2015

# Spectroscopic Probes of Protein Structure, Dynamics, Hydration and Electrostatics

Ileana Marquez Pazos
*University of Pennsylvania*, ipazos@sas.upenn.edu

# Spectroscopic Probes of Protein Structure, Dynamics, Hydration and Electrostatics


## Abstract

The structure, dynamics and function of a protein are intimately controlled by a large number of intermolecular and intramolecular interactions. Thus, achieving a quantitative and molecular-level understanding of how proteins fold and function requires experimental techniques that can â??senseâ?? and differentiate various molecular forces and, in many cases, in a site-specific manner. To that end, the focus of this thesis work is to develop non-natural amino acid-based infrared and fluorescence probes that can be used to assess the local hydration status and electrostatic environment of proteins. First, we expand the utility of a well-known site-specific spectroscopic probe, p-cyano-phenylalanine (PheCN), by showing that (1) its fluorescence is sensitive to the presence of various anions and can thus be used to measure the heterogeneity of the protein conformation, (2) when placed at the N-terminal end of a peptide this non-natural amino acid can be used as a pH sensor for a wide variety of applications, and (3) its nitrile stretching vibration is a microscopic reporter of how a co-solvent, such as urea and trimethylamine N-oxide, modulates the protein-water interactions. Secondly, we demonstrate that the ester carbonyl stretching vibration of the non-natural amino acids, L-aspartic acid 4-methyl ester and L-glutamic acid 5-methyl ester, can be used to site-specifically quantify the electrostatic environment of proteins as its vibrational frequency correlates linearly with the local electrostatic field. Application of this infrared probe to amyloids allows us to gain new insight into their structure and dynamics.



## Degree Type

Dissertation

## Degree Name

Doctor of Philosophy (PhD)

## Graduate Group

Chemistry

## First Advisor

Feng Gai

## Keywords

cosolvent, electric field in proteins, fluorescence spectroscopy, infrared spectroscopy, non-natural amino acid, site-specific probe

## Subject Categories

Chemistry

# SPECTROSCOPIC PROBES OF PROTEIN STRUCTURE, DYNAMICS, HYDRATION AND ELECTROSTATICS

Ileana M. Pazos

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

2015

Supervisor of Dissertation

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

Graduate Group Chairperson

_____
Gary Molander
Hirschmann-Makineni Professor of Chemistry

Dissertation Committee:

Dr. Zahra Fakhraai, (Chair) Assistant Professor of Chemistry
Dr. Jeffery Saven, Professor of Chemistry
Dr. Tobias Baumgart, Associate Professor of Chemistry

SPECTROSCOPIC PROBES OF PROTEIN STRUCTURE, DYNAMICS, HYDRATION AND ELECTROSTATICS

To the memory of my grandmother, Olga "Apa" Marquez.

Acknowledgements

First and foremost, I would like to extend my sincerest gratitude to my thesis advisor, Dr. Feng Gai. His wholehearted support, guidance, patience, and sagacity have been and will always be deeply appreciated. I am humbled by the opportunity I had to develop my skills under his mentorship.

I would also like to thank my committee members, Dr. Zahra Fakhraai, Dr. Jeffery Saven, and Dr. Tobias Baumgart for their guidance, constructive criticisms, and the encouragement that they have offered me over the years.

I thank Dr. Matthew J. Tucker and Dr. Ayanjeet Ghosh for many productive conversations during late evenings in lab or over dinner. My classmates, Dr. Ethan Alguire and Dr. Chris Von Bargen, always found time to offer their theoretical perspectives on my research during our regular lunches, for which I still owe them money. I appreciate all of their friendships.

My time in lab has been enriched by my friends and colleagues, both past and present. I would first like to recognize Dr. Robert Culik and Beatrice Markiewicz for always being generous with their time, both professionally and personally. Also, I have enjoyed my fruitful collaborations with Rachel Abaskharon, Dr. Thomas Measey, Dr. Jianqiang Ma, Dr. Lev Chuntonov, and Ismail Ahmed. The current staff and post-doctoral members, Dr. Jian-Xin Chen, Dr. Thomas Troxler, Dr. Wenkai Zhang, Dr. Bei Ding, and Dr. Kwang Im Oh have offered me a wealth of scientific knowledge and career advice. I thank my

other current labmates Chun-Wei Lin, Mary Rose Mintzer, Jeffrey Rodgers and Debopreeti Mukherjee for contributing to a collaborative and kind environment. Afternoon expeditions with my labmates to Starbucks, Dunkin Donuts or CVS provide fresh air and a boost to my daily step counter. I would also like to thank previous and successful graduates of the lab: Dr. Kathryn Smith-Dupont, Dr. Arnaldo Serrano, Dr. Lin Guo, Dr. Matthias Waegele and Dr. Julie Rogers. They have all contributed to my academic growth over the years, and have my sincere gratitude.

Most importantly, I want to thank my family: my mother, Ileana Martinez; my father, Jose Pazos; my partner, Robert C. Donnan; my sister, Christena; my brothers, Eduardo, Javier, and Miguel; my nephews, Rafael and Vincent; and my best friend, Carolina Pombo for their love, support, and the joy they bring to my life.

ABSTRACT

# SPECTROSCOPIC PROBES OF PROTEIN STRUCTURE, DYNAMICS, HYDRATION AND ELECTROSTATICS

Ileana M. Pazos

Feng Gai

The structure, dynamics and function of a protein are intimately controlled by a large number of intermolecular and intramolecular interactions. Thus, achieving a quantitative and molecular-level understanding of how proteins fold and function requires experimental techniques that can 'sense' and differentiate various molecular forces and, in many cases, in a site-specific manner. To that end, the focus of this thesis work is to develop non-natural amino acid-based infrared and fluorescence probes that can be used to assess the local hydration status and electrostatic environment of proteins. First, we expand the utility of a well-known site-specific spectroscopic probe, *p*-cyano-phenylalanine ($Phe_{CN}$), by showing that (1) its fluorescence is sensitive to the presence of various anions and can thus be used to measure the heterogeneity of the protein conformation, (2) when placed at the N-terminal end of a peptide this non-natural amino acid can be used as a pH sensor for a wide variety of applications, and (3) its nitrile stretching vibration is a

microscopic reporter of how a co-solvent, such as urea and trimethylamine N-oxide, modulates the protein-water interactions. Secondly, we demonstrate that the ester carbonyl stretching vibration of the non-natural amino acids, L-aspartic acid 4-methyl ester and L-glutamic acid 5-methyl ester, can be used to site-specifically quantify the electrostatic environment of proteins as its vibrational frequency correlates linearly with the local electrostatic field. Application of this infrared probe to amyloids allows us to gain new insight into their structure and dynamics.

**Table of Contents**

**List of Tables**

**List of Figures**

# 1 Introduction

Proteins are linear polymers made up of 20 naturally occurring amino acids whose side chains vary in size, charge, polarity, and hydrophobicity. For a given protein, the specific ordering of its constituent amino acids or sequence determines its three-dimensional native fold conformation. However, this folded structure is not static, but rather undergoes various spontaneous structural fluctuations, which, in many cases, are necessary for proteins to execute their functions. On the other hand, changing one or multiple environmental conditions can amplify some of these spontaneous fluctuations, leading to undesirable consequences such as unfolding and aggregation. Because of the large number of degrees of freedom associated with even a small-sized protein, it is experimentally challenging to probe protein conformational dynamics, including those related to folding and function, with high structural and temporal resolution. This is because protein conformational motions can occur both locally and globally and also on a wide range of timescales. In other words, multiple techniques and probes are necessary to provide complementary information for a more complete description of the underlying conformational energy landscape and dynamics of the protein system in question. Furthermore, studies concerning the structure-function relationship of proteins require knowledge of various structural and environmental factors as well as their changes at a specific region, such as the electrostatic field at the substrate binding site of an enzyme. Thus, there is still a great need for spectroscopic probes that can provide site-specific structural and/or environmental information in protein conformational studies.

In this regard, the research described in this thesis focuses on expanding the application of an existing site-specific probe, *p*-cyanophenylalanine, and demonstrating the utility of the ester carbonyl stretching vibration as an infrared (IR) probe of protein local electrostatics and hydration status. To provide background information that is relevant to studies described in this thesis, the upcoming sections are organized as follows: 1) a concise review of the necessary concepts, 2) non-natural amino acid-based site-specific vibrational and fluorescent probes, 3) a brief introduction to some of the spectroscopic techniques, and 4) the thesis outline.

## 1.1 Hydrogen bonding

A distinct and important physical property of water molecules is their ability and tendency to form hydrogen bonds (H-bonds) among themselves and with other H-bonding acceptors and donors.[1-4] Water and their corresponding H-bonds are believed to play a crucial role in the folding and function of biological molecules and assemblies. However, it is very difficult to directly assess the properties of these H-bonds because they are inherently heterogeneous and are made and broken on extremely fast timescales. In a complex system, when water interacts with biologically relevant cosolvents, such as urea or TMAO, the intermolecular H-bonding between water molecules is perturbed. In fact, the hydration of a protein, especially in the first hydration shell, is a key determinant

of the three-dimensional structure of the protein. Therefore, understanding H-bonding dynamics in biological systems is an important undertaking.

## 1.2 Cosolvents

Nature utilizes a wide variety of small molecules, such as salts, amino acids, sugars, methylamines and urea, as osmolytes to maintain cell volume and to protect the cell components against denaturing stresses.[5-8] These molecules have the capability to promote or disrupt specific macromolecular interactions. Because this has vast pharmaceutical potential which includes dissolving and re-folding insoluble recombinant proteins that can extend their shelf life, there is a great interest in understanding what drives and controls these interactions at the molecular-level. Two common osmolytes, urea and guanidine hydrochloride (GdnHCl), are protein denaturants that can cause a protein to unfold. On the other hand, trimethylamine N-oxide (TMAO) is able to enhance protein stability and counteract the denaturing effect of urea.[9-10] While there is a large body of research[11-38] on how such stabilizing and destabilizing effects might take place, a coherent view of the underlying mechanism of action at the molecular level has not been reached.

The mechanism of action of osmolytes has been explained by two different mechanisms: the "indirect" and "direct" mechanisms. For example, the indirect mechanism involves the disruption of the hydrogen bonding network of water. For a

denaturant, this is thought to weaken the hydrophobic interactions of the protein, in turn making the protein less compact and more easily solvated by water molecules.[39-42] However, it is under debate whether this indirect mechanism would affect the structure and dynamics of water enough to significantly reduce the strength of stabilizing interactions within proteins.[43-44] As such, many previous studies have focused on investigating whether a given osmolyte or cosolvent can affect the H-bond network of water molecules.[45-48] While such studies have provided significant insights into our understanding of how a given cosolvent interacts with and changes the H-bonding network of water, they have offered little, if any, direct information on how the cosolvent of interest mediates the strength of the H-bonds formed between water and the protein in solution. On the other hand, the direct mechanism suggests that urea unfolds proteins through interactions between the osmolyte and the protein backbone or sidechains. Specifically, the stronger H-bonds between the urea carbonyl and the backbone amides would be preferred over intra-backbone H-bonds of the folded protein.

In contrast, TMAO has a protecting or counteracting role to preserve the folded conformation of a protein. It has been suggested that the interaction between TMAO and the backbone are entropically unfavorable, resulting in a preferential depletion of TMAO from protein surfaces. This can in turn result in an increase of hydration of the surface of the protein.[49-54] While this view is consistent with a protecting osmolyte, it is difficult to experimentally examine the effect of TMAO on hydration dynamics of a protein.

Therefore, to gain a comprehensive understanding of the mechanism of action of TMAO or urea, an ideal experiment would report on the effect of a cosolvent on the H-bonding ability of water, near the protein surface. This could be achieved by monitoring the spectral diffusion of the vibrational transition of a IR probe, located either on a sidechain or the backbone of a protein, during the same timescale of H-bond forming and breaking. One of the methods to measure the dynamics of such spectral diffusion is two-dimensional (2D) IR spectroscopy. This experiment is capable of measuring the vibrational frequency-frequency correlation function (FFCF), which is a connection to the microscopic molecular dynamics of the system. The FFCF function is denoted by $C_1$:

$$C_1(t) = \left\langle \delta\omega_{1,0}(\tau_1)\delta\omega_{1,0}(0) \right\rangle = \frac{\delta(t)}{T_2} + \sum_i \Delta_i^2 \exp(-t/\tau_i) \qquad \text{Equation 1.1}$$

where $T_2$ is the pure-dephasing time, $\Delta_i$ is the frequency fluctuation amplitude and the $\tau_i$ is the correlation time of the $i$th component. Essentially, spectral diffusion is a measure of how quickly the system loses its memory and goes from an elongated shape to a circular shape. This spectral lineshape of the 2D IR spectrum can report important parameters of dynamics because the change of the spectral diffusion is related to the microscopic molecular dynamics. To quantify the change, one can analyze the nodal line slope, the photon echo peak shift or the center line slope (CLS) of the spectra.[55-56] The CLS is determined by taking the inverse of the maximum of slices that are parallel to pump frequency axis of the 2D IR spectrum. Fayer and coworkers showed analytically that the CLS, as a function of waiting time, is related to the FFCF.[57]

## 1.3 Crowding

Intercellular environments are congested due to a high concentration of macromolecules ranging between 80 – 400 g/L.[58] This high degree of crowding could cause a large difference between the thermodynamic or kinetic behaviors of, for example, a protein, obtained *in vitro* and *in vivo* measurements.[59] Quantitative reproduction of a cellular environment is difficult because *in vivo* conditions include heterogeneous and non-spherical particles. Thus, *in vitro* studies of the macromolecular crowding effect on biomolecules often use high concentrations of relatively inert spherical crowding agents, such as Ficoll and Dextran.[60] The macromolecular crowding effect is an entropic effect, which arises from the excluded volume of the crowding agents. Regardless of other attractive or repulsive interactions that might be present, the excluded volume is a nonspecific steric repulsion that is always present and depends on the size and concentration of the crowding agent. The volume accesible to a larger particle can be very restricted, which results in a relative reduction of configurational entropy of the biopolymer in question. Also, the volume available to the solvent is reduced which increases the effective concentration of the solute and results in an increase in its chemical activity. Consequently, for example, crowding can alter the rate and equilibrium constant of biochemical reactions.[61]

As suggested by a recent molecular dynamics simulation, much like macromolecular crowding, the driving force of the protecting mechanism of TMAO is entropic stabilization.[62] Specifically, TMAO is capable of forming H-bonds to the

solvent-exposed backbone nitrogen atoms of a protein. However those H-bonds compete with intramolecular H-bonds that are responsible for the secondary structure of the peptide. An increase in these intramolecular H-bonds would result in a reduction of the number of H-bonds between TMAO and the backbone producing a depletion of TMAO from the vicinity of the surface of the protein. An increase in intramolecular H-bonding results in a compaction of protein conformations. [33] In other words, TMAO limits the degrees of freedom of the unfolded state of the protein and thus entropically destabilizes it. Because this is similar to the crowding effect where the unfolded state of a protein is destabilized due to geometric restrictions,[63] Thirumalai and coworkers termed the depletion-induced protein structure formation of TMAO as nano-crowding.[62]

## 1.4 Electric fields and the Vibrational Stark Effect

Electric fields play an important role in protein function. Similar to H-bonding, the electric field produced by the environment modifies the electron density distribution of a given molecule. These fields are spatially non-uniform because they are determined by the immediate environment. This local environment consists of non-covalent interactions with solvent molecules, and charged, polar and polarizable cosolvents or nearby sidechains. It is important to understand the electrostatics of protein since it plays a central role in their structure and dynamics. For example, there can be dramatic changes of electric field during an enzymatic reaction, which are thought to have critical

contributions to the catalytic rate.[64-66] While it is possible for the electric field to be theoretically calculated, experimental measurements are scarce because of the lack of appropriate probes that report on electrostatic fields in a site-specific manner.

Other existing experimental approaches to measure electric field include measuring shifts in redox potential,[67] NMR chemical shifts,[68-72] and $pK_a$ shifts of ionizable residues.[73-75] Some of the most promising experimental approaches have involved IR probes because they can be site-specific and sensitive to electric fields in biological macromolecules.[76-77] However, the vibrational frequency shifts of many IR probes do not have a simple relationship to the electric field amplitude exerted by the surrounding molecules. To understand the solvent-induced vibrational frequency shifts, the probe can be dissolved in solvents of varying polarities, and a correlation between the electric field and frequency shifts is determined. Alternatively, the relationship can also be measured by applying a static electric field to an isotropic and immobilized sample.[78-79] This phenomenon can be described with linear Stark effect theory:

$$hc\Delta\bar{\nu}_{observed} = -\Delta\vec{\mu}_{tuning\ rate} \cdot \Delta\vec{F}_{environment} \qquad \text{Equation 1.2}$$

where the vibrational frequency shift ($\Delta\bar{\nu}_{observed}$) induced by the electric field is related to the electric field vector ($\Delta\vec{F}_{environment}$) by the vibrational Stark tuning rate ($\Delta\vec{\mu}_{tuning\ rate}$).[79] The Stark tuning rate is a measure of the sensitivity of the vibrational frequency to an external electric field, with a more sensitive probe having a larger tuning rate.[80] Classically, the Stark effect is expected to cause a small peak shift in the absorbance spectrum but no change in line shape. However, the center frequency of a

transition does not necessarily have a linear relationship with electric field. There is also a quadratic force constant of the normal mode that can be affected by an electric field which arises from the difference polarizability of the transition.

The nitrile moiety has proven to be a very convenient IR probe due to the fact that its vibrational frequency is located in an uncongested region of the infrared spectra and it can easily be incorporated into proteins.[81-82] However, the Stark effect for the nitrile group does not appear to be straight forward. Specifically, the frequency shift of the nitrile is believed to report either a H-bonding component[83] or on a more complex electrostatic sensitivity beyond the first order electric field term.[84] Thus, given the proven importance of electric field in various biological processes, it would be quite advantageous to find a localized and simple vibrational transition whose frequency exhibits a linear dependence on the electric field exerted by its surrounding environment. For example, the stretching mode frequencies of CO and CF are approximately proportional to the solvent electric field projected onto the bond axes, thus have been suggested to be able to serve this role.[85]

## 1.5 Experimental Methods

Computers have theoretically simulated biological systems that are composed of thousands to millions of individual atoms, where each frame of the trajectory contains the precise location of every atom. These simulations range from timescales smaller than a

chemical bond vibration ($10^{-15}$ s) to the timescales of protein folding ($10^{-3}$ s).[86-87] Molecular dynamics (MD) simulations are capable of this level of spatiotemporal resolution. However experimentally, there are trade-offs between spatial and temporal resolution. X-ray crystallography and nuclear magnetic resonance (NMR) are two methods that can be applied to the study of three-dimensional molecular structures of proteins at atomic resolution.[88] X-ray crystallography has the unparalleled ability to provide structural information of proteins at atomistic resolution of large proteins that are capable of forming ordered crystals. However this method requires protein crystallization which is a difficult task and only represents a static picture of the protein with no dynamical information. Protein structure determination by solution NMR spectroscopy relies on the isotopic enrichment of $^{13}C$ and $^{15}N$ to alleviate resonance overlap and to allow multiple distance and angular measurements. Additionally, dynamics from hydrogen-deuterium exchange experiments reflect the solvent accessibility of the protein surface and reveals information about regions that are buried or involved in hydrogen bonding.[89]

### 1.5.1 Fluorescence Spectroscopy

Absorption of a photon takes place on the femtosecond scale and fluorescence occurs in the range of 1–100 nanoseconds ($10^{-9}$–$10^{-7}$ s).[90] There are three amino acids with intrinsic fluorescence properties, phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) but only Tyr and Trp are used experimentally because their quantum yields are larger. Those residues can be used to follow protein folding because their quantum yields

are sensitive to their environment which changes when protein folds/unfolds. The intrinsic or extrinsic fluorophores can report on local environment or provide structural information such as distances through fluorescence resonance energy transfer (FRET) approaches. In a hydrophobic environment, Tyr and Trp have a high quantum yield but in a hydrophilic environment their quantum yield is decreased leading to lowered fluorescence intensity. For Trp, there is also a strong Stokes shift that is dependent on the solvent, meaning that the maximum emission wavelength of Trp will be at a larger wavelength in more polar environments. Moreover, various extrinsic fluorescent dyes have been used to report on protein conformational changes.[91-93] These extrinsic dyes can be covalently attached to a protein and often have larger quantum yields. On the other hand, they are often larger aromatic molecules that can be large and a bulky substitution for an amino acid which can in turn affect the native conformation of the protein.

Fluorescence spectroscopy can be used to explore the kinetics of a photophysical intermolecular deactivation process, using the Stern–Volmer relationship. In the past, quenching of protein fluorescence by either iodide, molecular oxygen or acrylamide were popular due to the ease of the experiments and the value of the structural information obtained.[94-98] For tryptophan, the experiment reports on the accessibility of the indole ring in the system of interest to the quencher. The Stern-Volmer equation relates the decrease in fluorescence in the absence or presence of a collisional quencher. The quenching can be measured as a decrease in either the fluorescence intensity or the fluorescence lifetime.

### 1.5.2 Infrared Spectroscopy

Infrared spectroscopy (IR) is one of the oldest and well-established experimental techniques for the analysis of secondary structure of polypeptides that began with the pioneering work of Elliot and Ambrose in the early 1950s.[99] It is convenient, non-destructive and can be used under a wide variety of conditions.[100] This powerful technique can probe the fundamental vibrations of molecules (200-4000 $cm^{-1}$). Polypeptides and proteins exhibit nine amide bands that represent delocalized vibrations of the peptide backbone. N-methylacetamide (NMA) has been used as a model to compute vibrational force field since it is the smallest molecule that contains a trans-peptide group.[101,102] It has therefore become the starting point for a normal mode analysis of polypeptide backbone vibrations to obtain a force field for the peptide group.

Infrared analysis of secondary structure typically makes use of the amide I vibrational mode (1600−1700 $cm^{-1}$) which arises mainly from the C=O stretching vibration. The location and direction of the transition dipole vector whose center is located close to the oxygen and close to the C=O bond is also shown. It points away from the C=O bond towards the C—N bond by 20˚. The localization of the amide I mode to the C=O stretch makes analysis of this vibrational band practical and much less complicated than the other amide modes. In addition to the amide I, other backbone modes include NH stretching vibrations termed the amide A and B (at ~3300 and ~3170 $cm^{-1}$) and the amide II (~1550$cm^{-1}$), which arises from the out-of-phase combination of the NH in-plane bend, the CN stretching vibration, and smaller contributions from CO in-

plane bend and CC and NC stretching vibrations. The backbone modes also include the amide III (1200-1400 $cm^{-1}$) and the skeletal stretch (880-1200 $cm^{-1}$).[102,103]

Some of the major advantages of IR is that the amide I is sensitive to secondary structure as a result of transition dipole coupling (TDC) among the backbone units and the absorption frequency also displays sensitivity to H-bonding and through bond coupling. In fact, IR spectroscopy is one of the few methods that can directly report on the strength of H-bonds. TDC is a resonance interaction between oscillating dipoles which depends on the relative orientations and separation distance of the dipoles. TDC is the fundamental mechanism that is responsible for the sensitivity of the amide I vibration to secondary structure. For this reason, IR spectroscopy can be particularly useful in the study of proteins because the vibrations can report on structure or environment.[103] This is because ultrafast IR techniques can measure dynamics on extremely short time scales, femtoseconds ($10^{-15}$ seconds) to nanoseconds ($10^{-9}$ seconds). However, the interpretation can be complicated since all the molecular vibrations from the solvent and protein will contribute to the IR spectrum. Therefore it can be difficult to directly observe specific region of the protein unless you introduce a functional group that does not already exist in the protein (nitrile, azide group, etc.).

While the backbone can provide global information about the protein conformation, isotopic labeling allows us to single out an individual residue for kinetic or structural information. Labeling the backbone with heavier isotopes changes the specific mass of an atom results in a vibrational shift (20-50 $cm^{-1}$) to lower wavenumbers, also,

this effectively uncouples that specific oscillator from the ensemble in the polypeptide system. Therefore, a common strategy is to isotopically label the backbone of a protein with $^{13}C$ or/and $^{18}O$ and monitor the amide I mode. IR is also sensitive to some chemical changes, such as the protonation or deprotonation of side chains. This is an important example as it is often essential for protein function.

## 1.6 Site-specific spectroscopic probes

It is difficult or even impossible to rely on the intrinsic vibrational absorbances or fluorescent emissions of a protein to report information of local environment. Optical spectroscopy in conjunction with extrinsic site-specific probes, such as non-natural amino acids, can report changes in a protein environment that are otherwise difficult to attain experimentally. However, there are several criteria for an IR probe to be useful as a reporter of the local environment. The probe should be sensitive to changes in local environment and contain a simple vibrational transition that is decoupled from the rest of the molecule. Also, the absorption band should have a relatively intense and narrow absorption (high extinction coefficient) in an uncongested region of the IR spectrum. Furthermore, the probe should be stable, un-reactive, and easy to incorporate into a protein using either *in vitro* or *in vivo* methods.[104-113] Additionally, it should be relatively small in size so that it is minimally perturbing to an amino acid side-chain.

Spectroscopy can be used to monitor specific events during protein folding such as structure, conformational transitions, local hydration states, electrostatic field changes, and binding interactions. However, this is a difficult task since the conformational free energy landscape of a protein depends on a large number of degrees of freedom. Therefore, the chosen optical probe should be sensitive to local environmental changes, such as a transition from mostly polar to a mostly hydrophobic environment.[114-123] However, side chain substitutions, with moieties such as C≡N, are relatively small and can be used as IR probes. The past few years have seen a rapid increase in the employment of various side chain–based IR probes to study a wide variety of biological questions, ranging from protein folding to enzymatic reactions. For easy comparison, basic spectroscopic properties of these probes are summarized in Table 1.

Specifically, *p*-cyanophenylalanine ($Phe_{CN}$) in Figure 1.1 is a nitrile-containing non-natural amino acid and is an ideal fluorescent probe of protein folding and binding.[124-130] $Phe_{CN}$ is sensitive to its immediate environment exemplified by the large fluorescence quantum yield in water and the smaller quantum yield in acetonitrile. For example, Tang *et al.* have investigated the binding/insertion/dimerization kinetics of a transmembrane peptide using the $Phe_{CN}$ as a fluorescent reporter.[131] Furthermore, $Phe_{CN}$ is useful as an IR conformational probe because the nitrile stretching frequency is sensitive to changes in environment such as dehydration and membrane interactions.[132-142] For example, Tucker *et al.* demonstrated that the nitrile stretching band of a series of $Phe_{CN}$ labeled membrane-binding peptides report on the location and structure of the

membrane-bound helical peptides.[143] Additionally, Webb and coworkers employed $Phe_{CN}$ as a local electric field probe providing an estimate of the electric field of a bilayer by inserting a $Phe_{CN}$-labeled transmembrane helix into bicelle model membranes.[144]

The nitrile functional group is a good candidate for a conformational probe not only because it is small with intermediate polarity, thereby making it a minimally perturbing, but also because it is sensitive to local environment. This nitrile-derivatized amino acid increases the fluorescence quantum efficiency of Phe by 5 times making it a good fluorescent probe. Despite the fact that the C≡N stretching vibration has become a valuable probe in a wide range of biological applications,[81,83,145-165] it cannot easily be used to quantify the local electrostatic field of proteins in both H-bonding and non-H-bonding environments. This is due to the dependence of the C≡N stretching frequency on the angle and position of the H-bond.[166] Specifically, when the H-bond angle between the water OH and the solute CN is linear, the nitrile stretch frequency is blue-shifted. On the other hand, at other angles, a H-bond can form between water and the π-orbial of the nitrile thereby decreasing the bond strength and inducing a frequency red-shift. This is in contrast to the simple relationship exhibited between H-bonding and vibrational frequency found for the C═O stretch present in an amide, ester, and ketone. Because H-bonding is a ubiquitous electrostatic interaction in biological molecules, a simple reporter of the electrostatic field is an important advantage in the investigation of molecular association and enzymatic reactions.[66,82,167-168] The ability to quantitatively assess the

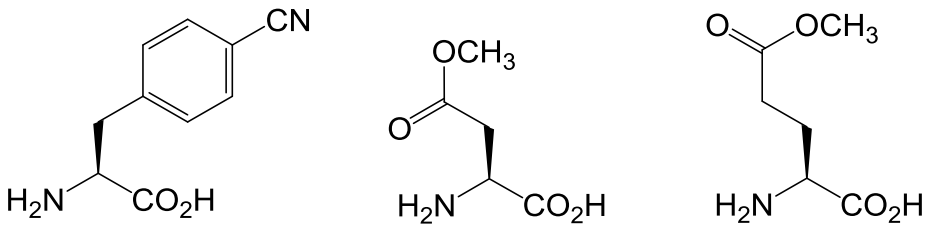
local electrostatic field and how it changes inside a protein will find great use in the study of protein function and dynamics.

A promising candidate to measure the local electrostatic field in biological molecules is a non-natural amino acid that contains a single carbonyl group (C═O) in its sidechain. A computational study from Cho *et al*. has predicted that the stretching mode of the C═O of acetone is localized and has a frequency that varies linearly with the electrostatic field for both H-bonding and non-H-bonding environments.[85] Boxer and coworkers[169] have recently shown that the C═O stretching frequency of *p*-acetyl-L-phenylalanine (p-Ac-Phe) can serve as a reporter for the local electrostatic field of proteins. However, this vibrational transition overlaps with the protein amide I band, so its application requires careful background subtraction using the wild-type protein. To circumvent this inconvenience, the C═O stretching vibration of an ester moiety can be used as an alternative. Previous studies[170-171] have shown that an ester carbonyl absorbs in a spectral region (1700-1800 $cm^{-1}$) where no protein IR bands are present at neutral pH[103] except those arising from protonated carboxylic groups.[172] As discussed later, the C═O stretching vibrations of ester functional groups correlate linearly with the local electrostatic field for both H-bonding and non-H-bonding solvents. This indicates that the ester-derivatized amino acids, in Figure 1.1, are promising candidates to measure the local electrostatic field in biological molecules.

## 1.7 Thesis Outline

The motivation of this Thesis work has been to develop new experimental tools to understand biological events of interest, such as proteins folding. As outlined below, this work applies multiple non-natural amino acids as probes to site-specifically interrogate the structure or environment of proteins using a variety of experimental methods including IR, fluorescence and two-dimensional IR (2D-IR) spectroscopies.

In the effort to expand the utility of $Phe_{CN}$ as a fluorescent marker, Chapter 2 reports the Stern-Volmer quenching constants of a series of commonly encountered anions. Also, the conformational heterogeneity of a mini-protein is shown using this method.[173] Additionally in Chapter 3, $Phe_{CN}$ was shown to conveniently serve as a pH sensor near the N-terminus of a peptide and it was used to directly measure the kinetics of a cell-penetrating peptide. In Chapter 4, urea, GdnHCl and TMAO are all shown to decrease the strength of the H-bonds formed between water and a nitrile derivatized non-natural amino acid $Phe_{CN}$, despite having different effects on protein stability.[174] The protein stabilizing mechanisms of TMAO using 2D-IR was further explored in Chapter 5.[175] In Chapter 6, ester-derivatized non-natural amino acids are introduced as a promising and versatile local electrostatic IR probe.[176] The first biological application of the ester probe to study the electric field and confined water within amyloid fibrils is described in Chapter 7. Finally, in Chapter 8 the summary and perspective of each of these projects are highlighted.




**Phe$_{CN}$** **D$_{M}$** **E$_{M}$**

Figure 1.1 Model nitrile- and ester-containing non-natural amino acids.

Table 1 Overview of site-specific infrared probes, modified from Ma *et al.*[142]

| | Name | Frequency range ($cm^{-1}$) | Extinction in $H_2O$ ($M^{-1}cm^{-1}$) | Reference |
|---|---|---|---|---|
| **Nitriles** | *p*-cyano-phenylalanine | 2,220–2,250 | ~220 | [177] |
| | Cyano-cysteine | 2,150–2,180 | ~120[a] | (148) |
| | 5-cyano-tryptophan | 2,210–2,240 | ~160[b] | [178] |
| | Cyanate | 2,220–2,300 | ~800[c] | (107) |
| **Azides** | Azidohomoalanine | 2,100–2,140 | 350–400[d] | [179] |
| | *p*-azido-phenylalanine | 2,100–2,140 | ~610 | (115) |
| | 3-picolyl azide adenine dinucleotide | 2,080–2,160 | ~2,000 | (149) |
| **Carbonyls** | Ketone carbonyl | 1,660–1,700 | ~1,800[e] | (126) |
| | Ester carbonyl | 1,690–1,770 | ~290 | (127) |
| | Carboxylic acid | 1,700–1,775 | ~280 | (138) |
| | Carboxylate | 1,555–1,600 | ~820 | (150) |
| **Metal carbonyls** | $CpRe(CO)_3$ | 2,010–2,030 | ~4,100 | (146) |
| | CORM-2 | 2,040–2,100 | NA | (143, 144) |
| **Others** | Carbon deuterium | 2,100–2,400 | 5–10 | (151) |
| | Cysteine thiol | 2,550–2,600 | 5–150 | (139) |
| | Phosphate | 1,200–1,300 | ~500 | (140) |
| | Fluorocarbon | 1,200 | ~700[e] | (141) |

Solvent conditions other than $H_2O$ are indicated: [a]50/50 (v/v) glycerol/water, [b]60/40 (v/v) water/THF, [c]THF, [d]$D_2O$, and [e]2-MeTHF. Abbreviation: NA, not applicable.

