} = 2.0$ ns.

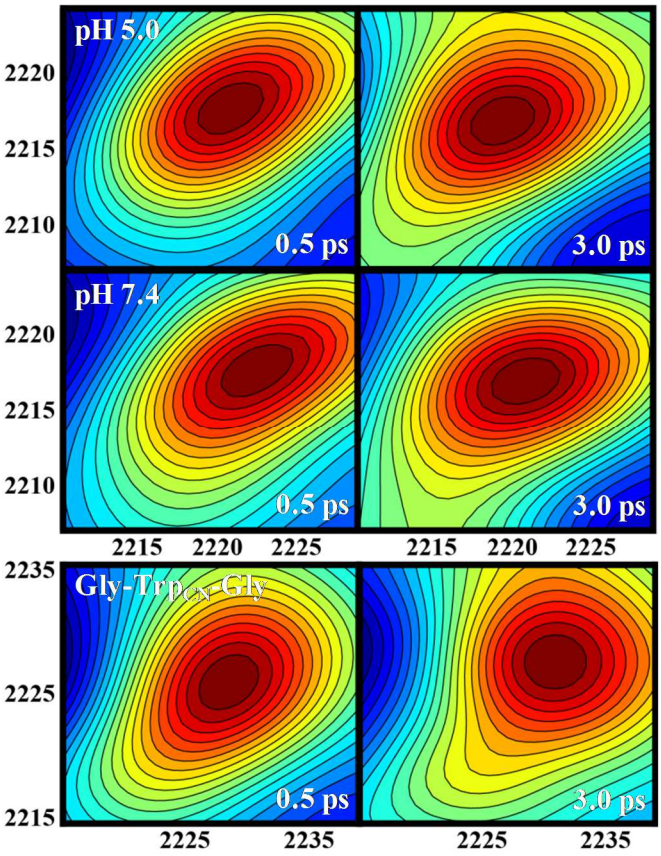


Figure 10.14 Representative 2D IR spectra of the C≡N vibrational mode of M2TM-W_{CN} (15 mM) in DPC micelles (1:35 ratio) at pH 5.0 (A) and pH 7.0 (B). Shown in (C) are the 2D IR spectra of Gly-Trp$_{CN}$-Gly (80 mM) in H_2O. In each case, the respective waiting time (*T*) is indicated.

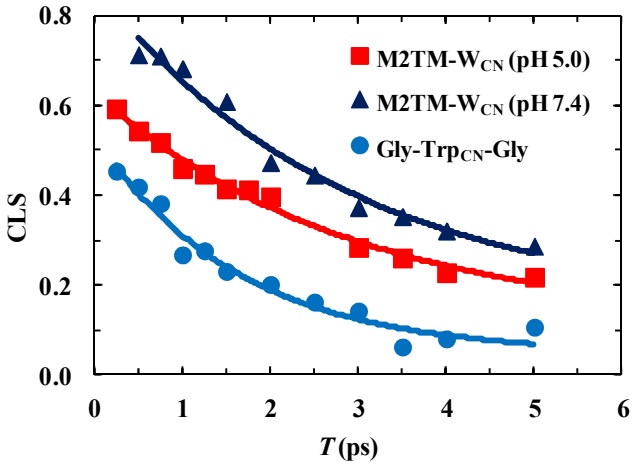


Figure 10.15 Center line slope (CLS) vs. waiting time (T) plots obtained from the 2D IR spectra of M2TM-W_{CN} in DPC micelles at pH 7.4 (red), pH 5.0 (green), and Gly-Trp_{CN}-Gly in H_2O (blue). The smooth lines are fit of this data to the following equation: $CLS(T) = A \cdot e^{-T/t} + B$, and the resulting parameters are listed in Table 10.3.

11 Summary and Future Directions

Protein molecules undergo a diverse set of intra and intermolecular interactions and atomic motions to achieve their biologically relevant structures. Overall, this thesis addresses how one can experimentally dissect protein dynamics and structure from a mechanistic perspective by chemically fine-tuning aspects of the molecular structure to interrogate configuration-dependent and/or site-specific details. More importantly however, the methods presented here only provide an initial stepping stone, and further study implementing the proposed techniques may be helpful in answering other interesting biological questions.

In Chapter 4, we demonstrate how a disulfide covalent constraint placed deliberately along a key hydrogen bonding site of a β-hairpin, can be used to trap a kinetically measureable transition-state analog. Fortunately, a large library of helix and β-sheet staples exists in the literature. Utilizing this principal to study the folding of intrinsically disordered proteins (IDPs) would be an interesting direction to pursue. Recent experimental evidence shows that IDP folding proceeds through a heterogeneous nucleation process so that binding can occur with multiple active sites.[543] Cross-linking particular regions of IDPs may lead to a more homogenous mechanism allowing for more specific binding, and also may speed up the rate of folding.

In Chapter 5, we show how a helix cross-linker can site-specifically increase the local density near a protein chain, thus inducing a measurable internal friction effect. The added landscape roughness retards the folding rate by a factor of ~5. This technique is not only useful to understand the physical principles behind internal friction, but can be a

useful tool to control the rate and direction of a folding pathway. Incorporation of such constraints in larger proteins can be used to tune the folding of certain secondary structural elements and thus regulate the order of folding events.

Chapters 6 and 7 discuss the application of using a light-sensitive non-natural amino acid, Lys(nvoc), which is a hydrophobic but converts to a natural lysine upon photocleavage, to explore aggregation mechanisms and manipulate the self assembly of fibril protein structures. The switchable nature of the Lys(nvoc) moiety allows one to site-specifically alter the chemical environment from hydrophobic to hydrophilic via light. As such, we have shown that short amyloid-forming peptides containing this Lys(nvoc) modification selectively incorporate into the hydrophobic core of wild-type fibrils, where the light-activated charge is strong enough to cause repulsive interactions and disassembly. Evidence in the literature proposes that even the fibril state is dynamic and peptide monomers can intercalate into the ordered structure.[390,544] Thus, it would be advantageous to test whether small amyloid motifs or low molecular weight gelators (i.e., dipeptides) containing Lys(nvoc), could spontaneously embed in the mature fibril after a long equilibration time. This could significantly increase the biological applicability of the method in question. Also important, this section also discussed the effectiveness of using Lys(nvoc) to control the morphology and degradation of peptide hydrogel matrices. Because these biomaterials are engineered for medical applications such as drug delivery, the next logical step would be to test the encapsulation activity of the designed photodegradable hydrogels.

In Chapters 8 and 9, we present new infrared and fluorescence spectroscopic properties of a nitrile-modified non-natural amino acid probe, 5-cyanotryotophan (Trp_{CN}). Through a solvent study, we show that the nitrile stretching frequency correlates well with Kamlet-Taft solvent parameters, describing the polarity and the degree of hydrogen bond accepting and donating ability. In other words, the IR frequency is tuned by interactions on the entire sidechain moiety, thus, demonstrating its sensitivity. Also important, we show that the fluorescence intensity, emission maximum, and lifetime of Trp_{CN}, is exceptionally responsive to hydration environment, even more so than the native sidechain. Thus, it is not only effective as an IR probe, but it is a suitable fluorescence probe of protein hydration. However, the nature of the quenching mechanism in aqueous environments of Trp_{CN} is not fully understood based on our study herein. While we hypothesize that the hydrogen bonding interactions between water and the π-system of the indole ring can induce new non-radiative decay pathways leading to the fast lifetime, we cannot completely rule out other possibilities without probing the relaxation processes in detail. As a result, future work will first involve performing pump-probe transient absorption studies. From here, we hope to identify if there are other factors at play, such as a competing intersystem crossing pathway to an excited triplet state or an excitation-induced proton transfer mechanism of the N-H group. The pK_a of 5-cyanoindole is ~16 in the ground state. A recent study showed that upon excitation to the first singlet state, the pK_a decreases to approximately 10.[545] This acidification upon excitation may facilitate deprotonation followed by non-radiative decay of the anion in aqueous environments. Also important, a recent study showcased the fluorescence

properties of 6- and 7-cyanotryptophan,[57] indicating that positioning of the nitrile functional group significantly dictates the electronic properties. As a result, future work will also aim to investigate how the position of the nitrile moiety affects not only its fluorescence properties, but also its vibrational spectroscopic properties. Preliminary work shows that both the extinction coefficient and frequency of the nitrile stretch is affected by ring position. Lastly, implementing an isotopically labeled Trp_{CN} probe, i.e., containing $^{13}C\equiv^{15}N$, may also yield more useful spectroscopic properties. Most notably, it will red-shift the nitrile stretching frequency by ~80 cm^{-1}, allowing simultaneous observation of the amide I mode and the site-specific label of the studied protein.

Finally, in Chapter 10, we employ the Trp_{CN} probe to study a functionally important tryptophan (Trp) residue in a membrane protein channel, namely the transmembrane domain of the M2 proton channel of the influenza A virus, and we interpret the mechanism of function based on the principles established in Chapter 8 and 9. The Trp residue acts as a gate that controls proton conduction of the channel, and mutation with Trp_{CN} allowed us to observe the local electrostatics, hydration status, and dynamics inside the pore at this site. We find surprisingly that the hydration environment stays relatively constant between the open and closed states, and that this may suggest that Trp acts as a bottleneck in the conduction mechanism. There are several future investigations that could further help dissect the proton transfer mechanism. Recently, we showed that a backbone isotope labeled histidine sidechain can spectroscopically report on changes in pH, tautomer structure, and hydration.[546] Since histidine functions as the proton shuttle in this M2 proton channel, and is thought to limit the rate of the proton

conduction, it would be advantageous to look at the pH dependent changes from this site. Furthermore, aspartic acid is located three residues below Trp. By using an isotope labeled carbonyl in the sidechain of aspartic acid, we could further probe the hydration status or hydrogen bonding interactions near that site using IR spectroscopy. Finally, the ultimate future goal to understanding this biological system and utilize the full potential of the Trp_{CN} probe, would be synthesize an amber codon suppressor tRNA, containing the necessary nitrile modification to be able to express the probe in the full length M2 protein sequence.

Bibliography

(1) Anfinsen, C. B. Principles That Govern Folding of Protein Chains. *Science* 1973, *181*, 223-230.

(2) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The Protein Folding Problem. *Annu. Rev. Biophys.* 2008, *37*, 289-316.

(3) Henzler-Wildman, K.; Kern, D. Dynamic Personalities of Proteins. *Nature* 2007, *450*, 964-972.

(4) Levinthal, C. How to Fold Graciously. *Mössbaun Spectroscopy in Biological Systems Proceedings* 1969, *67*, 22-24.

(5) Dill, K. A.; Chan, H. S. From Levinthal to Pathways to Funnels. *Nat. Struct. Biol.* 1997, *4*, 10-19.

(6) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The Energy Landscapes and Motions of Proteins. *Science* 1991, *254*, 1598-1603.

(7) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, Pathways, and the Energy Landscape of Protein-Folding - a Synthesis. *Proteins: Struct., Funct., Genet.* 1995, *21*, 167-195.

(8) Jackson, S. E. How Do Small Single-Domain Proteins Fold? *Folding Des.* 1998, *3*, R81-R91.

(9) Fersht, A. R.; Daggett, V. Protein Folding and Unfolding at Atomic Resolution. *Cell* 2002, *108*, 573-582.

(10) 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.

(11) Fersht, A. R. Nucleation Mechanisms in Protein Folding. *Curr. Opin. Struct. Biol.* 1997, *7*, 3-9.

(12) Kubelka, J. Time-Resolved Methods in Biophysics. 9. Laser Temperature-Jump Methods for Investigating Biomolecular Dynamics. *Photochem. Photobiol. Sci.* 2009, *8*, 499-512.

(13) Callender, R. H.; Dyer, R. B.; Gilmanshin, R.; Woodruff, W. H. Fast Events in Protein Folding: The Time Evolution of Primary Processes. *Annu. Rev. Phys. Chem.* 1998, *49*, 173-202.

(14) Eaton, W. A.; Muñoz, V.; Hagen, S. J.; Jas, G. S.; Lapidus, L. J.; Henry, E. R.; Hofrichter, J. Fast Kinetics and Mechanisms in Protein Folding1. *Annu. Rev. Biophys. Biomol. Struct.* 2000, *29*, 327-359.

(15) Williams, S.; Causgrove, T. P.; Gilmanshin, R.; Fang, K. S.; Callender, R. H.; Woodruff, W. H.; Dyer, R. B. Fast Events in Protein Folding: Helix Melting and Formation in a Small Peptide. *Biochemistry* 1996, *35*, 691-697.

(16) Olsen, K. A.; Fesinmeyer, R. M.; Stewart, J. M.; Andersen, N. H. Hairpin Folding Rates Reflect Mutations within and Remote from the Turn Region. *Proc. Natl. Acad. Sci. U. S. A.* 2005, *102*, 15483-15487.

(17) Yang, W. Y.; Gruebele, M. Folding Lambda-Repressor at Its Speed Limit. *Biophys. J.* 2004, *87*, 596-608.

(18) Yang, W. Y.; Gruebele, M. Folding at the Speed Limit. *Nature* 2003, *423*, 193-197.

(19) Kubelka, J.; Hofrichter, J.; Eaton, W. A. The Protein Folding 'Speed Limit'. *Curr. Opin. Struct. Biol.* 2004, *14*, 76-88.

(20) Borgia, A.; Wensley, B. G.; Soranno, A.; Nettels, D.; Borgia, M. B.; Hoffmann, A.; Pfeil, S. H.; Lipman, E. A.; Clarke, J.; Schuler, B. Localizing Internal Friction Along the Reaction Coordinate of Protein Folding by Combining Ensemble and Single-Molecule Fluorescence Spectroscopy. *Nat. Commun.* 2012, *3*, 1195.

(21) Soranno, A.; Buchli, B.; Nettels, D.; Cheng, R. R.; Muller-Spath, S.; Pfeil, S. H.; Hoffmann, A.; Lipman, E. A.; Makarov, D. E.; Schuler, B. Quantifying Internal Friction in Unfolded and Intrinsically Disordered Proteins with Single-Molecule Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 2012, *109*, 17800-17806.

(22) Dobson, C. M.; Sali, A.; Karplus, M. Protein Folding: A Perspective from Theory and Experiment. *Angew. Chem., Int. Ed.* 1998, *37*, 868-893.

(23) Dobson, C. M. Principles of Protein Folding, Misfolding and Aggregation. *Semin. Cell Dev. Biol.* 2004, *15*, 3-16.

(24) Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G. Theory of Protein Folding: The Energy Landscape Perspective. *Annu. Rev. Phys. Chem.* 1997, *48*, 545-600.

(25) Chiti, F.; Dobson, C. M. Protein Misfolding, Functional Amyloid, and Human Disease. *Annu. Rev. Biochem.* 2006, *75*, 333-366.

(26) Gazit, E. A Possible Role for Pi-Stacking in the Self-Assembly of Amyloid Fibrils. *FASEB Journal* 2002, *16*, 77-83.

(27) Azriel, R.; Gazit, E. Analysis of the Minimal Amyloid-Forming Fragment of the Islet Amyloid Polypeptide. An Experimental Support for the Key Role of the Phenylalanine Residue in Amyloid Formation. *J. Biol. Chem.* 2001, *276*, 34156-34161.

(28) Gazit, E. Mechanisms of Amyloid Fibril Self-Assembly and Inhibition. Model Short Peptides as a Key Research Tool. *FEBS Journal* 2005, *272*, 5971-5978.

(29) Inouye, H.; Gleason, K. A.; Zhang, D.; Decatur, S. M.; Kirschner, D. A. Differential Effects of Phe19 and Phe20 on Fibril Formation by Amyloidogenic Peptide a Beta 16-22 (Ac-KLVFFAE-Nh2). *Proteins* 2010, *78*, 2306-2321.

(30) Caughey, B.; Peter T. Lansbury, J. Protofibrils, Pores, Fibrils, and Neurodegeneration: Separating the Responsible Protein Aggregates from the Innocent Bystanders. *Annu. Rev. Neurosci.* 2003, *26*, 267-298.

(31) Butterfield, S.; Hejjaoui, M.; Fauvet, B.; Awad, L.; Lashuel, H. A. Chemical Strategies for Controlling Protein Folding and Elucidating the Molecular Mechanisms of Amyloid Formation and Toxicity. *J. Mol. Biol.* 2012, *421*, 204-236.

(32) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis. *Proteins: Struct., Funct., Bioinf.* 1995, *21*, 167-195.

(33) Dill, K. A. Theory for the Folding and Stability of Globular Proteins. *Biochemistry* 1985, *24*, 1501-1509.

(34) Schellman, J. A. The Thermodynamic Stability of Proteins. *Annu. Rev. Biophys. Biophys. Chem.* 1987, *16*, 115-137.

(35) Becktel, W. J.; Schellman, J. A. Protein Stability Curves. *Biopolymers* 1987, *26*, 1859-1877.

(36) Zwanzig, R. Two-State Models of Protein Folding Kinetics. *Proc. Natl. Acad. Sci. U. S. A.* 1997, *94*, 148-150.

(37) Houston, P. L. *Chemical Kinetics and Reaction Dynamics*; Dover Publications, 2006.

(38) Steinfeld, J. I.; Francisco, J. S.; Hase, W. L. *Chemical Kinetics and Dynamics*; Prentice Hall, 1999.

(39) McQuarrie, D. A.; Simon, J. D. *Physical Chemistry: A Molecular Approach*; University Science Books, 1997.

(40) Hänggi, P.; Talkner, P.; Borkovec, M. Reaction-Rate Theory: Fifty Years after Kramers. *Rev. Mod. Phys.* 1990, *62*, 251-341.

(41) Qiu, L. L.; Hagen, S. J. Internal Friction in the Ultrafast Folding of the Tryptophan Cage. *Chem. Phys.* 2004, *307*, 243-249.

(42) Ansari, A.; Jones, C. M.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. The Role of Solvent Viscosity in the Dynamics of Protein Conformational-Changes. *Science* 1992, *256*, 1796-1798.

(43) Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. Measuring the Rate of Intramolecular Contact Formation in Polypeptides. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 7220-7225.

(44) Cellmer, T.; Henry, E. R.; Hofrichter, J.; Eaton, W. A. Measuring Internal Friction of an Ultrafast-Folding Protein. *Proc. Natl. Acad. Sci. U. S. A.* 2008, *105*, 18320-18325.