## 2 Quenching of *p*-Cyanophenylalanine Fluorescence by Various Anions

### 2.1 Abstract

To expand the spectroscopic utility of the non-natural amino acid-cyanophenylalanine ($Phe_{CN}$), we examine the quenching efficiencies of a series of commonly encountered anions toward its fluorescence. We find that iodide exhibits an unusually large Stern-Volmer quenching constant, making it a convenient choice in $Phe_{CN}$ fluorescence quenching studies. Indeed, using the villin headpiece subdomain as a testbed we demonstrate that iodide quenching of $Phe_{CN}$fluorescence offers a convenient means to reveal protein conformational heterogeneity. Furthermore, we show that the amino group of $Phe_{CN}$strongly quenches its fluorescence, suggesting that $Phe_{CN}$could be used as a local pH sensor.

### 2.2 Introduction

Because of its distinctive spectroscopic characteristics, relatively small size, and ease of incorporation, the non-natural amino acid, *p*-cyanophenylalanine ($Phe_{CN}$), has recently emerged as a useful spectroscopic probe of protein structure, dynamics and electrostatics.[138,177,180-181] In particular, several photophysical properties of $Phe_{CN}$ make it an attractive fluorescence reporter of many biological processes, such as protein folding,[182] binding,[131,155,183-185] and aggregation.[186-187] This is because (1) the fluorescence quantum yield of $Phe_{CN}$ is sensitive to its environment[182,184], such as the degree of

hydration, (2) the fluorescence lifetime of $Phe_{CN}$ correlates linearly with the hydrogen bonding ability of many protic solvents[188], making it possible to use it to quantitatively determine local hydrogen bonding strength, (3) $Phe_{CN}$ fluorescence can be selectively excited even in the presence of other naturally occurring aromatic amino acids,[184] such as tyrosine (Tyr) and tryptophan (Trp), and (4) $Phe_{CN}$ can be used as a fluorescence resonance energy transfer (FRET) donor to Trp and other natural or non-natural amino acid fluorophores,[126,178,189-192] further demonstrating its versatility as a fluorescence reporter. In principle, the application of $Phe_{CN}$ fluorescence can be further extended to structural and dynamic studies by exploiting the mechanism of dynamic quenching by, for example, a halide ion. To facilitate this potential application and to identify practically useful anionic quenchers, herein we determine the quenching efficiencies of a series of sodium salts on the fluorescence of $Phe_{CN}$, either as a free amino acid or in a peptide environment. In addition, we demonstrate, using the villin headpiece subdomain (HP35) as a model, how quenching of $Phe_{CN}$ fluorescence can be used to scrutinize the conformation heterogeneity of proteins, even under native conditions.

The technique of fluorescence quenching has long been used in protein structural studies.[193] For example, quenching of Trp fluorescence[194-195] by either iodide or acrylamide[196] has been widely employed to reveal the relative accessibility of the indole ring in the system of interest, thus providing useful structural information. While we fully anticipate that $Phe_{CN}$ can be used in similar applications, we emphasize that the current study is not aimed to show that $Phe_{CN}$ can replace Trp, but rather to demonstrate another

potential utility of this versatile spectroscopic probe. Because for amino acid fluorophores, such as Trp and Tyr, quenching of fluorescence by many anions typically occurs through the mechanism of dynamic quenching,[193] without involving static complex formation, we expect that those quenchers that have been shown to be effective in decreasing the fluorescence quantum yields of Trp and Tyr would also be effective for $Phe_{CN}$ fluorescence. Thus, in the current study we have chosen the following quenchers: NaCl, NaBr, and NaI. In addition, to determine the effect of commonly used salts in biological studies on the fluorescence quantum yield of $Phe_{CN}$, we have also included $NaH_2PO_4$, $Na_2CO_3$, $NaHCO_3$, NaSCN and $Na_2S_2O_3$. Our results indicate that among these anions, iodide exhibits the largest Stern-Volmer quenching constant, making it a convenient dynamic quencher of $Phe_{CN}$ fluorescence in protein conformational studies. Indeed, using iodide quenching of $Phe_{CN}$ fluorescence we are able to show, in agreement with previous studies,[128,136] that HP35 samples at least two conformations in its folded basin.

## 2.3 Experimental

### 2.3.1 Sample preparation

*p*-Cyanophenylalanine was purchased from Bachem Americas (Torrance, CA). The sodium salts were purchased from either Fisher Scientific or Acros Organics. They were the highest grade available and thus used without further purification. All solution

samples were prepared by directly dissolving an appropriate amount of the targeted solute in Millipore water of appropriate volume. Protein synthesis was achieved on a PS3 automated peptide synthesizer (Protein Technologies, MA) using standard Fmoc protocols.

### 2.3.2 Spectroscopic measurements

The absorption spectra were collected on a Perkin Elmer Lambda 25 UV/Vis spectrometer at room temperature. The fluorescence spectra were collected on a Jobin Yvon Horiba Fluorolog 3.10 spectrofluorometer using a 1 cm quartz cuvette at 25 °C, with a spectral resolution of 1 nm and an integration time of 1 s/nm. The excitation wavelength was 275 nm. For the free $Phe_{CN}$ quenching experiments, the concentration of the fluorophore was maintained at a constant value of 165 μM, whereas for other quenching experiments the peptide and protein concentrations were in the range of 10 – 20 μM. In addition, the Raman band of water has been subtracted from the raw fluorescence spectra and, whenever necessary, the inner-filter effect was considered and corresponding corrections were made.

### 2.3.3 Data analysis

Unless otherwise specified, all fluorescence quenching data were fit to the standard Stern-Volmer equation,[197]

$$\frac{F_0}{F} = (1 + K_{SV}[Q])e^{V[Q]} \qquad \text{Equation 2.1}$$

where $F$ and $F_0$ represent the fluorescence intensities of $Phe_{CN}$ obtained with and without the presence of the quencher Q, respectively, and [Q] is the molar concentration of the quencher. In addition, $K_{SV}$ is the Stern-Volmer constant for the dynamic quenching process and $V$ is an effective volume constant representing the effect of sphere-of-action.[90]

## 2.4 Results and discussion

### 2.4.1 Quenching of free $Phe_{CN}$ fluorescence by halide ions

Halide ions, especially iodide, are commonly used in fluorescence quenching studies of biological systems.[198] For example, iodide is an efficient quencher of Trp fluorescence, with a Stern-Volmer quenching constant ($K_{SV}$) of 11.6 $M^{-1}$ at neutral pH.[198] As shown (Figure Figure 2.1 and Figure 2.2), the fluorescence quantum yield of free $Phe_{CN}$ in aqueous solution is also sensitive to the presence of iodide: a quantitative analysis of the fluorescence quenching data yields a $K_{SV}$ of 58.6 ± 2.6 $M^{-1}$. As expected (Figure 2.2 and Table 2.1), bromide is also an efficient quencher of $Phe_{CN}$ fluorescence ($K_{SV}$ = 54.9 ± 2.4 $M^{-1}$), whereas in comparison chloride is far less effective ($K_{SV}$ = 10.6 ± 2.0 $M^{-1}$). This trend is consistent with the well documented notion that dynamic fluorescence quenching often arises from the heavy-atom effect[90], which increases the

yield or rate of the intersystem crossing process.[199] Perhaps more importantly, in comparison to quenching experiments carried out with Trp, both iodide and bromide exhibit an unusually large quenching efficiency towards $Phe_{CN}$ fluorescence, suggesting that both are excellent candidates for quenching $Phe_{CN}$ fluorescence in practice. A previous study by Serrano *et al*.[128] has shown that solute-solvent interactions have a large effect on the non-radiative decay rate of the fluorescent state of free $Phe_{CN}$, suggesting that internal conversion is the major non-radiative decaying channel. On the other hand, it has been shown that Trp has a relatively large intrinsic triplet state quantum yield (>0.1).[200] Thus, it seems reasonable to attribute the difference between the iodide quenching efficiencies of $Phe_{CN}$ fluorescence and Trp fluorescence to the difference in their intrinsic quantum yields of triplet state formation.

Taken together, the above results suggest that, due to its large $K_{SV}$ value, iodide is potentially a very useful $Phe_{CN}$ fluorescence quencher in protein conformational studies. To validate this potential utility of iodide, we further conducted fluorescence quenching experiments using a tripeptide, Gly-$Phe_{CN}$-Gly. As shown (Equation 2.1 and Figure 2.3), the $Phe_{CN}$ fluorescence quenching efficiency of iodide obtained in this case, as judged by the value of $K_{SV}$, is almost identical to that obtained with free $Phe_{CN}$, indicating that the large effect of iodide on $Phe_{CN}$ fluorescence stems from its collisions with the sidechain of $Phe_{CN}$ and therefore can be used as a sensitive and convenient $Phe_{CN}$ fluorescence quencher in protein conformational studies.

### 2.4.2 Quenching of free $Phe_{CN}$ fluorescence by other anions

In order to better determine the conditions under which $Phe_{CN}$ fluorescence can be used as an effective spectroscopic reporter, we also studied the $Phe_{CN}$ fluorescence quenching effects of the following sodium salts: $NaH_2PO_4$, NaSCN, $Na_2S_2O_3$, and $Na_2CO_3$. These salts were chosen because they are commonly used in biological studies to serve either as a buffer, denaturant, or precipitant.[201] As shown (Figure 2.4 and Table 2.1), both $SCN^-$ and $S_2O_3^{2-}$ can effectively quench the fluorescence of free $Phe_{CN}$. On the other hand, phosphate, which is commonly used in biological buffers, has little, if any, effect on the fluorescence quantum yield of $Phe_{CN}$. What is surprising, however, is that $Na_2CO_3$ seems to show an unexpectedly larger quenching efficiency towards $Phe_{CN}$ fluorescence (Figure 2.5). Since $Na_2CO_3$ is a relatively strong base ($pK_b = 4.67$), which increases the pH of the aqueous $Phe_{CN}$ solution to about 10.2 at 1 mM concentration, it is possible that the large quenching effect observed arises from deprotonation of the amine group of the amino acid, which for phenylalanine has a $pK_a$ value of 9.13. Indeed, as shown (Figure 2.5), adjusting the pH of the $Phe_{CN}$ solution using NaOH to 9.9 achieved the same quenching effect, whereas at 1 mM concentration $NaHCO_3$, which is a weaker base than $Na_2CO_3$, only showed a small degree of fluorescence quenching. Thus taken together, these results validate the above hypothesis. In addition, the results obtained with Gly-$Phe_{CN}$-Gly and $Na_2CO_3$, which demonstrate that $Na_2CO_3$ shows appreciable fluorescence quenching only at high concentrations (data not shown), also corroborate this notion. Taken together, these results suggest that the $-NH_2$ group is an efficient

quencher of the fluorescence of free $Phe_{CN}$. This finding is particularly interesting as it suggests that $Phe_{CN}$ fluorescence could be potentially useful as a pH sensor, for example, by appending a $Phe_{CN}$ residue to the N-terminus of a peptide or protein.

### 2.4.3 Application to HP35

To demonstrate the utility of quenching $Phe_{CN}$ fluorescence by iodide, we used it to characterize the native state conformational heterogeneity of HP35. As shown (Figure 2.6), the NMR structure of HP35 indicates that it has a well organized hydrophobic cluster with Phe58 being situated inside the resultant hydrophobic core. However, two recent studies, one using two-dimensional infrared (2D IR) spectroscopy[136] and another employing time-resolved fluorescence energy transfer (FRET)[128], provided strong evidence suggesting that even in the native state of HP35 Phe58 can sample two different environments, one of which is accessible by water molecules. Thus, HP35 constitutes a good model system to test the feasibility of using iodide quenching of $Phe_{CN}$ fluorescence in protein conformational studies. To do so, we mutated the Phe58 residue to $Phe_{CN}$ and also Trp64 to alanine (the resultant doubled mutant is referred to as HP35-$Phe_{CN}$-WA) because Trp is known to quench $Phe_{CN}$ fluorescence through the Förster mechanism.[183] Since this double mutation has been shown to decrease the stability of the native fold,[128] we conducted the fluorescence quenching experiments in aqueous solutions that contain 20% (v/v) trifluoroethanol (TFE). As shown (Figure 2.6), it is clear that the fluorescence quenching data can not be fit by the standard Stern-Volmer equation, which suggests that the $Phe_{CN}$ residue samples two or more inequivalent environments. Indeed, the iodide

quenching data can be well described by the following model, which assumes, based on previous studies,[128] that HP35-Phe$_{CN}$-WA can sample two different conformations in the native basin:

$$\frac{F_0}{F} = \frac{1}{\frac{\alpha}{(1+K_{SV1}[Q])e^{V_1[Q]}} + \frac{1-\alpha}{(1+K_{SV2}[Q])e^{V_2[Q]}}} \qquad \text{Equation 2.2}$$

where $K_{SV1}$ (and $V_1$) and $K_{SV2}$ (and $V_2$) represent the corresponding Stern-Volmer quenching parameters for conformations 1 and 2, respectively, and α is the fractional fluorescence contribution of conformation 1. As shown (Figure 2.6), by fixing the $K_{SV1}$ value to the Stern-Volmer quenching constant of iodide obtained with the Gly-Phe$_{CN}$-Gly peptide, we were able to quantitatively fit the fluorescence quenching data of HP35-Phe$_{CN}$-AW to Equation 2.2. The resultant fitting parameters are consistent with previous studies.[128,136] For example, the Stern-Volmer constant for the second conformation is about 0.001 $M^{-1}$, indicating that the Phe$_{CN}$ residue in this conformation is far less solvent-accessible, in comparison to that in the first conformation. In addition, the value of α (0.51) is also in qualitative agreement with the study of Urbanek*et al.*[136], who have estimated, based on parameters obtained by analyzing both the linear and two dimensional infrared spectra, that the population of the solvent exposed Phe$_{CN}$ conformation is 56%. The difference between these two studies most likely arises from one or all of the following factors: (1) in the current case the parameter α does not directly report on the molar fraction of the population of the Phe$_{CN}$ exposed conformation as the fluorescence quantum yield of Phe$_{CN}$ is known to depend on hydration, (2) the

addition of 20% TFE changes the equilibrium constant of these two protein conformations, and (3) the two mutations in HP35-$Phe_{CN}$-WA alter the native population distribution. Regardless of the origin of this difference, our results nevertheless confirm that iodide quenching of $Phe_{CN}$ fluorescence is a useful tool in the study of protein conformational heterogeneity.

## 2.5 Conclusion

To further expand the spectroscopic utility of the non-natural amino acid $Phe_{CN}$, herein we investigate the quenching efficiencies of a series of anions to its fluorescence. Our results show that iodide has an unusually large Stern-Volmer quenching constant (i.e., $K_{SV}$ = 61.8 ± 3.0 $M^{-1}$) and thus can be used as an efficient quenching agent in conjunction with $Phe_{CN}$ fluorescence in protein conformational studies. We validate the applicability of this approach by applying it to reveal the native conformational heterogeneity of the villin headpiece subdomain (HP35). Furthermore, we demonstrate that the neutral amino group, but not the positive ammonium group, of the $Phe_{CN}$ amino acid strongly quenches its fluorescence, suggesting that a $Phe_{CN}$ residue located at the N-terminus of a polypeptide sequence could be used as a local pH sensor.

## 2.6 Acknowledgments

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978) and a National Science Foundation Graduate Research Fellowship to IMP (DGE-0822). We also thank Dr. Arnaldo L. Serrano for synthesizing the HP35 mutant.

## 2.7 Original Publication

This Chapter has been reprinted from Chemical Physics Letters, Ileana M. Pazos, Rachel M. Roesch, and Feng Gai, (2013) 563, 93–96. DOI: 10.1016/j.cplett.2013.02.015, with permission from Elsevier.

Table2.1 $Phe_{CN}$ fluorescence quenching constants for various salts

| **Salt** | **$K_{SV}$ ($M^{-1}$)** | **V ($M^{-1}$)** |
|---|---|---|
| NaI | 58.6 ± 2.6 | 4.6 ± 0.5 |
| NaBr | 54.9 ± 2.4 | 1.4 ± 0.5 |
| NaCl | 10.6 ± 2.0 | 0.0 |
| $Na_2S_2O_3$ | 39.6 ± 4.0 | 10.0 ± 2.0 |
| NaSCN | 35.4 ± 3.0 | 2.7 ± 0.5 |

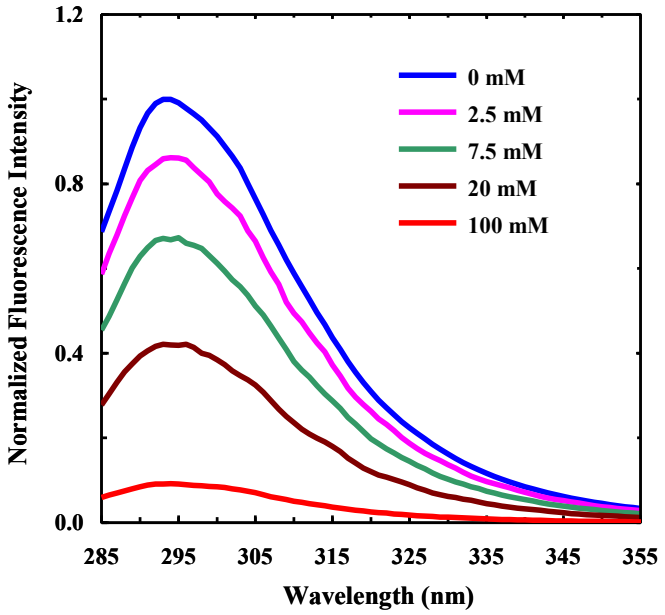



Figure 2.1 Normalized fluorescence spectra of free $Phe_{CN}$ vs [NaI], as indicated.

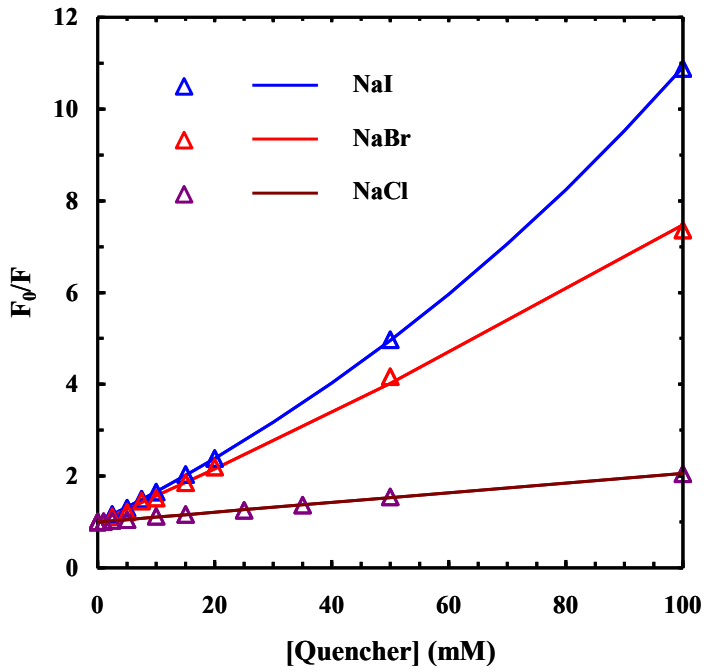



Figure 2.2 Relative fluorescence intensity of free $Phe_{CN}$ versus concentration of different halide quenchers, as indicated. The smooth lines are fits of these data to the Stern-Volmer equation and the resulting parameters are listed in Table 2.1.

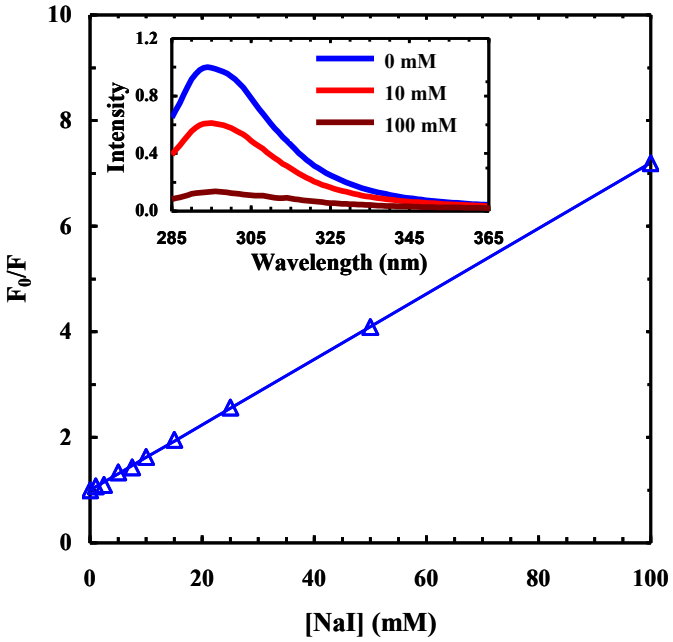



Figure 2.3 Relative fluorescence intensity of Gly-Phe$_{CN}$-Gly versus [NaI]. The smooth line is the fit of these data to the Stern-Volmer equation and the resulting parameters are $K_{SV}$=61.8 ± 3.0 $M^{-1}$ and V = 0.024 ± 0.010 $M^{-1}$. Shown in the inset are representative normalized fluorescence spectra of this tripeptide collected at different NaI concentrations, as indicated.

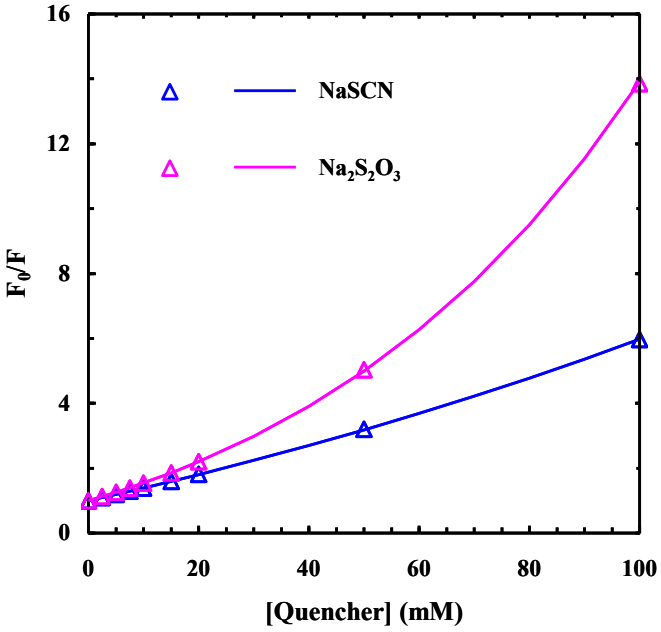



Figure 2.4 Relative fluorescence intensity of free $Phe_{CN}$ versus concentration of different quenchers, as indicated.

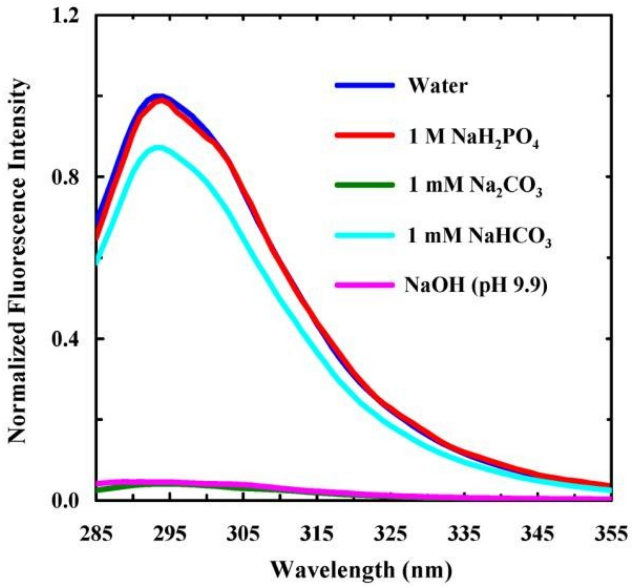



Figure 2.5 Spectra of free $Phe_{CN}$ collected with and without the presence of $NaH_2PO_4$, $Na_2CO_3$, $NaHCO_3$, or NaOH, as indicated.

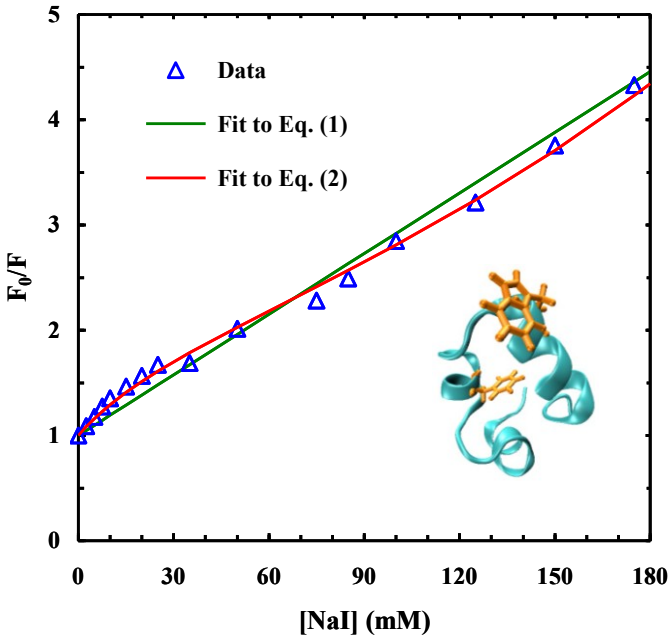



Figure 2.6 Relative fluorescence intensity of $Phe_{CN}$ in HP35-$Phe_{CN}$-WA versus [NaI]. The smooth lines are fits of these data to Equation 2.1 and Equation 2.2, as indicated. The fitting parameters obtained with Equation 2.2 are $K_{SV1}$= 61.8 ± 3.0 $M^{-1}$, $V_1$ = 3.2 ± 3.0 $M^{-1}$, $K_{SV2}$ = 0.001 ± 0.001 $M^{-1}$, $V_2$ = 4.8 ± 3.0 $M^{-1}$, and = 0.51 ± 0.10. Shown in the inset is the NMR structure of HP35 with the native Phe and Trp sidechains highlighted.

# 3 Sensing pH via *p*-Cyanophenylalanine Fluorescence: Application to Determine Peptide N-terminal p$K_a$ and Membrane-Penetration Kinetics

## 3.1 Abstract

We expand the spectroscopic utility of a well-known infrared and fluorescence probe, *p*-cyanophenylalanine, by showing that it can also serve as a pH sensor. This new application is based on the notion that the fluorescence quantum yield of this unnatural amino acid, when placed at or near the N-terminal end of a polypeptide, depends on the protonation status of the N-terminal amino group of the peptide. Using this pH sensor, we are able to determine the N-terminal p$K_a$ values of nine tripeptides and also the membrane penetration kinetics of a cell-penetrating peptide. Taken together, these examples demonstrate the applicability of using this unnatural amino acid fluorophore to study pH-dependent biological processes or events that accompany a pH change.

## 3.2 Introduction

Recently, *p*-cyanophenylalanine ($Phe_{CN}$) has emerged as a convenient and versatile site-specific fluorescence reporter for various biochemical and biophysical studies[181]. The broad spectroscopic utility of this unnatural amino acid, which is a structural derivative of tyrosine (Tyr) or phenylalanine (Phe),[202] stems from the fact that its fluorescence quantum yield and lifetime are sensitive to environment,[188] as well as its easy incorporation into peptides and proteins.[203] For example, dehydration leads to a

significant decrease in the fluorescence intensity of $Phe_{CN}$, thus making it a useful probe of processes that involve exclusion of water, such as protein folding,[138,182] binding,[131,184] aggregation,[186,204-205] and interaction with membranes.[155,206-208] In addition, the fluorescence quantum yield of $Phe_{CN}$ can be modulated by various metal ions [173] and several amino acid sidechains.[183,189-192] In particular, its fluorescence can be quenched by a nearby tryptophan (Trp) residue via the mechanism of fluorescence resonance energy transfer (FRET).[128,209-210] Since this FRET pair has a Förster radius of approximately 16 Å, it has become a very useful tool to probe protein conformational changes over a relatively short distance.[211] Furthermore, Raleigh and coworkers have shown that the fluorescence intensity of $Phe_{CN}$, when placed at or near the N-terminus of a peptide, depends on the protonation status of the N-terminal amine group of the peptide.[187] However, the potential utility of this pH dependence of the $Phe_{CN}$ fluorescence has not been demonstrated. Herein, we show that this unnatural amino acid can be used as a pH sensor to study pH-dependent biological processes or events that accompany a pH change. Specifically, we carry out two experiments to demonstrate the utility of $Phe_{CN}$ fluorescence as a pH sensor; in the first one, we use it to determine the N-terminal p$K_a$ values of a series of short peptides and, in the second one, we employ it to characterize the membrane penetration kinetics of a cell-penetrating peptide, the trans-activator of transcription (TAT) peptide derived from HIV-1.[212-215]

Several methods, including high voltage electrophoresis,[216] circular dichroism (CD)[217-220] and nuclear magnetic resonance (NMR) spectroscopy,[221-222] have been used to

determine the N-terminal p$K_a$ values of unstructured or α-helical peptides.[223] However, due to various limitations, for instance the CD method relies on the peptide of interest to form a well-folded structure such as an α-helix, applying these methods in practice is not always feasible, straightforward or convenient. Since fluorescence measurements are easier and also accessible to most, if not all, biochemical and biophysical researchers, it would be advantageous to devise a method that allows determination of N-terminal p$K_a$ values of peptides and proteins via fluorescence spectroscopy.

Cell-penetrating peptides (CPPs) are short cationic peptides which can spontaneously translocate across cell membranes and, as such, are ideal vehicles to deliver exogenous cargos into cells.[224] In spite of many previous efforts, however, several aspects of the translocation actions of CPPs are not well understood or characterized. For example, measurements of the intrinsic membrane penetration kinetics of CPPs are often done by measuring the fluorescence signal of a dye molecule attached to the CPP of interest. However, the fluorescent dyes used in this type of studies are typically large in size and, consequently, may have a significant effect on the CPP's penetration rate. Here we show, using TAT as an example, that this concern can be alleviated by using $Phe_{CN}$ fluorescence as a probe to follow the kinetics of CPP membrane penetration. Specifically, we replaced the N-terminal Tyr residue of TAT with $Phe_{CN}$, (the resultant peptide is hereinafter referred to as $F_{CN}$-TAT), which is expected to cause only a minimum perturbation to the peptide as Tyr and $Phe_{CN}$ are similar in size, and we exploited the pH-dependence of the $Phe_{CN}$ fluorescence to quantitatively assess the penetration kinetics of

TAT across a model membrane.[225] Interestingly, we find that under our experimental conditions TAT first translocates across the membrane on a timescale of minutes and then causes membrane leakage on a timescale of hours.

## 3.3 Experimental

### 3.3.1 Peptide synthesis and sample preparation

Peptides were synthesized on a PS3 automated peptide synthesizer (Protein Technologies, MA) using standard 9-fluorenylmethoxy-carbonyl (Fmoc) solid phase synthesis protocols and Fmoc-protected amino acids from either Bachem Americas (Torrence, CA) or AnaSpec (Fremont, CA). Before use, all peptide samples were further purified by reverse-phase high-performance liquid chromatography and verified by mass spectrometry. All peptide samples used in the pH titration measurements were prepared by dissolving lyophilized peptides into either acidic (25 mM $H_3PO_4$) or basic (25 mM NaOH) Millipore water, and the final concentration of each sample (20 μM) was determined optically using the absorbance of $Phe_{CN}$ at 280 nm and a molar extinction coefficient of 850 $M^{-1}$ $cm^{-1}$ [184].

### 3.3.2 Fluorescence measurements and pH titration

Fluorescence spectra of the $Phe_{CN}$-containing peptides were measured at 25 $^{o}$C on a Fluorolog 3.10 spectrofluorometer (Horiba, NJ) using a 1 cm quartz cuvette, a spectral

resolution of 1 nm, an excitation wavelength of 275 nm, and an integration time 1 s/nm. During a specific pH titration experiment, the concentration of the peptide under consideration was maintained at 20 μM and the pH of the solution was varied between 2 and 12. This was achieved by mixing two 20 μM peptide stock solutions, one prepared in 25 mM $H_3PO_4$ and the other in 25 mM NaOH aqueous solution, in a cuvette at an appropriate volume ratio to achieve the desired pH, which was further measured using a pH meter (Orion/ThermoScientific, MA) Additionally, in each case a background spectrum, obtained with the buffer solution, was subtracted. Because the shape of the $Phe_{CN}$ fluorescence spectrum is insensitive to environment, only the peak intensity was used to generate the pH titration curves. To determine the N-terminal p$K_a$ of each peptide, the corresponding fluorescence titration curve was fit to the following equation:

$$F(x) = \frac{(b_1 + m_1 x)(10^{-x}) + (b_2 + m_2 x)(10^{-pK_a})}{10^{-pK_a} + 10^{-x}}, \qquad \text{Equation 3.1}$$

where $F(x)$ is the fluorescence intensity at pH $x$, $b_i$ and $m_i$ are the intercept and slope of the linear base lines (i = 1 for acidic and i = 2 for basic).

### 3.3.3 Measurement of the membrane penetration kinetics of $F_{CN}$-TAT

The 100-nm, large unilamellar vesicles (LUVs) used in the peptide penetration experiments were composed of either 100% DOPG or a mixture of DOPC, DOPG and cholesterol (3:1:1) and prepared following previously published procedures [226] from stock solutions of the respective lipids (purchased from Avanti Polar Lipids Inc.,

AL). Briefly, the respective lipid solution in chloroform was first dried under a flow of nitrogen, allowing a lipid film to form, which was followed by a 30-minute lyophilization to remove any remaining solvent. The resultant lipid film was then rehydrated with Millipore water and the pH of the sample was adjusted to 3.5 - 4.0 using $H_3PO_4$ and NaOH. This sample was then subjected to seven rounds of slow vortexing, freezing and thawing to form LUVs. The resulting vesicle solution was then extruded 11 times through an extruder (Avanti Polar Lipids Inc., AL) equipped with a 100 nm membrane. After extrusion, the LUV solution was diluted to 100 μM (lipid concentration) with Millipore water and the pH was adjusted between 10.0 - 10.5 with $H_3PO_4$ and NaOH. This process resulted in LUV's with acidic pH inside the vesicle and basic pH outside. All LUV solutions were stored at 4 °C and used within one week after preparation.