(45) Sagnella, D. E.; Straub, J. E.; Thirumalai, D. Time Scales and Pathways for Kinetic Energy Relaxation in Solvated Proteins: Application to Carbonmonoxy Myoglobin. *J. Chem. Phys.* 2000, *113*, 7702-7711.

(46) Markiewicz, B. N.; Jo, H.; Culik, R. M.; DeGrado, W. F.; Gai, F. Assessment of Local Friction in Protein Folding Dynamics Using a Helix Cross-Linker. *J. Phys. Chem. B* 2013, *117*, 14688-14696.

(47) Greenfield, N. J. Using Circular Dichroism Collected as a Function of Temperature to Determine the Thermodynamics of Protein Unfolding and Binding Interactions. *Nat. Protocols* 2007, *1*, 2527-2535.

(48) Kelly, S. M.; Jess, T. J.; Price, N. C. How to Study Proteins by Circular Dichroism. *Biochim. Biophys. Acta, Proteins Proteomics* 2005, *1751*, 119-139.

(49) Berova, N.; Nakanishi, K.; Woody, R. *Circular Dichroism: Principles and Applications*; Wiley, 2000.

(50) Grishina, I. B.; Woody, R. W. Contributions of Tryptophan Sidechains to the Circular Dichroism of Globular Proteins: Exciton Couplets and Coupled Oscillators. *Faraday Discuss.* 1994, *99*, 245-262.

(51) Wu, L.; McElheny, D.; Takekiyo, T.; Keiderling, T. A. Geometry and Efficacy of Cross-Strand Trp/Trp, Trp/Tyr, and Tyr/Tyr Aromatic Interaction in a Beta-Hairpin Peptide. *Biochemistry* 2010, *49*, 4705-4714.

(52) Culik, R. M.; Jo, H.; DeGrado, W. F.; Gai, F. Using Thioamides to Site-Specifically Interrogate the Dynamics of Hydrogen Bond Formation in Beta-Sheet Folding. *J. Am. Chem. Soc.* 2012, *134*, 8026-8029.

(53) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; 3rd ed.; Springer: New York, NY, 2006.

(54) Royer, C. A. Probing Protein Folding and Conformational Transitions with Fluorescence. *Chem. Rev.* 2006, *106*, 1769-1784.

(55) Chen, Y.; Gai, F.; Petrich, J. W. Single-Exponential Fluorescence Decay of the Nonnatural Amino Acid 7-Azatryptophan and the Nonexponential Fluorescence Decay of Tryptophan in Water. *J. Phys. Chem.* 1994, *98*, 2203-2209.

(56) Markiewicz, B. N.; Mukherjee, D.; Troxler, T.; Gai, F. Utility of 5-Cyanotryptophan Fluorescence as a Sensitive Probe of Protein Hydration. *J. Phys. Chem. B* 2016, *120*, 936-944.

(57) Talukder, P.; Chen, S.; Roy, B.; Yakovchuk, P.; Spiering, M. M.; Alam, M. P.; Madathil, M. M.; Bhattacharya, C.; Benkovic, S. J.; Hecht, S. M. Cyanotryptophans as Novel Fluorescent Probes for Studying Protein Conformational Changes and DNA-Protein Interaction. *Biochemistry* 2015, *54*, 7457-7469.

(58) Serrano, A. L.; Troxler, T.; Tucker, M. J.; Gai, F. Photophysics of a Fluorescent Non-Natural Amino Acid: P-Cyanophenylalanine. *Chem. Phys. Lett.* 2010, *487*, 303-306.

(59) Wahl, M. Time-Correlated Single Photon Counting *Technical Note - PicoQuant GmbH* 2014, https://www.picoquant.com/scientific/technical-and-application-notes/category/technical_notes_techniques_and_methods.

(60) Schuler, B.; Eaton, W. A. Protein Folding Studied by Single-Molecule Fret. *Curr. Opin. Struct. Biol.* 2008, *18*, 16-26.

(61) Borgia, A.; Williams, P. M.; Clarke, J. Single-Molecule Studies of Protein Folding. *Annu. Rev. Biochem.* 2008, *77*, 101-125.

(62) Gruebele, M.; Sabelko, J.; Ballew, R.; Ervin, J. Laser Temperature Jump Induced Protein Refolding. *Acc. Chem. Res.* 1998, *31*, 699-707.

(63) Arrondo, J. L. R.; Muga, A.; Castresana, J.; Goñi, F. M. Quantitative Studies of the Structure of Proteins in Solution by Fourier-Transform Infrared Spectroscopy. *Prog. Biophys. Mol. Biol.* 1993, *59*, 23-56.

(64) Barth, A.; Zscherp, C. What Vibrations Tell Us About Proteins. *Q. Rev. Biophys.* 2002, *35*, 369-430.

(65) Surewicz, W. K.; Mantsch, H. H.; Chapman, D. Determination of Protein Secondary Structure by Fourier Transform Infrared Spectroscopy: A Critical Assessment. *Biochemistry* 1993, *32*, 389-394.

(66) Barth, A. Infrared Spectroscopy of Proteins. *Biochim. Biophys. Acta, Bioenerg.* 2007, *1767*, 1073-1101.

(67) Abe, Y.; Krimm, S. Normal Vibrations of Crystalline Polyglycine I. *Biopolymers* 1972, *11*, 1817-1839.

(68) Torii, H.; Tatsumi, T.; Tasumi, M. Effects of Hydration on the Structure, Vibrational Wavenumbers, Vibrational Force Field and Resonance Raman Intensities of N-Methylacetamide. *J. Raman Spectrosc.* 1998, *29*, 537-546.

(69) Hamm, P.; Lim, M.; Hochstrasser, R. M. Structure of the Amide I Band of Peptides Measured by Femtosecond Nonlinear-Infrared Spectroscopy. *J. Phys. Chem. B* 1998, *102*, 6123-6138.

(70) Lee, C.; Cho, M. H. Local Amide I Mode Frequencies and Coupling Constants in Multiple-Stranded Antiparallel Beta-Sheet Polypeptides. *J. Phys. Chem. B* 2004, *108*, 20397-20407.

(71) Lansbury, P. T., Jr.; Costa, P. R.; Griffiths, J. M.; Simon, E. J.; Auger, M.; Halverson, K. J.; Kocisko, D. A.; Hendsch, Z. S.; Ashburn, T. T.; Spencer, R. G.; et al. Structural Model for the Beta-Amyloid Fibril Based on Interstrand Alignment of an Antiparallel-Sheet Comprising a C-Terminal Peptide. *Nat. Struct. Biol.* 1995, *2*, 990-998.

(72) Reisdorf, W. C.; Krimm, S. Infrared Amide I‘ Band of the Coiled Coil. *Biochemistry* 1996, *35*, 1383-1386.

(73) Walsh, S. T. R.; Cheng, R. P.; Wright, W. W.; Alonso, D. O. V.; Daggett, V.; Vanderkooi, J. M.; DeGrado, W. F. The Hydration of Amides in Helices; a Comprehensive Picture from Molecular Dynamics, IR, and NMR. *Protein Sci.* 2003, *12*, 520-531.

(74) Bernath, P. F. *Spectra of Atoms and Molecules*; Oxford University Press, 2005.

(75) Barth, A. The Infrared Absorption of Amino Acid Sidechains. *Prog. Biophys. Mol. Biol.* 2000, *74*, 141-173.

(76) Tamm, L. K.; Tatulian, S. A. Infrared Spectroscopy of Proteins and Peptides in Lipid Bilayers. *Q. Rev. Biophys.* 1997, *30*, 365-429.

(77) Goormaghtigh, E.; Raussens, V.; Ruysschaert, J. M. Attenuated Total Reflection Infrared Spectroscopy of Proteins and Lipids in Biological Membranes. *Biochim. Biophys. Acta, Rev. Biomembr.* 1999, *1422*, 105-185.

(78) Chittur, K. K. FTIR/Atr for Protein Adsorption to Biomaterial Surfaces. *Biomaterials* 1998, *19*, 357-369.

(79) Vigano, C.; Manciu, L.; Buyse, F.; Goormaghtigh, E.; Ruysschaert, J. M. Attenuated Total Reflection IR Spectroscopy as a Tool to Investigate the Structure, Orientation and Tertiary Structure Changes in Peptides and Membrane Proteins. *Pept. Sci.* 2000, *55*, 373-380.

(80) Decatur, S. M. Elucidation of Residue-Level Structure and Dynamics of Polypeptides Via Isotope-Edited Infrared Spectroscopy. *Acc. Chem. Res.* 2006, *39*, 169-175.

(81) Frey, S.; Tamm, L. K. Orientation of Melittin in Phospholipid Bilayers. A Polarized Attenuated Total Reflection Infrared Study. *Biophys. J.* 1991, *60*, 922-930.

(82) Tucker, M. J.; Getahun, Z.; Nanda, V.; DeGrado, W. F.; Gai, F. A New Method for Determining the Local Environment and Orientation of Individual Sidechains of Membrane-Binding Peptides. *J. Am. Chem. Soc.* 2004, *126*, 5078-5079.

(83) Huang, C.-Y.; Klemke, J. W.; Getahun, Z.; DeGrado, W. F.; Gai, F. Temperature-Dependent Helix−Coil Transition of an Alanine Based Peptide. *J. Am. Chem. Soc.* 2001, *123*, 9235-9238.

(84) Dyer, R. B.; Gai, F.; Woodruff, W. H.; Gilmanshin, R.; Callender, R. H. Infrared Studies of Fast Events in Protein Folding. *Acc. Chem. Res.* 1998, *31*, 709-716.

(85) Petrovich, M.; Jonsson, A. L.; Ferguson, N.; Daggett, V.; Fersht, A. R. Phi-Analysis at the Experimental Limits: Mechanism of Beta-Hairpin Formation. *J. Mol. Biol.* 2006, *360*, 865-881.

(86) Serrano, A. L.; Tucker, M. J.; Gai, F. Direct Assessment of the Alpha-Helix Nucleation Time. *J. Phys. Chem. B* 2011, *115*, 7472-7478.

(87) Leduc, A.-M.; Trent, J. O.; Wittliff, J. L.; Bramlett, K. S.; Briggs, S. L.; Chirgadze, N. Y.; Wang, Y.; Burris, T. P.; Spatola, A. F. Helix-Stabilized Cyclic Peptides as Selective Inhibitors of Steroid Receptor–Coactivator Interactions. *Proc. Natl. Acad. Sci. U. S. A.* 2003, *100*, 11273-11278.

(88) Almeida, A. M.; Li, R.; Gellman, S. H. Parallel β-Sheet Secondary Structure Is Stabilized and Terminated by Interstrand Disulfide Cross-Linking. *J. Am. Chem. Soc.* 2011, *134*, 75-78.

(89) Blackwell, H. E.; Grubbs, R. H. Highly Efficient Synthesis of Covalently Cross-Linked Peptide Helices by Ring-Closing Metathesis. *Angew. Chem., Int. Ed.* 1998, *37*, 3281-3284.

(90) Schafmeister, C. E.; Po, J.; Verdine, G. L. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc.* 2000, *122*, 5891-5892.

(91) Osapay, G.; Taylor, J. W. Multicyclic Polypeptide Model Compounds. 2. Synthesis and Conformational Properties of a Highly .Alpha.-Helical Uncosapeptide Constrained by Three Side-Chain to Side-Chain Lactam Bridges. *J. Am. Chem. Soc.* 1992, *114*, 6966-6973.

(92) Phelan, J. C.; Skelton, N. J.; Braisted, A. C.; McDowell, R. S. A General Method for Constraining Short Peptides to an A-Helical Conformation. *J. Am. Chem. Soc.* 1997, *119*, 455-460.

(93) Verdine, G. L.; Hilinski, G. J. In *Methods in Enzymology: Protein Engineering for Therapeutics, Vol 203, Pt B*; Wittrup, K. D., Verdine, G. L., Eds.; Elsevier Academic Press Inc: San Diego, 2012; Vol. 503, p 3-33.

(94) Rao, T.; Ruiz-Gomez, G.; Hill, T. A.; Hoang, H. N.; Fairlie, D. P.; Mason, J. M. Truncated and Helix-Constrained Peptides with High Affinity and Specificity for the Cfos Coiled-Coil of Ap-1. *PLoS One* 2013, *8*.

(95) Cabezas, E.; Satterthwait, A. C. The Hydrogen Bond Mimic Approach: Solid-Phase Synthesis of a Peptide Stabilized as an A-Helix with a Hydrazone Link. *J. Am. Chem. Soc.* 1999, *121*, 3862-3875.

(96) Haney, C. M.; Loch, M. T.; Horne, W. S. Promoting Peptide [Small Alpha]-Helix Formation with Dynamic Covalent Oxime Side-Chain Cross-Links. *Chem. Commun.* 2011, *47*, 10915-10917.

(97) Walker, M. A.; Johnson, T. General Method for the Synthesis of Cyclic Peptidomimetic Compounds. *Tetrahedron Lett.* 2001, *42*, 5801-5804.

(98) Muppidi, A.; Wang, Z. Y.; Li, X. L.; Chen, J. D.; Lin, Q. Achieving Cell Penetration with Distance-Matching Cysteine Cross-Linkers: A Facile Route to Cell-Permeable Peptide Dual Inhibitors of Mdm2/Mdmx. *Chem. Commun.* 2011, *47*, 9396-9398.

(99) Jo, H.; Meinhardt, N.; Wu, Y. B.; Kulkarni, S.; Hu, X. Z.; Low, K. E.; Davies, P. L.; DeGrado, W. F.; Greenbaum, D. C. Development of Alpha-Helical Calpain Probes by Mimicking a Natural Protein-Protein Interaction. *J. Am. Chem. Soc.* 2012, *134*, 17704-17713.

(100) Patgiri, A.; Jochim, A. L.; Arora, P. S. A Hydrogen Bond Surrogate Approach for Stabilization of Short Peptide Sequences in Alpha-Helical Conformation. *Acc. Chem. Res.* 2008, *41*, 1289-1300.

(101) Wang, D.; Chen, K.; Kulp, J. L.; Arora, P. S. Evaluation of Biologically Relevant Short A-Helices Stabilized by a Main-Chain Hydrogen-Bond Surrogate. *J. Am. Chem. Soc.* 2006, *128*, 9248-9256.

(102) Clarke, J.; Fersht, A. R. Engineered Disulfide Bonds as Probes of the Folding Pathway of Barnase: Increasing the Stability of Proteins against the Rate of Denaturation. *Biochemistry* 1993, *32*, 4322-4329.

(103) Grantcharova, V. P.; Riddle, D. S.; Baker, D. Long-Range Order in the Src SH3 Folding Transition State. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 7084-7089.

(104) Du, D. G.; Gai, F. Understanding the Folding Mechanism of an Alpha-Helical Hairpin. *Biochemistry* 2006, *45*, 13131-13139.

(105) Shandiz, A. T.; Capraro, B. R.; Sosnick, T. R. Intramolecular Cross-Linking Evaluated as a Structural Probe of the Protein Folding Transition State. *Biochemistry* 2007, *46*, 13711-13719.

(106) Chung, H. S.; Shandiz, A.; Sosnick, T. R.; Tokmakoff, A. Probing the Folding Transition State of Ubiquitin Mutants by Temperature-Jump-Induced Downhill Unfolding. *Biochemistry* 2008, *47*, 13870-13877.

(107) Bosco, G. L.; Baxa, M.; Sosnick, T. R. Metal Binding Kinetics of Bi-Histidine Sites Used in Ψ Analysis: Evidence of High-Energy Protein Folding Intermediates†. *Biochemistry* 2009, *48*, 2950-2959.

(108) Tucker, M. J.; Courter, J. R.; Chen, J. X.; Atasoylu, O.; Smith, A. B.; Hochstrasser, R. M. Tetrazine Phototriggers: Probes for Peptide Dynamics. *Angew. Chem., Int. Ed.* 2010, *49*, 3612-3616.

(109) Hansen, K. C.; Rock, R. S.; Larsen, R. W.; Chan, S. I. A Method for Photoinitating Protein Folding in a Nondenaturing Environment. *J. Am. Chem. Soc.* 2000, *122*, 11567-11568.

(110) Bochet, C. G. Photolabile Protecting Groups and Linkers. *J. Chem. Soc., Perkin Trans. 1* 2002, 125-142.

(111) Loudwig, S.; Bayley, H. In *Dynamic Studies in Biology: Phototriggers, Photoswitches and Caged Biomolecules*; Goeldner, M., Givens, R., Eds.; WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2005, p 253-303.

(112) Mayer, G.; Heckel, A. Biologically Active Molecules with a "Light Switch". *Angew. Chem., Int. Ed.* 2006, *45*, 4900-4921.

(113) Li, Y.; Foss, C. A.; Summerfield, D. D.; Doyle, J. J.; Torok, C. M.; Dietz, H. C.; Pomper, M. G.; Yu, S. M. Targeting Collagen Strands by Photo-Triggered Triple-Helix Hybridization. *Proc. Natl. Acad. Sci. U. S. A.* 2012, *109*, 14767-14772.

(114) Hirota, S.; Fujimoto, Y.; Choi, J.; Baden, N.; Katagiri, N.; Akiyama, M.; Hulsker, R.; Ubbink, M.; Okajima, T.; Takabe, T.; Funasaki, N.; Watanabe, Y.; Terazima, M. Conformational Changes During Apoplastocyanin Folding Observed by Photocleavable Modification and Transient Grating. *J. Am. Chem. Soc.* 2006, *128*, 7551-7558.

(115) Chen, H. L.; Hsu, J. C. C.; Man, H. V.; Li, M. S.; Hu, C. K.; Liu, C. H.; Luh, F. Y.; Chen, S. S. W.; Chang, E. S. H.; Wang, A. H. J.; Hsu, M. F.; Fann, W.; Chen,

R. P. Y. Studying Submicrosecond Protein Folding Kinetics Using a Photolabile Caging Strategy and Time-Resolved Photoacoustic Calorimetry. *Proteins* 2010, *78*, 2973-2983.

(116) Guerrero, L.; Smart, O. S.; Woolley, G. A.; Allemann, R. K. Photocontrol of DNA Binding Specificity of a Miniature Engrailed Homeodomain. *J. Am. Chem. Soc.* 2005, *127*, 15624-15629.