The membrane penetration kinetics of $F_{CN}$-TAT was initiated by manually adding an appropriate aliquot of a concentrated peptide solution (350 μM, pH 7.0) to a 2.0 mL LUV solution described above. The final peptide concentration was 1.0 μM, resulting in a 1:100 peptide to lipid ratio. The time dependent fluorescence intensity of $F_{CN}$-TAT at 298 nm was collected at 25 °C using the time based acquisition function of the Fluorolog 3.10 spectrofluorometer with time intervals of 2 seconds for the first 1,000 seconds, then every 10 seconds for the next 10,000 seconds and then 100 seconds for the remainder of the experiment. These data were then binned and the initial kinetic phase corresponding to the membrane penetration process was fit to a single-exponential function. To remove the contribution of water's Raman scattering to the fluorescence signal, most of the kinetic

traces were collected using an excitation wavelength of 240 nm. In other cases, an excitation wavelength of 275 nm was used. The pH of the solution was recorded throughout the duration of the experiment.

The fluorescence quenching assay used to estimate the timescale of membrane leakage was modified from a liposome flux assay [227]. Briefly, the LUVs were prepared using the abovementioned procedures, with the exception that a membrane impermeable fluorescent dye, pyranine (Alfa Aaser, MA), was added to the lipid solution. The concentration of pyranine was 4 μM in phosphate buffer (pH 7.5, 30 mM sodium phosphate and 100 mM NaCl). After extrusion of the lipid solution, the dye molecules that were not encapsulated inside LUVs were separated from the LUVs by passing the mixture through a pD-10 column (GE Healthcare, NJ). The dye-encapsulated LUV solution was then diluted into the same phosphate buffer that also contains 25 μM *p*-xylene-bis-pyridinium bromide (DPX) (Sigma-Aldrich, MO), a pyranine fluorescence quencher, to yield a final lipid concentration of 100 μM. The pyranine fluorescence was excited at 417 nm (isosbestic point) and monitored at 515 nm with and without the presence of $F_{CN}$-TAT peptide.

## 3.4 Results and Discussion

### 3.4.1 Effect of the N-terminal amine group on $Phe_{CN}$ fluorescence

Previous studies has suggested that the fluorescence of a $Phe_{CN}$ residue, at or near the N-terminus of a peptide, can be significantly quenched by the deprotonated neutral form of the N-terminal amino group.[173,187] To further validate this notion, we measured the $Phe_{CN}$ fluorescence spectra of a tripeptide consisting of glycine (G) and $Phe_{CN}$ ($F_{CN}$) with a sequence of $GF_{CN}G\text{-}CONH_2$, having either a free amino N-terminal end (the corresponding peptide is hereafter referred to as $GF_{CN}G$) or an acetylated N-terminus (the corresponding peptide is hereafter referred to as $*GF_{CN}G$). As shown (Figure 3.1), the $Phe_{CN}$ fluorescence intensity of $GF_{CN}G$ shows a drastic decrease when the pH of the peptide solution is increased from 2.1 to 10.0. As the N-terminal p$K_a$ of peptides is typically smaller than 9.0 and, thus, at pH 10.0 the N-terminus of $GF_{CN}G$ is expected to be deprotonated, this result suggests that the N-terminal amino group is an effective quencher of the $Phe_{CN}$ fluorescence in this case. In support of this conclusion, the $Phe_{CN}$ fluorescence of $*GF_{CN}G$ does not show such a pH dependence within the pH range studied (Figure 3.1 inset). Taken together, these results indicate that a $Phe_{CN}$ residue, when placed at the second position in a peptide sequence, can be used as a pH sensor.

To demonstrate its application in this regard, we first used $Phe_{CN}$ fluorescence to determine the N-terminal p$K_a$ of $GF_{CN}G$. As shown (Figure 3.2), the intensity of $Phe_{CN}$ fluorescence of $GF_{CN}G$ displays a sigmoidal dependence on pH, characteristic of an acid-base titration. As expected, the peptide $*GF_{CN}G$, whose N-terminus is acetylated, does

not show such a transition. Further fitting the fluorescence titration curve of $GF_{CN}G$ to a two-state model (i.e., Equation 3.1) yielded a p$K_a$ of 7.99, which is similar to the N-terminal p$K_a$ (7.85) of the amino group of an antimicrobial peptide[221], melittin, which has a Gly residue at the N-terminus. Interestingly, as indicated (Figure 3.2), swapping of the first two amino acids in $GF_{CN}G$ (the resulting peptide is hereafter referred to as $F_{CN}GG$) leads to a decrease in not only the N-terminal p$K_a$ value (i.e., 6.97) but also the quantum yield of $Phe_{CN}$ fluorescence. The latter is suggestive of a fluorescence quenching mechanism that is electron transfer in nature.

### 3.4.2 Determining N-terminal $pK_a$ values of a series of tripeptides

The results obtained with $GF_{CN}G$ and $F_{CN}GG$ indicate that the dynamic range of the $Phe_{CN}$ as a pH sensor can be tuned by changing the identity of the first amino acid. To this end, we carried out acid-base titrations on eight additional peptides having the following sequence: $NH_2$-($XF_{CN}G$)-$CONH_2$, where X represents a different amino acid in each case. Specifically, we chose several amino acids with varying sidechain properties, including positively charged (K), negatively charged (D), polar (N, S, G), or nonpolar (G, M, P, V, A) sidechains, to demonstrate the sensitivity of the fluorescence of $Phe_{CN}$. As shown (Figure 3.3), the corresponding pH titration curves of these peptides measured via $Phe_{CN}$ fluorescence intensity have similar sigmoidal shapes but different mid-points (i.e., p$K_a$ values). Indeed, the N-terminal p$K_a$ values of these peptides, determined by fitting the respective acid-base titrations curves to Equation 3.1, are spread over two units of pH, ranging from 6.7 to 8.6 (Table 3.1). In particular, the N-terminal p$K_a$ (7.99) of the $AF_{CN}G$

tripeptide agrees well with that (8.0) determined for two alanine-based pentapeptides ($NH_2$-AAEAA-Ac and $NH_2$-AAHAA-Ac.[228] Because of the scarcity of N-terminal p$K_a$ values of short peptides, however, making more comparisons with other studies is not possible. On the other hand, it is clear that the N-terminal p$K_a$ value of the $XF_{CN}G$ peptide is, in most cases, smaller than that determined for an -helical peptide with the same N-terminal residue (Table 3.1), suggesting that the peptide structure has a significant effect on the electronic property of the N-terminal amino group. In addition, as indicated (Figure 3.4), except for the proline (P) and serine (S) variants, the N-terminal p$K_a$ values of other peptides show a weak linear correlation with those (i.e., p$K_2$) of the corresponding free amino acids. While further confirmation is needed, the large deviation observed for proline and serine likely results from their interactions with the $Phe_{CN}$ sidechain. Regardless of the exact mechanism that controls the N-terminal p$K_a$ values of these tripeptides, the data in Figure 3.3 indicates that by changing the N-terminal residue, the useful dynamic range of $Phe_{CN}$ fluorescence can be extended over five units of pH, from approximately 5 to 10.

### 3.4.3 Using $Phe_{CN}$ fluorescence to probe TAT membrane-penetrating kinetics

In this example, we seek to demonstrate that $Phe_{CN}$ can serve as a pH sensor to monitor the rate of membrane penetration of CPPs. Specifically, we employed a well-studied CPP, TAT (YGRKKRRQRRR-$CONH_2$) as our model peptide. As discussed above, the $Phe_{CN}$ pH sensor was introduced via a Tyr to $Phe_{CN}$ mutation. Previous studies have shown that the Tyr residue plays a minimal, if any, role in facilitating TAT

translocation across membranes. Thus, we expect that the mutant peptide (i.e., $F_{CN}$-TAT) will exhibit similar, if not identical, membrane penetration kinetics as the wild type peptide. As expected (Figure 3.5 inset), the $Phe_{CN}$ fluorescence intensity of $F_{CN}$-TAT depends on pH. To utilize this dependence to report the membrane penetration kinetics of $F_{CN}$-TAT, we prepared a DOPG vesicle solution where the pH inside the vesicle was 3.5 and the pH outside the vesicle was 10.0. As shown (Figure 3.5), upon mixing this vesicle solution with a $F_{CN}$-TAT solution at pH 10.0, the $Phe_{CN}$ fluorescence increases with time in two well separated kinetic phases. This increase in $Phe_{CN}$ fluorescence is consistent with the idea that a certain fraction of the peptide molecules has translocated across the DOPG membranes and, thus, are experiencing a significant decrease in pH. However, repetitive measurements indicate that while the first or fast kinetic phase is well defined and reproducible, the second phase appears to be ill-defined and occurs on a much longer timescale. These differences prompt us to assign the fast phase to peptide internalization in vesicles (or membrane penetration) and the slow one to membrane leakage. This is based on the notion that, unlike the penetration process, a number of peptide molecules need to work together, for example, through the formation of transmembrane pores [229-230], to cause the membrane to leak; as a result, the leakage kinetics are intrinsically more stochastic like and statistically more sensitive to the experimental uncertainties (e.g., variations in the peptide and vesicle concentrations). To confirm this assignment, we measured the pH of the peptide-vesicle solution at different reaction times and found that after the completion of the fast kinetics phase (e.g., at 50 minutes) the pH of the system

was practically unchanged (i.e., 10.0 to 9.9), whereas after 5 hours the pH of the solution was decreased to 7.4. On the other hand, adding the $GF_{CN}G$ tripeptide to the same DOPG vesicle solution did not cause any appreciable change in the pH value after 12 hours. Thus, taken together these results support the aforementioned assignment. Moreover, the notion that the slow kinetic phase reports on peptide induced membrane leakage is further corroborated by the data obtained from a dye leakage experiment. As shown (Figure 3.5), under similar experimental conditions (i.e., same peptide and lipid concentrations) the fluorescence intensity of a dye (pyranine) that is initially encapsulated inside the DOPG vesicle shows a significant decrease (approximately 50% without counting photobleaching) over 10 hours, due to membrane leakage and consequently fluorescence quenching by a quencher (DPX) initially existing outside the vesicles. As indicated (Figure 3.6), the fast or peptide penetration kinetics can be fit by a single-exponential function with a time constant of 10.5 ± 2.2 minutes. This value is similar to those reported in the literature for TAT,[225,231-232] providing further supporting evidence for our assignment.

Using 240 nm photons to excite the $Phe_{CN}$ fluorescence may lead to undesirable heating effect. To verify that this is not the case, we also carried out fluorescence kinetic measurements using an excitation wavelength of 275 nm. As shown (Figure 3.7), the resultant penetration kinetics has a time constant of 12.8 ± 2 minutes, indicating that the photo-excitation induced heating affects minimally, if any, the kinetic results. Similarly, to better mimic eukaryotic cell membranes, we also conducted the penetration experiment

using a ternary mixture of DOPC, DOPG, and cholesterol.[233-234] As indicated (Figure 3.7), the $F_{CN}$-TAT penetration kinetics obtained with the mixed lipid membrane are comparable to those measured with DOPG vesicle, with a time constant of 11.9 ± 1.5 minutes.

## 3.5 Conclusion

The non-natural amino acid, p-cyanophenylalanine (PheCN), has unique infrared (IR) and fluorescent properties. As such, it has been used as both IR and fluorescence probes in a wide range of applications. Here, we further show that, for peptides having a PheCN residue at or near the N-terminus, the neutral form of the N-terminal amino group is much more effective than the protonated form in quenching the PheCN fluorescence. As a result, such peptides can be used as pH sensors. By examining the pH-dependence of the PheCN fluorescence of nine tripeptides (i.e., X-PheCN-Gly, where X represents either Asn, Met, Lys, Pro, Val, Asp, Ala, Gly or Ser), we are able to show that the dynamic range of such pH sensors can be tuned to cover several units of pH as the N-terminal pKa of these peptides depends on the identity of the N-terminal amino acid. In addition, using TAT as an example, we demonstrate that the aforementioned pH-dependence of $Phe_{CN}$ fluorescence can be exploited to measure penetration kinetics of cell-penetrating peptides across model membranes.

## 3.6 Acknowledgements

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978).

## 3.7 Original Publication

At the time of the submission of this thesis, this article had been accepted for publication by Analytical Biochemistry, Ileana M. Pazos, Ismail A. Ahmed, Mariana I. León Berríos, Feng Gai (2015), Available online 29 April 2015, doi:10.1016/j.ab.2015.04.026.

Table3.1 The N-terminal p$K_a$ value of the $XF_{CN}G$ tripeptide determined from the current experiment, the p$K_2$ value of the corresponding free amino acid X[235], and the N-terminal p$K_a$ value ($pK_a^H$) of an α-helical peptide with X at the N-terminus.[218] The uncertainty of the measured p$K_a$ values is ±0.1.

| X | p$K_a$ | $pK_a^H$ | p$K_2$ |
|---|---|---|---|
| N | 6.69 | 7.07 | 8.8 |
| M | 7.04 | 7.83 | 9.2 |
| K | 7.25 | 8.12 | 9.2 |
| P | 7.28 | 8.85 | 10.6 |
| V | 7.29 | 8.14 | 9.6 |
| D | 7.70 | 8.25 | 10.0 |
| A | 7.99 | 8.35 | 9.7 |
| G | 7.99 | 8.51 | 9.8 |
| S | 8.59 | 7.63 | 9.2 |
| $F_{CN}GG$ | 6.97 | | 9.1 |

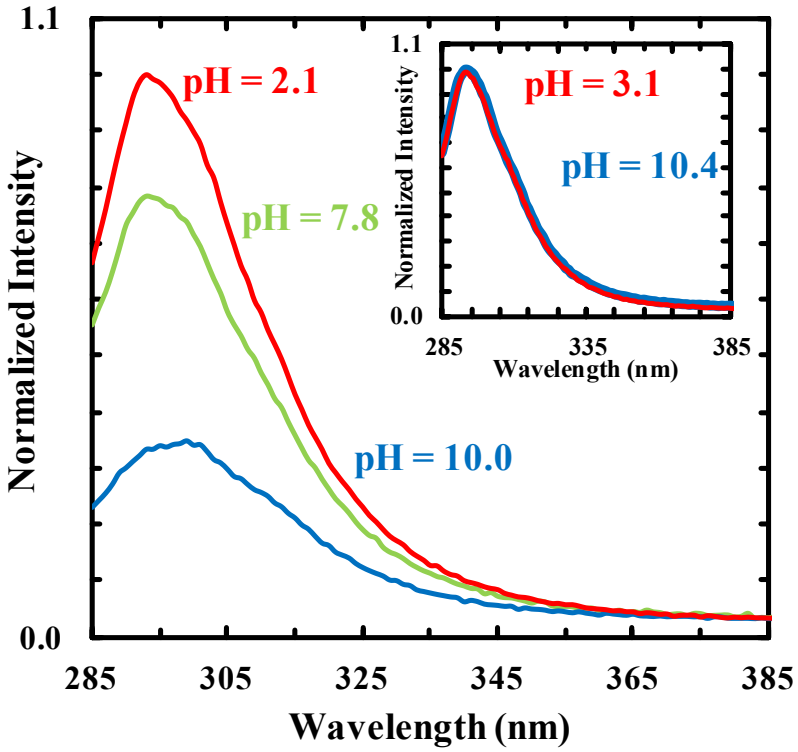



Figure 3.1 Normalized $Phe_{CN}$ fluorescence spectra of $GF_{CN}G$ obtained at different pH values, as indicated. Shown in the inset are the normalized $Phe_{CN}$ fluorescence spectra of $*GF_{CN}G$ (i.e., the N-terminal acetylated version of $GF_{CN}G$) collected under acidic and basic pH conditions.

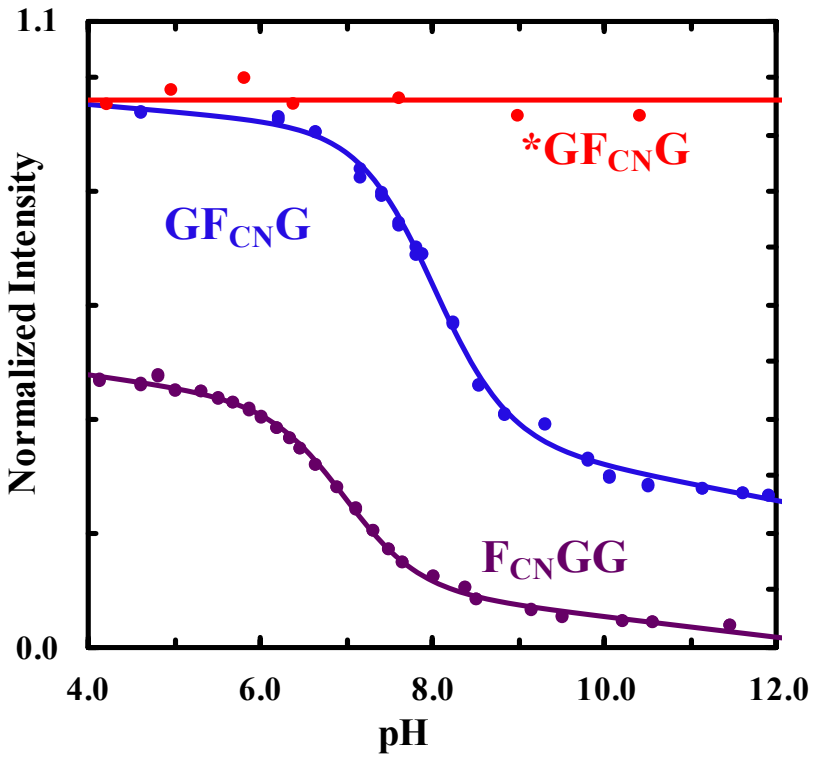



Figure 3.2 Normalized $Phe_{CN}$ fluorescence intensity versus pH of three tripeptides, as indicated. In each case, the smooth line represents the best fit of the corresponding fluorescence titration curve to Equation 3.1 and the resulting p$K_a$ value is listed in Table 3.1.

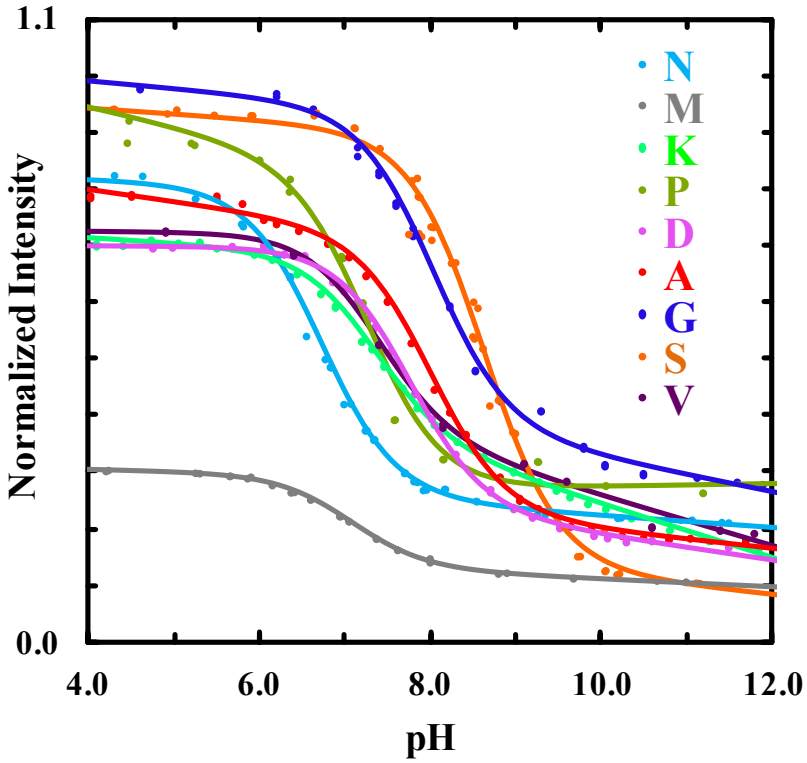



Figure 3.3 Normalized $Phe_{CN}$ fluorescence intensity versus pH of $XF_{CN}G$ peptides, as indicated by the X amino acid. Smooth lines correspond to the best fits of these fluorescence titration curves to Equation 3.1 and the resulting p$K_a$ values are given in Table 3.1. In addition, the significantly decreased fluorescence intensity of $MF_{CN}G$ is due to the previously verified quenching effect of methionine sidechain toward $Phe_{CN}$ fluorescence.[187]

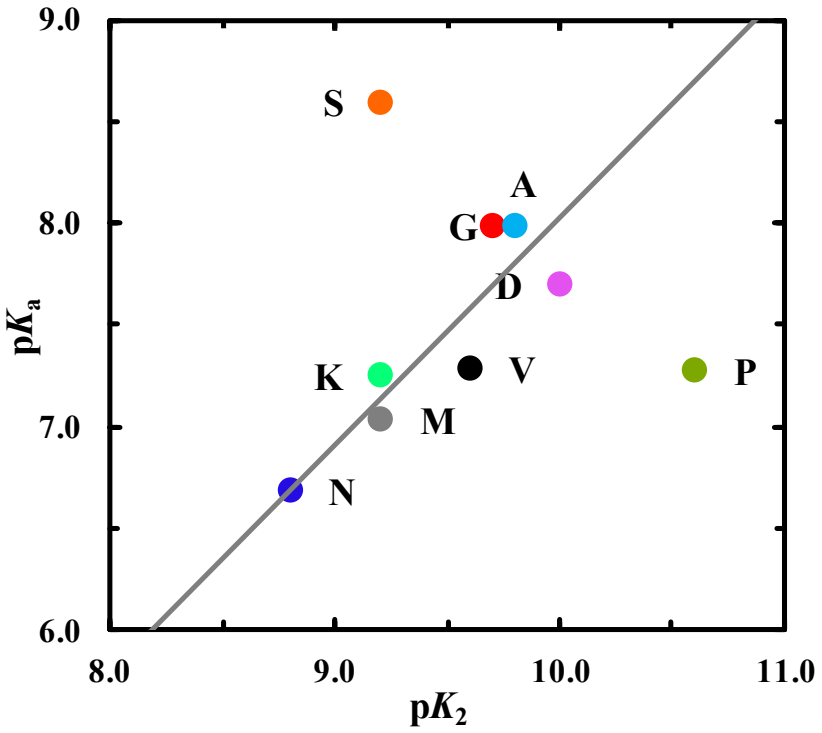



Figure 3.4 Comparison between the N-terminal p$K_a$ of $XF_{CN}G$ with the p$K_2$ of the corresponding free amino acid X. Except for proline and serine, a correlation ($R^2 = 0.98$) between the p$K_a$ and p$K_2$ values exists for the amino acids studied.

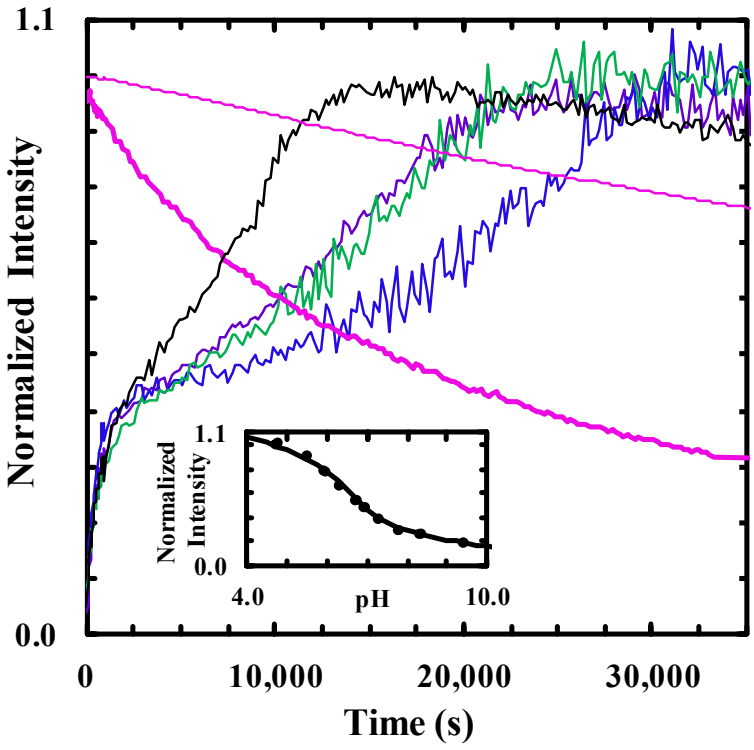



Figure 3.5 Four representative $Phe_{CN}$ fluorescence kinetic traces obtained upon mixing a $F_{CN}$-TAT solution with a DOPG LUV solution (black, purple, green and blue). The excitation wavelength was 240 nm. The pink lines are the time-dependent fluorescence intensities of the pyranine dye initially encapsulated inside the DOPG vesicles obtained in the presence (thick line) and absence (thin line) of the $F_{CN}$-TAT peptide and DPX quencher. In other words, the thin pink line measures the photobleaching rate of the pyranine dye, whereas the thick pink line reports the peptide induced membrane leakage kinetics. Shown in the inset is the normalized $Phe_{CN}$ fluorescence of $F_{CN}$-TAT as a function of pH.

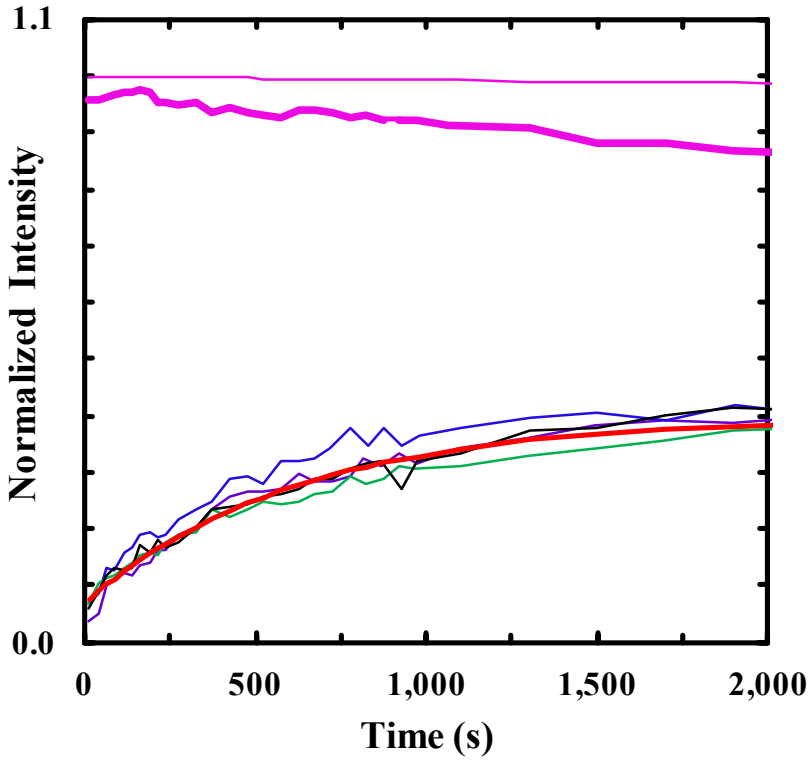



Figure 3.6 The first 2000-second portions of the same $Phe_{CN}$ fluorescence kinetic traces in Figure 3.5. The red line represents the best fit of the averaged data of these four curves to a single-exponential function with a time constant of 10.5 ± 2.2 minutes.

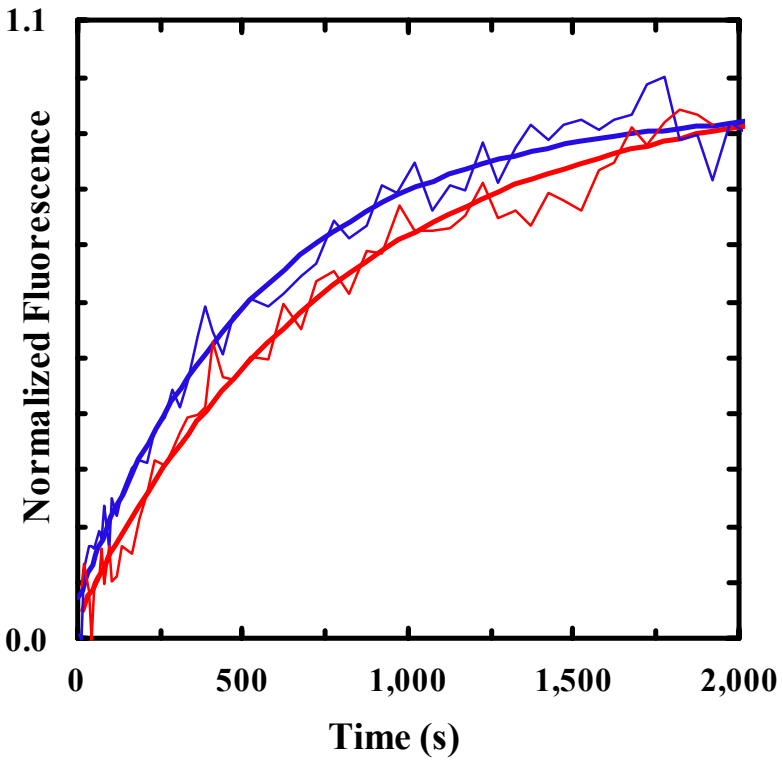



Figure 3.7 $F_{CN}$-TAT penetration kinetics. The thin red line represents the $F_{CN}$-TAT penetration kinetics obtained with DOPG LUVs and an excitation wavelength of 275 nm, and the thick and smooth red line represents the best fit of this kinetic trace to a single-exponential function with a time constant of 12.8 ± 2 minutes. The thin blue line represents the $F_{CN}$-TAT penetration kinetics obtained with the ternary lipid membrane described in the text, and the thick and smooth blue line represents the best fit of this kinetic trace to a single-exponential function with a time constant of 11.9 ± 1.5minutes.

# 4 Solute's Perspective on How Trimethylamine Oxide, Urea, and Guanidine Hydrochloride Affect Water's Hydrogen Bonding Ability

## 4.1 Abstract

While the thermodynamic effects of trimethylamine oxide (TMAO), urea, and guanidine hydrochloride (GdnHCl) on protein stability are well understood, the underlying mechanisms of action are less well characterized and, in some cases, even under debate. Herein, we employ the stretching vibration of two infrared (IR) reporters, i.e., nitrile (C≡N) and carbonyl (C═O), to directly probe how these cosolvents mediate the ability of water to form hydrogen bonds with the solute of interest, e.g., a peptide. Our results show that these three agents, despite having different effects on protein stability, all act to decrease the strength of the hydrogen bonds formed between water and the infrared probe. While the behavior of TMAO appears to be consistent with its protein-protecting ability, those of urea and GdnHCl are inconsistent with their role as protein denaturants. The latter is of particular interest as it provides strong evidence indicating that although urea and GdnHCl can perturb the hydrogen-bonding property of water their protein-denaturing ability does not arise from a simple indirect mechanism.

## 4.2 Introduction

It is well-known that both the stability and solubility of a protein in aqueous solution can be altered by the addition of certain cosolvents, such as trimethylamine

oxide (TMAO) and urea. While there is a large body of research[11-13,16-22,26-28,31,33-37,49,236-252] on how such stabilizing/destabilizing effects might take place, a coherent view of the underlying mechanism of action at the molecular level has not been reached. For example, it is still under debate whether urea and GdnHCl, both of which are commonly used as protein denaturants, affect the structure and dynamics of the solvent water molecules to such an extent that would significantly reduce the strength of stabilizing interactions in proteins.[253-254]

One of the distinctive physical properties of water is its strong ability to form hydrogen bonds (H-bonds), and these interactions are essential for protein folding. As such, many previous studies[20,40,54,255-257] have focused on investigating whether commonly encountered protectants (e.g., TMAO) and denaturants (e.g., urea and guanidine hydrochloride (GdnHCl)) affect the H-bond network of water molecules and, if so, to what extent. For example, Vanderkooi and co-workers have demonstrated, using IR spectroscopy and molecular dynamics (MD) simulations,[45] that the addition of TMAO results in a significant change in the vibrational spectrum of water, whereas the addition of urea has little effect. While such studies have provided significant insights into our understanding of how a given cosolvent interacts with water, and how such interactions might change the structure of water, they offered little, if any, direct information on how the cosolvent of interest mediates the strength of the H-bonds formed between water and a third molecular component (e.g., a protein) in solution. Because water–protein H-bonding interactions are known to play a critical role in determining not only protein

stability but also protein solubility, it would therefore be very useful if one could devise a method that allows a direct assessment of how a given cosolvent mediates the ability of water to form H-bonds with proteins. It is well-known that the vibrational frequency of nitriles is sensitive to local solvation environment[83,116,144,149,155,163,178,180,191,205,258-267] and, especially, H-bonding interactions[148,177,181] and thus is ideally suited to serve as a site-specific probe in this regard. Herein, we use this H-bonding probe to directly evaluate how the addition of TMAO, urea, or GdnHCl affects the water–solute H-bonds of interest.

## 4.3 Experimental Section

Reagent-grade urea, GdnHCl, TMAO, and NaCl were purchased from Aldrich (Saint Louis, MO) and used as received. Their concentrations were determined by the refractive index of the corresponding solution. The $Phe_{CN}$-Peptide was synthesized on a PS3 peptide synthesizer (Protein Technologies), purified by reverse-phase liquid chromatography, and verified by mass spectrometry. The peptide concentration was determined optically using the absorbance of the peptide solution at 280 nm and a molar extinction coefficient of 850 $cm^{-1}$ $M^{-1}$. All FTIR spectra at a resolution of 1 $cm^{-1}$ were collected on a Magna-IR 860 spectrometer (Nicolet, WI). The details of the setup have been described elsewhere.[268] The path length was 53 μm for all measurements, except when $^{13}C$-urea was used as a cosolvent, where a 14 μm sample cell was used.