(117) Jurt, S.; Aemissegger, A.; Guntert, P.; Zerbe, O.; Hilvert, D. A Photoswitchable Miniprotein Based on the Sequence of Avian Pancreatic Polypeptide. *Angew. Chem., Int. Ed.* 2006, *45*, 6297-6300.

(118) Pfister, R.; Ihalainen, J.; Hamm, P.; Kolano, C. Synthesis, Characterization and Applicability of Three Isotope Labeled Azobenzene Photoswitches. *Org. Biomol. Chem.* 2008, *6*, 3508-3517.

(119) Hoppmann, C.; Schmieder, P.; Heinrich, N.; Beyermann, M. Photoswitchable Click Amino Acids: Light Control of Conformation and Bioactivity. *ChemBioChem* 2011, *12*, 2555-2559.

(120) Ali, A. M.; Woolley, G. A. The Effect of Azobenzene Cross-Linker Position on the Degree of Helical Peptide Photo-Control. *Org. Biomol. Chem.* 2013, *11*, 5325-5331.

(121) Korbus, M.; Balasubramanian, G.; Muller-Plathe, F.; Kolmar, H.; Meyer-Almes, F. J. Azobenzene Switch with a Long-Lived Cis-State to Photocontrol the Enzyme Activity of a Histone Deacetylase-Like Amidohydrolase. *Biol. Chem.* 2014, *395*, 401-412.

(122) Sension, R. J.; Repinec, S. T.; Szarka, A. Z.; Hochstrasser, R. M. Femtosecond Laser Studies of the Cis-Stilbene Photoisomerization Reactions. *J. Chem. Phys.* 1993, *98*, 6291-6315.

(123) Erdelyi, M.; Karlen, A.; Gogoll, A. A New Tool in Peptide Engineering: A Photoswitchable Stilbene-Type Beta-Hairpin Mimetic. *Chem.- Eur. J.* 2006, *12*, 403-412.

(124) Varedian, M.; Erdelyi, M.; Persson, A.; Gogoll, A. Interplaying Factors for the Formation of Photoswitchable Beta-Hairpins: The Advantage of a Flexible Switch. *J. Pept. Sci.* 2009, *15*, 107-113.

(125) Lednev, I. K.; Ye, T. Q.; Hester, R. E.; Moore, J. N. Femtosecond Time-Resolved UV-Visible Absorption Spectroscopy of Trans-Azobenzene in Solution. *J. Phys. Chem.* 1996, *100*, 13338-13341.

(126) Sporlein, S.; Carstens, H.; Satzger, H.; Renner, C.; Behrendt, R.; Moroder, L.; Tavan, P.; Zinth, W.; Wachtveitl, J. Ultrafast Spectroscopy Reveals Subnanosecond Peptide Conformational Dynamics and Validates Molecular Dynamics Simulation. *Proc. Natl. Acad. Sci. U. S. A.* 2002, *99*, 7998-8002.

(127) Kumita, J. R.; Smart, O. S.; Woolley, G. A. Photo-Control of Helix Content in a Short Peptide. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 3803-3808.

(128) 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. *Proc. Natl. Acad. Sci. U. S. A.* 2005, *102*, 2379-2384.

(129) Waldauer, S. A.; Hassan, S.; Paoli, B.; Donaldson, P. M.; Pfister, R.; Hamm, P.; Caflisch, A.; Pellarin, R. Photocontrol of Reversible Amyloid Formation with a Minimal-Design Peptide. *J. Phys. Chem. B* 2012, *116*, 8961-8973.

(130) Deeg, A. A.; Schrader, T. E.; Pfizer, J.; Moroder, L.; Zinth, W. 2013; Vol. 7, p 209-213.

(131) Taniguchi, A.; Skwarczynski, M.; Sohma, Y.; Okada, T.; Ikeda, K.; Prakash, H.; Mukai, H.; Hayashi, Y.; Kimura, T.; Hirota, S.; Matsuzaki, K.; Kiso, Y. Controlled Production of Amyloid Beta Peptide from a Photo-Triggered, Water-Soluble Precursor "Click Peptide". *ChemBioChem* 2008, *9*, 3055-3065.

(132) Zhang, S. Fabrication of Novel Biomaterials through Molecular Self-Assembly. *Nat. Biotechnol.* 2003, *21*, 1171-1178.

(133) Matsuzawa, Y.; Ueki, K.; Yoshida, M.; Tamaoki, N.; Nakamura, T.; Sakai, H.; Abe, M. Assembly and Photoinduced Organization of Mono- and Oligopeptide Molecules Containing an Azobenzene Moiety. *Adv. Funct. Mater.* 2007, *17*, 1507-1514.

(134) Huang, Y. C.; Qiu, Z. J.; Xu, Y. M.; Shi, J. F.; Lin, H. K.; Zhang, Y. Supramolecular Hydrogels Based on Short Peptides Linked with Conformational Switch. *Org. Biomol. Chem.* 2011, *9*, 2149-2155.

(135) Lin, Y. Y.; Qiao, Y.; Tang, P. F.; Li, Z. B.; Huang, J. B. Controllable Self-Assembled Laminated Nanoribbons from Dipeptide-Amphiphile Bearing Azobenzene Moiety. *Soft Matter* 2011, *7*, 2762-2769.

(136) Qiu, Z. J.; Yu, H. T.; Li, J. B.; Wang, Y.; Zhang, Y. Spiropyran-Linked Dipeptide Forms Supramolecular Hydrogel with Dual Responses to Light and to Ligand-Receptor Interaction. *Chem. Comm.* 2009, 3342-3344.

(137) Haines, L. A.; Rajagopal, K.; Ozbas, B.; Salick, D. A.; Pochan, D. J.; Schneider, J. P. Light-Activated Hydrogel Formation Via the Triggered Folding and Self-Assembly of a Designed Peptide. *J. Am. Chem. Soc.* 2005, *127*, 17025-17029.

(138) Putnam, C. D.; Hammel, M.; Hura, G. L.; Tainer, J. A. X-Ray Solution Scattering (Saxs) Combined with Crystallography and Computation: Defining Accurate Macromolecular Structures, Conformations and Assemblies in Solution. *Q. Rev. Biophys.* 2007, *40*, 191-285.

(139) Cross, T. A.; Opella, S. J. Solid-State NMR Structural Studies of Peptides and Proteins in Membranes. *Curr. Opin. Struct. Biol.* 1994, *4*, 574-581.

(140) Owrutsky, J. C.; Raftery, D.; Hochstrasser, R. M. Vibrational-Relaxation Dynamics in Solutions. *Annu. Rev. Phys. Chem.* 1994, *45*, 519-555.

(141) Waegele, M. M.; Culik, R. M.; Gai, F. Site-Specific Spectroscopic Reporters of the Local Electric Field, Hydration, Structure, and Dynamics of Biomolecules. *J. Phys. Chem. Lett.* 2011, *2*, 2598-2609

(142) Ma, J.; Pazos, I. M.; Zhang, W.; Culik, R. M.; Gai, F. Site-Specific Infrared Probes of Proteins. *Annu. Rev. Phys. Chem.* 2015, *66*, 357-377.

(143) Kim, Y. S.; Hochstrasser, R. M. Applications of 2D IR Spectroscopy to Peptides, Proteins, and Hydrogen-Bond Dynamics. *J. Phys. Chem. B* 2009, *113*, 8231-8251.

(144) Tadesse, L.; Nazarbaghi, R.; Walters, L. Isotopically Enhanced Infrared Spectroscopy: A Novel Method for Examining Secondary Structure at Specific Sites in Conformationally Heterogeneous Peptides. *J. Am. Chem. Soc.* 1991, *113*, 7036-7037.

(145) Zimmermann, J.; Gundogdu, K.; Cremeens, M. E.; Bandaria, J. N.; Hwang, G. T.; Thielges, M. C.; Cheatum, C. M.; Romesberg, F. E. Efforts toward Developing Probes of Protein Dynamics: Vibrational Dephasing and Relaxation of Carbon–

Deuterium Stretching Modes in Deuterated Leucine. *J. Phys. Chem. B* 2009, *113*, 7991-7994.

(146) Schultz, K. C.; Supekova, L.; Ryu, Y. H.; Xie, J. M.; Perera, R.; Schultz, P. G. A Genetically Encoded Infrared Probe. *J. Am. Chem. Soc.* 2006, *128*, 13984-13985.

(147) Jo, H.; Culik, R. M.; Korendovych, I. V.; DeGrado, W. F.; Gai, F. Selective Incorporation of Nitrile-Based Infrared Probes into Proteins Via Cysteine Alkylation. *Biochemistry* 2010, *49*, 10354-10356.

(148) Connor, R. E.; Tirrell, D. A. Non-Canonical Amino Acids in Protein Polymer Design. *Polym. Rev.* 2007, *47*, 9-28.

(149) Nagarajan, S.; Taskent-Sezgin, H.; Parul, D.; Carrico, I.; Raleigh, D. P.; Dyer, R. B. Differential Ordering of the Protein Backbone and Sidechains During Protein Folding Revealed by Site-Specific Recombinant Infrared Probes. *J. Am. Chem. Soc.* 2011, *133*, 20335-20340.

(150) Pazos, I. M.; Ghosh, A.; Tucker, M. J.; Gai, F. Ester Carbonyl Vibration as a Sensitive Probe of Protein Local Electric Field. *Angew. Chem., Int. Ed.* 2014, *53*, 6080-6084.

(151) King, J. T.; Kubarych, K. J. Site-Specific Coupling of Hydration Water and Protein Flexibility Studied in Solution with Ultrafast 2D-IR Spectroscopy. *J. Am. Chem. Soc.* 2012, *134*, 18705-18712.

(152) Lindquist, B. A.; Furse, K. E.; Corcelli, S. A. Nitrile Groups as Vibrational Probes of Biomolecular Structure and Dynamics: An Overview. *Phys. Chem. Chem. Phys.* 2009, *11*, 8119-8132.

(153) Getahun, Z.; Huang, C.-Y.; Wang, T.; De León, B.; DeGrado, W. F.; Gai, F. Using Nitrile-Derivatized Amino Acids as Infrared Probes of Local Environment. *J. Am. Chem. Soc.* 2003, *125*, 405-411.

(154) Fafarman, A. T.; Webb, L. J.; Chuang, J. I.; Boxer, S. G. Site-Specific Conversion of Cysteine Thiols into Thiocyanate Creates an IR Probe for Electric Fields in Proteins. *J. Am. Chem. Soc.* 2006, *128*, 13356-13357.

(155) Waegele, M. M.; Tucker, M. J.; Gai, F. 5-Cyanotryptophan as an Infrared Probe of Local Hydration Status of Proteins. *Chem. Phys. Lett.* 2009, *478*, 249-253.

(156) Fafarman, A. T.; Sigala, P. A.; Herschlag, D.; Boxer, S. G. Decomposition of Vibrational Shifts of Nitriles into Electrostatic and Hydrogen-Bonding Effects. *J. Am. Chem. Soc.* 2010, *132*, 12811-12813.

(157) Sigala, P. A.; Fafarman, A. T.; Bogard, P. E.; Boxer, S. G.; Herschlag, D. Do Ligand Binding and Solvent Exclusion Alter the Electrostatic Character within the Oxyanion Hole of an Enzymatic Active Site? *J. Am. Chem. Soc.* 2007, *129*, 12104-12105.

(158) Walker, D. M.; Hayes, E. C.; Webb, L. J. Vibrational Stark Effect Spectroscopy Reveals Complementary Electrostatic Fields Created by Protein-Protein Binding at the Interface of Ras and Ral. *Phys. Chem. Chem. Phys.* 2013, *15*, 12241-12252.

(159) Marek, P.; Mukherjee, S.; Zanni, M. T.; Raleigh, D. P. Residue-Specific, Real-Time Characterization of Lag-Phase Species and Fibril Growth During Amyloid Formation: A Combined Fluorescence and IR Study of P-Cyanophenylalanine Analogs of Islet Amyloid Polypeptide. *J. Mol. Biol.* 2010, *400*, 878-888

(160) Markiewicz, B. N.; Yang, L.; Culik, R. M.; Gao, Y. Q.; Gai, F. How Quickly Can a Beta-Hairpin Fold from Its Transition State? *J. Phys. Chem. B* 2014, *118*, 3317-3325.

(161) Royer, C. A. The Nature of the Transition State Ensemble and the Mechanisms of Protein Folding: A Review. *Arch. Biochem. Biophys.* 2008, *469*, 34-45.

(162) Moran, L. B.; Schneider, J. P.; Kentsis, A.; Reddy, G. A.; Sosnick, T. R. Transition State Heterogeneity in GCN4 Coiled Coil Folding Studied by Using Multisite Mutations and Crosslinking. *Proc. Natl. Acad. Sci. U. S. A.* 1999, *96*, 10699-10704.

(163) Krantz, B. A.; Sosnick, T. R. Engineered Metal Binding Sites Map the Heterogeneous Folding Landscape of a Coiled Coil. *Nat. Struct. Biol.* 2001, *8*, 1042-1047.

(164) Geierhaas, C. D.; Salvatella, X.; Clarke, J.; Vendruscolo, M. Characterisation of Transition State Structures for Protein Folding Using 'High', 'Medium' and 'Low' Phi-Values. *Protein Eng. Des. Sel.* 2008, *21*, 215-222.

(165) Fersht, A. R.; Sato, S. Phi-Value Analysis and the Nature of Protein-Folding Transition States. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101*, 7976-7981.

(166) Pauling, L. Molecular Architecture and Biological Reactions. *Chem. Eng. News. Arch.* 1946, *24*, 1375-1377.

(167) Wolfende, R. Transition State Analogues for Enzyme Catalysis. *Nature* 1969, *223*, 704-705.

(168) Wolfende, R. Analog Approaches to Structure of Transition-State in Enzyme Reactions. *Acc. Chem. Res.* 1972, *5*, 10-18.

(169) Schramm, V. L. Transition States, Analogues, and Drug Development. *ACS Chem. Biol.* 2013, *8*, 71-81.

(170) Fedorov, A.; Shi, W.; Kicska, G.; Fedorov, E.; Tyler, P. C.; Furneaux, R. H.; Hanson, J. C.; Gainsford, G. J.; Larese, J. Z.; Schramm, V. L.; Almo, S. C. Transition State Structure of Purine Nucleoside Phosphorylase and Principles of Atomic Motion in Enzymatic Catalysis. *Biochemistry* 2001, *40*, 853-860.

(171) Schramm, V. L. Enzymatic Transition State Theory and Transition State Analogue Design. *J. Biol. Chem.* 2007, *282*, 28297-28300.

(172) Klimov, D. K.; Thirumalai, D. Mechanisms and Kinetics of Beta-Hairpin Formation. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 2544-2549.

(173) Kobayashi, N.; Honda, S.; Yoshii, H.; Munekata, E. Role of Side-Chains in the Cooperative Beta-Hairpin Folding of the Short C−Terminal Fragment Derived from Streptococcal Protein G. *Biochemistry* 2000, *39*, 6564-6571.

(174) Espinosa, J. F.; Syud, F. A.; Gellman, S. H. Analysis of the Factors That Stabilize a Designed Two-Stranded Antiparallel Beta-Sheet. *Protein Sci.* 2002, *11*, 1492-1505.

(175) Maness, S. J.; Franzen, S.; Gibbs, A. C.; Causgrove, T. P.; Dyer, R. B. Nanosecond Temperature Jump Relaxation Dynamics of Cyclic Beta-Hairpin Peptides. *Biophys. J.* 2003, *84*, 3874-3882.

(176) Baumketner, A.; Shea, J.-E. The Thermodynamics of Folding of a Beta Hairpin Peptide Probed through Replica Exchange Molecular Dynamics Simulations. *Theor. Chem. Acc.* 2006, *116*, 262-273.

(177) Best, R. B.; Mittal, J. Microscopic Events in Beta-Hairpin Folding from Alternative Unfolded Ensembles. *Proc. Natl. Acad. Sci. U. S. A.* 2011, *108*, 11087-11092.

(178) Xu, Y.; Du, D.; Oyola, R. Infrared Study of the Stability and Folding Kinetics of a Series of Beta-Hairpin Peptides with a Common NPDG Turn. *J. Phys. Chem. B* 2011, *115*, 15332-15338.

(179) Santiveri, C. M.; Leon, E.; Rico, M.; Jimenez, M. A. Context-Dependence of the Contiribution of Disulfide Bonds to Beta-Hairpin Stability. *Chem.- Eur. J.* 2008, *14*, 488-499.

(180) Carulla, N.; Woodward, C.; Barany, G. Synthesis and Characterization of a Beta-Hairpin Peptide That Represents a 'Core Module' of Bovine Pancreatic Trypsin Inhibitor (BPTI). *Biochemistry* 2000, *39*, 7927-7937.

(181) Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. Stability of Cyclic Beta-Hairpins: Asymmetric Contributions from Sidechains of a Hydrogen-Bonded Cross-Strand Residue Pair. *J. Am. Chem. Soc.* 2003, *125*, 388-395.

(182) Mirassou, Y.; Santiveri, C. M.; Perez de Vega, M. J.; Gonzalez-Muniz, R.; Jimenez, M. A. Disulfide Bonds Versus Trp-Trp Pairs in Irregular Beta-Hairpins: NMR Structure of Vammin Loop 3-Derived Peptides as a Case Study. *ChemBioChem* 2009, *10*, 902-910.

(183) 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. *Proc. Natl. Acad. Sci. U. S. A.* 2008, *105*, 9588-9593.

(184) Wang, T.; Lau, W. L.; DeGrado, W. F.; Gai, F. T-Jump Infrared Study of the Folding Mechanism of Coiled-Coil GCN4-P1. *Biophys. J.* 2005, *89*, 4180-4187.

(185) Munoz, V.; Thompson, P. A.; Hofrichter, J.; Eaton, W. A. Folding Dynamics and Mechanism of Beta-Hairpin Formation. *Nature* 1997, *390*, 196-199.

(186) Chen, R. P.-Y.; Huang, J. J.-T.; Chen, H.-L.; Jan, H.; Velusamy, M.; Lee, C.-T.; Fann, W.; Larsen, R. W.; Chan, S. I. Measuring the Refolding of Beta-Sheets with Different Turn Sequences on a Nanosecond Time Scale. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101*, 7305-7310.