## 4.4 Results and Discussion

Acetonitrile (ACN) is the simplest organic nitrile compound, whose stretching mode has been extensively studied and characterized in a large number of solvents.[132,269-271] It is found that the nitrile group readily forms H-bonds with protic solvents, and more importantly for the current study, its stretching frequency shows a correlation with the H-bonding ability of the solvent, with the relation that a higher frequency corresponds to a stronger C≡N–solvent H-bond.[272] For example, in water the C≡N stretching band of ACN is centered at 2259.8 $cm^{-1}$, whereas in methanol it shifts to 2259.0 $cm^{-1}$.[132] Therefore, ACN constitutes a convenient model system to probe whether the addition of a third molecular component will affect the strength of the C≡N–$H_2O$ H-bond(s).

As shown (Figure 4.1), the nitrile stretching vibrational band of ACN in water collected at room temperature clearly shows a red-shift upon addition of TMAO (4.8 M), from 2260.0 to 2258.4 $cm^{-1}$. This result provides the first direct indication that TMAO, which by itself cannot form H-bonds with the nitrile group, can effectively reduce the strength of the C≡N–$H_2O$ H-bond(s). To further confirm this finding, we also repeated the measurements at other temperatures. As indicated (Figure 4.2), the above-mentioned red-shift is maintained over the entire temperature range of the experiment. Furthermore, in both pure water and the TMAO solution, the center frequency of the nitrile stretching band decreases with increasing temperature, due presumably to thermally induced weakening of the corresponding C≡N–$H_2O$ H-bond(s).[273]

What is more interesting to note is that urea and GdnHCl also induce a similar red-shift in the nitrile stretching band of ACN (Figure 4.2). This result is rather unexpected as both cosolvents are well-known to have an opposite effect on protein solubility and stability compared to TMAO. On the other hand, addition of NaCl leads to a slight but measurable increase in the nitrile stretching frequency (Figure 4.2). Thus, taken together, these results corroborate our expectation that the nitrile stretching vibration can be used to probe cosolvent-induced changes in the local H-bond environment.

It is well recognized that the solute–solvent interaction energy is influenced by many factors and thus is difficult to determine, especially based on spectroscopic data alone.[274] On the other hand, it is possible to quantify the change in local electric field and thus the change in solvent–solute electrostatic interaction energy using experimentally determined vibrational frequency shift of an appropriate vibrator and its Stark tuning rate. However, application of the C≡N stretching vibration in this regard is not straightforward as Cho and co-workers have shown that the nitrile stretching frequency shift caused by C≡N–$H_2O$ H-bonding interactions cannot be quantitatively described by a simple linear Stark effect relationship.[84] Therefore, below, we only seek to obtain a qualitative estimate of the extent to which the ACN–water electrostatic interaction energy changes in response to addition of a given cosolvent, by assuming that the vibrational frequency shift of the nitrile moiety can be described by its Stark tuning rate. For ACN, this assumption may not be totally unreasonable as the computational results of Cho and co-workers[260]

have suggested that the electric field vector components parallel to the C≡N molecular axis are the dominant factors of the nitrile stretching frequency shift induced by C≡N–$H_2O$ H-bonding interactions. The Stark tuning rate of the nitrile stretching vibration of ACN has been determined to be 0.43 $cm^{-1}$/(MV/cm).[78,275] By assuming the permanent dipole moment of ACN (3.92 D) to be independent of solvent conditions, a simple calculation indicates that the addition of TMAO (4.8 M) results in a decrease in the ACN–water electrostatic interaction energy due to decreased H-bonding interactions by approximately 0.69 kcal/mol. While this value cannot be simply interpreted as the TMAO-induced change in the solvation energy of ACN, its trend is nevertheless consistent with the role of TMAO as a protein protectant.

The ability of urea and GdnHCl to decrease the nitrile stretching frequency of ACN seems to be inconsistent with their protein-denaturing abilities. To verify the findings obtained with ACN, we further measured the effect of the above-mentioned cosolvents on the nitrile stretching frequency of a pentapeptide that contains an non-natural amino acid, $Phe_{CN}$. Similar to that of ACN, the nitrile stretching mode of $Phe_{CN}$ has been shown to be sensitive to the local hydration status of peptides and proteins and, hence, can be used as a local H-bond reporter.[143,177,273] The primary reason that we employ a short peptide (sequence: GK-$Phe_{CN}$-TV, hereafter referred to as $Phe_{CN}$-Peptide), instead of a model protein, is to avoid the denaturant-induced conformational transitions of the latter, which could complicate data interpretation.

As shown (Figure 4.3), similar to that observed for ACN, addition of TMAO induces a red-shift in the nitrile stretching band of the $Phe_{CN}$-Peptide, in comparison with that obtained in pure water. This result further corroborates the above notion that TMAO, with its inability to directly form a H-bond with the nitrile group, must mediate the nitrile–water interaction by reducing the corresponding H-bond strength. Perhaps more important is that, consistent with the results obtained with ACN, both urea and GdnHCl induce a similar, albeit smaller, red-shift in the nitrile stretching frequency of the $Phe_{CN}$-Peptide (Figure 4.4). In comparison, addition of NaCl results in a blue-shift in the nitrile stretching frequency and hence a strengthening of the corresponding C≡N–$H_2O$ H-bond(s). While the current spectroscopic results do not allow a direct assessment of the effect of the added cosolvents on water–water H-bonding interactions, those obtained with NaCl appear to be consistent with several previous studies,[276] which demonstrated that with increasing NaCl concentration, the strength of water–water H-bonds decreases.[4,277]

Despite having distinctively different effects on protein stability, our results suggest that TMAO and urea both act to decrease the C≡N–$H_2O$ H-bond strength. To verify that this effect is not specific to the nitrile probe, we further studied the effect of TMAO and urea on the amide I′ band of *N*-methylacetamide (NMA), a compound that contains a single amide unit and has been widely used as the simplest model of a protein backbone.[278-279] The amide I band of polypeptides arises mainly from the backbone C═O stretching vibrations and is an established IR probe of protein conformations,[138] due to its

sensitivity to various structural determinants, such as H-bonding interactions. Unlike nitriles, forming a stronger H-bond with the solvent will lead to a decrease in the amide I′ vibrational frequency of NMA.

As shown (Figure 4.5), the amide I′ band of NMA obtained in TMAO solution is shifted to a higher wavenumber compared to that obtained in $D_2O$, akin to the dehydration-induced blue-shift of the amide I′ band of helical peptides.[280] Thus, this result indicates that addition of TMAO decreases the strength of the H-bond(s) formed between the carbonyl group and water molecules, in agreement with the simulation study of Gao and co-workers.[281] As indicated (Figure 4.5), $^{13}$C-urea shows a similar effect (but to a lesser extent). It should be noted that because $^{12}$C-urea absorbs strongly in the amide I′ region $^{13}$C-urea was used herein to allow the measurement of part of the amide I′ band of interest. This result is consistent with the results of Cremer and co-workers,[282] which demonstrated that urea and its methyl derivatives can cause the amide stretching vibration of poly(*N*-isopropylacrylamide) to shift to higher frequencies. Taken together, the C≡N and C═O results therefore argue strongly that the observed decrease/increase in the strength of solute–$H_2O$ H-bonds is not specific to the spectroscopic reporter used; rather, it intimately reports the perturbation of the cosolvent of interest on water's hydrogen bonding ability.

Our observation that TMAO decreases the solute–water hydrogen bonding interactions is consistent with its ability to protect proteins from unfolding. Water always seeks to form additional H-bonds, e.g., with any exposed protein amide units; hence,

decreasing the protein–water H-bonding interaction strength is expected to strengthen the intraprotein H-bonds, as observed in a recent MD simulation,[283] consequently increasing the stability of the folded state. In addition, our results are in agreement with several previous studies.[25,254,284] For example, Vanderkooi and co-workers have shown that addition of TMAO to a mixture of 5% $H_2O$ and 95% $D_2O$ leads to a significant broadening, primarily at the low-frequency side of the OH stretching band, indicative of an enhancement of the population of a more strongly H-bonded water species.[45] It is reasonable to assume that such water species would form weaker H-bonds with other solute molecules, as observed in the present case. Similarly, our TMAO results are corroborated by the molecular dynamics (MD) simulation study of Gao and co-workers,[280] which indicated that TMAO enhances the tetrahedral water structure and weakens the interactions between the amide carbonyl group and water molecules.[10] In addition, our results suggest that when considering the fact that proteins normally contain a large number of sites capable of forming hydrogen bonds with water the TMAO-induced weakening of protein–water H-bonds could play an important role in the overall stabilizing effect of TMAO, which is entirely consistent with the notion that TMAO acts to make water a poor solvent for the protein backbone.[40,62,285]

Many studies have shown that urea and GdnHCl are only weakly hydrated and, consequently, have little effect on the water–water H-bonding network.[143,286-287] Thus, it is surprising that our results indicate that these cosolvents not only have a significant effect on the C≡N–$H_2O$ H-bond(s) but also act to reduce the strength of such H-bonds. In

addition, the data obtained with NMA (Figure 4.5) suggest that urea could also decrease the interaction between water and protein backbone carbonyls, though probably to a lesser extent compared to TMAO. Considering the fact that both urea and GdnHCl are strong protein denaturants, these results are rather intriguing. In other words, based on these results alone one might, for example, predict that urea is a protecting osmolyte, instead of a denaturing one. Thus, these results provide strong evidence in support of the notion that urea and GdnHCl destabilize proteins via specific binding interactions[36,287] and, in particular, the conclusion reached by Zhou and co-workers that the electrostatic (hydrogen-bonded) interactions only play a relatively minor (even negative) role in urea-induced protein denaturation.[252] Furthermore, our results are consistent with the viewpoint that for a given denaturing agent a direct correlation may not exist between its denaturation efficiency and its ability to modify water structure.[276,288]

Since both urea and guanidinium are hydrogen bond donors, it is possible that the observed vibrational frequency shifts do not report a change in the corresponding solute–water H-bonding environment. Instead, they correspond to the formation of a new type of H-bond, i.e., that formed between the nitrile moiety and denaturant. While we cannot entirely rule out this possibility, we believe this is an unlikely scenario, at least in the case of the $Phe_{CN}$-Peptide. This is because ample evidence in the literature suggests that urea and guanidinium show preferential binding to hydrophobic side chains ($Phe_{CN}$ in the current case), in a parallel or side-to-side fashion,[286] which is incompatible with a hydrogen-bonded configuration in which the denaturant and $Phe_{CN}$ are aligned in a head-

to-tail manner. In addition, if the denaturants participate directly to form either new or additional H-bonds, the nitrile stretching frequency is expected to increase, which is not observed.

In light of the current findings, it would be quite useful to design new experiments that would allow a more direct visualization of the specific interactions between denaturants and proteins. In principle, such interactions could be studied in great detail by multidimensional infrared spectroscopy, for example, by examining the vibrational couplings between two well-chosen vibrators, one on the peptide (such as a nitrile group on the side chain or an amide carbonyl on the backbone) and another on the cosolvent (such as urea's carbonyl group).

## 4.5 Conclusions

In summary, we have investigated how TMAO, urea, and GdnHCl affect the hydrogen bonding interactions between a solute molecule of interest and water via the vibrational frequency shift of a nitrile IR probe, aiming to shed new insight into the protecting or denaturing mechanism of these cosolvents. Our results demonstrated that all three cosolvents act to decrease the strength of the hydrogen bonds formed between water and the nitrile moiety. We believe that this is an important finding as it suggests that (1) both urea and GdnHCl are capable of perturbing the structure of water, at least the water–water interactions in the hydration shell of the solute, (2) simply knowing how a given

cosolvent affects the structure and dynamics of water is probably insufficient to allow an accurate prediction of its protecting or denaturing efficiency, and (3) more specifically, urea and GdnHCl denature proteins through specific interactions with the targets.

## 4.6 Acknowledgments

We gratefully acknowledge financial support from the National Institutes of Health (GM-065978) and a National Science Foundation Graduate Research Fellowship to IMP (DGE-0822).

## 4.7 Original Publication

This Chapter has been reprinted from Journal of Physical Chemistry B, Ileana M. Pazos, and Feng Gai, (2012) 116, 41 12473–12478. DOI: 10.1021/jp307414s, with permission from American Chemical Society.

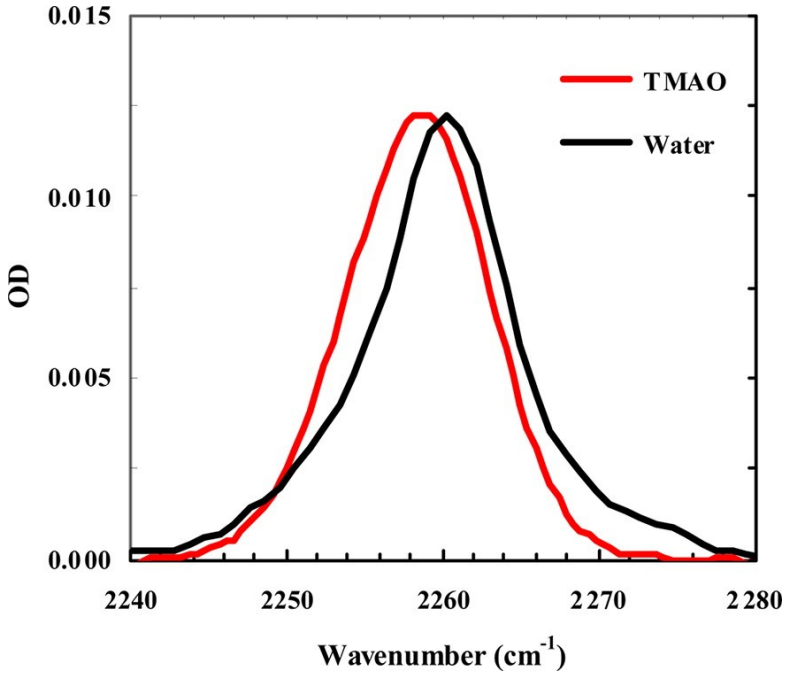



Figure 4.1 C≡N stretching vibrational bands of ACN (~10 mM) in water and 4.8 M TMAO aqueous solution, as indicated. These data were collected at 25.9 and 25.1 °C, respectively. For easy comparison, the spectrum obtained in TMAO solution has been scaled.

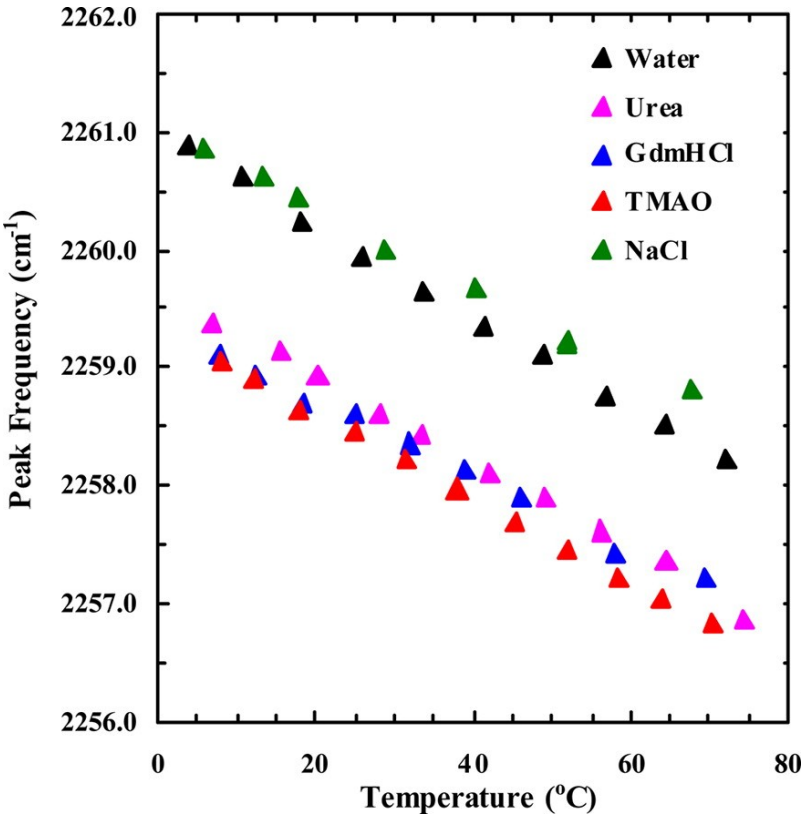



Figure 4.2 Temperature dependence of the peak frequency of the C≡N stretching vibrational band of ACN (~50 mM) obtained under different solvent conditions, as indicated. The concentrations of the cosolvent or cosolute in each case were: [TMAO] = 4.8 M, [Urea] = 5.2 M, [GdnHCl] = 5.1 M, and [NaCl] = 5.0 M.

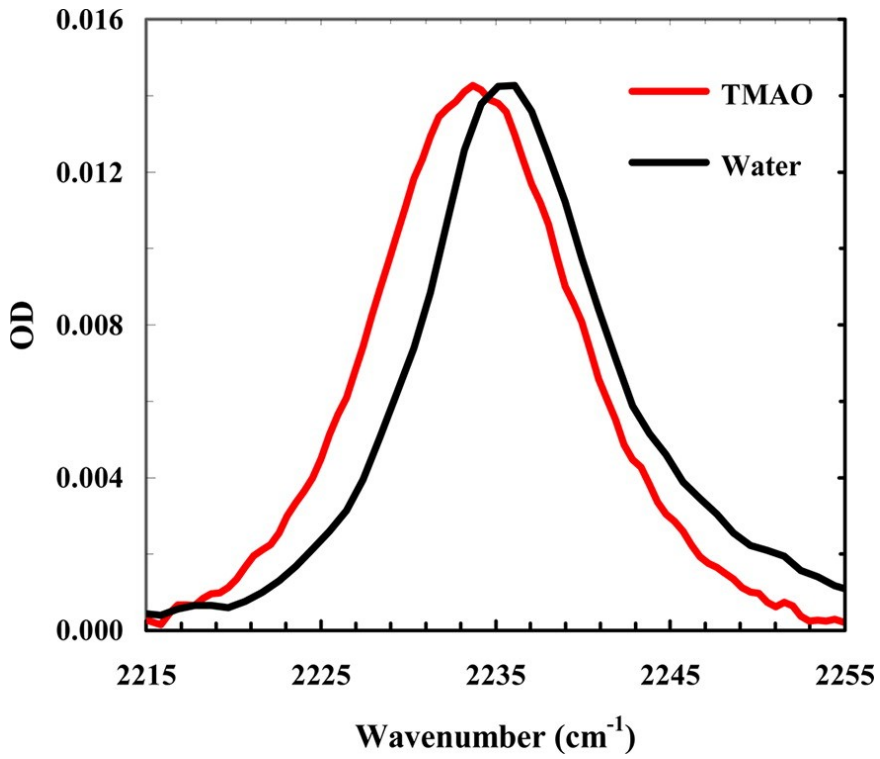



Figure 4.3 C≡N stretching vibrational bands of the $Phe_{CN}$ peptide (~15 mM) in water and 4.8 M TMAO aqueous solution, as indicated. These data were collected at 27.7 and 25.2 °C, respectively. For easy comparison, the spectrum obtained in TMAO solution has been normalized to that obtained in water.

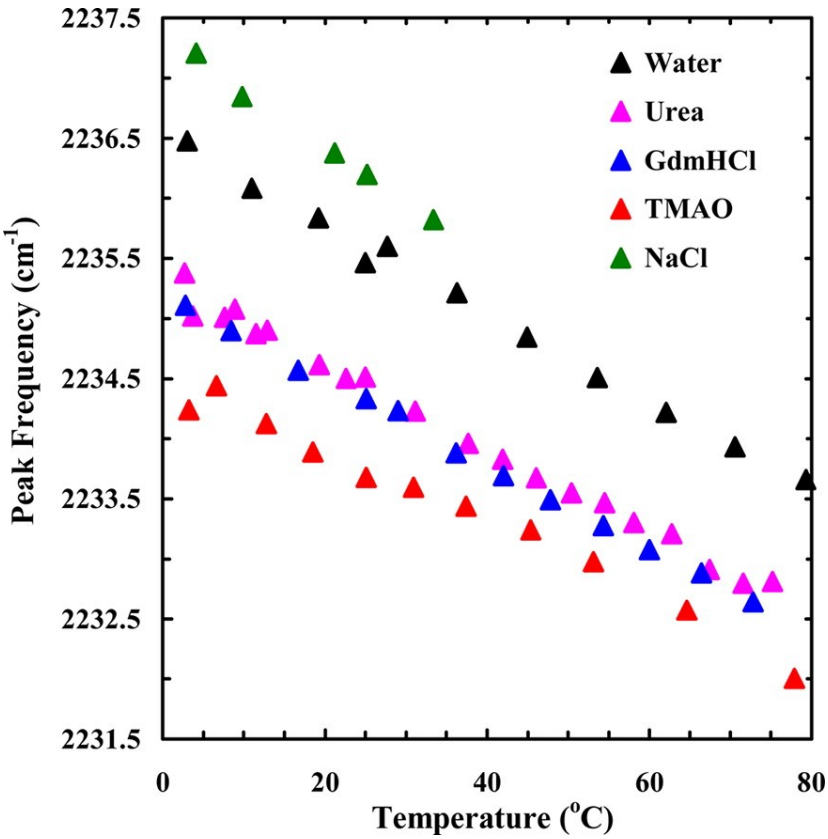



Figure 4.4 Temperature dependence of the peak frequency of the C≡N stretching vibrational band of the $Phe_{CN}$-Peptide obtained under different solvent conditions, as indicated. The concentrations of the cosolvent or cosolute in each case were: [TMAO] = 4.8 M, [Urea] = 5.2 M, [GdnHCl] = 5.1 M, and [NaCl] = 5.0 M.

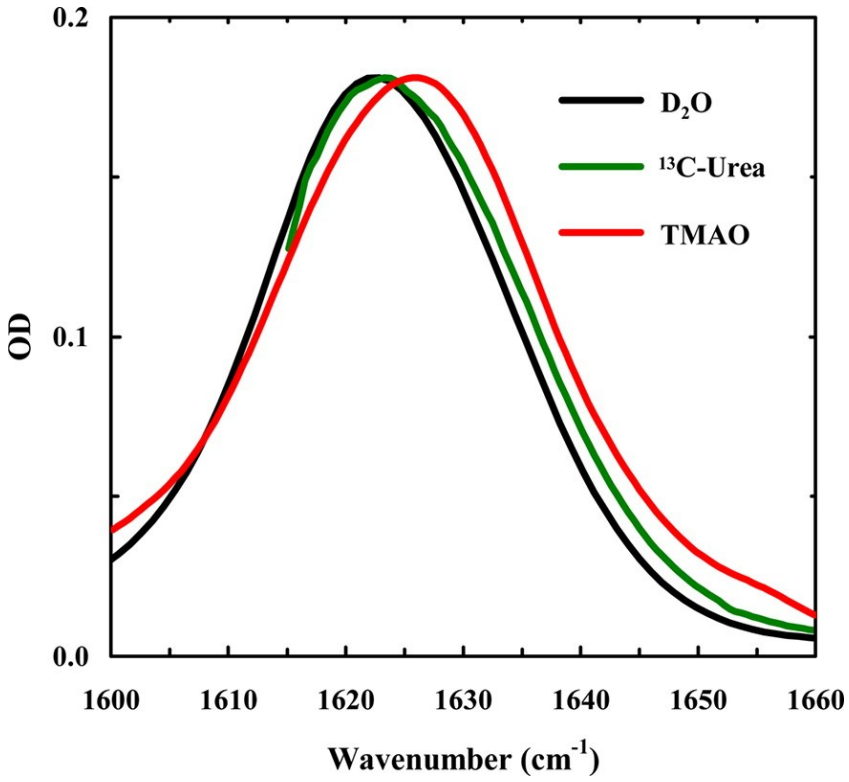



Figure 4.5 FTIR spectra of NMA in the amide I′ region at 25.0 °C under different solvent conditions, as indicated. The concentrations of TMAO and 13C-urea were 4.7 and 5.1 M, respectively. Since the C═O stretching vibrational band of12C-urea overlaps strongly with that of NMA, $^{13}$C-urea was used in the current case. The NMA concentrations were ~50 mM for the $^{13}$C-urea experiment and ~10 mM for other measurements. For easy comparison, the spectra obtained in TMAO and $^{13}$C-urea solutions have been normalized to that obtained in $D_2O$.

# 5 Microscopic insights into the protein-stabilizing effect of trimethylamine N-oxide (TMAO)

## 5.1 Significance

The ability of trimethylamine N-oxide (TMAO) to counteract osmotic stress along with the deleterious effects of denaturants such as urea is fascinating and has inspired many studies. To further our understanding of how TMAO acts to stabilize folded proteins, we carry out an infrared experiment designed to probe the microscopic details of this action explicitly from the perspective of the solute molecule. Our results reveal that the protein-stabilizing effect of TMAO originates from two contributions: One is entropic and the other is enthalpic in nature. Thus, this study provides not only microscopic details underlying the stabilizing action of TMAO, but also a method that can be used to study the stability-perturbing effect of other cosolvents.

## 5.2 Abstract

Although it is widely known that trimethylamine N-oxide (TMAO), an osmolyte used by nature, stabilizes the folded state of proteins, the underlying mechanism of action is not entirely understood. To gain further insight into this important biological phenomenon, we use the C≡N stretching vibration of an non-natural amino acid, *p*-cyano-phenylalanine, to directly probe how TMAO affects the hydration and conformational dynamics of a model peptide and a small protein. By assessing how the lineshape and

spectral diffusion properties of this vibration change with cosolvent conditions, we are able to show that TMAO achieves its protein-stabilizing ability through the combination of (at least) two mechanisms: (*i*) It decreases the hydrogen bonding ability of water and hence the stability of the unfolded state, and (*ii*) it acts as a molecular crowder, as suggested by a recent computational study, that can increase the stability of the folded state via the excluded volume effect.

## 5.3 Introduction

Nature employs a variety of small organic molecules to cope with osmotic stress. Trimethylamine N-oxide (TMAO) is one such naturally occurring osmolyte that protects intracellular components against disruptive stress conditions.[289] In particular, previous studies have shown that TMAO is able to enhance protein stability and to counteract the denaturing effect of urea.[10,290] TMAO (Figure 5.1, Inset) adopts a skewed tetrahedral structure with a charged oxygen capable of accepting hydrogen bonds (H bonds) and three hydrophobic (methyl) groups. This amphiphilic structural arrangement makes TMAO a rather special cosolvent, because it can form H bonds with water, self-associate in a manner similar to surfactants, and show preferential interactions with or exclusion[45,245,254-255,257,291-294] from certain protein functional groups.[22,33,241,248,281,295-300] Indeed, these molecular properties of TMAO have been used, either individually or in combination, to rationalize its biological activities. For example, a prevailing view about

TMAO is that its osmotic and protective role is caused by the molecule's tendency to be preferentially depleted from protein surfaces, as suggested by physicochemical measurements.[49,51,54,240,301] This thermodynamic picture, as described by Bolen and coworkers[53], implies that the protein is preferentially hydrated. Although this notion is consistent with TMAO's being a protecting osmolyte, to the best of our knowledge no experimental studies have been carried out to directly examine the effect of TMAO on protein hydration dynamics, an aspect that is also important to protein function. Herein, we use two-dimensional infrared (2D IR) spectroscopy and a site-specific IR probe to explore this critical issue and gain insight toward achieving a microscopic understanding of the molecular mechanism of the protecting action of TMAO.

2D IR spectroscopy is capable of assessing the frequency–frequency correlation function (FFCF) of a given IR probe, which reports on the underlying dynamics of events that lead to fluctuations of its vibrational frequency.[56] For an IR probe that is able to interact with water via H bonding, measurement of its FFCF can provide, sometimes in a site-specific manner, detailed information about the hydration dynamics of the protein molecule of interest. For example, this approach has been used to identify the existence of mobile water molecules inside Aβ40 amyloid fibrils[302-303] and to interrogate the water-assisted drug-binding mechanism of HIV-1 reverse transcriptase[304], among many other applications.[305-310] In the current study, we capitalize on the established sensitivity of the nitrile stretching vibration (C≡N) to local hydration and electrostatic environment[177] and use the non-natural amino acid *p*-cyano-phenyalanine ($Phe_{CN}$) as a local IR reporter. The

advantages of using the C≡N stretching vibration of $Phe_{CN}$ are that (*i*) it is located in a spectrally uncongested region (2,000−2,400 $cm^{-1}$), where water has a relatively low absorbance; (*ii*) it has a reasonably large extinction coefficient; and (*iii*) it is, in most cases, a simple transition that is decoupled from other vibrational modes of the molecule.[181]

Recently, we have shown that upon addition of TMAO, the C≡N stretching vibration of $Phe_{CN}$ in a short peptide shifts to a lower frequency compared with that obtained in pure water.[174] As shown (Figure 5.1), the peptide studied in the current case, Gly-$Phe_{CN}$-Gly (hereafter referred to as $GF_{CN}G$), exhibits the same trend. Taken together, these results suggest that TMAO weakens the strength of the H bond formed between the nitrile group and water molecules. Whereas this finding is consistent with the ability of TMAO to increase protein stability, a microscopic picture regarding how TMAO alters protein–water interactions has yet to be established. To provide new insights into the molecular mechanism of the protecting action of TMAO, herein we measure the FFCF of the nitrile probe in two model systems, the tripeptide $GF_{CN}G$ and the villin headpiece subdomain (HP35) with $Phe_{CN}$ mutations under different cosolvent conditions. In particular, we are able to extract both the homogeneous and the inhomogeneous contributions to the total lineshape of the C≡N stretching band and, perhaps more importantly, to provide a dynamic view, from the perspective of the $Phe_{CN}$ sidechain, of how a given solution condition (i.e., in the presence of TMAO, urea, or both) affects the hydration dynamics of the peptide in question.[20] Our results suggest that TMAO

molecules produce a local protein environment that strengthens backbone–backbone H bonds and also reduces the conformational entropy of the system. In other words, TMAO increases protein stability by acting as a nano-crowder, as proposed by Thirumalai, Straub and coworkers based on computer simulations.[62]

## 5.4 Results

Consistent with our previous study, TMAO affects the C≡N stretching band of $GF_{CN}G$ by shifting its peak to a lower frequency and broadening its width.[174] Although these changes in lineshape indicate a change in the local electrostatic and H-bonding environment of the C≡N probe, and thus contain information about the molecular effect of TMAO, a quantitative assessment of the factors underlying these changes, based on linear IR measurements alone, is not feasible. Thus, herein we combine data from linear and nonlinear IR measurements and use response function theory to fully characterize the lineshape of the C≡N stretching vibrations obtained under different solvent conditions. This analysis allows us to dissect the homogeneous and inhomogeneous contributions to the linear IR spectrum of interest, which provides insights into the mechanism of action of TMAO.

### 5.4.1 2D IR Photon Echo Spectra of $GF_{CN}G$.

As indicated (Figure 5.2), the absorptive 2D IR spectra of the nitrile group of $GF_{CN}G$, collected at different waiting times ($T$) and solution conditions, clearly show that

the 2D lineshapes and the spectral diffusion dynamics of the C≡N stretching vibration depend on the cosolvent. For example, at $T$ = 500 fs, the 2D IR lineshape of this vibration in water becomes more circular than that in 6 M TMAO, suggesting that adding TMAO significantly changes contributions from the homogenous part, as well as the inhomogenous part, to the total lineshape of the vibrational transition. To provide a more quantitative assessment of how a particular cosolvent affects the microscopic environment of the nitrile probe and hence its vibrational transition, below we examine how the key spectral and dynamic properties of this vibrator change with solution condition.

First, we compare the vibrational lifetimes ($T_1$) of the nitrile probe obtained under different solution conditions. As shown (Table 5.1), the $T_1$ of the C≡N stretching vibration, consistent with previous studies[136,148], is insensitive to solvent and, thus, is not a good reporter of its local environment. This finding is consistent with the study of Cho and coworkers[311], which showed that the vibrational relaxation of a nitrile moiety attached to a large molecule in water is dominated by intramolecular pathways. In addition, the measured $T_1$ value (~4 ps) indicates that in this case the lifetime-broadening contribution (~1.3 $cm^{-1}$) to the spectral width (10−15 $cm^{-1}$) is small.

Second, we compare the effects of cosolvent on the spectral diffusion of this probe, using the center line slope (CLS) method developed by Fayer and coworkers.[57] A center line is defined as the line that connects the maxima of the peaks of a series of cuts through a particular 2D IR spectrum that are parallel to the $\omega_\tau$frequency axis. It has been

shown that, within the short time approximation, the waiting-time (i.e., *T*) dependence of the CLS directly reports on the normalized FFCF of the vibrational probe in question.[57] As indicated (Figure 5.3), the CLS (*T*) of the nitrile probe, obtained under all solution conditions, can be described by the following equation:

$$\mathrm{CLS}(T) = A \cdot e^{-t/\tau} + B \qquad \text{Equation 5.1}$$

where *A*, *B*, and $\tau$ are constants. The exponential term in Equation 5.1 is often referred to as the Kubo term, which, in the current case, most likely reports on the dynamics of H-bond formation and breaking between the nitrile group and water, which cause the vibrational frequencies to fluctuate.[312-314] As shown (Table 5.2), the decay time constant of this Kubo term determined for $GF_{CN}G$ in water is about 2.2 ps, which is in good agreement with previously reported values for the nitrile stretching vibration in aqueous solutions.[55,136,308,314] In addition, the offset term *B* obtained in water is practically zero. However, adding a cosolvent, either TMAO or urea, results in a measurable decrease in $\tau$ and an appreciable increase in *B* and, interestingly, TMAO is more effective than urea at increasing this offset. Fayer and coworkers[57] have shown, assuming that the system follows Bloch dynamics, that this offset term is related to the total inhomogeneous component of the lineshape. Thus, these results suggest that TMAO can effectively slow down or even suppress some motions that lead to local environmental changes of the nitrile probe and, hence, its spectral diffusion.