(187) Dyer, R. B.; Maness, S. J.; Peterson, E. S.; Franzen, S.; Fesinmeyer, R. M.; Andersen, N. H. The Mechanism of β-Hairpin Formation. *Biochemistry* 2004, *43*, 11560-11566.

(188) Du, D.; Zhu, Y.; Huang, C. Y.; Gai, F. Understanding the Key Factors That Control the Rate of Beta-Hairpin Folding. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101*, 15915-15920.

(189) Yang, W. Y.; Pitera, J. W.; Swope, W. C.; Gruebele, M. Heterogeneous Folding of the Trpzip Hairpin: Full Atom Simulation and Experiment. *J. Mol. Biol.* 2004, *336*, 241-251.

(190) Du, D.; Tucker, M. J.; Gai, F. Understanding the Mechanism of Beta-Hairpin Folding Via Phi-Value Analysis. *Biochemistry* 2006, *45*, 2668-2678.

(191) Pitera, J. W.; Haque, I.; Swope, W. C. Absence of Reptation in the High-Temperature Folding of the Trpzip2 Beta-Hairpin Peptide. *J. Chem. Phys.* 2006, *124*.

(192) Zhang, J.; Qin, M.; Wang, W. Folding Mechanism of β-Hairpins Studied by Replica Exchange Molecular Simulations. *Proteins: Struct., Funct., Bioinf.* 2006, *62*, 672-685.

(193) Narayanan, R.; Pelakh, L.; Hagen, S. J. Solvent Friction Changes the Folding Pathway of the Tryptophan Zipper TZ2. *J. Mol. Biol.* 2009, *390*, 538-546.

(194) Roy, S.; Jansen, T. L. C.; Knoester, J. Structural Classification of the Amide I Sites of a B-Hairpin with Isotope Label 2DIR Spectroscopy. *Phys. Chem. Chem. Phys.* 2010, *12*, 9347-9357.

(195) Smith, A. W.; Lessing, J.; Ganim, Z.; Peng, C. S.; Tokmakoff, A.; Roy, S.; Jansen, T. L. C.; Knoester, J. Melting of a β-Hairpin Peptide Using Isotope-Edited 2D IR Spectroscopy and Simulations. *J. Phys. Chem. B* 2010, *114*, 10913-10924.

(196) 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. *Chem. Commun.* 2012, *48*, 487-497.

(197) Deeg, A. A.; Rampp, M. S.; Popp, A.; Pilles, B. M.; Schrader , T. E.; Moroder, L.; Hauser, K.; Zinth, W. Isomerization- and Temperature-Jump-Induced Dynamics of a Photoswitchable β-Hairpin. *Chem. - Eur. J.* 2013, 1-11.

(198) Serrano, A. L.; Waegele, M. M.; Gai, F. Spectroscopic Studies of Protein Folding: Linear and Nonlinear Methods. *Protein Sci.* 2012, *21*, 157-170.

(199) Tam, J. P.; Wu, C. R.; Liu, W.; Zhang, J. W. Disulfide Bond Formation in Peptides by Dimethyl Sulfoxide. Scope and Applications. *J. Am. Chem. Soc.* 1991, *113*, 6657-6662.

(200) 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. *Angew. Chem., Int. Ed.* 2011, *50*, 10884-10887.

(201) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A 2nd Generation Force-Field for the Simulation of Proteins, Nucleic-Acids, and Organic-Molecules. *J. Am. Chem. Soc.* 1995, *117*, 5179-5197.

(202) Onufriev, A.; Bashford, D.; Case, D. A. Exploring Protein Native States and Large-Scale Conformational Changes with a Modified Generalized Born Model. *Proteins* 2004, *55*, 383-394.

(203) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. Numerical-Integration of Cartesian Equations of Motion of a System with Constraints - Molecular-Dynamics of N-Alkanes. *J. Comput. Phys.* 1977, *23*, 327-341.

(204) Gao, Y. Q. An Integrate-over-Temperature Approach for Enhanced Sampling. *J. Chem. Phys.* 2008, *128*.

(205) Shao, Q. A.; Gao, Y. Q. Temperature Dependence of Hydrogen-Bond Stability in Beta-Hairpin Structures. *J. Chem. Theory Comput.* 2010, *6*, 3750-3760.

(206) Yang, L. J.; Shao, Q.; Gao, Y. Q. Enhanced Sampling Method in Molecular Simulations. *Prog. Chem.* 2012, *24*, 1199-1213.

(207) Streicher, W. W.; Makhatadze, G. I. Calorimetric Evidence for a Two-State Unfolding of the Beta-Hairpin Peptide Trpzip4. *J. Am. Chem. Soc.* 2006, *128*, 30-31.

(208) Shao, Q.; Wei, H.; Gao, Y. Q. Effects of Turn Stability and Side-Chain Hydrophobicity on the Folding of β-Structures. *J. Mol. Biol.* 2010, *402*, 595-609.

(209) Hwang, S.; Hilty, C. Folding of a Tryptophan Zipper Peptide Investigated on the Basis of the Nuclear Overhauser Effect and Thermal Denaturation. *J. Phys. Chem. B* 2011, *115*, 15355-15361.

(210) Juraszek, J.; Vreede, J.; Bolhuis, P. G. Transition Path Sampling of Protein Conformational Changes. *Chem. Phys.* 2012, *396*, 30-44.

(211) Liao, C. Y.; Zhou, J. Replica Exchange Molecular Dynamics Simulations on the Folding of Trpzip4 Beta-Hairpin. *Acta Chim. Sin.* 2013, *71*, 593-601.

(212) Cochran, A. G.; Skelton, N. J.; Starovasnik, M. A. Tryptophan Zippers: Stable, Monomeric β-Hairpins. *Proc. Natl. Acad. Sci. U. S. A.* 2001, *98*, 5578-5583.

(213) Guvench, O.; Brooks, C. L. Tryptophan Sidechain Electrostatic Interactions Determine Edge-to-Face Vs Parallel-Displaced Tryptophan Sidechain Geometries in the Designed β-Hairpin "Trpzip2". *J. Am. Chem. Soc.* 2005, *127*, 4668-4674.

(214) Wu, L.; McElheny, D.; Huang, R.; Keiderling, T. A. Role of Tryptophan−Tryptophan Interactions in Trpzip β-Hairpin Formation, Structure, and Stability. *Biochemistry* 2009, *48*, 10362-10371.

(215) Abkevich, V. I.; Shakhnovich, E. I. What Can Disulfide Bonds Tell Us About Protein Energetics, Function and Folding: Simulations and Bioninformatics Analysis. *J. Mol. Biol.* 2000, *300*, 975-985.

(216) Fesinmeyer, R. M.; Hudson, F. M.; Andersen, N. H. Enhanced Hairpin Stability through Loop Design: The Case of the Protein G B1 Domain Hairpin. *J. Am. Chem. Soc.* 2004, *126*, 7238-7243.

(217) 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. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101*, 4077-4082.

(218) Dyer, R. B. Ultrafast and Downhill Protein Folding. *Curr. Opin. Struct. Biol.* 2007, *17*, 38-47.

(219) Munoz, V.; Sadqi, M.; Naganathan, A. N.; de Sancho, D. Exploiting the Downhill Folding Regime Via Experiment. *Hfsp J.* 2008, *2*, 342-353.

(220) Davis, C. M.; Xiao, S.; Raleigh, D. P.; Dyer, R. B. Raising the Speed Limit for β-Hairpin Formation. *J. Am. Chem. Soc.* 2012, *134*, 14476-14482.

(221) Kolano, C.; Helbing, J.; Kozinski, M.; Sander, W.; Hamm, P. Watching Hydrogen-Bond Dynamics in a Beta-Turn by Transient Two-Dimensional Infrared Spectroscopy. *Nature* 2006, *444*, 469-472.

(222) Cheng, R. R.; Uzawa, T.; Plaxco, K. W.; Makarov, D. E. Universality in the Timescales of Internal Loop Formation in Unfolded Proteins and Single-Stranded Oligonucleotides. *Biophys. J.* 2010, *99*, 3959-3968.

(223) Xu, Y.; Bunagan, M. R.; Tang, J.; Gai, F. Probing the Kinetic Cooperativity of β-Sheet Folding Perpendicular to the Strand Direction. *Biochemistry* 2008, *47*, 2064-2070.

(224) Socci, N. D.; Onuchic, J. N.; Wolynes, P. G. Diffusive Dynamics of the Reaction Coordinate for Protein Folding Funnels. *J. Chem. Phys.* 1996, *104*, 5860-5868.

(225) Shea, J. E.; Onuchic, J. N.; Brooks, C. L. Exploring the Origins of Topological Frustration: Design of a Minimally Frustrated Model of Fragment B of Protein A. *Proc. Natl. Acad. Sci. U. S. A.* 1999, *96*, 12512-12517.

(226) Ferreiro, D. U.; Hegler, J. A.; Komives, E. A.; Wolynes, P. G. Localizing Frustration in Native Proteins and Protein Assemblies. *Proc. Natl. Acad. Sci. U. S. A.* 2007, *104*, 19819-19824.

(227) Hills, R. D.; Kathuria, S. V.; Wallace, L. A.; Day, I. J.; Brooks, C. L.; Matthews, C. R. Topological Frustration in Beta Alpha-Repeat Proteins: Sequence Diversity Modulates the Conserved Folding Mechanisms of Alpha/Beta/Alpha Sandwich Proteins. *J. Mol. Biol.* 2010, *398*, 332-350.

(228) Wolynes, P. G.; Eaton, W. A.; Fersht, A. R. Chemical Physics of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* 2012, *109*, 17770-17771.

(229) Jacob, M.; Geeves, M.; Holtermann, G.; Schmid, F. X. Diffusional Barrier Crossing in a Two-State Protein Folding Reaction. *Nat. Struct. Biol.* 1999, *6*, 923-926.

(230) Hagen, S. J.; Qiu, L. L.; Pabit, S. A. Diffusional Limits to the Speed of Protein Folding: Fact or Friction? *J. Phys.: Condens. Matter* 2005, *17*, S1503-S1514.

(231) Frauenfelder, H.; Fenimore, P. W.; Chen, G.; McMahon, B. H. Protein Folding Is Slaved to Solvent Motions. *Proc. Natl. Acad. Sci. U. S. A.* 2006, *103*, 15469-15472.

(232) Plaxco, K. W.; Baker, D. Limited Internal Friction in the Rate-Limiting Step of a Two-State Protein Folding Reaction. *Proc. Natl. Acad. Sci. U. S. A.* 1998, *95*, 13591-13596.

(233) Wagner, C.; Kiefhaber, T. Intermediates Can Accelerate Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* 1999, *96*, 6716-6721.

(234) Hinsen, K.; Petrescu, A. J.; Dellerue, S.; Bellissent-Funel, M. C.; Kneller, G. R. Harmonicity in Slow Protein Dynamics. *Chem. Phys.* 2000, *261*, 25-37.

(235) Shea, J. E.; Onuchic, J. N.; Brooks, C. L. Energetic Frustration and the Nature of the Transition State in Protein Folding. *J. Chem. Phys.* 2000, *113*, 7663-7671.

(236) Portman, J. J.; Takada, S.; Wolynes, P. G. Microscopic Theory of Protein Folding Rates. II. Local Reaction Coordinates and Chain Dynamics. *J. Chem. Phys.* 2001, *114*, 5082-5096.

(237) Kaya, H.; Chan, H. S. Solvation Effects and Driving Forces for Protein Thermodynamic and Kinetic Cooperativity: How Adequate Is Native-Centric Topological Modeling? *J. Mol. Biol.* 2003, *326*, 911-931.

(238) Zagrovic, B.; Pande, V. Solvent Viscosity Dependence of the Folding Rate of a Small Protein: Distributed Computing Study. *J. Comput. Chem.* 2003, *24*, 1432-1436.

(239) Ellison, P. A.; Cavagnero, S. Role of Unfolded State Heterogeneity and En-Route Ruggedness in Protein Folding Kinetics. *Protein Sci.* 2006, *15*, 564-582.

(240) Chahine, J.; Oliveira, R. J.; Leite, V. B. P.; Wang, J. Configuration-Dependent Diffusion Can Shift the Kinetic Transition State and Barrier Height of Protein Folding. *Proc. Natl. Acad. Sci. U. S. A.* 2007, *104*, 14646-14651.

(241) Alexander-Katz, A.; Wada, H.; Netz, R. R. Internal Friction and Nonequilibrium Unfolding of Polymeric Globules. *Phys. Rev. Lett.* 2009, *103*, 028102.

(242) Best, R. B. How Well Does a Funneled Energy Landscape Capture the Folding Mechanism of Spectrin Domains? *J. Phys. Chem. B* 2013.

(243) Schulz, J. C.; Schmidt, L.; Best, R. B.; Dzubiella, J.; Netz, R. R. Peptide Chain Dynamics in Light and Heavy Water: Zooming in on Internal Friction. *J. Am. Chem. Soc.* 2012, *134*, 6273-6279.

(244) Jas, G. S.; Eaton, W. A.; Hofrichter, J. Effect of Viscosity on the Kinetics of Alpha-Helix and Beta-Hairpin Formation. *J. Phys. Chem. B* 2001, *105*, 261-272.

(245) Liu, F.; Nakaema, M.; Gruebele, M. The Transition State Transit Time of WW Domain Folding Is Controlled by Energy Landscape Roughness. *J. Chem. Phys.* 2009, *131*.

(246) Hagen, S. J. Solvent Viscosity and Friction in Protein Folding Dynamics. *Curr. Protein Pept. Sci.* 2010, *11*, 385-395.

(247) Wensley, B. G.; Batey, S.; Bone, F. A.; Chan, Z. M.; Tumelty, N. R.; Steward, A.; Kwa, L. G.; Borgia, A.; Clarke, J. Experimental Evidence for a Frustrated Energy Landscape in a Three-Helix-Bundle Protein Family. *Nature* 2010, *463*, 685-688.

(248) Wensley, B. G.; Kwa, L. G.; Shammas, S. L.; Rogers, J. M.; Browning, S.; Yang, Z. Q.; Clarke, J. Separating the Effects of Internal Friction and Transition State Energy to Explain the Slow, Frustrated Folding of Spectrin Domains. *Proc. Natl. Acad. Sci. U. S. A.* 2012, *109*, 17795-17799.

(249) Sekhar, A.; Vallurupalli, P.; Kay, L. E. Defining a Length Scale for Millisecond-Timescale Protein Conformational Exchange. *Proc. Natl. Acad. Sci. U. S. A.* 2013, *110*, 11391-11396.

(250) Pabit, S. A.; Roder, H.; Hagen, S. J. Internal Friction Controls the Speed of Protein Folding from a Compact Configuration. *Biochemistry* 2004, *43*, 12532-12538.

(251) Buscaglia, M.; Lapidus, L. J.; Eaton, W. A.; Hofrichter, J. Effects of Denaturants on the Dynamics of Loop Formation in Polypeptides. *Biophys. J.* 2006, *91*, 276-288.

(252) Waldauer, S. A.; Bakajin, O.; Lapidus, L. J. Extremely Slow Intramolecular Diffusion in Unfolded Protein L. *Proc. Natl. Acad. Sci. U. S. A.* 2010, *107*, 13713-13717.

(253) Kawakami, M.; Byrne, K.; Brockwell, D. J.; Radford, S. E.; Smith, D. A. Viscoelastic Study of the Mechanical Unfolding of a Protein by AFM. *Biophys. J.* 2006, *91*, L16-L18.

(254) Khatri, B. S.; Byrne, K.; Kawakami, M.; Brockwell, D. J.; Smith, D. A.; Radford, S. E.; McLeish, T. C. B. Internal Friction of Single Polypeptide Chains at High Stretch. *Faraday Discuss.* 2008, *139*, 35-51.

(255) Hinczewski, M.; von Hansen, Y.; Netz, R. R. Deconvolution of Dynamic Mechanical Networks. *Proc. Natl. Acad. Sci. U.S.A.* 2010, *107*, 21493-21498.

(256) Lannon, H.; Haghpanah, J. S.; Montclare, J. K.; Vanden-Eijnden, E.; Brujic, J. Force-Clamp Experiments Reveal the Free-Energy Profile and Diffusion Coefficient of the Collapse of Protein Molecules. *Phys. Rev. Lett.* 2013, *110*.

(257) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-Residue Protein. *Nat. Struct. Biol.* 2002, *9*, 425-430.

(258) Barua, B.; Lin, J. C.; Williams, V. D.; Kummler, P.; Neidigh, J. W.; Andersen, N. H. The Trp-Cage: Optimizing the Stability of a Globular Miniprotein. *Protein Eng. Des. Sel.* 2008, *21*, 171-185.

(259) Chen, E. F.; 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. *J. Am. Chem. Soc.* 2003, *125*, 12443-12449.

(260) Paoli, B.; Seeber, M.; Backus, E. H. G.; Ihalainen, J. A.; Hamm, P.; Caflisch, A. Bulky Sidechains and Non-Native Salt Bridges Slow Down the Folding of a Cross-Linked Helical Peptide: A Combined Molecular Dynamics and Time-Resolved Infrared Spectroscopy Study. *J. Phys. Chem. B* 2009, *113*, 4435-4442.

(261) Paoli, B.; Pellarin, R.; Caflisch, A. Slow Folding of Cross-Linked Alpha-Helical Peptides Due to Steric Hindrance. *J. Phys. Chem. B* 2010, *114*, 2023-2027.

(262) Halabis, A.; Zmudzinska, W.; Liwo, A.; Oldziej, S. Conformational Dynamics of the Trp-Cage Miniprotein at Its Folding Temperature. *J. Phys. Chem. B* 2012, *116*, 6898-6907.

(263) Meuzelaar, H.; Marino, K. A.; Huerta-Viga, A.; Panman, M. R.; Smeenk, L. E. J.; Kettelarij, A. J.; van Maarseveen, J. H.; Timmerman, P.; Bolhuis, P. G.; Woutersen, S. Folding Dynamics of the Trp-Cage Miniprotein: Evidence for a Native-Like Intermediate from Combined Time-Resolved Vibrational Spectroscopy and Molecular Dynamics Simulations. *J. Phys. Chem. B* 2013.

(264) Snow, C. D.; Zagrovic, B.; Pande, V. S. The Trp Cage: Folding Kinetics and Unfolded State Topology Via Molecular Dynamics Simulations. *J. Am. Chem. Soc.* 2002, *124*, 14548-14549.

(265) Chowdhury, S.; Lee, M. C.; Xiong, G. M.; Duan, Y. Ab Initio Folding Simulation of the Trp-Cage Mini-Protein Approaches NMR Resolution. *J. Mol. Biol.* 2003, *327*, 711-717.

(266) Nikiforovich, G. V.; Andersen, N. H.; Fesinmeyer, R. M.; Frieden, C. Possible Locally Driven Folding Pathways of Tc5b, a 20-Residue Protein. *Proteins: Struct., Funct., Genet.* 2003, *52*, 292-302.

(267) Ding, F.; Buldyrev, S. V.; Dokholyan, N. V. Folding Trp-Cage to NMR Resolution Native Structure Using a Coarse-Grained Protein Model. *Biophys. J.* 2005, *88*, 147-155.

(268) Linhananta, A.; Boer, J.; MacKay, I. The Equilibrium Properties and Folding Kinetics of an All-Atom Go Model of the Trp-Cage. *J. Chem. Phys.* 2005, *122*.

(269) Chen, J. H.; Im, W. P.; Brooks, C. L. Balancing Solvation and Intramolecular Interactions: Toward a Consistent Generalized Born Force Field. *J. Am. Chem. Soc.* 2006, *128*, 3728-3736.

(270) Kentsis, A.; Gindin, T.; Mezei, M.; Osman, R. Calculation of the Free Energy and Cooperativity of Protein Folding. *PLoS One* 2007, *2*.

(271) Paschek, D.; Nymeyer, H.; Garcia, A. E. Replica Exchange Simulation of Reversible Folding/Unfolding of the Trp-Cage Miniprotein in Explicit Solvent: On the Structure and Possible Role of Internal Water. *J. Struct. Biol.* 2007, *157*, 524-533.

(272) Piana, S.; Laio, A. A Bias-Exchange Approach to Protein Folding. *J. Phys. Chem. B* 2007, *111*, 4553-4559.

(273) Yang, L. J.; Grubb, M. P.; Gao, Y. Q. Application of the Accelerated Molecular Dynamics Simulations to the Folding of a Small Protein. *J. Chem. Phys.* 2007, *126*.

(274) Juraszek, J.; Bolhuis, P. G. Rate Constant and Reaction Coordinate of Trp-Cage Folding in Explicit Water. *Biophys. J.* 2008, *95*, 4246-4257.

(275) Paschek, D.; Hempel, S.; Garcia, A. E. Computing the Stability Diagram Trp-Cage Miniprotein of The. *Proc. Natl. Acad. Sci. U. S. A.* 2008, *105*, 17754-17759.

(276) Xu, W. X.; Mu, Y. G. Ab Initio Folding Simulation of Trpcage by Replica Exchange with Hybrid Hamiltonian. *Biophys. Chem.* 2008, *137*, 116-125.

(277) Cerny, J.; Vondrasek, J.; Hobza, P. Loss of Dispersion Energy Changes the Stability and Folding/Unfolding Equilibrium of the Trp-Cage Protein. *J. Phys. Chem. B* 2009, *113*, 5657-5660.

(278) Kannan, S.; Zacharias, M. Folding Simulations of Trp-Cage Mini Protein in Explicit Solvent Using Biasing Potential Replica-Exchange Molecular Dynamics Simulations. *Proteins* 2009, *76*, 448-460.

(279) Marinelli, F.; Pietrucci, F.; Laio, A.; Piana, S. A Kinetic Model of Trp-Cage Folding from Multiple Biased Molecular Dynamics Simulations. *PLoS Comput. Biol.* 2009, *5*.

(280) Best, R. B.; Mittal, J. Balance between Alpha and Beta Structures in Ab Initio Protein Folding. *J. Phys. Chem. B* 2010, *114*, 8790-8798.

(281) Shao, Q.; Shi, J. Y.; Zhu, W. L. Enhanced Sampling Molecular Dynamics Simulation Captures Experimentally Suggested Intermediate and Unfolded States in the Folding Pathway of Trp-Cage Miniprotein. *J. Chem. Phys.* 2012, *137*.

(282) Juraszek, J.; Bolhuis, P. G. Effects of a Mutation on the Folding Mechanism of a β-Hairpin. *J. Phys. Chem. B* 2009, *113*, 16184-16196.

(283) Lee, I. H.; Kim, S. Y. Dynamic Folding Pathway Models of the Trp-Cage Protein. *BioMed Res. Int.* 2013.

(284) Marinelli, F. Following Easy Slope Paths on a Free Energy Landscape: The Case Study of the Trp-Cage Folding Mechanism. *Biophys. J.* 2013, *105*, 1236-1247.

(285) Huang, C.-Y.; Getahun, Z.; Zhu, Y.; Klemke, J. W.; DeGrado, W. F.; Gai, F. Helix Formation Via Conformation Diffusion Search. *Proc. Natl. Acad. Sci. U. S. A.* 2002, *99*, 2788-2793.

(286) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.* 1999, *151*, 283-312.

(287) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* 1983, *79*, 926-935.

(288) Williams, D. V.; Byrne, A.; Stewart, J.; Andersen, N. H. Optimal Salt Bridge for Trp-Cage Stabilization. *Biochemistry* 2011, *50*, 1143-1152.

(289) Patchornik, A.; Degani, Y.; Neumann, H. Selective Cyanylation of Sulfhydryl Groups. *J. Am. Chem. Soc.* 1970, *92*, 6969-6971.

(290) Streicher, W. W.; Makhatadze, G. I. Unfolding Thermodynamics of Trp-Cage, a 20 Residue Miniprotein, Studied by Differential Scanning Calorimetry and Circular Dichroism Spectroscopy. *Biochemistry* 2007, *46*, 2876-2880.

(291) Ahmed, Z.; Beta, I. A.; Mikhonin, A. V.; Asher, S. A. UV−Resonance Raman Thermal Unfolding Study of Trp-Cage Shows That It Is Not a Simple Two-State Miniprotein. *J. Am. Chem. Soc.* 2005, *127*, 10943-10950.

(292) Hu, Z.; Tang, Y.; Wang, H.; Zhang, X.; Lei, M. Dynamics and Cooperativity of Trp-Cage Folding. *Arch. Biochem. Biophys.* 2008, *475*, 140-147.

(293) Day, R.; Paschek, D.; Garcia, A. E. Microsecond Simulations of the Folding/Unfolding Thermodynamics of the Trp-Cage Miniprotein. *Proteins* 2010, *78*, 1889-1899.

(294) Mukherjee, S.; Waegele, M. M.; Chowdhury, P.; Guo, L.; Gai, F. Effect of Macromolecular Crowding on Protein Folding Dynamics at the Secondary Structure Level. *J. Mol. Biol.* 2009, *393*, 227-236.

(295) Culik, R. M.; Annavarapu, S.; Nanda, V.; Gai, F. Using D-Amino Acids to Delineate the Mechanism of Protein Folding: Application to Trp-Cage. *Chem. Phys.* 2013, *422*, 131-134.

(296) Zwanzig, R. Diffusion in a Rough Potential. *Proc. Natl. Acad. Sci. U. S. A.* 1988, *85*, 2029-2030.

(297) Hyeon, C. B.; Thirumalai, D. Can Energy Landscape Roughness of Proteins and Rna Be Measured by Using Mechanical Unfolding Experiments? *Proc. Natl. Acad. Sci. U. S. A.* 2003, *100*, 10249-10253.

(298) Best, R. B.; Hummer, G. Diffusion Models of Protein Folding. *Phys. Chem. Chem. Phys.* 2011, *13*, 16902-16911.

(299) Erbaş, A.; Netz, Roland R. Confinement-Dependent Friction in Peptide Bundles. *Biophys. J.* 2013, *104*, 1285-1295.

(300) Juraszek, J.; Bolhuis, P. G. Sampling the Multiple Folding Mechanisms of Trp-Cage in Explicit Solvent. *Proc. Natl. Acad. Sci. U. S. A.* 2006, *103*, 15859-15864.

(301) Markiewicz, B. N.; Oyola, R.; Du, D.; Gai, F. Aggregation Gatekeeper and Controlled Assembly of Trpzip Beta-Hairpins. *Biochemistry* 2014, *53*, 1146-1154.

(302) Kelly, J. W. Alternative Conformations of Amyloidogenic Proteins Govern Their Behavior. *Curr. Opin. Struct. Biol.* 1996, *6*, 11-17.

(303) Dobson, C. M. Protein Folding and Misfolding. *Nature* 2003, *426*, 884-890.

(304) Selkoe, D. J. Folding Proteins in Fatal Ways. *Nature* 2003, *426*, 900-904.

(305) RamirezAlvarado, M.; Blanco, F. J.; Serrano, L. De Novo Design and Structural Analysis of a Model Beta-Hairpin Peptide System. *Nat. Struct. Biol.* 1996, *3*, 604-612.

(306) Richardson, J. S.; Richardson, D. C. Natural Beta-Sheet Proteins Use Negative Design to Avoid Edge-to-Edge Aggregation. *Proc. Natl. Acad. Sci. U. S. A.* 2002, *99*, 2754-2759.

(307) Wang, W.; Hecht, M. H. Rationally Designed Mutations Convert De Novo Amyloid-Like Fibrils into Monomeric Beta-Sheet Proteins. *Proc. Natl. Acad. Sci. U. S. A.* 2002, *99*, 2760-2765.

(308) Hecht, M. H.; Das, A.; Go, A.; Bradley, L. H.; Wei, Y. N. De Novo Proteins from Designed Combinatorial Libraries. *Protein Sci.* 2004, *13*, 1711-1723.

(309) Ventura, S.; Zurdo, J.; Narayanan, S.; Parreño, M.; Mangues, R.; Reif, B.; Chiti, F.; Giannoni, E.; Dobson, C. M.; Aviles, F. X.; Serrano, L. Short Amino Acid Stretches Can Mediate Amyloid Formation in Globular Proteins: The Src Homology 3 (SH3) Case. *Proc. Natl. Acad. Sci. U. S. A.* 2004, *101*, 7258-7263.

(310) Pokala, N.; Handel, T. M. Energy Functions for Protein Design: Adjustment with Protein–Protein Complex Affinities, Models for the Unfolded State, and Negative Design of Solubility and Specificity. *J. Mol. Biol.* 2005, *347*, 203-227.

(311) Monsellier, E.; Chiti, F. Prevention of Amyloid-Like Aggregation as a Driving Force of Protein Evolution. *EMBO Rep.* 2007, *8*, 737-742.

(312) Marqusee, S.; Baldwin, R. L. Helix Stabilization by Glu-...Lys+ Salt Bridges in Short Peptides of De Novo Design. *Proc. Natl. Acad. Sci. U. S. A.* 1987, *84*, 8898-8902.

(313) Marqusee, S.; Robbins, V. H.; Baldwin, R. L. Unusually Stable Helix Formation in Short Alanine-Based Peptides. *Proc. Natl. Acad. Sci. U. S. A.* 1989, *86*, 5286-5290.

(314) Serag, A. A.; Altenbach, C.; Gingery, M.; Hubbell, W. L.; Yeates, T. O. Arrangement of Subunits and Ordering of Beta-Strands in an Amyloid Sheet. *Nat. Struct. Biol.* 2002, *9*, 734-739.

(315) Thirumalai, D.; Klimov, D. K.; Dima, R. I. Emerging Ideas on the Molecular Basis of Protein and Peptide Aggregation. *Curr. Opin. Struct. Biol.* 2003, *13*, 146-159.

(316) Pawar, A. P.; Dubay, K. F.; Zurdo, J.; Chiti, F.; Vendruscolo, M.; Dobson, C. M. Prediction of "Aggregation-Prone" and "Aggregation-Susceptible" Regions in Proteins Associated with Neurodegenerative Diseases. *J. Mol. Biol.* 2005, *350*, 379-392.

(317) Munoz, V.; Ghirlando, R.; Blanco, F. J.; Jas, G. S.; Hofrichter, J.; Eaton, W. A. Folding and Aggregation Kinetics of a Beta-Hairpin. *Biochemistry* 2006, *45*, 7023-7035.

(318) Nagarkar, R. P.; Hule, R. A.; Pochan, D. J.; Schneider, J. P. De Novo Design of Strand-Swapped β-Hairpin Hydrogels. *J. Am. Chem. Soc.* 2008, *130*, 4466-4474.

(319) Smith, M. H.; Miles, T. F.; Sheehan, M.; Alfieri, K. N.; Kokona, B.; Fairman, R. Polyglutamine Fibrils Are Formed Using a Simple Designed Beta-Hairpin Model. *Proteins* 2010, *78*, 1971-1979.

(320) Williams, A. D.; Portelius, E.; Kheterpal, I.; Guo, J.-t.; Cook, K. D.; Xu, Y.; Wetzel, R. Mapping Aβ Amyloid Fibril Secondary Structure Using Scanning Proline Mutagenesis. *J. Mol. Biol.* 2004, *335*, 833-842.

(321) Senguen, F. T.; Doran, T. M.; Anderson, E. A.; Nilsson, B. L. Clarifying the Influence of Core Amino Acid Hydrophobicity, Secondary Structure Propensity, and Molecular Volume on Amyloid-Beta 16-22 Self-Assembly. *Mol. BioSyst.* 2011, *7*, 497-510.

(322) Lakshmanan, A.; Cheong, D. W.; Accardo, A.; Di Fabrizio, E.; Riekel, C.; Hauser, C. A. E. Aliphatic Peptides Show Similar Self-Assembly to Amyloid Core Sequences, Challenging the Importance of Aromatic Interactions in Amyloidosis. *Proc. Natl. Acad. Sci. U. S. A.* 2013, *110*, 519-524.

(323) 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. *J. Phys. Chem. B* 2009, *113*, 5661-5674.

(324) Popp, A.; Wu, L.; Keiderling, T. A.; Hauser, K. Impact of Beta-Turn Sequence on Beta-Hairpin Dynamics Studied with Infrared-Detected Temperature Jump. *Spectrosc. Int. J.* 2012, *27*, 557-564.

(325) Jones, K. C.; Peng, C. S.; Tokmakoff, A. Folding of a Heterogeneous β-Hairpin Peptide from Temperature-Jump 2D IR Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 2013, *110*, 2828-2833.

(326) Takekiyo, T.; Wu, L.; Yoshimura, Y.; Shimizu, A.; Keiderling, T. A. Relationship between Hydrophobic Interactions and Secondary Structure Stability for Trpzip β-Hairpin Peptides†. *Biochemistry* 2009, *48*, 1543-1552.

(327) Smith, A. W.; Chung, H. S.; Ganim, Z.; Tokmakoff, A. Residual Native Structure in a Thermally Denatured Beta-Hairpin. *J. Phys. Chem. B* 2005, *109*, 17025-17027.

(328) Krejtschi, C.; Huang, R.; Keiderling, T. A.; Hauser, K. Time-Resolved Temperature-Jump Infrared Spectroscopy of Peptides with Well-Defined Secondary Structure: A Trpzip Beta-Hairpin Variant as an Example. *Vib. Spectrosc.* 2008, *48*, 1-7.

(329) Nymeyer, H. Energy Landscape of the Trpzip2 Peptide. *J. Phys. Chem. B* 2009, *113*, 8288-8295.

(330) Yang, L. J.; Shao, Q.; Gao, Y. Q. Thermodynamics and Folding Pathways of Trpzip2: An Accelerated Molecular Dynamics Simulation Study. *J. Phys. Chem. B* 2009, *113*, 803-808.

(331) Humphrey, W.; Dalke, A.; Schulten, K. Vmd: Visual Molecular Dynamics. *J. Mol. Graphics* 1996, *14*, 33-38.

(332) Dunbrack, R. L.; Cohen, F. E. Bayesian Statistical Analysis of Protein Side-Chain Rotamer Preferences. *Protein Sci.* 1997, *6*, 1661-1681.

(333) Merkel, J. S.; Sturtevant, J. M.; Regan, L. Sidechain Interactions in Parallel β Sheets: The Energetics of Cross-Strand Pairings. *Structure* 1999, *7*, 1333-1343.

(334) Patchorn.A; Amit, B.; Woodward, R. B. Photosensitive Protecting Groups. *J. Am. Chem. Soc.* 1970, *92*, 6333-&.

(335) Rusiecki, V. K.; Warne, S. A. Synthesis of Nα-Fmoc-Nε-Nvoc-Lysine and Use in the Preparation of Selectively Functionalized Peptides. *Bioorg. Med. Chem. Lett.* 1993, *3*, 707-710.

(336) Measey, T. J.; Gai, F. Light-Triggered Disassembly of Amyloid Fibrils. *Langmuir* 2012, *28*, 12588-12592.

(337) Measey, T. J.; Markiewicz, B. N.; Gai, F. Amide I Band and Photoinduced Disassembly of a Peptide Hydrogel. *Chem. Phys. Lett.* 2013, *580*, 135-140.

(338) Wood, J. S.; Koszelak, M.; Liu, J.; Lawrence, D. S. A Caged Protein Kinase Inhibitor. *J. Am. Chem. Soc.* 1998, *120*, 7145-7146.

(339) Solomek, T.; Mercier, S.; Bally, T.; Bochet, C. G. Photolysis of Ortho-Nitrobenzylic Derivatives: The Importance of the Leaving Group. *Photochem. Photobiol. Sci.* 2012, *11*, 548-555.

(340) Kubelka, J.; Keiderling, T. A. Differentiation of Beta-Sheet-Forming Structures: Ab Initio-Based Simulations of IR Absorption and Vibrational CD for Model Peptide and Protein Beta-Sheets. *J. Am. Chem. Soc.* 2001, *123*, 12048-12058.

(341) Lomakin, A.; Chung, D. S.; Benedek, G. B.; Kirschner, D. A.; Teplow, D. B. On the Nucleation and Growth of Amyloid Beta-Protein Fibrils: Detection of Nuclei and Quantitation of Rate Constants. *Proc. Natl. Acad. Sci. U. S. A.* 1996, *93*, 1125-1129.

(342) Petty, S. A.; Decatur, S. M. Intersheet Rearrangement of Polypeptides During Nucleation of Beta-Sheet Aggregates. *Proc. Natl. Acad. Sci. U. S. A.* 2005, *102*, 14272-14277.