To provide a more quantitative assessment of the cosolvent-induced changes in the spectral and dynamic properties of the nitrile stretching vibration, we calculate the

homogeneous and inhomogeneous contributions to the total lineshape, by analyzing the linear IR spectrum using the linear response function.[56] Specifically, we fit the IR spectrum, $S_{IR}(\omega)$, to the following function[312]:

$$S_{IR}(\omega) \propto \Re \int_0^t e^{i(\omega-\omega_0)t} \cdot e^{-g(t)} \cdot e^{-t/2T_1} dt \qquad \text{Equation 5.2}$$

where $T_1$ is the vibrational lifetime, $\omega_0$ is the center frequency, and $g(t)$ is the vibrational lineshape function, defined as

$$g(t) = \int_0^t \int_0^{\tau'} d\tau' d\tau'' C(\tau'') \qquad \text{Equation 5.3}$$

where C(t) represents the full FFCF of the IR probe and is described by the following equation (53):

$$C(t) = \frac{2\delta(t)}{T_2^*} + \Delta_s^2 + \Delta^2 \cdot e^{-t/\tau} \qquad \text{Equation 5.4}$$

where the first term represents the motional narrowing component determined by the pure dephasing time ($T_2$*), the second term represents a static or slowly varying component, and the third term is the aforementioned Kubo term. As shown (Figure 5.1), all of the linear IR spectra can be reasonably well described by  and the resultant fitting parameters (i.e., $T_2$*, $\Delta_s$, and $\Delta$) are listed in Table 5.1. In particular, the parameters obtained for water are in good agreement with those previously reported for this probe.[308] With the knowledge of $T_2$*, we further calculated the pure dephasing linewidth ($\Gamma$) for each case using the relationship of $\Gamma = 1/\pi T_2$*.

In the current case, we have neglected the contribution from $T_{rot}$, which represents the rotational relaxation time of the probe[315] and is expected to be much longer than $T_1$ or $T_2$*. As indicated (Table 5.1), the spectral parameters produced from the above analysis clearly show that the pure dephasing linewidth ($\Gamma$) of the C≡N stretching vibration decreases in the following order: $H_2O$ > 6 M urea > 1 M TMAO > TMAO/urea (3 M/3 M) > 6 M TMAO, whereas the total inhomogeneous component (i.e., $\sqrt{\Delta^2 + \Delta_s^2}$) follows the exact opposite trend (i.e., $H_2O$ < 6 M urea < 1 M TMAO < TMAO/urea (3 M/3 M) < 6 M TMAO). Moreover, the contribution from the Kubo term (i.e., $\Delta^2$) decreases upon adding cosolvents to the system, which, consistent with the above CLS analysis, indicates that cosolvents weaken the H bonding between the C≡N probe and water molecules.

As shown in Table 5.1, the most striking differences for the C≡N stretching vibration under different solution conditions are in the values of the static term (Δ2s). The origin of this static term, for large molecules such as proteins, is generally attributed to motions that can affect the local electrostatic environment of the IR probe but occur on a timescale that is much slower than the vibrational lifetime or the effective time window of the experiment, such as backbone conformational changes.[316-318] Interestingly, in the current case, the static component obtained in water is nearly zero, indicating that during the time window of the experiment the nitrile probe can exhaustively sample, statistically speaking, all possible environments. This is consistent with the notion that the spectral diffusion of the nitrile probe in pure water is dominated by water H-bonding dynamics, which occur on the timescale of 1−3 ps.[319] Furthermore, and perhaps more interestingly,

adding TMAO to the peptide solution leads to a significant increase in the static component, which indicates that TMAO not only broadens the frequency distribution of the C≡N stretching vibration (Figure 5.1), but also "freezes" a certain number of motions leading to vibrational frequency fluctuations. As shown in Table 5.1, although urea exhibits a similar effect, the net influence, in comparison with that of TMAO at the same concentration, is much smaller. Garcia and coworkers[320] have shown that the osmotic pressure of TMAO exhibits a positive deviation from its ideal value, whereas that of urea shows a negative deviation. Thus, a more rigorous comparison between the effects of TMAO and urea would require measurements on TMAO and urea samples that have identical osmotic pressures.

### 5.4.2 Linear IR and 2D IR Spectra of HP35 Mutants.

To confirm that the conclusions reached above are not limited to short peptides, we also performed similar experiments on two $Phe_{CN}$ mutants of a well-studied helical protein, HP35.[321-322] The first variant corresponds to a Phe76/$Phe_{CN}$ mutation at the C terminus of the protein (referred to as HP35-P76), wherein the $Phe_{CN}$ residue is exposed to solvent. Because several studies have suggested that protecting osmolytes, such as TMAO, could directly interact with aromatic sidechains[22], the second mutant we chose to study contains a Phe58/$Phe_{CN}$ mutation at the core of the protein (referred to as HP35-P58). Because the $Phe_{CN}$ residue in HP35-P58 is, to a large extent, buried in a hydrophobic environment[308,322-323], it is unlikely that TMAO will show any direct interactions with the nitrile probe. Thus, results obtained from both mutants will allow us

to gain a more complete assessment of the effect of TMAO on protein stability. Unfortunately, we found that HP35 readily aggregates even in 1 M TMAO, which prevented us from making a direct evaluation of the effect of TMAO on the C≡N stretching vibration of both HP35 variants. To circumvent this problem, we evaluated the effect of TMAO in the presence of urea, which is known to help solubilize the protein. Specifically, we used a solution condition that was used in the study of $GF_{CN}G$, i.e., TMAO/urea (3 M/3 M). As indicated, both the linear and 2D IR spectra of HP35-P76 (Figure 5.5a, 5.4a and Table 5.3) show that the cosolvent-induced changes in the spectral and dynamic properties of the C≡N stretching vibration of HP35-P76 are very similar to those observed for the $GF_{CN}G$ peptide under the same solution conditions. For example, the 2D lineshape, at $T$ = 500 fs, obtained in pure water is distinctly more circular than that measured in the presence of cosolvent. A more quantitative analysis using the method discussed above reveals that the cosolvent-induced spectral property changes in this case are similar to those obtained with $GF_{CN}G$ (Table 5.3), indicating that the findings for $GF_{CN}G$, in terms of the effect of TMAO, are not unique to short peptides, but rather are generally applicable to proteins. Similar to that observed for HP35-P76, addition of the same cosolvent mixture results in a shift of the C≡N stretching frequency of HP35-P58 to lower wavenumbers (Figure 5.5b and Table 5.3). However, the CLS of HP35-P58 obtained from its 2D IR spectra (Figure 5.5b) in water shows a slower decay and also a significant offset (Figure 5.6 and Table 5.3). This is consistent with the study of Fayer and coworkers,[308,322-323] which showed that the frequency fluctuation dynamics of the

C≡N probe inside the hydrophobic core of HP35 are slowed down. However, addition of TMAO/urea (3 M/3 M) leads to a further increase in the amplitude of this offset (Figure 5.6), indicating that the nitrile probe samples a more slowly varying or rigid environment.[323] Because, in this case, the $Phe_{CN}$ residue is unlikely to directly interact with cosolvent molecules, this increase in protein rigidity thus most likely arises from the TMAO-induced crowding effect, as discussed above.

## 5.5 Discussion

The current view on how TMAO and urea affect protein stability has been recently reviewed by Canchi and García.[324] Owing to various quantitative thermodynamic measurements, such as vapor pressure osmometry,[35,325-326] volumetric measurements,[327] and equilibrium dialysis,[328] we now know a great deal about why urea and TMAO act differently. Urea's protein denaturation effect is generally believed to arise from the molecule's tendency to accumulate around both the protein backbone and sidechains, owing, for example, to a stronger dispersion interaction between urea and the protein backbone,[246] which preferentially stabilizes the unfolded state. However, TMAO molecules are found to be excluded from various protein backbone and sidechain units, thus destabilizing the unfolded state.[320] Moreover, many studies have shown that both TMAO and urea can alter the structure and dynamics of the H-bonding network of water, thus resulting in a change in protein stability.[249] Although previous studies have

significantly enhanced our understanding of the mechanisms by which TMAO/urea increase/decrease protein stability, especially from a thermodynamic point of view, many important microscopic details have yet to be revealed or verified by experiments. Herein, we use linear and nonlinear IR measurements to directly probe, from a sidechain's perspective, how TMAO and urea affect the H-bonding dynamics of water at the protein–solvent interface.

Two key findings that emerged from the current study are worthy of further discussion. First, the spectral diffusion data of the C≡N stretching vibration clearly show that addition of either urea or TMAO makes the local H-bonding environment fluctuate faster (i.e., the smaller time constant of the Kubo term in Table 5.2). This picture is consistent with our early observation that TMAO and urea can weaken the H-bond strength between the nitrile probe and water molecules.[174] It is obvious that a weakened H-bonding interaction will result in a shallower potential energy surface along the H-bond coordinate, thus making the diffusion dynamics along this coordinate faster (at the same temperature). In addition, weakening the H-bonding interactions will lead to a smaller vibrational energy splitting, which, in turn, would lead to a smaller contribution to the total FFCF amplitude by the corresponding H-bonding dynamics, as observed (*A* in Table 5.2 and $\Delta^2$ in Table 5.1). Furthermore, we can rule out the possibility that the observed changes in the H-bonding dynamics arise from an increase in viscosity (e.g., 6 M urea increases the solution viscosity by only 40%), because Kubarych and coworkers have recently shown that increasing bulk viscosity slows down spectral diffusion.[329]

Because water always competes, whenever possible, with a protein's intrinsic H-bonding donors/acceptors to form additional H bonds with, for example, the backbone amides, the decreased strength of H bonds involving water would preferentially destabilize the unfolded state of proteins, where more amide units are exposed to the solvent. Thus, this is an important finding because it provides, to the best of our knowledge, the first experimental evidence indicating that TMAO can enhance protein stability by attenuating the strength of H bonds formed between protein polar groups and water. Interestingly, our IR measurements suggest, although to a lesser degree in comparison with TMAO, that urea can also weaken the H-bonding ability of water with proteins.[174] At first glance, this finding seems to be inconsistent with the denaturing role of urea. However, upon consideration of the fact that urea, unlike TMAO, shows preferential accumulation in the vicinity of protein backbones and sidechains[22], this result provides strong evidence in support of the notion that urea's ability to denature proteins does not arise from its effect on the structure of water, but is instead achieved through a more direct mechanism. This conclusion is in agreement with that reached by Pielak and coworkers[288], who showed that there is no correlation between a solute's impact on water structure and its effect on protein stability.

The second major finding is that the static component ($\Delta 2s$) in the FFCF, which is absent (or insignificant) in pure water, increases significantly upon addition of TMAO. This static component, together with the aforementioned Kubo term, determines the inhomogeneous broadening of the C≡N stretching vibration. Thus, the fact that TMAO

increases the amplitude of Δ2s provides strong evidence indicating that it either slows down the peptide's conformational motions, creates new and slowly moving conformational states that are not populated in pure water, or both. This picture is entirely consistent with the NMR study of Loria and coworker[330], which showed that TMAO can significantly rigidify the protein backbone, and with the simulations of Thirumalai and coworkers[62], which demonstrated that TMAO can act as a nano-crowder, thus increasing protein stability via the excluded volume effect. Despite this consistency, we note that a study by Qu and Bolen[331] indicated that the efficacy of the macromolecular crowding in forcing proteins to fold may be modest and may also depend on the crowding agents. Thus, it would be very useful if new experiments can be designed to directly quantify the crowding efficacy of TMAO.

Taken together, the effect of TMAO on the Kubo and static terms of the FFCF of the nitrile probe provides insights into the mechanism of action of TMAO in increasing the stability of proteins. Primarily, our results indicate that the protein stabilizing ability of TMAO arises from both enthalpic and entropic contributions. The enthalpic factor results from a decreased H-bonding ability of water, whereas the entropic factor stems from the crowding effect of TMAO; both act to destabilize the unfolded state ensemble. Because the tripeptide used in the current study is relatively small, it is possible to use molecular dynamics simulations[28,246,332], in conjunction with vibrational frequency calculations[84], to directly calculate the FFCFs of the nitrile probe under different

cosolvent conditions and to provide further molecular insights. We hope that the current study will inspire new computational efforts in this regard.

## 5.6 Materials and Methods

### 5.6.1 Sample Preparation.

Reagent-grade urea and TMAO were purchased from Aldrich (Saint Louis, MO) and used as received. Fmoc-$Phe_{CN}$-OH was purchased from Bachem Americas (Torrance, CA). Syntheses of $GF_{CN}G$, HP35-P76 and HP35-P58 were carried out on a PS3 peptide synthesizer (Protein Technologies), purified by reverse-phase high-performance liquid chromatography, and verified by mass spectrometry. HP35-P76 is a double mutant of HP35 with Trp64 and Phe76 replaced by 7-azatryptophan (7AW) and $Phe_{CN}$, respectively. HP35-P58 is also a double mutant of HP35 where Phe58 is replaced with $Phe_{CN}$ and Trp64 is mutated to 7AW. Peptide and protein solutions were prepared by dissolving lyophilized $GF_{CN}G$, HP35-P76 or HP35-P58 samples in either Millipore water or aqueous solutions of urea and/or TMAO, as indicated. The concentration of $GF_{CN}G$ was 50-80 mM and the concentrations of HP35-P76 and HP35-P58 were 8-12 mM, as determined optically using the molar extinction coefficients at 280 nm of 850 $cm^{-1}M^{-1}$ for $GF_{CN}G$ and 6820 $cm^{-1}M^{-1}$ for both HP35 variants. No signs of aggregation were observed based on the amide I′ bands of $GF_{CN}G$ and HP35 variants in neat $D_2O$ at the same concentrations.

### 5.6.2 Linear and 2D IR Measurements.

The home-made IR sample holder consists of two $CaF_2$ windows and a 25 μm Teflon spacer. FTIR spectra were collected at a resolution of 1 $cm^{-1}$ on a Magna-IR 860 spectrometer (Nicolet, WI). The details of the 2D IR instrument have been described elsewhere.[303] Briefly, three femtosecond IR pulses ($k_1$, $k_2$, and $k_3$) were focused on the sample to induce a third order response (photon echo), which was heterodyned by a fourth IR pulse (the local oscillator), dispersed by a monochromator, and detected with a 64-pixel array detector (Infrared Associates). Rephasing and nonrephasing data were collected by scanning the time delay ($\tau$) between $k_1$ and $k_2$ from -6 to +6.5 ps at 2 fs intervals. The waiting time ($T$) between pump ($k_1$/$k_2$) and probe ($k_3$) pulses was varied from 0 to 6 ps. The final spectrum, $S$ ($\omega_\tau$, $T$, $\omega_t$), was obtained from three Fourier transforms of $S$ ($\tau$, $T$, $\lambda$), and only the real part of the result was shown. In addition, no background subtraction was made, as the signals from the solvent were found to not appreciably affect the 2D lineshape of the nitrile probe. Finally, the isotropic transient grating experiment was performed using the same 2D IR photon echo setup where the delay ($\tau$) was set to zero.

## 5.7 Acknowledgments

We thank Dr. Arnaldo Serrano for synthesizing the HP35 variants. This work was supported by National Institutes of Health (NIH) Grant GM012592. The 2D IR

instrumentation was developed with NIH Research Resource Grant P41 GM104605. I.M.P. was supported by National Science Foundation Graduate Research Fellowship DGE-0822.

## 5.8 Original Publication

This Chapter has been reprinted from the Proceeding of the National Academy of Sciences, Jianqiang Ma, Ileana M. Pazos, and Feng Gai, (2014) vol. 111 no. 2 8476–8481. DOI: 10.1073/pnas.1403224111, with permission.

Table 5.1 Spectroscopic parameters and lifetimes ($T_1$) of the C≡N stretching vibration of $GF_{CN}G$ in different solvent conditions.

| | $\omega_0$ ($cm^{-1}$) | FWHM ($cm^{-1}$) | $T_1$ (ps) | Γ ($cm^{-1}$) | Δ ($cm^{-1}$) | $\Delta_s$ ($cm^{-1}$) |
|---|---|---|---|---|---|---|
| $Phe_{CN}$* | 2236.3 | 12.2 | 4.5 | 5.8±0.1 | 2.0±0.1 | - |
| $H_2O$ | 2236.3 | 10.4 | 4.2±0.1 | 5.8±0.1 | 2.1±0.2 | 0 |
| 6M Urea | 2234.7 | 11.2 | 4.0±0.2 | 5.2±0.2 | 1.8±0.1 | 1.0±0.2 |
| 1M TMAO | 2236.0 | 10.6 | 4.1±0.2 | 5.1±0.2 | 1.7±0.1 | 1.2±0.2 |
| TMAO/Urea (3M/3M) | 2234.2 | 13.3 | 3.8±0.2 | 5.0±0.2 | 2.0±0.2 | 2.2±0.1 |
| 6M TMAO | 2232.3 | 13.4 | 3.9±0.2 | 4.8±0.2 | 2.0±0.2 | 2.9±0.3 |

*The $Phe_{CN}$ parameters are taken from reference[308], which were measured in water.

The center frequency ($\omega_0$) and width (FWHM) were determined by fitting the linear IR spectra to a Gaussian function. The $T_1$ values were determined by fitting the isotropic transient grating signals to an exponential function. Other lineshape parameters were obtained by fitting the C≡N stretching bands in Figure 5.1 to Equation 5.2 to Equation 5.4.

Table 5.2 Parameters obtained from fitting the CLS curves in Figure 5.3 to Equation 5.1.

| | A | $\tau$ (ps) | B |
|---|---|---|---|
| $H_2O$ | 0.26 ± 0.02 | 2.17 ± 0.43 | 0.01 ± 0.02 |
| 6M Urea | 0.20 ± 0.01 | 1.44 ± 0.30 | 0.15 ± 0.30 |
| 1M TMAO | 0.22 ± 0.02 | 1.30 ± 0.43 | 0.18 ± 0.02 |
| TMAO/Urea (3M/3M) | 0.31 ± 0.02 | 1.85 ± 0.39 | 0.24 ± 0.02 |
| 6M TMAO | 0.18± 0.02 | 1.46 ± 0.43 | 0.51 ± 0.02 |

Table 5.3 Spectroscopic parameters obtained from fitting the C≡N stretching bands of HP35-P76 and P58 in Figure 5.4, fit to Equation 5.2 to Equation 5.4.

| | $\omega_0$ ($cm^{-1}$) | FWHM ($cm^{-1}$) | $\Gamma$ ($cm^{-1}$) | $\Delta$ ($cm^{-1}$) | $\Delta_s$ ($cm^{-1}$) |
|---|---|---|---|---|---|
| **HP35-P76** | | | | | |
| $H_2O$ | 2235.1 | 10.8 | 5.6 ± 0.4 | 1.9 ± 0.3 | 0.4 ± 0.3 |
| TMAO/Urea (3M/3M) | 2233.4 | 13.8 | 5.1 ± 0.4 | 2.2 ± 0.3 | 2.3 ± 0.4 |
| **HP35-P58** | | | | | |
| $H_2O$ | 2234.2 | 13.4 | 4.8 ± 0.6 | 4.0 ± 0.4 | 2.0 ± 0.4 |
| TMAO/Urea (3M/3M) | 2232.1 | 14.0 | 4.2 ± 0.4 | 2.3 ± 0.4 | 4.1 ± 0.6 |

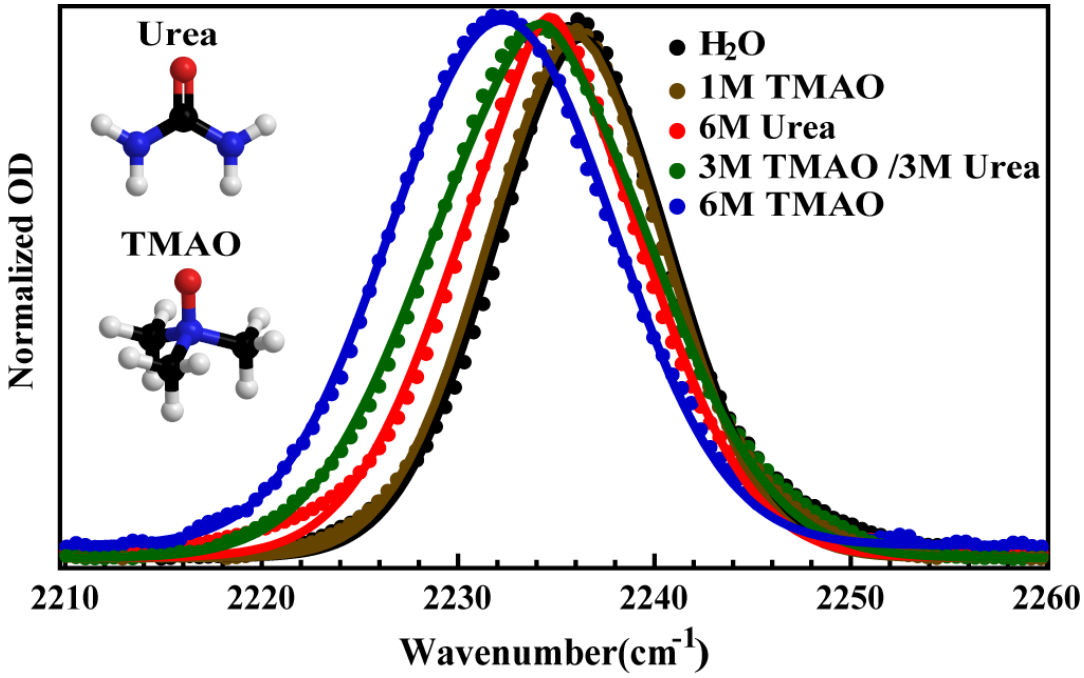



Figure 5.1 The C≡N stretch bands of $GF_{CN}G$ obtained under different solvent conditions, as indicated. The solid lines are fits of these spectra to the linear response function using Equation 5.2 to Equation 5.4 in the text and the resulting fitting parameters are listed in Table 5.1. The structures of urea and TMAO are shown as insets.

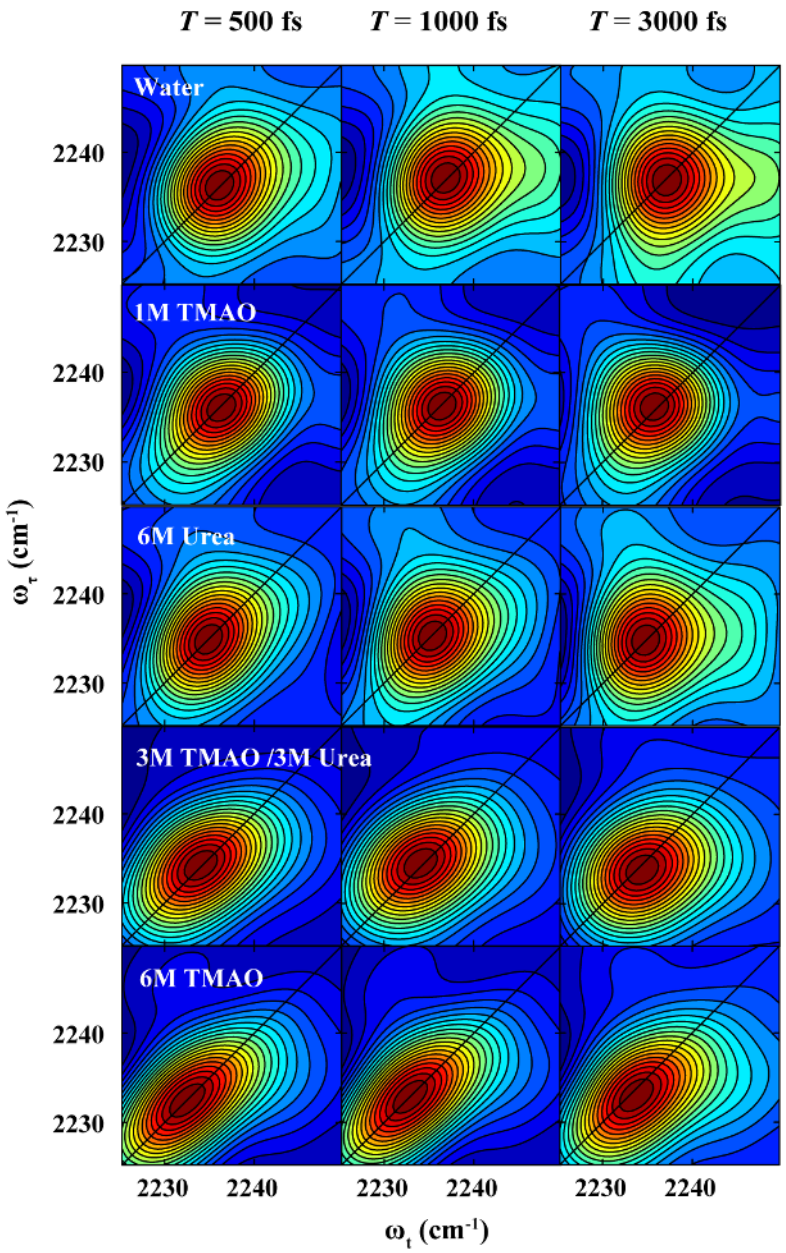



Figure 5.2 Representative absorptive 2D IR photon echo spectra of $GF_{CN}G$ obtained under different solvent conditions and at different waiting times ($T$), as indicated.

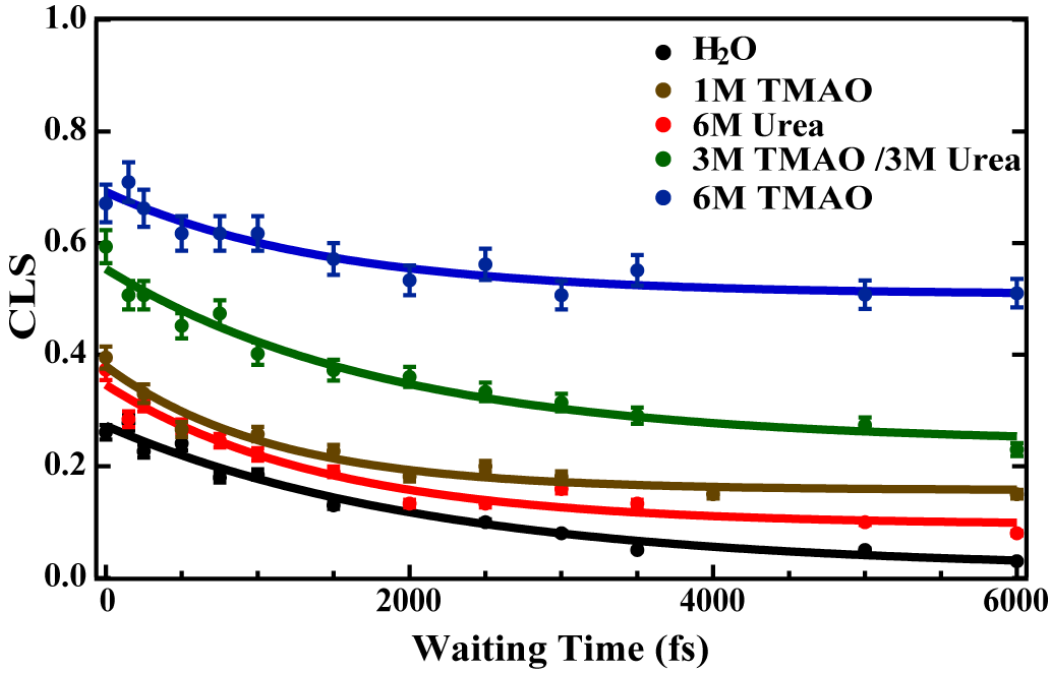



Figure 5.3 CLS versus *T* plots of $GF_{CN}G$ obtained under different solvent conditions, as indicated. The solid lines are fits of these data to Equation 5.1with those fitting parameters listed in Table 5.2.

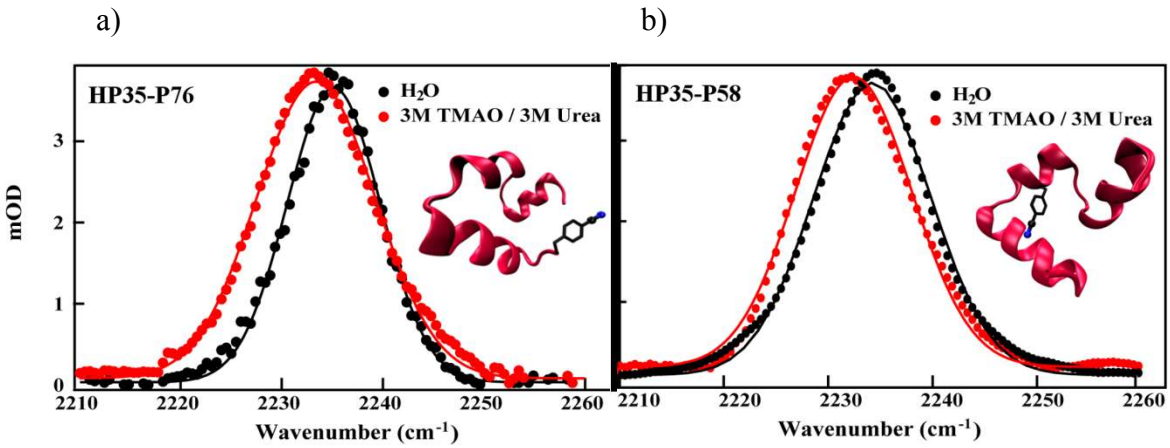



Figure 5.4 Normalized FTIR Spectra of a) HP35-P76 and b) HP35-P58 obtained in water and a mixture of TMAO/Urea (3M/3M), as indicated. Fitting these spectra to a Gaussian function (smooth line) yielded the spectroscopic parameters listed in Table 5.3. Also shown is the location of the $Phe_{CN}$ residue, generated using VMD based on the NMR structure of HP35 (PDB code: 1VII).

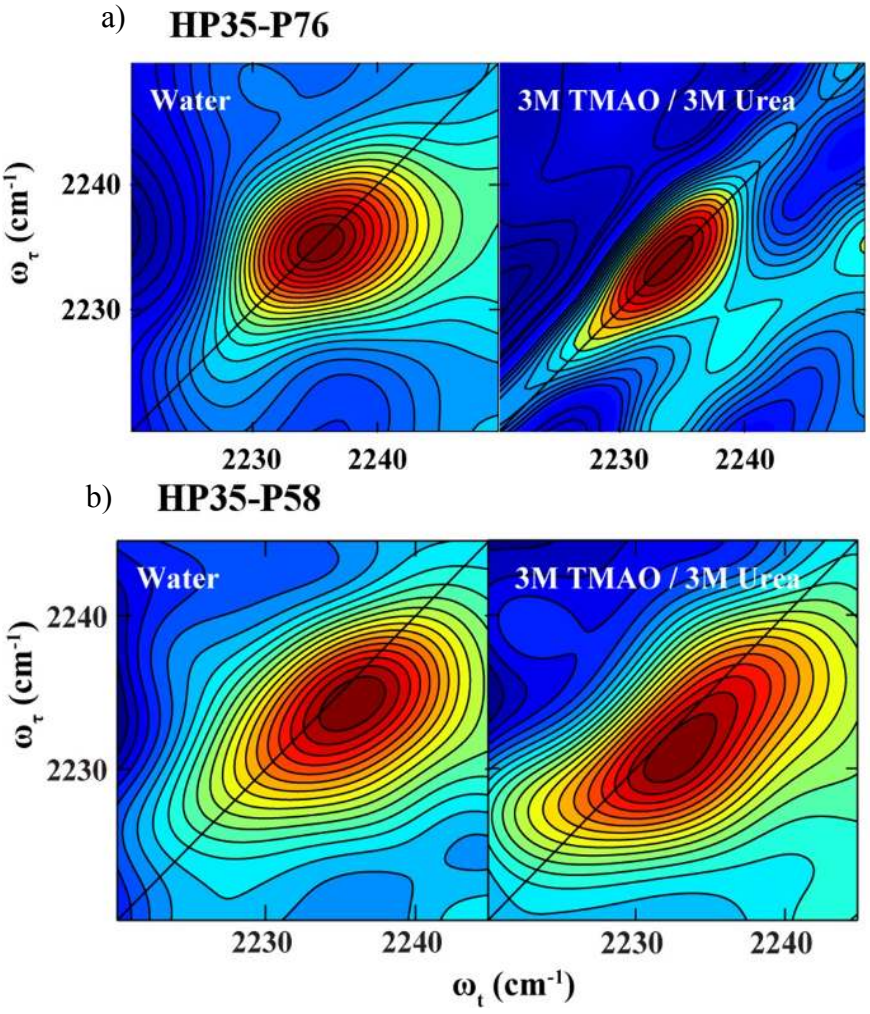



Figure 5.5 Pure absorptive 2D IR photon echo spectra of a) HP35-P76 and b) HP35-P58 obtained at $T$ = 500 fs and in water and in TMAO/Urea (3M/3M). Pure absorptive 2D IR photon echo spectra of HP35-P58 obtained at $T$ = 500 fs and in water and in TMAO/Urea (3M/3M).