(343) Perálvarez-Marín, A.; Barth, A.; Gräslund, A. Time-Resolved Infrared Spectroscopy of pH-Induced Aggregation of the Alzheimer Aβ1–28 Peptide. *J. Mol. Biol.* 2008, *379*, 589-596.

(344) Jimenez, J. L.; Nettleton, E. J.; Bouchard, M.; Robinson, C. V.; Dobson, C. M.; Saibil, H. R. The Protofilament Structure of Insulin Amyloid Fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 2002, *99*, 9196-9201.

(345) Petkova, A. T.; Ishii, Y.; Balbach, J. J.; Antzutkin, O. N.; Leapman, R. D.; Delaglio, F.; Tycko, R. A Structural Model for Alzheimer's Beta-Amyloid Fibrils Based on Experimental Constraints from Solid State NMR. *Proc. Natl. Acad. Sci. U. S. A.* 2002, *99*, 16742-16747.

(346) Zandomeneghi, G.; Krebs, M. R.; McCammon, M. G.; Fandrich, M. FTIR Reveals Structural Differences between Native Beta-Sheet Proteins and Amyloid Fibrils. *Protein Sci.* 2004, *13*, 3314-3321.

(347) Yan, S.; Gawlak, G.; Makabe, K.; Tereshko, V.; Koide, A.; Koide, S. Hydrophobic Surface Burial Is the Major Stability Determinant of a Flat, Single-Layer β-Sheet. *J. Mol. Biol.* 2007, *368*, 230-243.

(348) Larini, L.; Shea, J.-E. Role of β-Hairpin Formation in Aggregation: The Self-Assembly of the Amyloid-β(25–35) Peptide. *Biophys. J.* 2012, *103*, 576-586.

(349) Qiang, W.; Yau, W.-M.; Luo, Y.; Mattson, M. P.; Tycko, R. Antiparallel β-Sheet Architecture in Iowa-Mutant β-Amyloid Fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 2012, *109*, 4443-4448.

(350) Barrow, C. J.; Yasuda, A.; Kenny, P. T. M.; Zagorski, M. G. Solution Conformations and Aggregational Properties of Synthetic Amyloid Beta-Peptides of Alzheimers-Disease - Analysis of Circular-Dichroism Spectra. *J. Mol. Biol.* 1992, *225*, 1075-1093.

(351) Nasir, S.; Ramirez, P.; Ali, M.; Ahmed, I.; Fruk, L.; Mafe, S.; Ensinger, W. Nernst-Planck Model of Photo-Triggered, pH-Tunable Ionic Transport through Nanopores Functionalized with "Caged" Lysine Chains. *J. Chem. Phys.* 2013, *138*.

(352) Abello, N.; Kerstjens, H. A. M.; Postma, D. S.; Bischoff, R. Selective Acylation of Primary Amines in Peptides and Proteins. *J. Proteome Res.* 2007, *6*, 4770-4776.

(353) Basle, E.; Joubert, N.; Pucheault, M. Protein Chemical Modification on Endogenous Amino Acids. *Chem. Biol.* 2010, *17*, 213-227.

(354) Clark, A. H.; Saunderson, D. H.; Suggett, A. Infrared and Laser-Raman Spectroscopic Studies of Thermally-Induced Globular Protein Gels. *Int. J. Pept. Protein Res.* 1981, *17*, 353-364.

(355) Casal, H. L.; Kohler, U.; Mantsch, H. H. Structural and Conformational Changes of Beta-Lactoglobulin B: An Infrared Spectroscopic Study of the Effect of pH and Temperature. *Biochim. Biophys. Acta* 1988, *957*, 11-20.

(356) Naumann, D.; Schultz, C.; Gorne-Tschelnokow, U.; Hucho, F. Secondary Structure and Temperature Behavior of the Acetylcholine Receptor by Fourier Transform Infrared Spectroscopy. *Biochemistry* 1993, *32*, 3162-3168.

(357) Fabian, H.; Schultz, C.; Naumann, D.; Landt, O.; Hahn, U.; Saenger, W. Secondary Structure and Temperature-Induced Unfolding and Refolding of

Ribonuclease T1 in Aqueous Solution. A Fourier Transform Infrared Spectroscopic Study. *J. Mol. Biol.* 1993, *232*, 967-981.

(358) Rughani, R. V.; Schneider, J. P. Molecular Design of Beta-Hairpin Peptides for Material Construction. *MRS Bull.* 2008, *33*, 530-535.

(359) Peng, C. S.; Jones, K. C.; Tokmakoff, A. Anharmonic Vibrational Modes of Nucleic Acid Bases Revealed by 2D IR Spectroscopy. *J. Am. Chem. Soc.* 2011, *133*, 15650-15660.

(360) Knowles, T. P. J.; Buehler, M. J. Nanomechanics of Functional and Pathological Amyloid Materials. *Nat. Nanotechnol.* 2011, *6*, 469-479.

(361) Dobson, C. M. Protein Misfolding, Evolution and Disease. *Trends Biochem.Sci.* 1999, *24*, 329-332.

(362) Loo, Y.; Zhang, S.; Hauser, C. A. From Short Peptides to Nanofibers to Macromolecular Assemblies in Biomedicine. *Biotechnol. Adv.* 2012, *30*, 593-603.

(363) Dobson, C. M. Experimental Investigation of Protein Folding and Misfolding. *Methods* 2004, *34*, 4-14.

(364) Bowerman, C. J.; Ryan, D. M.; Nissan, D. A.; Nilsson, B. L. The Effect of Increasing Hydrophobicity on the Self-Assembly of Amphipathic Beta-Sheet Peptides. *Mol. BioSyst.* 2009, *5*, 1058-1069.

(365) Noor, H.; Cao, P.; Raleigh, D. P. Morin Hydrate Inhibits Amyloid Formation by Islet Amyloid Polypeptide and Disaggregates Amyloid Fibers. *Protein Sci.* 2012, *21*, 373-382.

(366) Cao, P.; Raleigh, D. P. Analysis of the Inhibition and Remodeling of Islet Amyloid Polypeptide Amyloid Fibers by Flavanols. *Biochemistry* 2012, *51*, 2670-2683.

(367) Porat, Y.; Abramowitz, A.; Gazit, E. Inhibition of Amyloid Fibril Formation by Polyphenols: Structural Similarity and Aromatic Interactions as a Common Inhibition Mechanism. *Chem. Biol. Drug Des.* 2006, *67*, 27-37.

(368) Cheng, P. N.; Liu, C.; Zhao, M.; Eisenberg, D.; Nowick, J. S. Amyloid β-Sheet Mimics That Antagonize Protein Aggregation and Reduce Amyloid Toxicity. *Nat. Chem.* 2012, *4*, 927-933.

(369) Tjernberg, L. O.; Naslund, J.; Lindqvist, F.; Johansson, J.; Karlstrom, A. R.; Thyberg, J.; Terenius, L.; Nordstedt, C. Arrest of Beta-Amyloid Fibril Formation by a Pentapeptide Ligand. *J. Biol. Chem.* 1996, *271*, 8545-8548.

(370) Kapurniotu, A.; Schmauder, A.; Tenidis, K. Structure-Based Design and Study of Non-Amyloidogenic, Double N-Methylated IAPP Amyloid Core Sequences as Inhibitors of IAPP Amyloid Formation and Cytotoxicity. *J. Mol. Biol.* 2002, *315*, 339-350.

(371) Deiters, A. Principles and Applications of the Photochemical Control of Cellular Processes. *ChemBioChem* 2010, *11*, 47-53.

(372) Lemke, E. A. Precision Control of Cellular Pathways with Light. *ChemBioChem* 2010, *11*, 1825-1827.

(373) Tang, X.; Dmochowski, I. J. Regulating Gene Expression with Light-Activated Oligonucleotides. *Mol. BioSyst.* 2007, *3*, 100-110.

(374) Dmochowski, I. J.; Tang, X. Taking Control of Gene Expression with Light-Activated Oligonucleotides. *Biotechniques* 2007, *43*, 161, 163, 165 passim.

(375) Kramer, R. H.; Fortin, D. L.; Trauner, D. New Photochemical Tools for Controlling Neuronal Activity. *Curr. Opin. Neurobiol.* 2009, *19*, 544-552.

(376) Rock, R. S.; Hansen, K. C.; Larsen, R. W.; Chan, S. I. Rapid Photochemical Triggering of Protein Unfolding in a Nondenaturing Environment. *Chem. Phys.* 2004, *307*, 201-208.

(377) Damjanovic, A.; Garcia-Moreno, B.; Lattman, E. E.; Garcia, A. E. Molecular Dynamics Study of Water Penetration in Staphylococcal Nuclease. *Proteins* 2005, *60*, 433-449.

(378) Takayama, Y.; Castaneda, C. A.; Chimenti, M.; Garcia-Moreno, B.; Iwahara, J. Direct Evidence for Deprotonation of a Lysine Sidechain Buried in the Hydrophobic Core of a Protein. *J. Am. Chem. Soc.* 2008, *130*, 6714-6715.

(379) Tatsu, Y.; Shigeri, Y.; Ishida, A.; Isamu, K.; Fujisawa, H.; Yumoto, N. Synthesis of Caged Peptides Using Caged Lysine: Application to the Synthesis of Caged AIP, a Highly Specific Inhibitor of Calmodulin-Dependent Protein Kinase II. *Bioorg. Med. Chem. Lett.* 1999, *9*, 1093-1096.

(380) Watai, Y.; Sase, I.; Shiono, H.; Nakano, Y. Regulation of Nuclear Import by Light-Induced Activation of Caged Nuclear Localization Signal in Living Cells. *FEBS Lett.* 2001, *488*, 39-44.

(381) Chen, P. R.; Groff, D.; Guo, J.; Ou, W.; Cellitti, S.; Geierstanger, B. H.; Schultz, P. G. A Facile System for Encoding Unnatural Amino Acids in Mammalian Cells. *Angew. Chem., Int. Ed.* 2009, *48*, 4052-4055.

(382) Baumann, L.; Beck-Sickinger, A. G. Identification of a Potential Modification Site in Human Stromal Cell-Derived Factor-1. *Biopolymers* 2010, *94*, 771-778.

(383) Tenidis, K.; Waldner, M.; Bernhagen, J.; Fischle, W.; Bergmann, M.; Weber, M.; Merkle, M. L.; Voelter, W.; Brunner, H.; Kapurniotu, A. Identification of a Penta- and Hexapeptide of Islet Amyloid Polypeptide (IAPP) with Amyloidogenic and Cytotoxic Properties. *J. Mol. Biol.* 2000, *295*, 1055-1071.

(384) Mazor, Y.; Gilead, S.; Benhar, I.; Gazit, E. Identification and Characterization of a Novel Molecular-Recognition and Self-Assembly Domain within the Islet Amyloid Polypeptide. *J. Mol. Biol.* 2002, *322*, 1013-1024.

(385) Balbach, J. J.; Ishii, Y.; Antzutkin, O. N.; Leapman, R. D.; Rizzo, N. W.; Dyda, F.; Reed, J.; Tycko, R. Amyloid Fibril Formation by a Beta 16-22, a Seven-Residue Fragment of the Alzheimer's Beta-Amyloid Peptide, and Structural Characterization by Solid State NMR. *Biochemistry* 2000, *39*, 13748-13759.

(386) Chen, P. Self-Assembly of Ionic-Complementary Peptides: A Physicochemical Viewpoint. *Colloids Surf., A* 2005, *261*, 3-24.

(387) Zhang, S. Emerging Biological Materials through Molecular Self-Assembly. *Biotechnol. Adv.* 2002, *20*, 321-339.

(388) Mukherjee, S.; Chowdhury, P.; Gai, F. Effect of Dehydration on the Aggregation Kinetics of Two Amyloid Peptides. *J. Phys. Chem. B* 2009, *113*, 531-535.

(389) Kuipers, B. J. H.; Gruppen, H. Prediction of Molar Extinction Coefficients of Proteins and Peptides Using UV Absorption of the Constituent Amino Acids at 214 nm to Enable Quantitative Reverse Phase High-Performance Liquid Chromatography−Mass Spectrometry Analysis. *J. Agric. Food Chem.* 2007, *55*, 5445-5451.

(390) Carulla, N.; Caddy, G. L.; Hall, D. R.; Zurdo, J.; Gairi, M.; Feliz, M.; Giralt, E.; Robinson, C. V.; Dobson, C. M. Molecular Recycling within Amyloid Fibrils. *Nature* 2005, *436*, 554-558.

(391) Middleton, C. T.; Marek, P.; Cao, P.; Chiu, C. C.; Singh, S.; Woys, A. M.; de Pablo, J. J.; Raleigh, D. P.; Zanni, M. T. Two-Dimensional Infrared Spectroscopy Reveals the Complex Behaviour of an Amyloid Fibril Inhibitor. *Nat. Chem.* 2012, *4*, 355-360.

(392) Nguyen, P. H.; Li, M. S.; Stock, G.; Straub, J. E.; Thirumalai, D. Monomer Adds to Preformed Structured Oligomers of Abeta-Peptides by a Two-Stage Dock-Lock Mechanism. *Proc. Natl. Acad. Sci. U. S. A.* 2007, *104*, 111-116.

(393) O'Brien, E. P.; Okamoto, Y.; Straub, J. E.; Brooks, B. R.; Thirumalai, D. Thermodynamic Perspective on the Dock−Lock Growth Mechanism of Amyloid Fibrils. *J. Phys. Chem. B* 2009, *113*, 14421-14430.

(394) Reddy, G.; Straub, J. E.; Thirumalai, D. Dynamics of Locking of Peptides onto Growing Amyloid Fibrils. *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 11948-11953.

(395) Aggeli, A.; Bell, M.; Carrick, L. M.; Fishwick, C. W. G.; Harding, R.; Mawer, P. J.; Radford, S. E.; Strong, A. E.; Boden, N. pH as a Trigger of Peptide Beta-Sheet Self-Assembly and Reversible Switching between Nematic and Isotropic Phases. *J. Am. Chem. Soc.* 2003, *125*, 9619-9628.

(396) Collier, J. H.; Hu, B. H.; Ruberti, J. W.; Zhang, J.; Shum, P.; Thompson, D. H.; Messersmith, P. B. Thermally and Photochemically Triggered Self-Assembly of Peptide Hydrogels. *J. Am. Chem. Soc.* 2001, *123*, 9463-9464.

(397) Ozbas, B.; Kretsinger, J.; Rajagopal, K.; Schneider, J. P.; Pochan, D. J. Salt-Triggered Peptide Folding and Consequent Self-Assembly into Hydrogels with Tunable Modulus. *Macromolecules* 2004, *37*, 7331-7337.

(398) Raeburn, J.; McDonald, T. O.; Adams, D. J. Dipeptide Hydrogelation Triggered Via Ultraviolet Light. *Chem. Commun.* 2012, *48*, 9355-9357.

(399) Caplan, M. R.; Schwartzfarb, E. M.; Zhang, S.; Kamm, R. D.; Lauffenburger, D. A. Control of Self-Assembling Oligopeptide Matrix Formation through Systematic Variation of Amino Acid Sequence. *Biomaterials* 2002, *23*, 219-227.

(400) Marini, D. M.; Hwang, W.; Lauffenburger, D. A.; Zhang, S.; Kamm, R. D. Left-Handed Helical Ribbon Intermediates in the Self-Assembly of a β-Sheet Peptide. *Nano Lett.* 2002, *2*, 295-299.

(401) MacKintosh, F. C.; Käs, J.; Janmey, P. A. Elasticity of Semiflexible Biopolymer Networks. *Phys. Rev. Lett.* 1995, *75*, 4425-4428.

(402) Tomatsu, I.; Peng, K.; Kros, A. Photoresponsive Hydrogels for Biomedical Applications. *Adv. Drug Deliv. Rev.* 2011, *63*, 1257-1266.

(403) Zhang, W.; Markiewicz, B. N.; Doerksen, R. S.; Smith III, A. B.; Gai, F. C≡N Stretching Vibration of 5-Cyanotryptophan as an Infrared Probe of Protein Local Environment: What Determines Its Frequency? *Phys. Chem. Chem. Phys.* 2015, *18*, 7027-7034.

(404) Edelstein, L.; Stetz, M. A.; McMahon, H. A.; Londergan, C. H. The Effects of Alpha-Helical Structure and Cyanylated Cysteine on Each Other. *J. Phys. Chem. B* 2010, *114*, 4931-4936.

(405) de la Arada, I.; Andraka, N.; Pacios, M. G.; Arrondo, J. L. R. A Conventional and 2dcos Infrared Approach to the Kinetics of Protein Misfolding. *Curr. Protein Pept. Sci.* 2011, *12*, 181-187.

(406) Waegele, M. M.; Gai, F. Infrared Study of the Folding Mechanism of a Helical Hairpin: Porcine Pyy. *Biochemistry* 2010, *49*, 7659-7664.

(407) Bogan, A. A.; Thorn, K. S. Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* 1998, *280*, 1-9.

(408) Spier, A. D.; Lummis, S. C. R. The Role of Tryptophan Residues in the 5-Hydroxytryptamine3 Receptor Ligand Binding Domain. *J. Biol. Chem.* 2000, *275*, 5620-5625.

(409) Shimazaki, Y.; Yajima, T.; Takani, M.; Yamauchi, O. Metal Complexes Involving Indole Rings: Structures and Effects of Metal-Indole Interactions. *Coord. Chem. Rev.* 2009, *253*, 479-492.

(410) Saikumar, P.; Murali, R.; Reddy, E. P. Role of Tryptophan Repeats and Flanking Amino Acids in Myb-DNA Interactions. *Proc. Natl. Acad. Sci. U. S. A.* 1990, *87*, 8452-8456.

(411) Hsu, C. H.; Chen, C.; Jou, M. L.; Lee, A. Y. L.; Lin, Y. C.; Yu, Y. P.; Huang, W. T.; Wu, S. H. Structural and DNA-Binding Studies on the Bovine Antimicrobial

Peptide, Indolicidin: Evidence for Multiple Conformations Involved in Binding to Membranes and DNA. *Nucleic Acids Res.* 2005, *33*, 4053-4064.

(412) Samanta, U.; Pal, D.; Chakrabarti, P. Environment of Tryptophan Sidechains in Proteins. *Proteins: Struct., Funct., Bioinf.* 2000, *38*, 288-300.

(413) Bartlett, G. J.; Porter, C. T.; Borkakoti, N.; Thornton, J. M. Analysis of Catalytic Residues in Enzyme Active Sites. *J. Mol. Biol.* 2002, *324*, 105-121.

(414) Hong, M.; DeGrado, W. F. Structural Basis for Proton Conduction and Inhibition by the Influenza M2 Protein. *Protein Sci.* 2012, *21*, 1620-1633.

(415) de Jesus, A. J.; Allen, T. W. The Role of Tryptophan Sidechains in Membrane Protein Anchoring and Hydrophobic Mismatch. *Biochim. Biophys. Acta, Biomembr.* 2013, *1828*, 864-876.

(416) Fabian, H.; Schultz, C.; Backmann, J.; Hahn, U.; Saenger, W.; Mantsch, H. H.; Naumann, D. Impact of Point Mutations on the Structure and Thermal Stability of Ribonuclease T1 in Aqueous Solution Probed by Fourier Transform Infrared Spectroscopy. *Biochemistry* 1994, *33*, 10725-10730.

(417) Asplund, M. C.; Zanni, M. T.; Hochstrasser, R. M. Two-Dimensional Infrared Spectroscopy of Peptides by Phase-Controlled Femtosecond Vibrational Photon Echoes. *Proc. Natl. Acad. Sci. U. S. A.* 2000, *97*, 8219-8224.

(418) Allerhand, A.; Schleyer, P. V. Nitriles and Isonitriles as Proton Acceptors in Hydrogen Bonding - Correlation of $\Delta\upsilon$OH with Acceptor Structure. *J. Am. Chem. Soc.* 1963, *85*, 866-870.

(419) Kamlet, M. J.; Abboud, J. L. M.; Abraham, M. H.; Taft, R. W. Linear Solvation Energy Relationships. 23. A Comprehensive Collection of the Solvatochromic Parameters, .Pi.*, .Alpha., And .Beta., and Some Methods for Simplifying the Generalized Solvatochromic Equation. *J. Org. Chem.* 1983, *48*, 2877-2887.

(420) Marcus, Y. The Properties of Organic Liquids That Are Relevant to Their Use as Solvating Solvents. *Chem. Soc. Rev.* 1993, *22*, 409-416.

(421) Oh, K. I.; Choi, J. H.; Lee, J. H.; Han, J. B.; Lee, H.; Cho, M. Nitrile and Thiocyanate IR Probes: Molecular Dynamics Simulation Studies. *J. Chem. Phys.* 2008, *128*, 10.

(422) Getahun, Z.; Huang, C. Y.; Wang, T.; De Leon, B.; DeGrado, W. F.; Gai, F. Using Nitrile-Derivatized Amino Acids as Infrared Probes of Local Environment. *J. Am. Chem. Soc.* 2003, *125*, 405-411.

(423) Johnston, A. J.; Zhang, Y.; Busch, S.; Pardo, L. C.; Imberti, S.; McLain, S. E. Amphipathic Solvation of Indole: Implications for the Role of Tryptophan in Membrane Proteins. *J. Phys. Chem. B* 2015, *119*, 5979-5987.

(424) Aschaffenburg, D. J.; Moog, R. S. Probing Hydrogen Bonding Environments: Solvatochromic Effects on the CN Vibration of Benzonitrile. *J. Phys. Chem. B* 2009, *113*, 12736-12743.

(425) Hamm, P.; Zanni, M. *Concepts and Methods of 2D Infrared Spectroscopy*; Cambridge University Press, 2011.

(426) Ma, J. Q.; Pazos, I. M.; Gai, F. Microscopic Insights into the Protein-Stabilizing Effect of Trimethylamine N-Oxide (TMAO). *Proc. Natl. Acad. Sci. U. S. A.* 2014, *111*, 8476-8481.

(427) Hilaire, M. R.; Abaskharon, R. M.; Gai, F. Biomolecular Crowding Arising from Small Molecules, Molecular Constraints, Surface Packing, and Nano-Confinement. *J. Phys. Chem. Lett.* 2015, *6*, 2546-2553.

(428) Chung, J. K.; Thielges, M. C.; Lynch, S. R.; Fayer, M. D. Fast Dynamics of HP35 for Folded and Urea-Unfolded Conditions. *J. Phys. Chem. B* 2012, *116*, 11024-11031.

(429) Lee, K. K.; Park, K. H.; Choi, J. H.; Ha, J. H.; Jeon, S. J.; Cho, M. Ultrafast Vibrational Spectroscopy of Cyanophenols. *J. Phys. Chem. A* 2010, *114*, 2757-2767.

(430) Banno, M.; Kotani, A.; Ohta, K.; Tominaga, K. Solute-Solvent Interactions of Benzonitrile in Solutions Studied by Sub-Picosecond Infrared Pump-Probe Spectroscopy. *Bull. Chem. Soc. Jpn.* 2014, *87*, 470-478.

(431) Steiner, T.; Koellner, G. Hydrogen Bonds with Π-Acceptors in Proteins: Frequencies and Role in Stabilizing Local 3D Structures1. *J. Mol. Biol.* 2001, *305*, 535-557.

(432) Steiner, T. Hydrogen Bonds from Water Molecules to Aromatic Acceptors in Very High-Resolution Protein Crystal Structures. *Biophys. Chem.* 2002, *95*, 195-201.

(433) Perutz, M. F. The Role of Aromatic Rings as Hydrogen-Bond Acceptors in Molecular Recognition. *Philos. Trans. R. Soc., A* 1993, *345*, 105-112.

(434) Acharya, R.; Carnevale, V.; Fiorin, G.; Levine, B. G.; Polishchuk, A. L.; Balannik, V.; Samish, I.; Lamb, R. A.; Pinto, L. H.; DeGrado, W. F.; Klein, M. L. Structure and Mechanism of Proton Transport through the Transmembrane Tetrameric M2 Protein Bundle of the Influenza a Virus. *Proc. Natl. Acad. Sci. U. S. A.* 2010, *107*, 15075-15080.

(435) Williams, J. K.; Zhang, Y.; Schmidt-Rohr, K.; Hong, M. pH-Dependent Conformation, Dynamics, and Aromatic Interaction of the Gating Tryptophan Residue of the Influenza M2 Proton Channel from Solid-State NMR. *Biophys. J.* 2013, *104*, 1698-1708.

(436) Maruyama, T.; Takeuchi, H. Water Accessibility to the Tryptophan Indole N-H Sites of Gramicidin a Transmembrane Channel: Detection of Positional Shifts of Tryptophans 11 and 13 Along the Channel Axis Upon Cation Binding. *Biochemistry* 1997, *36*, 10993-11001.

(437) Chaudhuri, A.; Haldar, S.; Sun, H.; Koeppe Ii, R. E.; Chattopadhyay, A. Importance of Indole NH Hydrogen Bonding in the Organization and Dynamics of Gramicidin Channels. *Biochim. Biophys. Acta, Biomembr.* 2014, *1838*, 419-428.

(438) Vostrikov, V. V.; Daily, A. E.; Greathouse, D. V.; Koeppe, R. E. Charged or Aromatic Anchor Residue Dependence of Transmembrane Peptide Tilt. *J. Biol. Chem.* 2010, *285*, 31723-31730.

(439) Pogozheva, I. D.; Mosberg, H. I.; Lomize, A. L. Life at the Border: Adaptation of Proteins to Anisotropic Membrane Environment. *Protein Sci.* 2014, *23*, 1165-1196.

(440) Haynes, W. M. *CRC Handbook of Chemistry and Physics*; 96th ed.; CRC Press: Boca Raton, FL, 2015.

(441) Vivian, J. T.; Callis, P. R. Mechanisms of Tryptophan Fluorescence Shifts in Proteins. *Biophys. J.* 2001, *80*, 2093-2109.

(442) Pal, S. K.; Peon, J.; Zewail, A. H. Biological Water at the Protein Surface: Dynamical Solvation Probed Directly with Femtosecond Resolution. *Proc. Natl. Acad. Sci. U. S. A.* 2002, *99*, 1763-1768.

(443) Yuan, T.; Weljie, A. M.; Vogel, H. J. Tryptophan Fluorescence Quenching by Methionine and Selenomethionine Residues of Calmodulin: Orientation of Peptide and Protein Binding. *Biochemistry* 1998, *37*, 3187-3195.

(444) Qiu, W. H.; Li, T. P.; Zhang, L. Y.; Yang, Y.; Kao, Y. T.; Wang, L. J.; Zhong, D. P. Ultrafast Quenching of Tryptophan Fluorescence in Proteins: Interresidue and Intrahelical Electron Transfer. *Chem. Phys.* 2008, *350*, 154-164.

(445) Alston, R. W.; Lasagna, M.; Grimsley, G. R.; Scholtz, J. M.; Reinhart, G. D.; Pace, C. N. Peptide Sequence and Conformation Strongly Influence Tryptophan Fluorescence. *Biophys. J.* 2008, *94*, 2280-2287.

(446) Lehrer, S. Solute Perturbation of Protein Fluorescence. Quenching of the Tryptophyl Fluorescence of Model Compounds and of Lysozyme by Iodide Ion. *Biochemistry* 1971, *10*, 3254-3263.

(447) Eftink, M. R.; Ghiron, C. A. Fluorescence Quenching Studies with Proteins. *Anal. Biochem.* 1981, *114*, 199-227.

(448) Zelent, B.; Kuśba, J.; Gryczynski, I.; Johnson, M. L.; Lakowicz, J. R. Time-Resolved and Steady-State Fluorescence Quenching of N-Acetyl-L-Tryptophanamide by Acrylamide and Iodide. *Biophys. Chem.* 1998, *73*, 53-75.

(449) Muiño, P. L.; Callis, P. R. Solvent Effects on the Fluorescence Quenching of Tryptophan by Amides Via Electron Transfer. Experimental and Computational Studies. *J. Phys. Chem. B* 2009, *113*, 2572-2577.

(450) Zhang, L.; Kao, Y.-T.; Qiu, W.; Wang, L.; Zhong, D. Femtosecond Studies of Tryptophan Fluorescence Dynamics in Proteins: Local Solvation and Electronic Quenching. *J. Phys. Chem. B* 2006, *110*, 18097-18103.

(451) Chen, R. F. Fluorescence Quantum Yields of Tryptophan and Tyrosine. *Anal. Lett.* 1967, *1*, 35-42.

(452) Petrich, J. W.; Chang, M. C.; McDonald, D. B.; Fleming, G. R. On the Origin of Nonexponential Fluorescence Decay in Tryptophan and Its Derivatives. *J. Am. Chem. Soc.* 1983, *105*, 3824-3832.

(453) Gudgin, E.; Lopez-Delgado, R.; Ware, W. R. The Tryptophan Fluorescence Lifetime Puzzle. A Study of Decay Times in Aqueous Solution as a Function of pH and Buffer Composition. *Can. J. Chem.* 1981, *59*, 1037-1044.

(454) Chen, Y.; Liu, B.; Yu, H. T.; Barkley, M. D. The Peptide Bond Quenches Indole Fluorescence. *J. Am. Chem. Soc.* 1996, *118*, 9271-9278.

(455) Adams, P. D.; Chen, Y.; Ma, K.; Zagorski, M. G.; Sönnichsen, F. D.; McLaughlin, M. L.; Barkley, M. D. Intramolecular Quenching of Tryptophan Fluorescence by the Peptide Bond in Cyclic Hexapeptides. *J. Am. Chem. Soc.* 2002, *124*, 9278-9286.

(456) Guharay, J.; Sengupta, P. K. Characterization of the Fluorescence Emission Properties of 7-Azatryptophan in Reverse Micellar Environments. *Biochem. Biophys. Res. Commun.* 1996, *219*, 388-392.

(457) Wong, C.-Y.; Eftink, M. R. Biosynthetic Incorporation of Tryptophan Analogues into Staphylococcal Nuclease: Effect of 5-Hydroxytryptophan and 7-Azatryptophan on Structure and Stability. *Protein Sci.* 1997, *6*, 689-697.

(458) Broos, J.; Maddalena, F.; Hesp, B. H. In Vivo Synthesized Proteins with Monoexponential Fluorescence Decay Kinetics. *J. Am. Chem. Soc.* 2004, *126*, 22-23.

(459) Xu, J. H.; Chen, B. B.; Callis, P.; Muino, P. L.; Rozeboom, H.; Broos, J.; Toptygin, D.; Brand, L.; Knutson, J. R. Picosecond Fluorescence Dynamics of Tryptophan and 5-Fluorotryptophan in Monellin: Slow Water-Protein Relaxation Unmasked. *J. Phys. Chem. B* 2015, *119*, 4230-4239.

(460) Gosavi, P. M.; Moroz, Y. S.; Korendovych, I. V. Beta-(1-Azulenyl)-L-Alanine - a Functional Probe for Determination of Pka of Histidine Residues. *Chem. Commun.* 2015, *51*, 5347-5350.

(461) Jennings, P.; Jones, A. C.; Mount, A. R. Fluorescence Properties of Electropolymerised 5-Substituted Indoles in Solution. *J. Chem. Soc., Faraday Trans.* 1998, *94*, 3619-3624.

(462) Hill, P. A.; Wei, Q.; Troxler, T.; Dmochowski, I. J. Substituent Effects on Xenon Binding Affinity and Solution Behavior of Water-Soluble Cryptophanes. *J. Am. Chem. Soc.* 2009, *131*, 3069-3077.

(463) Oeltermann, O.; Brand, C.; Engels, B.; Tatchen, J.; Schmitt, M. The Structure of 5-Cyanoindole in the Ground and the Lowest Electronically Excited Singlet States, Deduced from Rotationally Resolved Electronic Spectroscopy and Ab Initio Theory. *Phys. Chem. Chem. Phys.* 2012, *14*, 10266-10270.

(464) Kirby, E. P.; Steiner, R. F. Influence of Solvent and Temperature Upon the Fluorescence of Indole Derivatives. *J. Phys. Chem.* 1970, *74*, 4480-4490.

(465) Chen, Y.; Liu, B.; Barkley, M. D. Trifluoroethanol Quenches Indole Fluorescence by Excited-State Proton Transfer. *J. Am. Chem. Soc.* 1995, *117*, 5608-5609.

(466) Yu, H. T.; Colucci, W. J.; McLaughlin, M. L.; Barkley, M. D. Fluorescence Quenching in Indoles by Excited-State Proton Transfer. *J. Am. Chem. Soc.* 1992, *114*, 8449-8454.

(467) Zhang, R. B.; Somers, K. R. F.; Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T.; Ceulemans, A. Hydrogen Bonding to Pi-Systems of Indole and 1-Methylindole: Is There Any OH···Phenyl Bond? *J. Phys. Chem. A* 2005, *109*, 8028-8034.

(468) Othon, C. M.; Kwon, O.-H.; Lin, M. M.; Zewail, A. H. Solvation in Protein (Un)Folding of Melittin Tetramer–Monomer Transition. *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 12593-12598.

(469) Muñoz, M. a. A.; Carmona, C.; Balón, M. The Ground and Singlet Excited-State Hydrogen-Bonding Interactions of N-Methylindole with Trifluoroethanol in N-Hexane: A Model to Explain the Anomalous Fluorescence of Indole in Polar Protic Solvents. *Chem. Phys. Lett.* 2004, *393*, 217-221.

(470) Gervasio, F. L.; Chelli, R.; Marchi, M.; Procacci, P.; Schettino, V. Determination of the Potential of Mean Force of Aromatic Amino Acid Complexes in Various Solvents Using Molecular Dynamics Simulations: The Case of the Tryptophan−Histidine Pair. *J. Phys. Chem. B* 2001, *105*, 7835-7846.

(471) Albani, J. R. Origin of Tryptophan Fluorescence Lifetimes Part 1. Fluorescence Lifetimes Origin of Tryptophan Free in Solution. *J. Fluoresc.* 2014, *24*, 93-104.

(472) Kiefer, J.; Noack, K.; Kirchner, B. Hydrogen Bonding in Mixtures of Dimethyl Sulfoxide and Cosolvents. *Curr. Phys. Chem.* 2011, *1*, 340-351.

(473) Kaatze, U.; Pottel, R.; Schäfer, M. Dielectric Spectrum of Dimethyl Sulfoxide/Water Mixtures as a Function of Composition. *J. Phys. Chem.* 1989, *93*, 5623-5627.

(474) Johnson, M. E.; Malardier-Jugroot, C.; Head-Gordon, T. Effects of Co-Solvents on Peptide Hydration Water Structure and Dynamics. *Phys. Chem. Chem. Phys.* 2010, *12*, 393-405.

(475) Arakawa, T.; Kita, Y.; Timasheff, S. N. Protein Precipitation and Denaturation by Dimethyl Sulfoxide. *Biophys. Chem.* 2007, *131*, 62-70.

(476) Roy, S.; Jana, B.; Bagchi, B. Dimethyl Sulfoxide Induced Structural Transformations and Non-Monotonic Concentration Dependence of Conformational Fluctuation around Active Site of Lysozyme. *J. Chem. Phys.* 2012, *136*, 115103.

(477) Gasymov, O. K.; Abduragimov, A. R.; Glasgow, B. J. Site-Directed Circular Dichroism of Proteins: 1lb Bands of Trp Resolve Position-Specific Features in Tear Lipocalin. *Anal. Biochem.* 2008, *374*, 386-395.

(478) Lu, Z.; Manias, E.; Macdonald, D. D.; Lanagan, M. Dielectric Relaxation in Dimethyl Sulfoxide/Water Mixtures Studied by Microwave Dielectric Relaxation Spectroscopy. *J. Phys. Chem. A* 2009, *113*, 12207-12214.

(479) Pignataro, B. *Ideas in Chemistry and Molecular Sciences: Advances in Nanotechnology, Materials and Devices*; WILEY-VCH Verlag GmbH & Co. KGaA: Weinheim, 2010; Vol. 1.

(480) Tucker, M. J.; Oyola, R.; Gai, F. Conformational Distribution of a 14-Residue Peptide in Solution: A Fluorescence Resonance Energy Transfer Study. *J. Phys. Chem. B* 2005, *109*, 4788-4795.

(481) Bunagan, M. R.; Yang, X.; Saven, J. G.; Gai, F. Ultrafast Folding of a Computationally Designed Trp-Cage Mutant: Trp(2)-Cage. *J. Phys. Chem. B* 2006, *110*, 3759-3763.