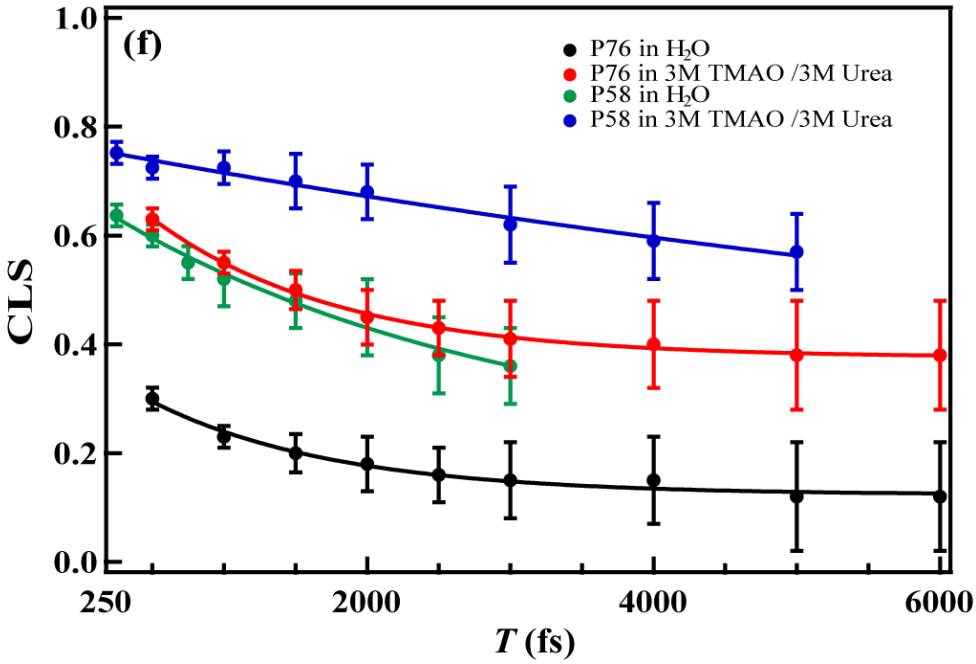



Figure 5.6 CLS versus *T* plots of HP35-P76 and HP35-P58 obtained under different solvent conditions, as indicated. The solid lines are fits of these data to Equation 5.1 with those fitting parameters listed in Table 5.3. The relatively large uncertainties in the CLS data and the apparent asymmetric 2D lineshape in the spectra (Figure 5.5) are due to distortions from background noise, because the absorbance of the nitrile group in these experiments was relatively low (~2 mOD).

# 6 Ester Carbonyl Vibration as a Sensitive Probe of Protein Local Electric Field

## 6.1 Abstract

The ability to quantify the local electrostatic environment of proteins and protein/peptide assemblies is key to gaining a microscopic understanding of many biological interactions and processes. Herein, we show that the ester carbonyl stretching vibration of two non-natural amino acids, l-aspartic acid 4-methyl ester and l-glutamic acid 5-methyl ester, is a convenient and sensitive probe in this regard, since its frequency correlates linearly with the local electrostatic field for both hydrogen-bonding and non-hydrogen-bonding environments. We expect that the resultant frequency–electric-field map will find use in various applications. Furthermore, we show that, when situated in a non-hydrogen-bonding environment, this probe can also be used to measure the local dielectric constant ($\varepsilon$).

## 6.2 Introduction

Electrostatic interactions are ubiquitous in biological molecules and, in many cases, play a key role in molecular association and enzymatic reactions.[81,333] However, quantification of the local electric field or how it changes inside a protein, especially in a site-specific manner and/or as a function of time, still remains a challenging task. One promising method in this regard is vibrational Stark spectroscopy,[259] which capitalizes on the intrinsic dependence of vibrational transitions on the local electrostatic environment.

This method is based on the use of an infrared (IR) probe that has a well-defined, localized vibrational mode to sense the amplitude of the local electric field through the frequency response.[334] For example, the vibrational Stark effect has been used to determine the local electric field at protein interfaces and to monitor protein conformational transitions and dynamics.[76,323,335-336] Although the theoretical underpinning of this methodology is straightforward, in practice the application of vibrational Stark spectroscopy to biological systems is currently limited by the availability of suitable vibrational probes. Herein, we show, by the use of linear and nonlinear IR spectroscopic measurements and molecular-dynamics (MD) simulations, that the ester carbonyl vibration in two non-natural amino acids can be used to quantitatively and site-specifically probe the electric fields of proteins, including those arising from hydrogen-bond (H-bond) interactions.

The utility of a vibrational probe to reliably and conveniently measure local electric fields in proteins is evaluated by how well it meets several criteria. First and foremost, its frequency must show a sensitivity to and quantifiable dependence on the local electric field. Also, a chemical or biological method must exist to incorporate the probe into a protein. Furthermore, it must minimally perturb the native chemical and structural environment of interest. Finally, its vibration must be a localized mode with a large cross-section and, ideally, be located in an uncongested region of the IR spectrum of proteins (e.g., 1700–2400 $cm^{-1}$).

For naturally occurring proteins, the amide I vibration arising from backbone amide units offers the largest IR intensity (molar extinction coefficient: ca. 800 $M^{-1}$ $cm^{-1}$) and shows a strong dependence on the local electrostatic environment, such as hydration. As a result, it has been widely used to investigate protein conformational transitions.[138,181,337-339] However, the amide I transition is generally delocalized and also contains contributions from other vibrational modes, thus making its use as a stand-alone probe of local electric field rather challenging. One viable strategy to overcome this limitation is to incorporate a single carbonyl group (C═O) into an amino acid side chain. A computational study by Choi and Cho predicted that the stretching mode of such a carbonyl group is not only localized, but its frequency also varies linearly with the electrostatic field for both H-bonding and non-H-bonding environments,[84] thus making it an ideal candidate for the aforementioned applications. Indeed, Boxer and co-workers[340] recently showed that the C═O stretching frequency of *p*-acetyl-L-phenylalanine (*p*-Ac-Phe) can serve as a reporter of the local electrostatic field of proteins. However, this vibrational transition (at ca. 1673 $cm^{-1}$) overlaps with the protein amide I band, and thus its application requires careful background subtraction by using the wild-type protein. To circumvent this inconvenience, we propose the use of the C═O stretching vibration of an ester moiety as an alternative probe. Previous studies[341-342] have shown that the ester carbonyl group absorbs in a spectral region (1700–1800 $cm^{-1}$) in which no other protein IR bands are present at a neutral pH value,[103] except those arising from protonated carboxylic groups.[343] Specifically, we tested the utility of two ester-containing non-

natural amino acids, l-aspartic acid 4-methyl ester (hereafter referred to as $D_M$) and l-glutamic acid 5-methyl ester (hereafter referred to as $E_M$). Although, to the best of our knowledge, $D_M$ and $E_M$ have not previously been introduced in proteins, we expected that it would not be a challenging task, as other ester-containing side chains have been successfully incorporated into proteins by genetic methods.[203,344]

## 6.3 Experimental Section

### 6.3.1 Sample Preparation

Methyl acetate (acetic acid methyl ester, MA) and methyl propionate (propanoic acid methyl ester, MP) were purchased from Acros Organics (Fair Lawn, NJ) and used as received. Fmoc-L-Asp(Me)-OH was purchased from Chem-Impex International Inc. (Wood Dale, IL) and Fmoc-L-Glu(Me)-OH was purchased from Santa Cruz Biotechnology Inc. (Fmoc=9-fluorenylmethoxycarbonyl). Peptides were synthesized by the use of standard Fmoc solid-phase methods on a PS3 peptide synthesizer from Protein Technologies (Tucson, AZ).

### 6.3.2 Spectroscopic Measurements

All Fourier transform infrared (FTIR) spectra were collected on a Thermo Nicolet Magna 860 FTIR spectrometer at 1 $cm^{-1}$ resolution equipped with a MCT detector at 25 ºC. The home-made IR sample holder consists of two $CaF_2$ windows and a 52 μm Teflon spacer divided into two compartments, one containing the sample solution and the other

containing only the solvent (i.e., reference). A motorized translation stage was used to alternately move the sample and reference sides of the sample holder in and out of the IR beam; each time a single-beam spectrum, corresponding to an average of 8 scans was collected for both the sample ($I_S$) and reference ($I_R$) sides; and the absorption spectrum of the solute, Ss(ν), was then calculated using $SS(\nu) = \log(I_R/I_S)$. The final reported result corresponds to an average of 32 such spectra and contains the absorbance of the solute only.[268,345]

2D IR spectra were collected using a setup that has been described in detail previously.[346] Briefly, three femtosecond IR pulses ($k_1$, $k_2$, and $k_3$) were brought to focus in the sample to induce a third order response (photon echo), which was heterodyned with the local oscillator, dispersed by a monochromator, and detected with a 64-element IR array detector. Rephasing and nonrephasing data were collected by scanning the time delay (τ) between $k_1$ and $k_2$ from -6 to +6.5 ps at 2 fs intervals. The waiting time (T) between pump ($k_1/k_2$) and probe ($k_3$) pulses was varied from 0 to 6 ps. The final spectrum, $S(\omega_\tau, T, \omega_t)$, was obtained from a double Fourier transform of S(τ, T, λ). For the current study only the real part of the result at T = 0 was shown.

### 6.3.3 Molecular Dynamic Simulations

MD simulations were performed using the molecular dynamics program NAMD2.7[347] and force field parameters from CHARMM22[348] and CGenFF CHARMM36 v. 2b7 for Small Molecule Drug Design.[349] Specifically, the solute of

interest was first immersed in a 35 Å cubic box containing 125 solvent molecules. Following an initial 1 ns equilibration run at 298 K and 1 atm in the NPT ensemble, a production run of 10 ns at 298 K in the NVT ensemble was carried out. The trajectory was saved every 500 fs, which resulted in 20,000 frames for each simulation. The subsequent EF calculation was performed using the VMD package.[350]

## 6.4 Results and Discussion

As a quantitative assessment of the electric-field dependence of the ester carbonyl stretching vibrations of $D_M$ and $E_M$, we first performed detailed vibrational solvatochromism studies on the respective side-chain mimics, methyl acetate (MA) and methyl propionate (MP). We chose MA and MP in these experiments, instead of the non-natural amino acids, because the model compounds are soluble in a wide range of polar and nonpolar solvents. The ester carbonyl stretching frequencies of MA and MP exhibited a strong dependence on the chosen solvent, ranging from hexane, an aprotic solvent with a very low dielectric constant (1.89 at 20 °C), to water (Figure 6.1; see also Table 6.1). For example, when the solvent was changed from hexane to DMSO, the center frequency of MA was red-shifted by 18.1 $cm^{-1}$, as compared to the 14.4 $cm^{-1}$ shift observed for p-Ac-Phe.[340] The MP in mixtures of $D_2O$ and acetonitrile (ACN) solutions (Figure 6.2), shows a series of spectra of MP experiencing several non-H-bonding and H-bonding environments. There is a gradual red shift with increasing water concentration in

the water/ACN solution, where a high ACN percentage, the peak arises from a population where the ester C═O is non-H-bonded. Additionally, it is important to note the dependence of the ester on the pH (Figure 6.3). At pH of 10.0 and above, the well-known base-catalyzed hydrolysis of an ester occurs. Thus, these results substantiate the utility of these ester carbonyl stretching vibrations as sensitive probes of the local electric field, provided that a quantitative relationship between the electric field and frequency can be determined.

Interestingly, in aprotic solvents the ester carbonyl stretching vibration results in a single absorption band, whereas in protic solvents, in which H-bonding between the vibrator and solvent is possible, the linear IR spectra contained more than one resolvable feature, thus suggesting that differently solvated or H-bonded species are present. Such spectral features have also been observed for nitrile and amide modes in protic solvents, such as methanol.[55,351] For example, in $D_2O$ the IR spectrum of MA contains to two distinct peaks, centered at 1703.6 and 1727.0 $cm^{-1}$, which is consistent with the two-dimensional IR spectroscopic study of Righini and co-workers.[352] In agreement with the study of Tominaga and co-workers,[353] the ester carbonyl stretching band of MA in methanol consists of three resolvable spectral features, centered at 1748.1, 1729.6, and 1708.1 $cm^{-1}$. For MP, however, the spectrum obtained in $D_2O$ is broad and almost featureless. To be able to better discern the underlying spectral contributions, we further carried out 2D IR spectroscopic measurements on MP in $D_2O$ and methanol. The 2D IR spectrum indicates that under the linear IR profile of MP in $D_2O$, two peaks, at 1703.1

and 1721.1 $cm^{-1}$, are present and represent two distinct species (Figure 6.4). With this information at hand, we then tried to determine how the ester C=O stretching frequencies of MA and MP vary with the local electric field.

### 6.4.1 Dependence on the Onsager Field

The Onsager reaction field model offers a convenient way to quantitatively describe the solvent-induced vibrational frequency shift of a solute molecule. The model consists of a spherical cavity in a continuous dielectric. The point dipole is situated in the center of the cavity. The calculated reaction field includes the potential in the cavity due both to the dipole itself and to the interaction of the dipole with the surrounding dielectric.

$$\nu_i(\mathrm{cm}^{-1}) \propto F^i_{\mathrm{Onsager}} \qquad \text{Equation 6.1}$$

where $\nu_i$ is the vibrational frequency of interest determined in solvent i and $F^i_{\mathrm{Onsager}}$ is the Onsager field of this solvent, which is defined as:

$$F_{\mathrm{Onsager}} = \left(\frac{\mu_0}{a^3}\right)\left[\frac{2(\varepsilon-1)(n^2+2)}{3(2\varepsilon+n^2)}\right] \qquad \text{Equation 6.2}$$

where $\varepsilon$ is the dielectric constant of the solvent, $n$ is the refractive index of the solute, and $\mu_0$ is the permanent dipole moment of the solute molecule in the gas phase. The constant $a$ is the Onsager cavity radius, which is related to the molecular volume of the solute molecule. In the current study, the following parameters were used: $n$(MA) = 1.361 and

$n$(MP) = 1.376 at 20 °C, $a^3$(MA) = 132 Å$^3$ mol$^{-1}$ and $a^3$(MP) = 160 Å$^3$ mol$^{-1}$, $\mu_0$(MA) = 1.9536 D and $\mu_0$(MP) = 1.7975 D (These gas-phase dipole moment parameters were calculated using DFT and the basis set B3LYP/6-31+G**). As shown (Figure 6.5), Equation 6.2 provides a satisfactory description of the vibrational frequencies obtained in aprotic solvents. As discussed in the main text, the frequency-Onsager field relationship thus obtained can be used to estimate the local dielectric constant of proteins.

This commonly used Onsager reaction field model[354] described fairly well the trend observed with aprotic solvents (Figure 6.5a), consistent with the study of Asbury and co-workers,[355] but failed to predict the frequency shifts induced by protic solvents (see Figure 6.5b). However, as shown (Figure 6.5b), Equation 6.1 is unable to describe the solvent-induced frequency shift of the ester carbonyl stretching vibration when all solvents are considered. This is expected, as the Onsager model does not account for specific electrostatic interactions, such as H-bonding, thus it is not applicable for protic solvents.

### 6.4.2 Electric Field Calculations from MD simulations

Therefore, following the work of Boxer and co-workers,[356] we used MD simulations to directly quantify the electric field experienced by the ester carbonyl vibration and to help assign the two C═O stretching bands observed in protic solvents (Figure 6.6). Briefly, for aprotic solvents the electric field was directly calculated by averaging the values obtained from about 20,000 frames from each MD simulation, whereas for protic solvents, because of the possibility of different H-bonding patterns, we

first divided the frames from each MD simulation into different clusters, according to a set of geometric criteria for H-bond formation, and then calculated the average electric field for each cluster.

For each aprotic solvent, the average local EF experienced by the ester C=O was calculated using the corresponding MD results and the method of Boxer and coworkers.[340] Specifically, for a given MD frame, the net electric fields on the carbonyl C and O atoms ($E_C$ and $E_O$), arising from all solvent partial charges within a sphere of 20 Å radius, were first determined. Afterwards, the average electric field ($\overline{E}$) along the carbonyl bond was calculated using: $\overline{E} = E_C + E_O$. Finally, the ensemble averaged electric field, $\langle \overline{E} \rangle$, was obtained by averaging values of $\overline{E}$ obtained from all MD frames. The results obtained for all aprotic solvents are summarized in Table 6.1.

For protic solvents, differently H-bonded ester C=O species are present. Thus, before calculating the EF, additional analysis was required to separate the MD frames into clusters, each of which corresponds to a specific H-bonding pattern. Specifically, we used the VMD H-bond plugin[357] to sort MD frames according to the following geometric H-bonding criteria: the donor-hydrogen-acceptor angle (<40˚) and distance (<3.2 Å) for the ester carbonyl ($O^1$) and a shorter distance (<2.6 Å) for the ester oxygen ($O^2$). Afterwards, the $\langle \overline{E} \rangle$ of each cluster was evaluated using the method described above. As shown (Figure 6.6), in methanol the MD frames for both MA and MP can be divided into three groups according to the number of H-bonds formed between the ester C=O and the solvent, i.e., 2 H-bonded ester C=O (8%), 1 H-bonded ester C=O (60%), and non-H-

bonded ester C=O (32%). This distribution is consistent with the IR spectra of MA and MP obtained in methanol, which contain two main bands and a shoulder that are well separated. As expected (Table 6.1), the H-bonded ester carbonyls experience a larger local electric field compared to non-H-bonded species. For water, the situation becomes more complicated as water can also form H-bonds with $O^2$ (Figure 6.6), however the relative percentages of these species, which serve to broaden the IR band, are small. Thus, for both MA and MP in water, only the electric fields of the two major species are reported in Table 6.1.

### 6.4.3 Ester Carbonyl Assignments

For MA in water, the majority of the carbonyl groups form either one (51 %) or two H-bonds (44 %) to water (Figure 6.6). Thus, we propose that the two peaks observed in the linear IR spectrum arise predominantly from these two species. Because the H-bonding of water to a carbonyl group induces a redshift in the carbonyl stretching vibration, we attribute the lower-frequency component to the doubly H-bonded species and the higher-frequency peak to the singly H-bonded species. Furthermore, the percentages of the lower- and higher-frequency components of the C═O stretching vibration of MA in $D_2O$ (Figure 6.1), as calculated on the basis of the integrated peak areas, are 55 and 45 %. This result provides further evidence in support of the above assignment. MD simulations revealed the existence of less-populated but differently H-bonded species, which were assumed to contribute to the broad width of the spectrum. A similar observation and assignment was made for MP in water. For both compounds in

methanol, the majority of the carbonyl groups were found to be either non-H-bonded or singly H-bonded, with relative percentages in agreement with the ratio of the integrated areas of the two IR peaks. Thus, we assigned the two major IR bands to these species, with the H-bonded carbonyl groups vibrating at a lower frequency. For MA, a minor band at approximately 1710 $cm^{-1}$ is clearly observable. On the basis of MD simulations, this band was attributed to doubly H-bonded carbonyl groups. Finally, the width of the calculated electric-field distribution for each differently solvated species shows a correlation with the corresponding spectral width of the C═O stretching vibration (see Figure 6.7), thus suggesting that the MD simulations are able to exhaustively sample the heterogeneous electrostatic environments of the probe.

The center frequencies of the ester carbonyl stretching vibrations of MA and MP show a linear dependence on the calculated electric field for both protic and aprotic solvents (Figure 6.8), thus indicating that an ester moiety, such as that in $D_M$ and $E_M$, could be used to quantitatively determine the local electrostatic field of proteins by the use of this frequency–field map.

### 6.4.4 Ester in Peptide Environment

To demonstrate the utility of this ester vibrational mode in biological applications, we first used $D_M$ and $E_M$ to probe the local electrostatic and/or hydration environment of two short peptides, Ac-Y$D_M$K-$NH_2$ (hereafter referred to as $D_M$-P) and Ac-Y$E_M$K-$NH_2$ (hereafter referred to as $E_M$-P). The ester carbonyl stretching bands of these peptides in $D_2O$ indicate, when compared to those of MA and MP, that the population of the doubly

H-bonded species (i.e., the spectral intensity at ca. 1705 $cm^{-1}$) is significantly decreased (Figure 6.9). This result is not surprising as, in comparison to their respective model compounds, the side chains of $D_M$ and $E_M$ are expected to be situated in a more crowded environment, thus limiting the accessibility of water molecules to the ester carbonyl group and hence decreasing the probability of the formation of two H-bonds. Furthermore, and perhaps more convincingly, $E_M$-P, the ester carbonyl group of which is expected to be further extended into the solvent than that of $D_M$-P, shows a smaller decrease in this regard. Thus, these results provide further validation of the sensitivity of the C═O stretching vibration of the ester moiety to its local electrostatic environment. In support of this notion, Xie and co-workers have shown that the native structural analogues of $D_M$ and $E_M$, that is, the protonated carboxylic acid side chains of Asp and Glu, can be used to sense H-bond formation in proteins.[343] The spectra of these two peptides in DMSO and methanol also support this notion. For example, in DMSO the center frequency of the ester C═O stretching band of EM-P is red-shifted by 3.4 $cm^{-1}$ from that of MP, whereas that of $D_M$-P is similar to that of MA. This redshift results from the addition of the peptide environment around the vibrational probe. Besides the solvent-induced electric field, the ester carbonyl group will also experience electrostatic forces arising from the peptide backbone and other amino acid side chains, thus making the vibrational frequency dependent on the position of the amino acid in the peptide. In other words, the redshift observed for $E_M$-P is most likely due to the closer (as compared to $D_M$-P) proximity between the ester carbonyl group and the polar amine group of the

lysine side chain. The observed non-hydrogen-bonded peaks of these peptides in methanol also corroborate this picture (Figure 6.9). Thus, these results demonstrate the ability of the ester carbonyl stretching vibration to sense minimal changes in its local electrostatic environment.

## 6.5 Conclusion

In conclusion, we have established that the ester carbonyl stretching frequencies of two non-natural amino acids, l-aspartic acid 4-methyl ester ($D_M$) and l-glutamic acid 5-methyl ester ($E_M$), show a linear dependence on the local electric field and, thus, can be used to quantify, in a site-specific manner, the local electrostatic environment of proteins. In comparison to commonly used nitrile-, azide-, and CD-based IR probes,[116,145,177-178,180,205,358-363] the ester carbonyl stretching vibration offers one distinct advantage: its large dynamic range makes it more useful for probing small changes in the local electric field. For example, it is sensitive enough to probe the difference in the electric fields between two points in a peptide environment that are separated by a single methylene unit. Furthermore, the size of $D_M$ is similar to that of asparagine, aspartic acid, and leucine, whereas the size of $E_M$ is similar to that of glutamine and glutamic acid. Taken together, these attributes of $D_M$ and $E_M$ suggest that they are two of the most promising local electrostatic IR probes of proteins. We have also devised a new method that enables the determination of the local dielectric constant of proteins. Because the C═O stretching

vibration of esters is also Raman-active,[364] we expect that the frequency–field relationships devised herein can also be used to study relevant biochemical and biophysical problems in conjunction with techniques based on Raman spectroscopy.[365]

## 6.6 Original Publication

This Chapter has been adapted and reprinted from Angewandte Chemie International Edition, Ileana M. Pazos, Dr. Ayanjeet Ghosh, Prof. Dr. Matthew J. Tucker, and Prof. Dr. Feng Gai, (2014) 53, 24 6080–6084. DOI: 10.1002/anie.201402011, with permission from John Wiley and Sons.

Table 6.1. Summary of the center frequencies and electric fields of MA and MP in different solvents.

| | MA / MP | | |
|---|---|---|---|
| | Center Frequency ($cm^{-1}$) | Electric Field ($MV\ cm^{-1}$) | Standard Deviation |
| Hexane | 1755.9 / 1751.8 | -0.8 / -0.7 | 0.7 / 0.8 |
| Diethyl Ether | 1751.2 / 1747.8 | -4.9 / -4.5 | 1.4 / 1.8 |
| Acetonitrile (ACN) | 1742.0 / 1738.8 | -11.4 / -10.4 | 2.3 / 2.6 |
| DMSO | 1738.2 / 1736.0 | -14.0 / -13.2 | 3.5 / 3.8 |
| $d_4$-Methanol (non-H-bonded) | 1748.1 / 1744.7 | -8.6 / -7.7 | 5.4 / 5.3 |
| $d_4$-Methanol (1-H-bonded) | 1729.6 / 1724.6 | -22.9 / -21.2 | 6.2 / 5.7 |
| $d_4$-Methanol (2-H-bonded) | 1708.1 / 1709.9 | -37.2 / -34.5 | 7.8 / 5.4 |
| $D_2O$ (1-H-bonded) | 1727.0 / 1721.1 | -26.3 / -25.3 | 9.5 / 8.9 |
| $D_2O$ (2-H-bonded) | 1703.6 / 1704.9 | -40.7 / -36.8 | 8.1 / 8.0 |

Table6.2. Summary of the full-width at half maximum (FWHM) of the ester C═O IR band and standard deviations the calculated electric field ($\sigma\langle_{\vec{F}}\rangle$) of MA and MP in a series of solvents.

| | MA / MP | |
|---|---|---|
| | FWHM ($cm^{-1}$) | $\sigma_{\langle\vec{F}\rangle}$ (MV $cm^{-1}$) |
| Hexane | 12.0 ± 1.0 / 12.0 ± 0.4 | 3.1 ± 0.6 / 2.8 ± 0.6 |
| Diethyl Ether | 14.0 ± 1.0 / 12.5 ± 0.7 | 4.0 ± 0.8 / 3.4 ± 0.7 |
| Acetonitrile (ACN) | 12.4 ± 1.2 / 13.6 ± 0.1 | 5.6 ± 1.1 / 5.7 ± 1.1 |
| DMSO | 14.3 ± 1.0 / 16.5 ± 0.7 | 7.0 ± 1.4 / 6.9 ± 1.4 |
| $d_4$-Methanol (non-H-bonded) | 12.5 ± 1.5 / 14.0 ± 1.3 | 5.3 ± 1.1 / 5.1 ± 1.0 |
| $d_4$-Methanol (1-H-bonded) | 15.6 ± 1.3 / 17.7 ± 1.0 | 6.2 ± 1.2 / 5.9 ± 1.2 |
| $D_2O$ (1-H-bonded) | 19.2 ± 1.5 / 32.5 ± 1.6 | 8.4 ± 1.7 / 8.0 ± 1.6 |
| $D_2O$ (2-H-bonded) | 12.0 ± 1.1 / 28.4 ± 1.3 | 10.0 ± 1.8 / 8.7 ± 1.7/ |

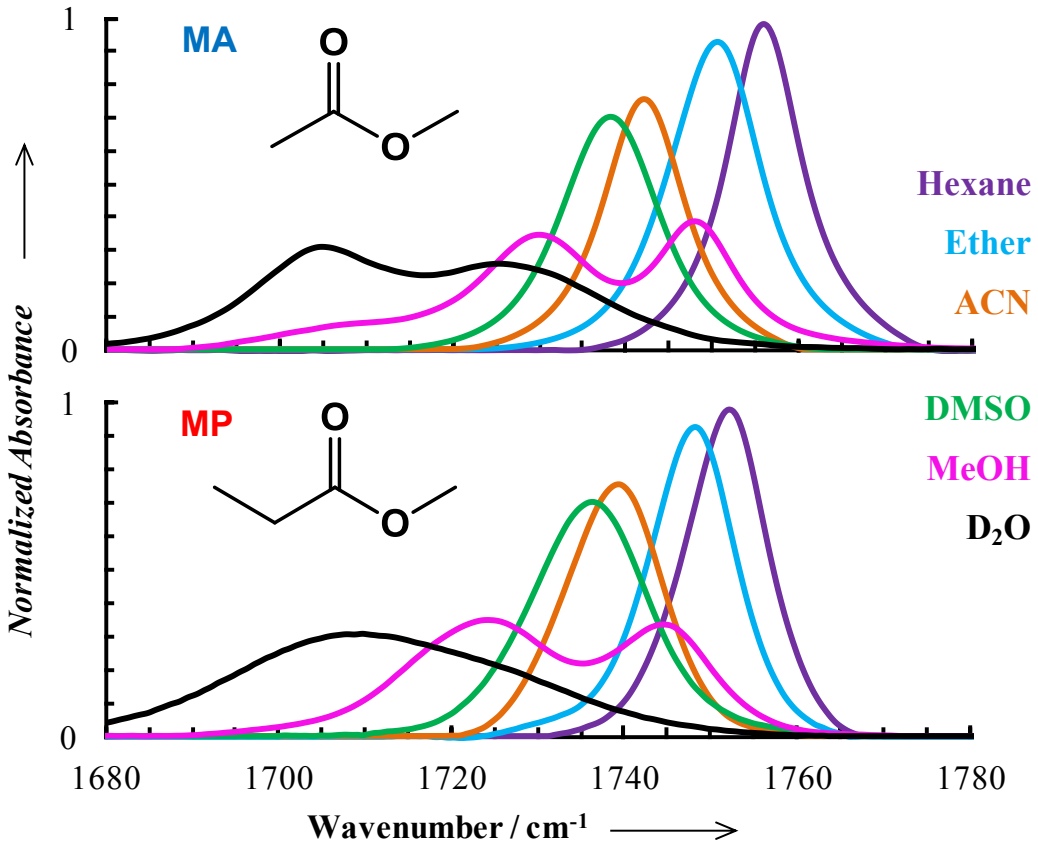



Figure 6.1 Normalized FTIR spectra of MA and MP in different solvents, as indicated. The concentration of the solute in each case was 20 mM, and normalization is based on the integrated area of the band observed in hexane (i.e., the spectra collected in other solvents were scaled so that their integrated areas were equal to that observed in hexane). For MA in hexane, the peak absorbance was measured to be 0.0715, which corresponds to a molar extinction coefficient of 650 $M^{-1}$ $cm^{-1}$.

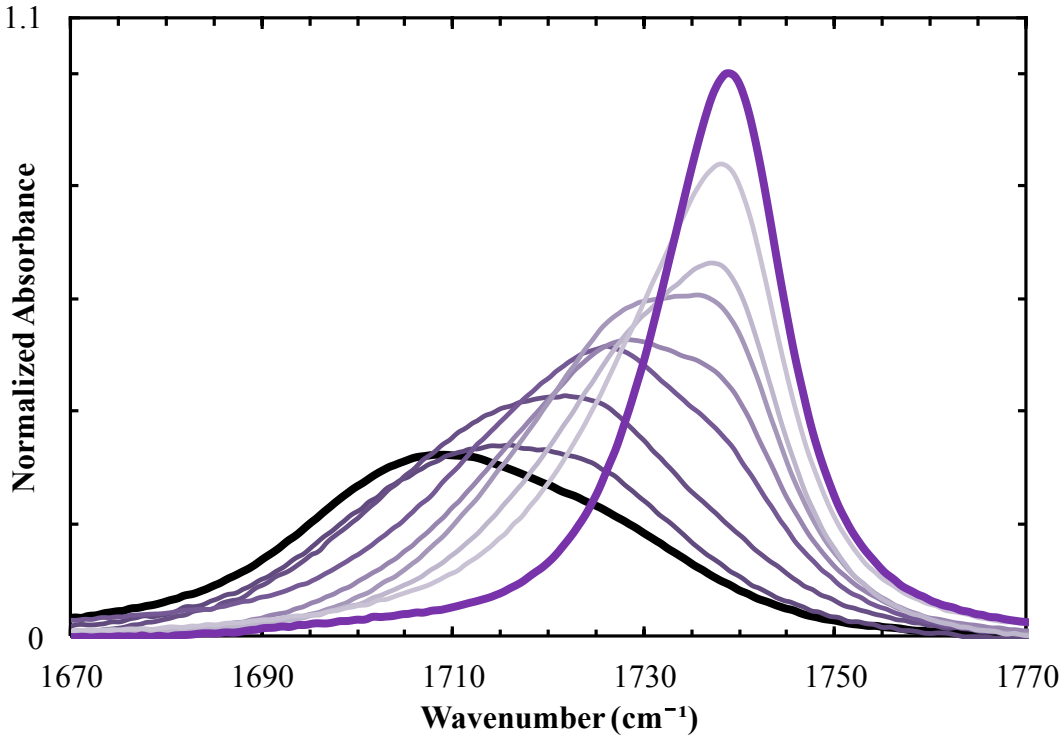



Figure 6.2 Normalized FTIR spectra of MP in mixtures of $D_2O$ and acetonitrile. The thick purple is MP in acetonitrile while the thick black is MP in $D_2O$. The thinner lines are MP in 25, 40, 60, 75, 90 and 95% (v/v) $D_2O$-acetonitrile mixtures.