(482) Shimbo, K.; Brassard, D. L.; Lamb, R. A.; Pinto, L. H. Ion Selectivity and Activation of the M2 Ion Channel of Influenza Virus. *Biophys. J.* 1996, *70*, 1335-1346.

(483) Pinto, L. H.; Lamb, R. A. The M2 Proton Channels of Influenza a and B Viruses. *J. Biol. Chem.* 2006, *281*, 8997-9000.

(484) Schnell, J. R.; Chou, J. J. Structure and Mechanism of the M2 Proton Channel of Influenza a Virus. *Nature* 2008, *451*, 591-595.

(485) Pielak, R. M.; Schnell, J. R.; Chou, J. J. Mechanism of Drug Inhibition and Drug Resistance of Influenza a M2 Channel. *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 7379-7384.

(486) Cady, S. D.; Schmidt-Rohr, K.; Wang, J.; Soto, C. S.; Degrado, W. F.; Hong, M. Structure of the Amantadine Binding Site of Influenza M2 Proton Channels in Lipid Bilayers. *Nature* 2010, *463*, 689-692.

(487) Balannik, V.; Carnevale, V.; Fiorin, G.; Levine, B. G.; Lamb, R. A.; Klein, M. L.; DeGrado, W. F.; Pinto, L. H. Functional Studies and Modeling of Pore-Lining Residue Mutants of the Influenza a Virus M2 Ion Channel. *Biochemistry* 2010, *49*, 696-708.

(488) Wu, Y. B.; Canturk, B.; Jo, H.; Ma, C. L.; Gianti, E.; Klein, M. L.; Pinto, L. H.; Lamb, R. A.; Fiorin, G.; Wang, J.; DeGrado, W. F. Flipping in the Pore: Discovery of Dual Inhibitors That Bind in Different Orientations to the Wild-Type Versus the Amantadine-Resistant S31n Mutant of the Influenza a Virus M2 Proton Channel. *J. Am. Chem. Soc.* 2014, *136*, 17987-17995.

(489) Sharma, M.; Yi, M. G.; Dong, H.; Qin, H. J.; Peterson, E.; Busath, D. D.; Zhou, H. X.; Cross, T. A. Insight into the Mechanism of the Influenza a Proton Channel from a Structure in a Lipid Bilayer. *Science* 2010, *330*, 509-512.

(490) Tang, Y.; Zaitseva, F.; Lamb, R. A.; Pinto, L. H. The Gate of the Influenza Virus M2 Proton Channel Is Formed by a Single Tryptophan Residue. *J. Biol. Chem.* 2002, *277*, 39880-39886.

(491) Luo, W.; Mani, R.; Hong, M. Side-Chain Conformation of the M2 Transmembrane Peptide Proton Channel of Influenza a Virus from 19f Solid-State NMR. *J. Phys. Chem. B* 2007, *111*, 10825-10832.

(492) Witter, R.; Nozirov, F.; Sternberg, U.; Cross, T. A.; Ulrich, A. S.; Fu, R. Solid-State 19f NMR Spectroscopy Reveals That Trp41 Participates in the Gating Mechanism of the M2 Proton Channel of Influenza a Virus. *J. Am. Chem. Soc.* 2008, *130*, 918-924.

(493) Chizhmakov, I. V.; Ogden, D. C.; Geraghty, F. M.; Hayhurst, A.; Skinner, A.; Betakova, T.; Hay, A. J. Differences in Conductance of M2 Proton Channels of Two Influenza Viruses at Low and High pH. *J. Physiol.* 2003, *546*, 427-438.

(494) Ma, C. L.; Fiorin, G.; Carnevale, V.; Wang, J.; Lamb, R. A.; Klein, M. L.; Wu, Y. B.; Pinto, L. H.; DeGrado, W. F. Asp44 Stabilizes the Trp41 Gate of the M2 Proton Channel of Influenza a Virus. *Structure* 2013, *21*, 2033-2041.

(495) Sansom, M. S. P.; Kerr, I. D.; Smith, G. R.; Son, H. S. The Influenza a Virus M2 Channel: A Molecular Modeling and Simulation Study. *Virology* 1997, *233*, 163-173.

(496) Smondyrev, A. M.; Voth, G. A. Molecular Dynamics Simulation of Proton Transport through the Influenza a Virus M2 Channel. *Biophys. J.* 2002, *83*, 1987-1996.

(497) Kass, I.; Arkin, I. T. How pH Opens a H+ Channel: The Gating Mechanism of Influenza a M2. *Structure* 2005, *13*, 1789-1798.

(498) Pinto, L. H.; Dieckmann, G. R.; Gandhi, C. S.; Papworth, C. G.; Braman, J.; Shaughnessy, M. A.; Lear, J. D.; Lamb, R. A.; DeGrado, W. F. A Functionally Defined Model for the M-2 Proton Channel of Influenza a Virus Suggests a Mechanism for Its Ion Selectivity. *Proc. Natl. Acad. Sci. U. S. A.* 1997, *94*, 11301-11306.

(499) Gandhi, C. S.; Shuck, K.; Lear, J. D.; Dieckmann, G. R.; DeGrado, W. F.; Lamb, R. A.; Pinto, L. H. Cu(II) Inhibition of the Proton Translocation Machinery of the Influenza a Virus M-2 Protein. *J. Biol. Chem.* 1999, *274*, 5474-5482.

(500) Schweighofer, K. J.; Pohorille, A. Computer Simulation of Ion Channel Gating: The M2 Channel of Influenza a Virus in a Lipid Bilayer. *Biophys. J.* 2000, *78*, 150-163.

(501) Fiorin, G.; Carnevale, V.; DeGrado, W. F. The Flu's Proton Escort. *Science* 2010, *330*, 456-458.

(502) Hu, F. H.; Luo, W. B.; Hong, M. Mechanisms of Proton Conduction and Gating in Influenza M2 Proton Channels from Solid-State NMR. *Science* 2010, *330*, 505-508.

(503) Khurana, E.; Peraro, M. D.; DeVane, R.; Vemparala, S.; DeGrado, W. F.; Klein, M. L. Molecular Dynamics Calculations Suggest a Conduction Mechanism for the M2 Proton Channel from Influenza a Virus. *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 1069-1074.

(504) Mould, J. A.; Li, H. C.; Dudlak, C. S.; Lear, J. D.; Pekosz, A.; Lamb, R. A.; Pinto, L. H. Mechanism for Proton Conduction of the M-2 Ion Channel of Influenza a Virus. *J. Biol. Chem.* 2000, *275*, 8592-8599.

(505) Lin, T. I.; Schroeder, C. Definitive Assignment of Proton Selectivity and Attoampere Unitary Current to the M2 Ion Channel Protein of Influenza a Virus. *J. Virol.* 2001, *75*, 3647-3656.

(506) Moffat, J. C.; Vijayvergiya, V.; Gao, P. F.; Cross, T. A.; Woodbury, D. J.; Busath, D. D. Proton Transport through Influenza a Virus M2 Protein Reconstituted in Vesicles. *Biophys. J.* 2008, *94*, 434-445.

(507) DeCoursey, T. E. Voltage-Gated Proton Channels and Other Proton Transfer Pathways (Vol 83, Pg 475, 2003). *Physiol. Rev.* 2003, *83*, 1067-1067.

(508) Chen, H.; Wu, Y.; Voth, G. A. Proton Transport Behavior through the Influenza a M2 Channel: Insights from Molecular Simulation. *Biophys. J.* 2007, *93*, 3470-3479.

(509) DiFrancesco, M. L.; Hansen, U. P.; Thiel, G.; Moroni, A.; Schroeder, I. Effect of Cytosolic pH on Inward Currents Reveals Structural Characteristics of the Proton Transport Cycle in the Influenza a Protein M2 in Cell-Free Membrane Patches of Xenopus Oocytes. *PLoS One* 2014, *9*, e107406.

(510) Hu, F.; Schmidt-Rohr, K.; Hong, M. NMR Detection of pH-Dependent Histidine–Water Proton Exchange Reveals the Conduction Mechanism of a Transmembrane Proton Channel. *J. Am. Chem. Soc.* 2012, *134*, 3703-3713.

(511) Zhou, H. X. A Theory for the Proton Transport of the Influenza Virus M2 Protein: Extensive Test against Conductance Data. *Biophys. J.* 2011, *100*, 912-921.

(512) Rasaiah, J. C.; Garde, S.; Hummer, G. Water in Nonpolar Confinement: From Nanotubes to Proteins and Beyond. *Annu. Rev. Phys. Chem.* 2008, *59*, 713-740.

(513) Voth, G. A. Computer Simulation of Proton Solvation and Transport in Aqueous and Biomolecular Systems. *Acc. Chem. Res.* 2006, *39*, 143-150.

(514) Manor, J.; Mukherjee, P.; Lin, Y.-S.; Leonov, H.; Skinner, J. L.; Zanni, M. T.; Arkin, I. T. Gating Mechanism of the Influenza a M2 Channel Revealed by 1d and 2D IR Spectroscopies. *Structure* 2009, *17*, 247-254.

(515) Manor, J.; Feldblum, E. S.; Zanni, M. T.; Arkin, I. T. Environment Polarity in Proteins Mapped Noninvasively by FTIR Spectroscopy. *J. Phys. Chem. Lett.* 2012, *3*, 939-944.

(516) Hong, M.; Fritzsching, K. J.; Williams, J. K. Hydrogen-Bonding Partner of the Proton-Conducting Histidine in the Influenza M2 Proton Channel Revealed from H-1 Chemical Shifts. *J. Am. Chem. Soc.* 2012, *134*, 14753-14755.

(517) Liang, R. B.; Li, H.; Swanson, J. M. J.; Voth, G. A. Multiscale Simulation Reveals a Multifaceted Mechanism of Proton Permeation through the Influenza a M2 Proton Channel. *Proc. Natl. Acad. Sci. U. S. A.* 2014, *111*, 9396-9401.

(518) Ghosh, A.; Wang, J.; Moroz, Y. S.; Korendovych, I. V.; Zanni, M.; DeGrado, W. F.; Gai, F.; Hochstrasser, R. M. 2D IR Spectroscopy Reveals the Role of Water in the Binding of Channel-Blocking Drugs to the Influenza M2 Channel. *J. Chem. Phys.* 2014, *140*, 235105.

(519) Ghosh, A.; Qiu, J.; DeGrado, W. F.; Hochstrasser, R. M. Tidal Surge in the M2 Proton Channel, Sensed by 2D IR Spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* 2011, *108*, 6115-6120.

(520) Okada, A.; Miura, T.; Takeuchi, H. Protonation of Histidine and Histidine−Tryptophan Interaction in the Activation of the M2 Ion Channel from Influenza a Virus. *Biochemistry* 2001, *40*, 6053-6060.

(521) Stouffer, A. L.; Acharya, R.; Salom, D.; Levine, A. S.; Di Costanzo, L.; Soto, C. S.; Tereshko, V.; Nanda, V.; Stayrook, S.; DeGrado, W. F. Structural Basis for the Function and Inhibition of an Influenza Virus Proton Channel. *Nature* 2008, *451*, 596-599.

(522) Waegele, M. M.; Gai, F. Computational Modeling of the Nitrile Stretching Vibration of 5-Cyanoindole in Water. *J. Phys. Chem. Lett.* 2010, *1*, 781-786.

(523) Ma, C.; Polishchuk, A. L.; Ohigashi, Y.; Stouffer, A. L.; Schön, A.; Magavern, E.; Jing, X.; Lear, J. D.; Freire, E.; Lamb, R. A.; DeGrado, W. F.; Pinto, L. H. Identification of the Functional Core of the Influenza a Virus a/M2 Proton-Selective Ion Channel. *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 12283-12288.

(524) Stouffer, A. L.; Nanda, V.; Lear, J. D.; DeGrado, W. F. Sequence Determinants of a Transmembrane Proton Channel: An Inverse Relationship between Stability and Function. *J. Mol. Biol.* 2005, *347*, 169-179.

(525) Yi, M. G.; Cross, T. A.; Zhou, H. X. Conformational Heterogeneity of the M2 Proton Channel and a Structural Model for Channel Activation. *Proc. Natl. Acad. Sci. U. S. A.* 2009, *106*, 13311-13316.

(526) Cady, S. D.; Mishanina, T. V.; Hong, M. Structure of Amantadine-Bound M2 Transmembrane Peptide of Influenza a in Lipid Bilayers from Magic-Angle-Spinning Solid-State NMR: The Role of Ser31 in Amantadine Binding. *J. Mol. Biol.* 2009, *385*, 1127-1141.

(527) Rogers, J. M. G.; Polishchuk, A. L.; Guo, L.; Wang, J.; DeGrado, W. F.; Gai, F. Photoinduced Electron Transfer and Fluorophore Motion as a Probe of the Conformational Dynamics of Membrane Proteins: Application to the Influenza a M2 Proton Channel. *Langmuir* 2011, *27*, 3815-3821.

(528) Liao, S. Y.; Yang, Y.; Tietze, D.; Hong, M. The Influenza M2 Cytoplasmic Tail Changes the Proton-Exchange Equilibria and the Backbone Conformation of the Transmembrane Histidine Residue to Facilitate Proton Conduction. *J. Am. Chem. Soc.* 2015, *137*, 6067-6077.

(529) Purkayastha, P.; Klemke, J. W.; Lavender, S.; Oyola, R.; Cooperman, B. S.; Gai, F. A1-Antitrypsin Polymerization: A Fluorescence Correlation Spectroscopic Study. *Biochemistry* 2005, *44*, 2642-2649.

(530) Kochendoerfer, G. G.; Salom, D.; Lear, J. D.; Wilk-Orescan, R.; Kent, S. B. H.; DeGrado, W. F. Total Chemical Synthesis of the Integral Membrane Protein Influenza a Virus M2: Role of Its C-Terminal Domain in Tetramer Assembly. *Biochemistry* 1999, *38*, 11905-11913.

(531) Pielak, R. M.; Chou, J. J. Kinetic Analysis of the M2 Proton Conduction of the Influenza Virus. *J. Am. Chem. Soc.* 2010, *132*, 17695-17697.

(532) Kukol, A.; Adams, P. D.; Rice, L. M.; Brunger, A. T.; Arkin, I. T. Experimentally Based Orientational Refinement of Membrane Protein Models: A Structure for the Influenza a M2 H+ Channel. *J. Mol. Biol.* 1999, *286*, 951-962.

(533) Cady, S. D.; Luo, W.; Hu, F.; Hong, M. Structure and Function of the Influenza a M2 Proton Channel. *Biochemistry* 2009, *48*, 7356-7364.

(534) Suydam, I. T.; Boxer, S. G. Vibrational Stark Effects Calibrate the Sensitivity of Vibrational Probes for Electric Fields in Proteins. *Biochemistry* 2003, *42*, 12050-12055.

(535) Bagchi, S.; Fried, S. D.; Boxer, S. G. A Solvatochromic Model Calibrates Nitriles' Vibrational Frequencies to Electrostatic Fields. *J. Am. Chem. Soc.* 2012, *134*, 10373-10376.

(536) Phongphanphanee, S.; Rungrotmongkol, T.; Yoshida, N.; Hannongbua, S.; Hirata, F. Proton Transport through the Influenza a M2 Channel: Three-Dimensional Reference Interaction Site Model Study. *J. Am. Chem. Soc.* 2010, *132*, 9782-9788.

(537) Nanda, V.; Cristian, L.; Toptygin, D.; Brand, L.; DeGrado, W. F. Nanosecond Dynamics of Influenza a/M2tm and an Amantadine Resistant Mutant Probed by Time-Dependent Red Shifts of a Native Tryptophan. *Chem. Phys.* 2013, *422*, 73-79.

(538) Czabotar, P. E.; Martin, S. R.; Hay, A. J. Studies of Structural Changes in the M2 Proton Channel of Influenza a Virus by Tryptophan Fluorescence. *Virus Res.* 2004, *99*, 57-61.

(539) Hamm, H. E.; Meier, S. M.; Liao, G.; Preininger, A. M. Trp Fluorescence Reveals an Activation-Dependent Cation-Π Interaction in the Switch II Region of Gα(I) Proteins. *Protein Sci.* 2009, *18*, 2326-2335.

(540) Piletic, I. R.; Moilanen, D. E.; Spry, D. B.; Levinger, N. E.; Fayer, M. D. Testing the Core/Shell Model of Nanoconfined Water in Reverse Micelles Using Linear and Nonlinear IR Spectroscopy. *J. Phys. Chem. A* 2006, *110*, 4985-4999.

(541) Kwak, K.; Park, S.; Finkelstein, I. J.; Fayer, M. D. Frequency-Frequency Correlation Functions and Apodization in Two-Dimensional Infrared Vibrational Echo Spectroscopy: A New Approach. *J. Chem. Phys.* 2007, *127*, 124503.

(542) Wang, J.; Ma, C.; Fiorin, G.; Carnevale, V.; Wang, T.; Hu, F.; Lamb, R. A.; Pinto, L. H.; Hong, M.; Klein, M. L.; DeGrado, W. F. Molecular Dynamics Simulation Directed Rational Design of Inhibitors Targeting Drug-Resistant Mutants of Influenza a Virus M2. *J. Am. Chem. Soc.* 2011, *133*, 12834-12841.

(543) Toto, A.; Camilloni, C.; Giri, R.; Brunori, M.; Vendruscolo, M.; Gianni, S. Molecular Recognition by Templated Folding of an Intrinsically Disordered Protein. *Sci. Rep.* 2016, *6*, 21994.

(544) Ma, J.; Komatsu, H.; Kim, Y. S.; Liu, L.; Hochstrasser, R. M.; Axelsen, P. H. Intrinsic Structural Heterogeneity and Long-Term Maturation of Amyloid β Peptide Fibrils. *ACS Chem. Neurosci.* 2013, *4*, 1236-1243.

(545) Stuhlmann, B.; Gra; Schmitt, M. Determination of the Geometry Change of 5-Cyanoindole Upon Electronic Excitation from a Combined Franck-Condon/Rotational Constants Fit. *Phys. Chem. Chem. Phys.* 2014, *16*, 899-905.

(546) Ghosh, A.; Tucker, M. J.; Gai, F. 2D IR Spectroscopy of Histidine: Probing Sidechain Structure and Dynamics Via Backbone Amide I Vibration. *J. Phys. Chem. B* 2014, *118*, 7799-7805.