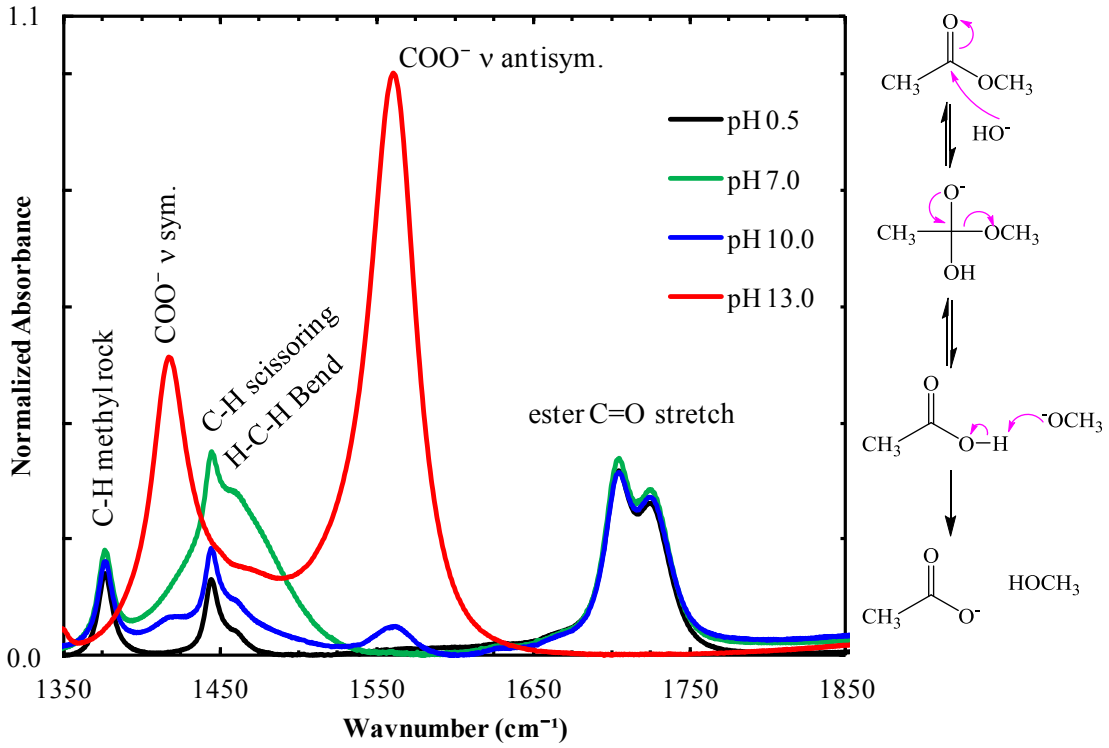



Figure 6.3 Normalized FTIR spectra of MA in of $D_2O$ at different pH levels. At or below pH 7.0 (black and green lines), the spectra are very similar and are characteristic to the spectra of MA in $D_2O$. However, at pH 10.0 (blue line) a new peak at ~1550 $cm^{-1}$ that can be assigned to a $COO^-$ stretching transition appears. At pH 13.0 (red line), methyl acetate reacts to acetic acid via the suggested mechanism.

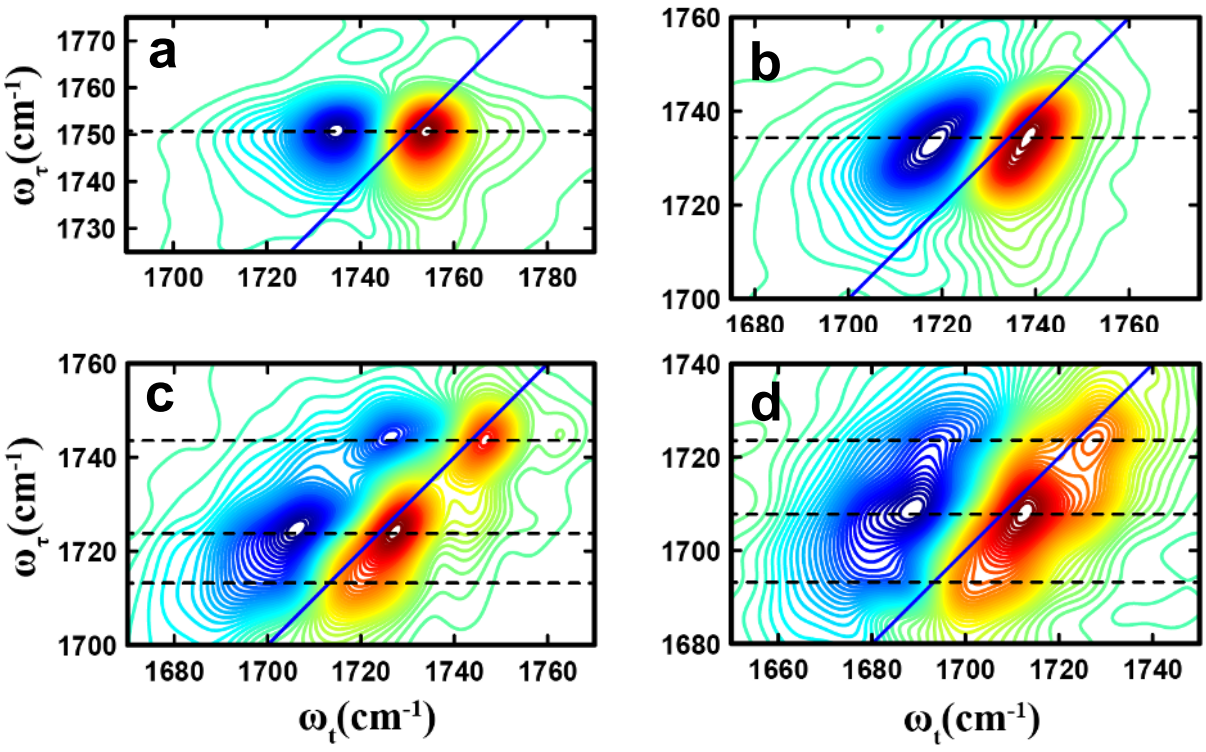



Figure 6.4 2D IR spectra of MP in a) hexane, b) DMSO, c) methanol, and d) water collected at zero waiting time (T). It is clear that for MP in hexane and DMSO, there is only one resolvable peak, consistent with the linear IR measurements. Similarly, the 2D IR spectrum in methanol, which consists of three peaks, is also in agreement with the FTIR spectrum. On the other hand, there are three peaks found in the 2D spectrum of MP in water. While the two stronger peaks at higher frequencies coincide with those observed in the linear IR spectrum, the low frequency transition is due to MP dimerization and thus is not discussed in the text.

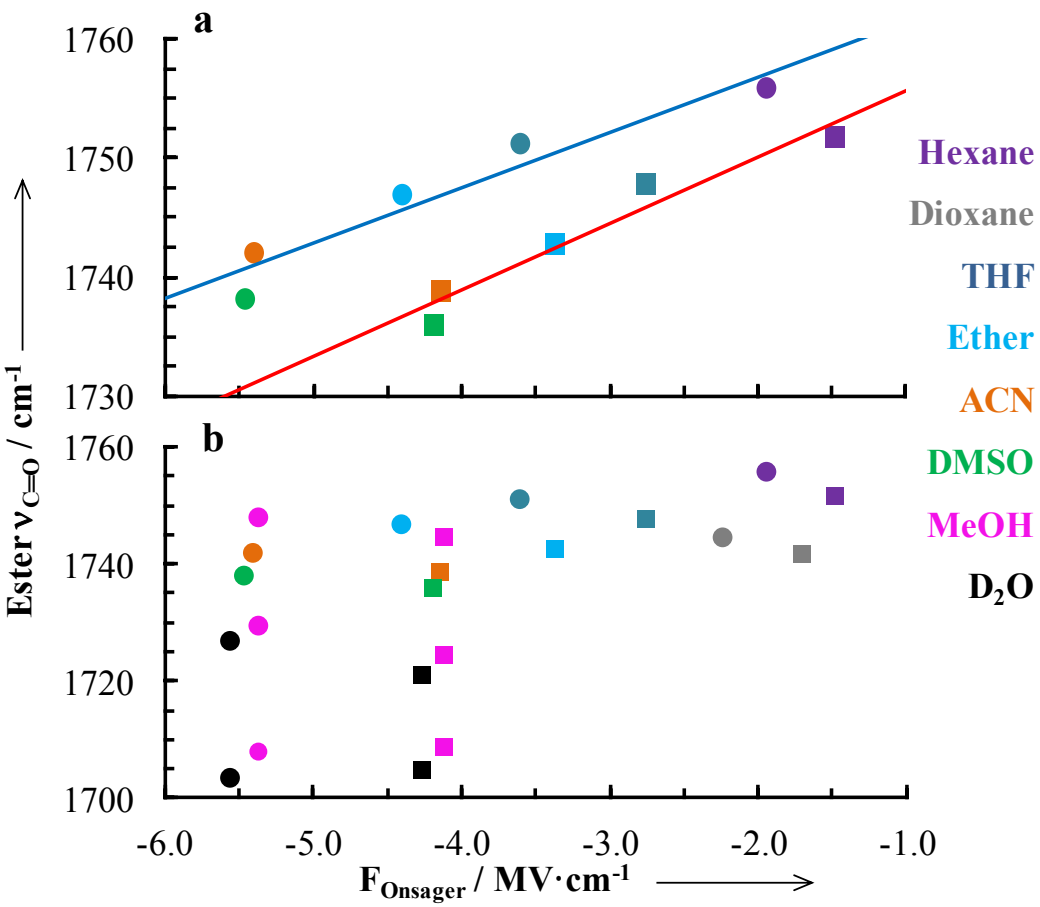



Figure 6.5 The C═O stretching frequencies of MA (circles) and MP (squares) versus the Onsager field of the solvent, as indicated. Fitting each data set to a straight line yields $\nu_{C=O}$ (cm$^{-1}$) = 4.66 $F_{Onsager}$ + 1766.3 for MA and $\nu_{C=O}$ (cm$^{-1}$) = 5.57 $F_{Onsager}$ + 1761.2 for MP.

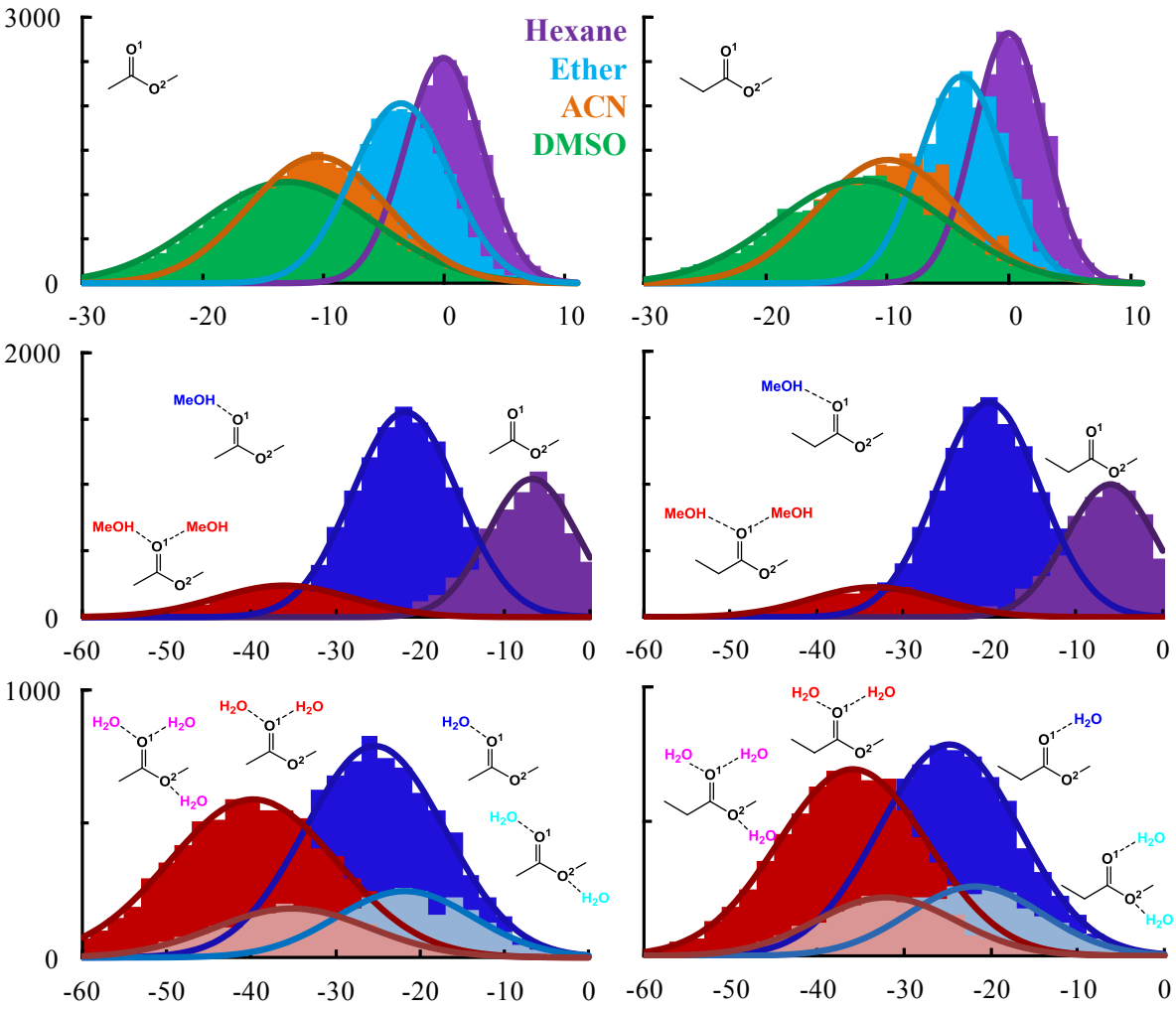



Figure 6.6 Calculated electric field distributions along the ester C═O bond of MA and MP in different solvents, as indicated. Calculated distributions of the electric field along the ester C═O bond for differently H-bonded species of MA (a & c) and MP (b & d) in water and methanol, as indicated.

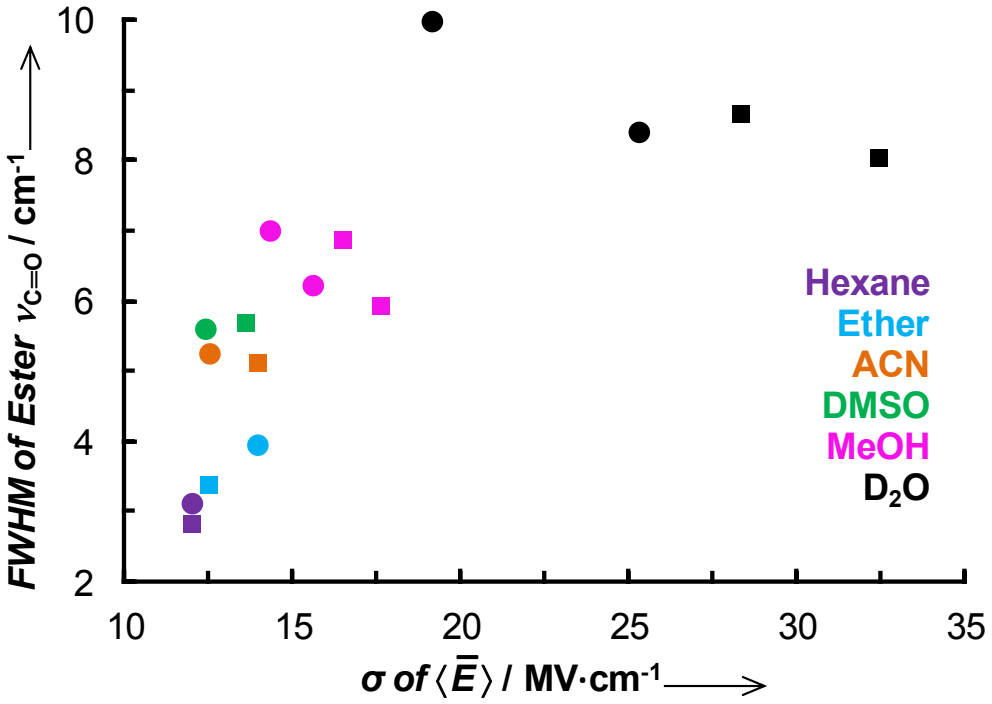



Figure 6.7 Correlation between the spectral width (FWHM) of the ester C=O of MA (circles) and MP (squares) is related to the standard deviation of the calculated electric field distribution for solvents, as indicated.

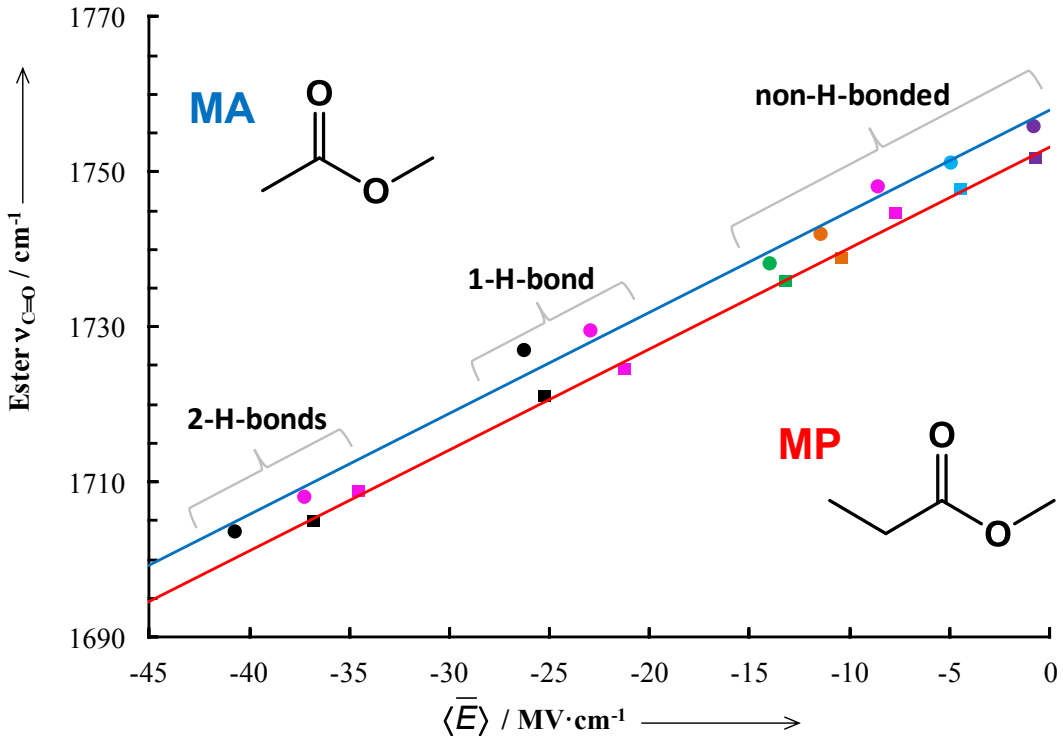



Figure 6.8 Center frequencies of the carbonyl stretching vibrations of model ester compounds versus the calculated local electric field for different solvents MA (circles) and MP (squares) are represented by the same colors as those used in Figure 6.1. The solid lines are the best fits of these data to the linear equations indicated in the figure.

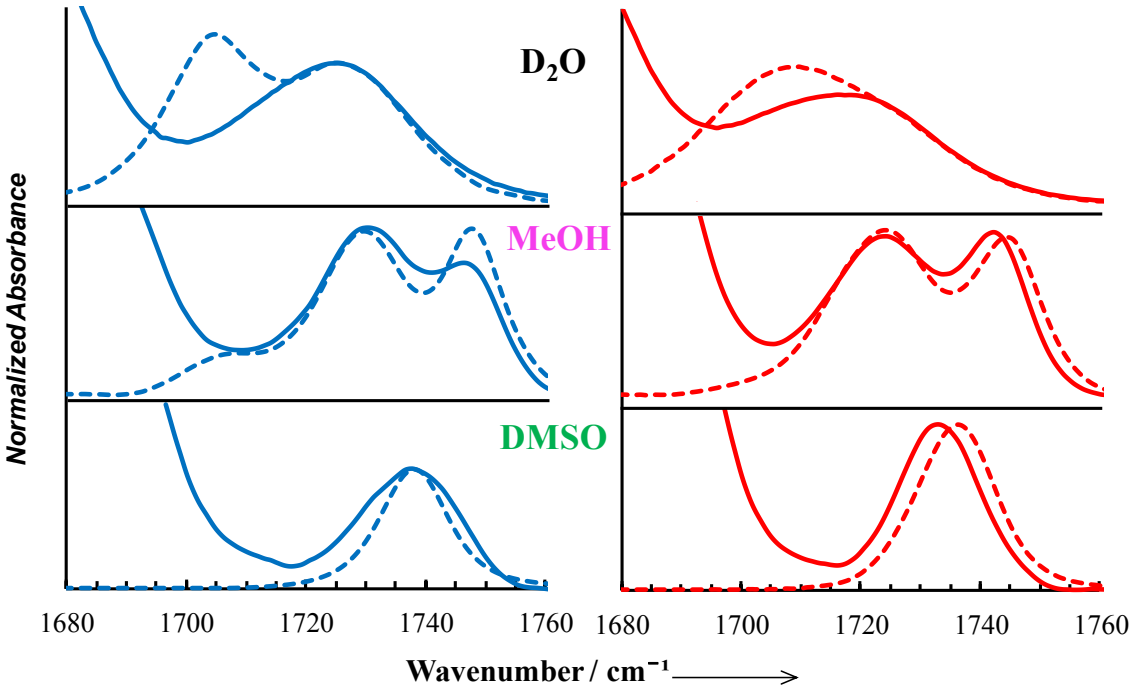



Figure 6.9 FTIR spectra in the ester C═O stretching region of $D_M$-P, MA, $E_M$-P, and MP in different solvents, as indicated. The peptide concentration was 2 mM, and in each case, the spectrum of the model compound has been normalized with respect to that of the corresponding peptide.

# 7 Structural and Dynamic Study of Amyloids using the Ester Probe

## 7.1 Introduction

Amyloids are self-aggregated long fibrils that are the consequence of misfolded proteins.[366] The accumulation of amyloid fibrils in organs is connected with over twenty fatal human diseases such as Alzheimer's, type 2 diabetes, and transthyretin amyloidosis.[367-369] However, for material science, these nanostructures can be useful for technological applications since they are known to be strong, flexible, and resistant to degradation.[370-371] Although proteins of different structures and sequences are capable of aggregating, the resultant fibrils often have common characteristics and are mostly composed of stacked β-sheets perpendicular to the fibril axis. It has been suggested that the first step of aggregation is a dehydration process.[372-376] This is accompanied by the formation of the interfaces that include hydrophobic or polar sidechains between tightly interdigitated β-sheets, thus they are referred to as a dry interface or a steric zipper.[377-378] Both *in vitro* and *in vivo* studies have been aimed to understand the structure and molecular mechanism(s) of amyloid formation.[379-391] While various factors, like solvent, temperature, pH, concentration, and sequence, play an important role in the structure and rate of aggregation,[392-401] the precise role of water is not fully understood. Therefore, while it is unlikely for amyloid fibrils to contain bulk water,[402-403] there is evidence of trapped water molecules.[302] This can be difficult to detect site-specifically especially inside the core of the fibril which are tightly packed. An ideal approach to this problem

would be the incorporation of a probe during the aggregation process that is sensitive to hydration status.

Therefore, a site-specific probe, such as L-aspartic acid 4-methyl ester ($D_M$), described in Chapter 6, is an appropriate candidate to study the structure and dynamics of amyloids. The objective is to use this IR probe to study key kinetic events in amyloid formation, for example dehydration.[404] It is well known that the amide I' band is sensitive to peptide/protein aggregation[405] and that it can be utilized to follow the aggregation kinetics of several amyloidogenic peptides.[406] We have shown that the ester C=O stretching vibration can be used to monitor the local hydration status. This is because the frequency and lineshape of the ester transition is sensitive to the number and strength of hydrogen bonds (H-bonds) that are formed between the H-bond donor and the ester group. In contrast, the IR signal arising from the ester probe will allow us to monitor the aggregation and fibrilization kinetics of these mutants in a site-specific manner. Therefore, incorporating the ester probe,[176] that is sensitive to electric field and hydration, into amyloid systems will allow us to obtain structural information and microscopic dehydration rate constants which are critical to a quantitative description of the mechanism of amyloid formation.

## 7.2 Experimental

### 7.2.1 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, $K^*$, was purchased from Anaspec, Inc. (Fremont, CA) and used without further purification. Fmoc-protected amino acids were purchased from Advanced Chem Tech (Louisville, KY). Rink amide resin was purchased from Novabiochem (San Diego, CA). All peptidess were synthesized using a standard fluoren-9-ylmethoxycarbonyl (Fmoc) chemistry protocol on a PS3 peptide synthesizer from Protein Technologies (Tuscon, AZ), and purified by reverse-phase HPLC (1100 series; Agilent Technologies, Santa Clara, CA). The identity of the peptides was verified using MALDI-TOF mass spectrometry (Voyager-DE RP, Applied Biosystems, Foster City, CA).

### 7.2.2 Linear IR Measurments

All Fourier transform infrared (FTIR) spectra were collected on a Thermo Nicolet Magna 860 FTIR spectrometer at 1 $cm^{-1}$ resolution equipped with a MCT detector at 25 ºC. The home-made IR sample holder consists of two $CaF_2$ windows and a 52 μm Teflon spacer divided into two compartments, one containing the sample solution and the other containing only the solvent (i.e., reference). A motorized translation stage was used to alternately move the sample and reference sides of the sample holder in and out of the IR beam; each time a single-beam spectrum, corresponding to an average of 8 scans was

collected for both the sample ($I_S$) and reference ($I_R$) sides; and the absorption spectrum of the solute, $S_s(\nu)$, was then calculated using $S_s(\nu) = \log(I_R/I_S)$. The final reported result corresponds to an average of 32 such spectra and contains the absorbance of the solute only.

### 7.2.3 Attenuated Total Reflectance (ATR) Measurements

The ATR IR spectrum of the dry film was collected on a Thermo Nicolet 6700 FTIR spectrometer using a Harrick Horizon multiple-reflection attachment. The film was prepared by spreading 10 μL of a 15 mM aggregated $D_2O$ solution over the surface area of the Harrick germanium crystal with the specifications of 45º, 50 x 10 x 2 mm (Model #EJ2122). The films were dried under a flow of nitrogen for 7 days. An average of 128 scans was taken with a resolution of 1 $cm^{-1}$.

### 7.2.4 Imaging

AFM experiments were performed in air at room temperature, using a multimode atomic force microscope (Bruker Dimension Icon AFM, Billerica, MA). Five microliters of sample solution was applied to a freshly-cleaved mica surface for ~5 s, 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. Multiple images were obtained for the sample at different locations on the mica substrate.

### 7.2.5 2D IR

2D IR spectra were collected using a setup that has been described in detail previously.[346] Briefly, three femtosecond IR pulses ($k_1$, $k_2$, and $k_3$) were brought to focus in the sample to induce a third order response (photon echo), which was heterodyned with the local oscillator, dispersed by a monochromator, and detected with a 64-element IR array detector. Rephasing and nonrephasing data were collected by scanning the time delay ($\tau$) between $k_1$ and $k_2$ from -6 to +6.5 ps at 2 fs intervals. The waiting time (T) between pump ($k_1/k_2$) and probe ($k_3$) pulses was varied from 0 to 6 ps. The final spectrum, $S(\omega_\tau, T, \omega_t)$, was obtained from a double Fourier transform of $S(\tau, T, \lambda)$.

## 7.3 Results and Discussion

Two amyloid systems were chosen for this study, a segment of the β-amyloid protein KLVFFAE (i.e., $A\beta_{16–22}$) and a segment of the transthyretin protein YTIAALLSPYS (i.e., $TTR_{105\text{-}115}$). The non-natural amino acid $D_M$ was used to characterize the electric fields and hydration environments in model amyloid fibrils.

### 7.3.1 Aβ-$D_M$

In this first example, we used $D_M$ to quantify the electrostatic environment in amyloid fibrils formed by a short segment of the β-amyloid peptide, KLVFFAE (i.e., $A\beta_{16–22}$), which is related to Alzheimer's disease. Although it is generally assumed that

the interior of amyloid has a low dielectric constant, to the best of our knowledge, no experiments have been attempted to directly measure the electrostatic properties of such β-sheet assemblies. In this study, we mutated the leucine residue at position 17 of $A\beta_{16–22}$ to $D_M$ (the resultant peptide is referred to as $A\beta\text{-}D_M$), since the side chains are similar in size. Prior to full onset of peptide aggregation (as judged by the amide I band at 1625 $cm^{-1}$), the ester band had a peak at approximately 1725 $cm^{-1}$ (Figure 7.1), thus indicating that the $D_M$ side chain was mostly hydrated and formed one H-bond with water, as expected. Upon further aggregation, the ester band became broader and was blueshifted, thus indicating that the population of the H-bonded ester carbonyl groups decreased. Further analysis indicated that this band can be decomposed into two Gaussian functions with center frequencies at 1727.2 and 1743.7 $cm^{-1}$ (Figure 7.1b). This result is consistent with the formation of β-sheet fibrils, which, according to the $A\beta_{16–22}$ fibrillar structures determined by Eisenberg and co-workers,[357] would lead to the creation of two distinct environments for the $D_M$ side chain: one in which the side chain is sequestered in a dehydrated interface, and the other in which it is exposed to solvent (Figure 7.2).

This structural model, proposed based on the crystal structures of $A\beta_{16\text{-}21}$ reported by Eisenberg and coworkers[357] (PDB ID code 2Y2A), showed that the basic fibril unit consists of two β-strands that are packed together via hydrophobic interactions, thus forming a dry interface. Assuming that this basic unit further assembles to form protofibrils that have an interior that contains water, as discussed by Thirumalai, Straub and coworkers,[407-408] we would expect the $D_M$ sidechains (or the native leucine

sidechains) in the fibrils to be distributed in three different environments, i.e., 50% in the dry interface, 25% exposed to bulk solvent, and 25% in the water-filled interior of the fibril. Upon removal of the bulk water, this model suggests that only approximately 25% of the $D_M$ sidechains can still experience interactions with water (i.e., those trapped in the interior of the fibril), as observed (Figure 7.1). However, as revealed by AFM measurements (Figure 7.3), the fibrils/aggregates thus formed are rather heterogeneous. This heterogeneity not only leads to a broad ester C═O stretching band, but also prevents a more quantitative assessment of the structural features of the fibrils on the basis of their IR signals.

To directly observe the buried ester sidechains within the fibril, it is necessary to remove the signal due to the hydrated surface-exposed ester sidechain which can be done by removing all the bulk water thus drying the sample. An aliquot of the abovementioned aggregated Aβ-$D_M$ sample was placed on the surface of a germanium crystal of an ATR unit and allowed it to dry under a gentle flow of $N_2$ for 7 days. Ester C═O stretching vibrations were observed as three well-resolved peaks at 1722.3, 1736.0, and 1747.6 $cm^{-1}$ in the IR spectrum of the resulting film (Figure 7.1). The lowest-frequency peak (at 1722.3 $cm^{-1}$) coincides with that due to singly H-bonded ester carbonyl groups, thus indicating that water is indeed present in the amyloid fibrils. On the other hand, the peak at 1747.6 $cm^{-1}$ must arise from non-H-bonded ester carbonyl groups, or those situated at the aforementioned dry interfaces (Figures 7.1 and 7.2), whereas the peak at 1736.0 $cm^{-1}$ most likely corresponds to outward-facing $D_M$ side chains that are H-bonded with water

prior to sample drying, and its frequency reflects the local electrostatic environment of the dry air/fibril interfaces. In support of these assignments, the relative percentages of these peaks, as determined from their integrated areas, are 28 (1722.3 $cm^{-1}$), 31 (1736.0 $cm^{-1}$), and 41 % (1747.6 $cm^{-1}$), which are similar to those (25, 25, and 50 %) calculated on the basis of the structural model of Eisenberg and co-workers.[357]

By using the frequency–field relationship of an ester model compound,[176] the three peaks of the Aβ-$D_M$ (in order of increasing frequency) give rise to the following local electric fields: −26.3, −16.2, and −8.0 MV $cm^{-1}$. More importantly, using the frequency–Onsager field relationship[176] the dielectric constant is experimentally determined to be 5.6 for the dry interior of the well-packed fibrils.

### 7.3.2 Aβ-$D_M$-Cage

In this work, we used a double mutations of the $A\beta_{16–22}$ motif, $KD_MVK^*FAE$ (hereafter refer as Aβ-$D_M$-Cage) to form fibrils. Incorporation of the cage ($K^*$), a photolabile lysine analog, has been shown to promote nucleation through hydrophobic interactions and results in very long and uniform fibrils.[409-410] This study aims to examine the side chain packing by characterizing the ester carbonyl stretch using both linear IR and 2D IR measurements. As shown in Figure 7.4, several features are observed in the solution and dry FTIR measurements of the Aβ-$D_M$-Cage. The fibril was formed at acidic pD conditions, but the absorbance at ~1715 $cm^{-1}$ indicates that most glutamic acid (E) side chains are deprotonated, which strongly suggest a presence of the salt bridge in the fibril.[411] This means that the ester carbonyl stretch region is nearly free from the

contamination of the protonated side chain. To confirm the presence of long uniform fibrils, an AFM image, Figure 7.5, of the fibrils was obtained. This dry film shows very homogenous and uniform fibril morphology. The height of the fibrils range between 15 to 20 nm (traces 1 and 2). There is also a periodic pitch that is observed along the fibril that has an average pitch distance of ~150 nm (traces 3 and 4).

The ester carbonyl stretch region of the fibril, Figure 7.4b, shows significant local environment changes in the fibril from solution to dry film. In solution, the ester carbonyl shows a single transition that is centered at 1723.8 $cm^{-1}$, which indicates that most ester side chains are forming a single H-bond. This is in contrast to the dry film, where the ester carbonyl shows two transitions that are centered at 1725.5 $cm^{-1}$ and 1739.4 $cm^{-1}$. The time dependent FTIR of the Aβ-$D_M$-Cage fibrils during the drying process is shown in Figure 7.6. The isosbestic point observed near 1734 $cm^{-1}$ during the drying process of the fibril suggests that the singly H-bonded ester side chain, which is dominant in the solution state, slowly shifts to a non-H-bonded ester group. After the film is dry, there is a 1:1 ratio of populations of singly- and non-H-bonded ester. The correlation between the ester band lineshape and the drying process, which is associated with the water molecules being expelled from the fibril, strongly indicates that there is a structural change in side chain packing when the fibril is dried. The amount of water molecules in the dry film fibril is undetectable, given the absorbance of the OD stretching region, however, the ester group shows a distinct singly-H-bonded population. It is unlikely that the majority of the singly-H-bonded population originates from a H-bond between water and the ester

probe. Therefore, the ester is assumed to be H-bonding to a nearby side chain, most likely a group from the cage. Also, there is an observed blueshift of the ester in solution compared to the dry film, which suggests a weaker single H-bond. This could be interpreted as H-bonds of different origins meaning, while in solution the ester is primarily forming H-bonds with water while in the film the ester is forming weaker H-bonds with the cage.

### 7.3.3 TTR-110$D_M$

The second amyloid system studied here is an 11-residue peptide corresponding to residues 105−115 of transthyretin.[412] Dobson *et al*. has shown that this peptide, $TTR_{105-115}$ (sequence: YTIAALLSPYS), can form fibrils with three possible polymorphs that all contain nanochannels of confined water.[407,413-415] All the polymorphs consist of tightly packed cross-β spine interfaces that are devoid of water, and a small water channel exists between the slightly hydrophilic protofilament surfaces as shown in Figure 7.7.[412] This trapped water is predicted to exhibit dynamics that are different from bulk water.[416] To probe the H-bonding dynamics of the confined water inside fibrils, two ester mutants of the $TTR_{105-115}$ were examined, TTR-110$D_M$ and TTR-111$D_M$ (Figure 7.7). To do this, the spectral diffusion dynamics were measured of the ester C=O vibration. Based on the fibril structures of $TTR_{105-115}$,[412] the TTR-110$D_M$ mutation will place the $D_M$ sidechain in both a dehydrated environment and at a position facing the exterior of the fibril. Meanwhile, the TTR-111$D_M$ also has an ester sidechain in the dehydrated environment but an ester sidechain is also facing the water channel (Figure 7.7). As shown in the AFM

image (Figure 7.8), this mutant forms fibrils that are similar to those formed by the wild-type peptide and the ester C=O band clearly indicates, as expected, that the $D_M$ residue samples two distinct environments, one with the presence of water (1730 $cm^{-1}$) and the other one without (1747 $cm^{-1}$).

As shown in Figure 7.9, the early stage of aggregation of the TTR-110$D_M$ is mostly comprised of the monomer that is solvated by water, indicated by a broad ester band centered at 1706.7 $cm^{-1}$, representing a doubly-H-bonded species. The changes in the 2D IR spectra of the doubly-H-bonded ester (1706.7 $cm^{-1}$), obtained at different waiting times (T), indicates the fluctuations of the solvation shell around those ester sidechains. It is common to quantify the rate of spectral diffusion dynamics using the center line slope (CLS) as a function of T. The fast decay (red trace) shown in Figure 7.9 shows evidence of bulk water near the ester probe. On the other hand, there is also a small population of ester side chains whose transition is centered at 1734.2 $cm^{-1}$ which is assigned to non-H-bonded esters. The 2D IR spectra of the non-H-bonded ester group do not illustrate a decay (black trace, Figure 7.9) meaning there is no significant spectral diffusion. This is due to the notion that the ester is sequestered in the dry hydrophobic interface inside the fibril and is not exposed to solvent. This experiment demonstrates the strategy of using the ester probe to study the water environments, aggregation kinetics, and structural changes during fibril formation.

The mature TTR-110$D_M$ fibrils, in Figure 7.10, have two distinct ester transitions in solutions that are centered at 1727.9 and 1734.2 $cm^{-1}$. These are assigned to singly-H-

bonded and non-H-bonded ester species. The singly-H-bonded ester group shows a clear spectral diffusion with a time constant of 0.75 picoseconds but there is no significant spectral diffusion for the non-H-bonded species. This observation implies that the singly-H-bonded ester group is solvated by water molecules, most likely in the bulk. Similarly, in the dry film, the ester transition is broad and is centered at 1730.2 $cm^{-1}$, but a more careful analysis of the 2D IR lineshape and the CLS reveal two transitions centered at 1729 and 1734 $cm^{-1}$. The 1729 $cm^{-1}$ transition shows a spectral diffusion with a time constant of 1.1 ps, while the 1734 $cm^{-1}$ transition does not show any spectral diffusion. They are assigned to a singly-H-bonded and non-H-bonded species, respectively, similar to solution fibril case. This information gives us a window into understanding the structure of the interior of the fibrils because half of the esters are in the "dry zipper" interface and the other half is the interface between fibril outer surface and the bulk. In solution, the ester group at the dry interface is non-H-bonded and contributes to the high frequency band while the ester group at the outer interface, on the other hand, could be singly-H-bonded or doubly-H-bonded. The doubly-H-bonded ester group would show a broad featureless transition at even lower frequencies than the singly-H-bonded ester. Upon drying, the outer surface loses contact with water molecules; however, the ester group at the outer surface can still form H-bond with the nearby sidechains, specifically the serine residue at position 112. The small frequency shift and difference in spectral diffusion between the solution and dried sample of the singly-H-bonded ester can be

explained by H-bonding with different species, the water molecules or the serine sidechain OH group.

The preliminary IR and AFM measurements of the TTR-110$D_M$ mutant are promising but additional work, especially for the second mutant is necessary for a complete story. The second mutant, TTR-111$D_M$ has an ester sidechain sequestered in the water channel. We expect to have two distinct ester populations, one that is non-H-bonded and inside the dry zipper and the second that is forming H-bonds with confined water molecules that are in the nanochannel. Because confined water is relatively well-structured and has stronger H-bonds, the fluctuations are expected to be slower which could result in a slower rate of spectral diffusion.

## 7.4 Conclusions

These three examples, Aβ-$D_M$, Aβ-$D_M$-Cage and TTR-$D_M$ showcase promising preliminary results of the ester probe reporting on amyloid structure and dynamics. In the first section describing Aβ-$D_M$, the ester probe provided the electric fields at three different locations of the amyloid fibril. Additionally, the dielectric constant inside the fibril was quantitatively assessed, which can be a challenging task in proteins. The Aβ-$D_M$-Cage experiment provides structural information, specifically the side chain packing within the fibril. Also, the ester provides information about the number and strength of the H-bonds inside the fibrils. And finally, the TTR-Dm fibrils provided can distinguish

between different H-bonding environments within the fibril. The ester probe is ideally suited to probe the electric field and H-bonding environments in a wide variety of biological system besides amyloid fibrils, such as transmembrane ion channels and enzyme binding cavities.

## 7.5 Original Publication

Part of this Chapter, the portion regarding the Aβ-$D_M$ peptide, has been adapted and reprinted from Angewandte Chemie International Edition, Ileana M. Pazos, Dr. Ayanjeet Ghosh, Dr. Matthew J. Tucker, and Dr. Feng Gai, (2014) 53, 24 6080–6084. DOI: 10.1002/anie.201402011, with permission from John Wiley and Sons. Other portions of this chapter, the Aβ-$D_M$-Cage and TTR-$D_M$, are currently unpublished and this work was done in collaboration with Dr. Jianqiang Ma.

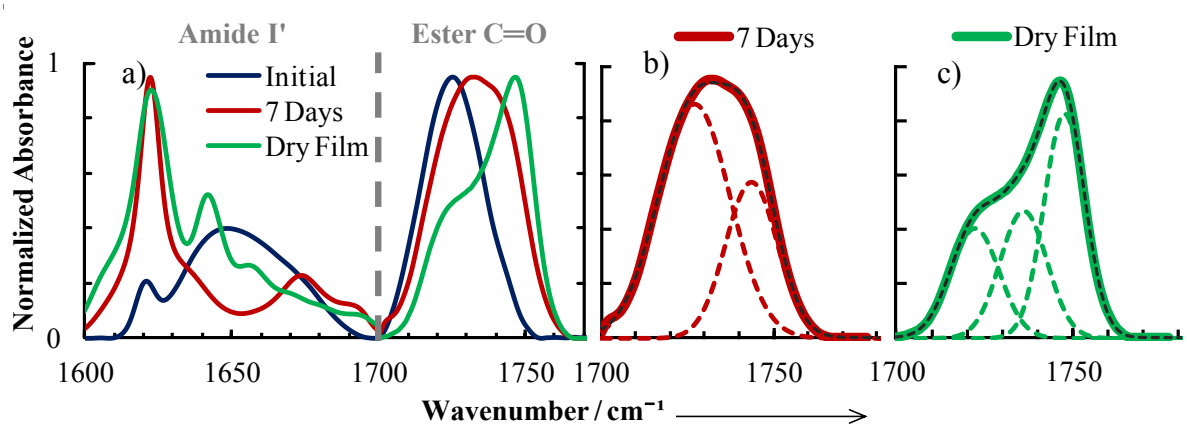



Figure 7.1. Normalized stretching vibrational bands of the ester carbonyl group of Aβ-$D_M$, as observed immediately after the preparation of a 15 mM sample of Aβ-$D_M$, after incubation of the sample for 7 days in $D_2O$, and in a dry film of Aβ-$D_M$ (dried under a flow of nitrogen for 7 days), as indicated. The band observed for the dry film can be decomposed into three Gaussian functions (gray dashed lines; the center frequencies are given in the main text). It should be noted that in each case a baseline that contains contributions from the amide I′ band has been subtracted.

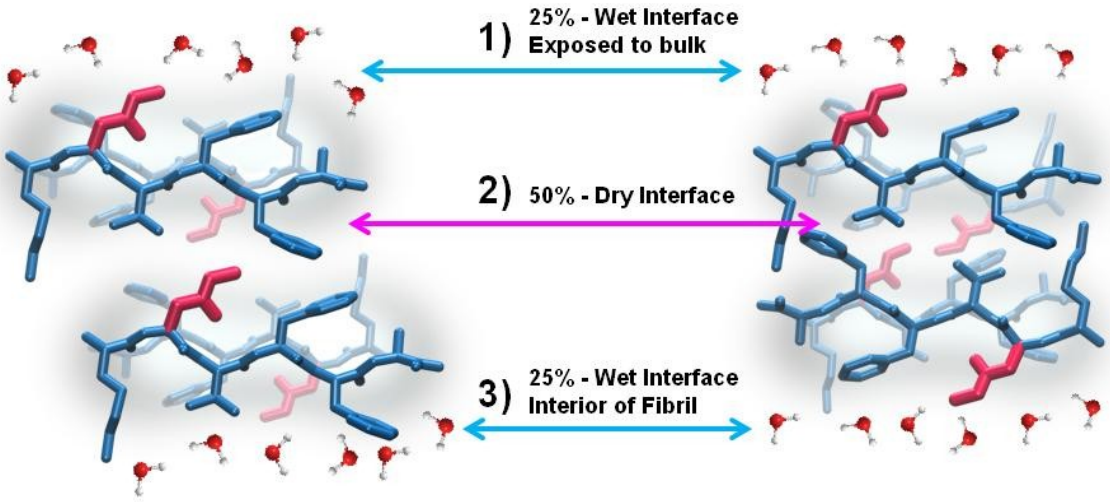



Figure 7.2 Cartoon representation of the possible local environments sampled by the $D_M$ sidechains (red) in the fibrils formed by the Aβ-$D_M$ peptide.

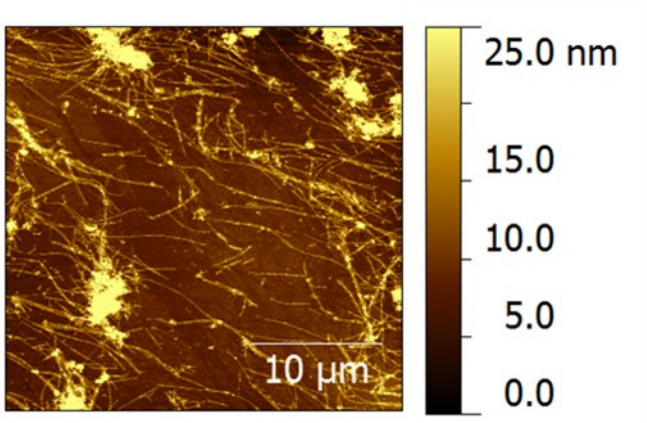



**Figure** 7.3 AFM image of Aβ-$D_M$, showing the heterogeneous morphology of the fibrils.

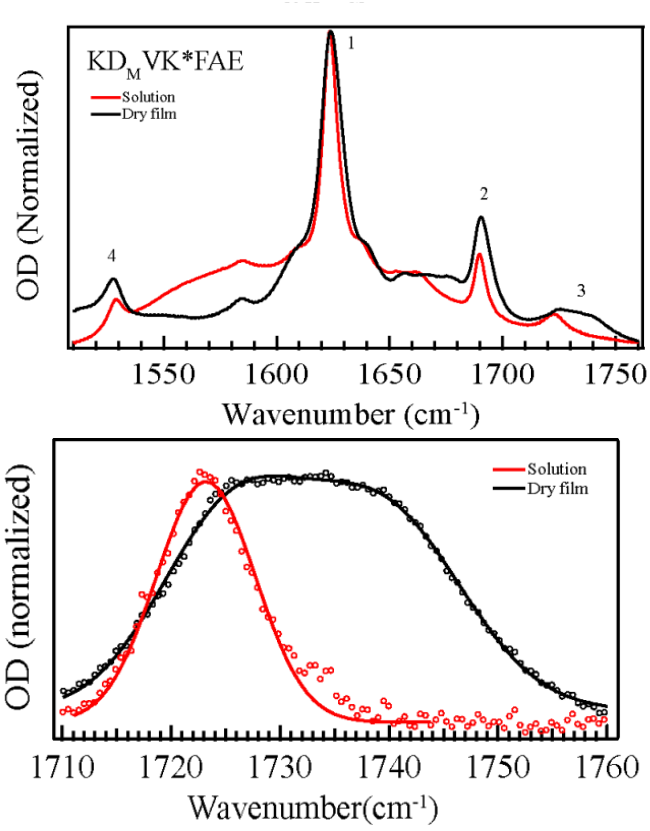



Figure 7.4 FTIR spectra of Aβ-$D_M$-Cage fibrils in solution and in dry film. (a) The spectra has markers 1-4 that are assigned to the following: 1 is the amide I' main band of the anti-parallel β-sheet. 2 has two contributions, one is from the amide I' main band of the anti-parallel β-sheet, the other is from the K* side chain carbonyl stretch reported by previous work.[409] 3 is the side chain ester carbonyl stretch from $D_M$. 4 is another K* side chain mode, which is likely a ring mode or the N=O stretch motion. (b) Ester region of FTIR spectra of the ester band in the fibril (open circles) and fittings with Gaussian functions (solid line).

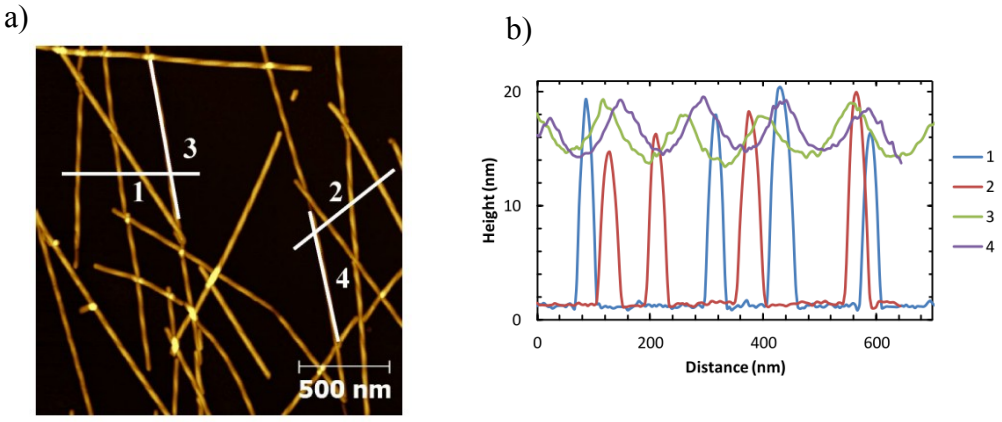



Figure 7.5 a) AFM images of Aβ-$D_M$-Cage fibrils and b) height measurements of the fibril dry film.

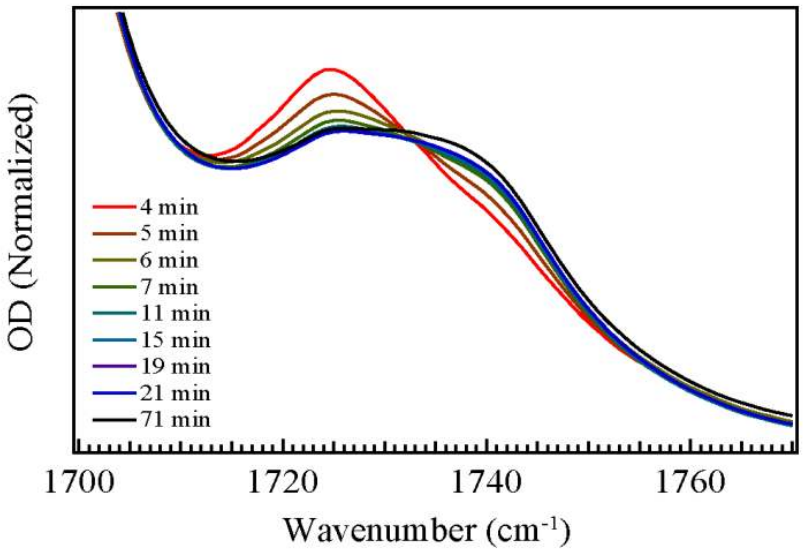



Figure 7.6 Time dependent FTIR of the Aβ-$D_M$-Cage fibril during drying process with $N_2$.

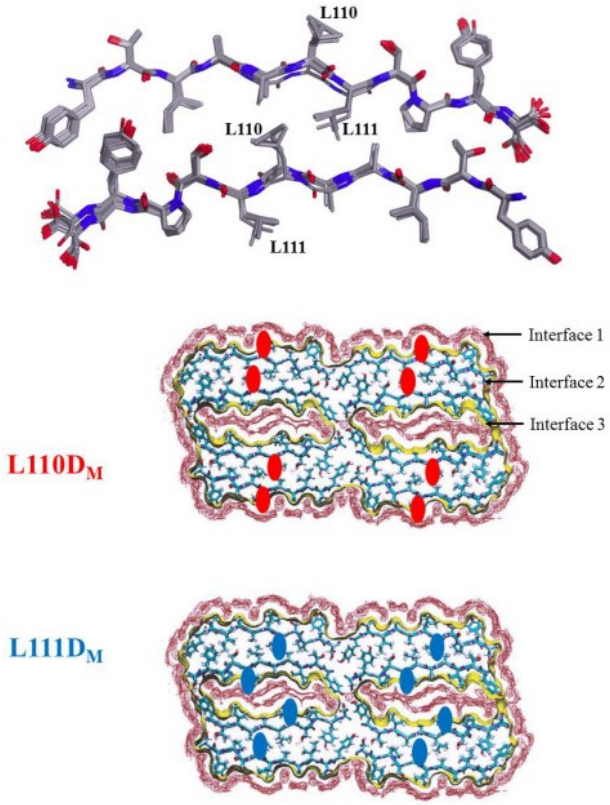



Figure 7.7 Proposed structures of $TTR_{105-115}$, with mutant locations are highlighted (TTR-$110D_M$ and TTR-$111D_M$), modified from Dobson et al.[417]

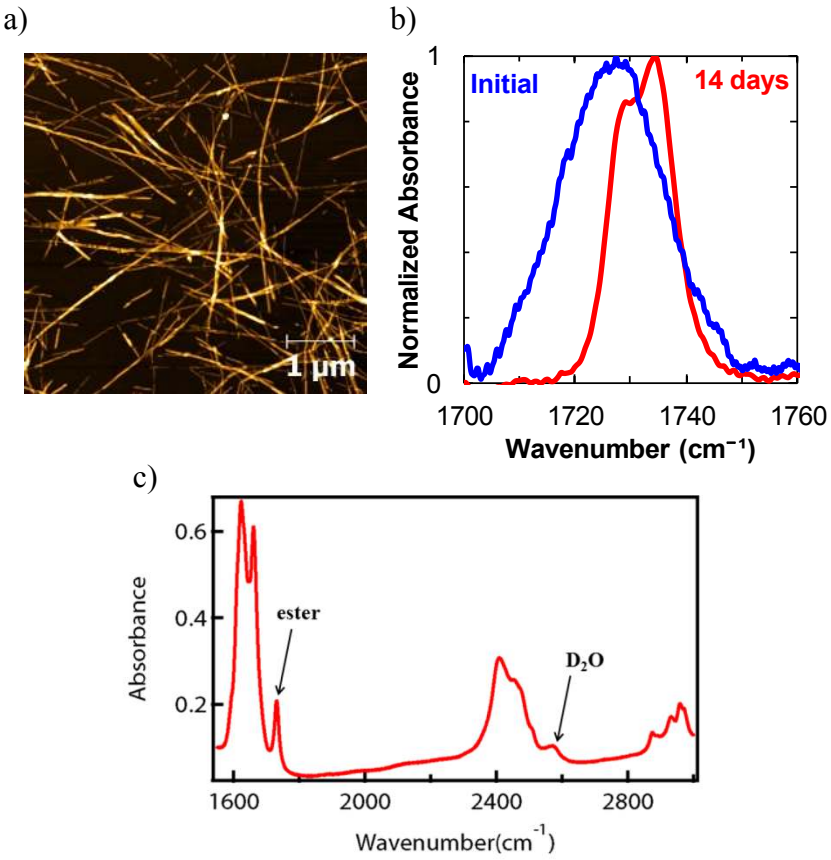



Figure 7.8 TTR-110D$_M$ mutant. a) AFM image of mature fibrils. b) FTIR of ester C=O stretch of the initial dissolved sample and after a 14 day aggregation period. c) FTIR of dried film of the matured fibril.

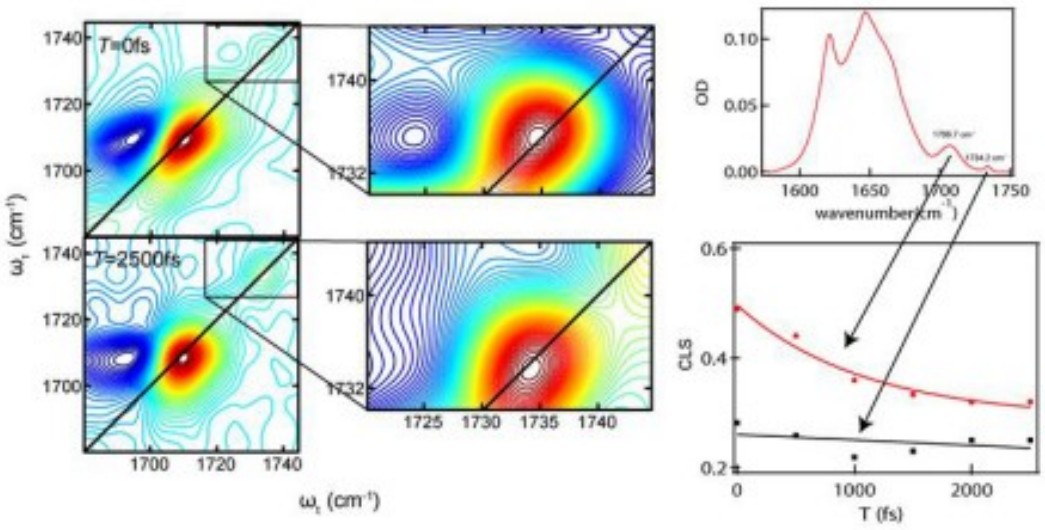



Figure 7.9 The amide I' and ester region 2D IR spectra of a partially aggregated TTR-L110D$_M$ mutant in solution.

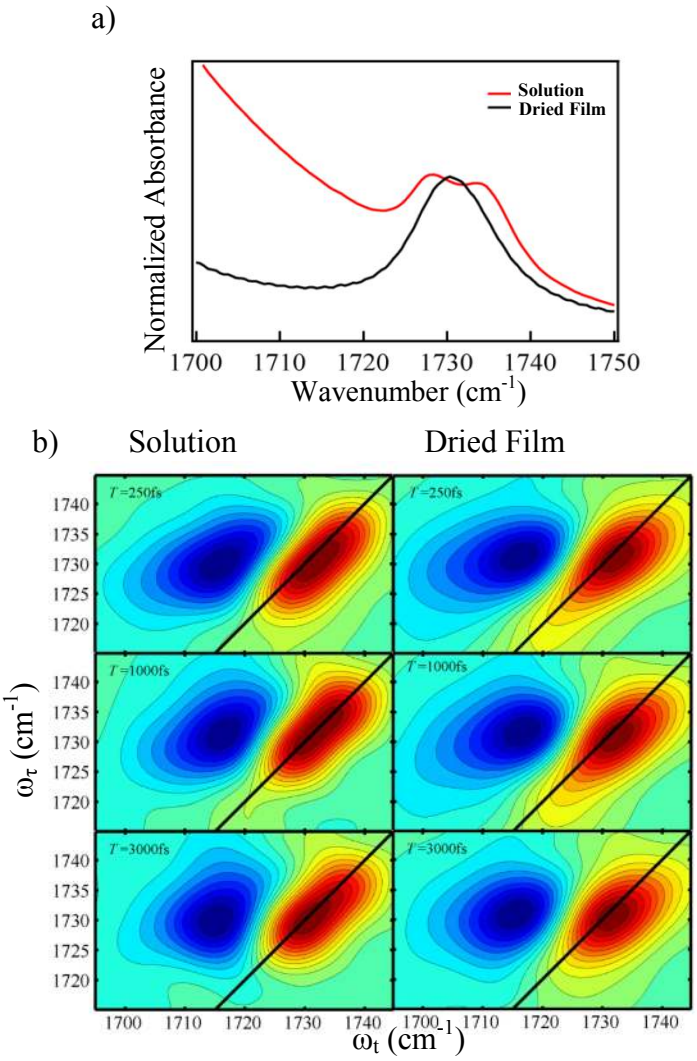



Figure 7.10 Mature TTR-110D$_M$ fibrils in solution and dry film. a) FTIR and b)2D IR photon echo measurements.

## 8 Summary and Perspective

A central dogma of biology concisely states that 'sequence determines structure determines function'. While this is generally considered an over-simplification, it does assert that it should be possible to deduce the function of a protein from its structure. However, in many cases merely knowing the structure of the biological molecules (e.g., proteins) in question is insufficient to reveal how its function is executed. This is because proteins are subject to various conformational motions and these protein dynamics are known to play an important and sometimes key role in protein functions. Furthermore, functional actions, such as binding or catalysis, always take place at a specific region or site of the proteins. Thus, in order to use spectroscopic techniques to monitor how proteins fluctuate in space, especially those in or near the sites where chemistry or function occurs, requires site-specific conformational probes. In addition, site-specific spectroscopic probes that are capable of 'recording' various physical properties of proteins, such as local electrostatic field and hydration, are needed to help determine factors that are key to function. Motivated by these demands, in this Thesis work we have expanded the utility of an existing spectroscopic probe, *p*-cyanophenylalanine ($Phe_{CN}$), and introduced a new non-natural amino acid-based IR probe. In addition, we have applied these probes to several outstanding biophysical questions, including protein protecting/denaturing mechanisms of several commonly used cosolvents and amyloid formations.

Two chapters of this Thesis are dedicated to expanding the utility of $Phe_{CN}$ as a fluorescent probe of protein conformation and conformational changes. Approximately ten years ago, the fluorescence quantum yield of $Phe_{CN}$ was first established to be sensitive to solvent and could be exploited to study protein folding and binding interactions.[184] While $Phe_{CN}$ has indeed found broad applications in this regard, one of its potential utilities, namely using fluorescence quenching to reveal protein structural information, has not been explored. Therefore, as detailed in Chapter 2, we first examined the $Phe_{CN}$ fluorescence quenching efficiencies of a series of anions. We found, among the anions studied, that iodide ($I^-$) exhibits the largest Stern-Volmer quenching constant (58.6 $M^{-1}$), suggesting that it is a suitable fluorescence quencher of $Phe_{CN}$ in protein conformational studies. Indeed, using this fluorophore-quencher pair, we were able to show, consistent with a previous study,[128] that the headpiece domain of chicken villin can sample at least two conformations in the native state potential well.[173] Based on the fact that tryptophan has been widely used in protein structural studies via fluorescence-quenching measurements,[95-98] we believe that the $Phe_{CN}$-iodide pair will find similar applications. As $Phe_{CN}$ is a structure analog of phenylalanine and tyrosine, this pair would be ideally suited to provide structural information, such as solvent accessibility, of a phenylalanine or tyrosine residue in the protein system of interest.

A previous study indicated that the N-terminal amino group ($-NH_2$) could quench the fluorescence of a nearby $Phe_{CN}$.[187] Motivated by this observation, in Chapter 3, we explored the possibility of a using $Phe_{CN}$ fluorescence as a pH reporter. This is based on

the fact that the pKa of the N-terminal amino group depends on the identity of the first amino acid and,[218] thus, it is possible to use a series of short peptides, for example, those with the following sequence $NH_2$−X−$Phe_{CN}$−Gly−Ac (where the X represents a different amino acid in each case), to cover a broad pH range. To test this potential utility of $Phe_{CN}$, we carried out fluorescence pH titrations on nine of the aforementioned tri-peptides. Our results showed that $NH_2$–Asn–$Phe_{CN}$−Gly−Ac has the lowest N-terminal pKa value (6.7), whereas $NH_2$−Ser−$Phe_{CN}$−Gly−Ac has the highest N-terminal pKa (8.6). Taken together, these results thus demonstrated that peptides or proteins with a $Phe_{CN}$ placed at the second position can be used to report pH or pH change in a range determined by the first amino acid. In addition, perhaps more useful, this type of pH sensors could be used to monitor pH-dependent biological processes or events that are accompanied by a pH change. To test this idea, we then measured the rate of a cell-penetrating peptide across model membranes and, as expected, the resultant fluorescence kinetics corroborates the proposed application. One immediate extension of this study would be to use this pH sensor to investigate the membrane penetration kinetics of pH (low) insertion peptide (pHLIP).[418] Engelman and coworkers have shown that pHLIP is unstructured and water soluble at high pH (>7) but able to cross the lipid bilayers a low pH (<3). Therefore, by inserting a $Phe_{CN}$ at the second position of the pHLIP sequence, one can monitor this kinetic event using a pH-drop stopped-flow fluorescence technique. Another potential application of this pH sensor is to use it to monitor the interaction of the N-terminal nitrogen atom with various metal ions.[419] For example, it has been

suggested that $Cu^{+2}$ can coordinate with the N-terminus of uncapped Aβ peptides and thus modulate its aggregation propensity.[420-421] Since coordination with a metal ion is expected to change the ability of the N-terminal amino group to quench a $Phe_{CN}$ fluorescence, we hypothesis that at a given pH, $Phe_{CN}$ would also be used to monitor metal ion binding.

As described in Chapters 4 and 5, in a different but related project, we used the established IR utility of $Phe_{CN}$ to study a longstanding unresolved problem, that is how commonly used biological cosolvents, such as TMAO, GdnHCl, and urea, act to deliver their respective protecting or denaturing effect on proteins. While many studies have been dedicated to this topic and offered many insights, a comprehensive understanding of the mechanism of action has not been achieved. To further our understanding of this problem, we devised both linear and non-linear IR experiments that would allow us to examine the actions of these cosolvents from a different perspective. In other words, using $Phe_{CN}$ as a site-specific IR probe, we tried to determine how addition of a specific cosolvent affects the ability of water to form H-bonds with protein polar groups and to verify the proposed role of these cosolvents as nano-crowders. As detailed in Chapter 4, we showed, using linear IR measurements, that these cosolvents all act to decrease the strength of the H-bonds formed between water and the nitrile group of $Phe_{CN}$. This finding is surprising since TMAO, GdnHCl, and urea have different effects on protein stability. While the behavior of TMAO appears to be consistent with its protein-protecting ability, those of urea and GdnHCl seem to contradict their role as protein

denaturants. This is because a reduced H-bonding ability of water would promote intramolecular H-bond formation, therefore stabilizing the folded state. Thus this result supports the so-called direct mechanism which suggests that urea and GdnHCl preferentially bind to the polypeptide backbone units, causing the unfolded state to be stabilized.[174]

To further verify the findings described in Chapter 4 and assess the effects of these cosolvents on the spontaneous conformational dynamics of proteins, we conducted two-dimensional IR measurements on the nitrile stretching vibrational of a $Phe_{CN}$ residue in both a peptide and a protein.[175] These experiments were designed based on the principle that the fluctuations dynamics of the electrostatic environment of the nitrile group of $Phe_{CN}$, including H-bonding dynamics arising from water and protein conformational dynamics, would be reflected in the vibrational spectral diffusion dynamics of the $Phe_{CN}$ probe. As detailed in Chapter 5, the 2D-IR results not only confirmed the finding we obtained in Chapter 4, namely both TMAO and urea decrease the strength of H-bonds formed between water and the nitrile group, but also showed that TMAO can increase protein stability through the excluded volume effect. In other words, TMAO is depleted from the polypeptide backbone as indicated from many previous studies, which allows it to behave as a crowding agent. Taken together, these studies not only yielded new microscopic information regarding the mechanism of action of the aforementioned cosolvents, but also provided new examples that site-specific IR probes are useful or even necessary to elucidate certain details of the biophysical or biochemical

questions of interest. An immediate extension of this study would be to investigate whether other biological osmolytes, such as glycine betaine, trehalose, sucrose, proline, taurine, glycine, and sarcosine, share similar mechanisms. Another application of the 2D-IR method utilized in this chapter would be to study the effect of macromolecular crowding or confinement on protein dynamics.[422-425]

Despite the extensive efforts dedicated in the past to develop and introduce non-natural amino acid-based site-specific IR probes, there is still a great need in spectroscopic probes that can be used to quantitatively and conveniently assess protein electric fields in a site-specific manner. This is because electric fields play an important role in many biological processes such as protein-protein interactions, ligand binding and catalysis and currently there are only a couple of options in this regard. Thus in Chapter 6, we devoted our focus to show that ester-derivatized non-natural amino acids are promising candidates to serve as a local electric field reporter of proteins as its carbonyl stretching frequency is sensitive to solvents.[176] To reveal the dependence of this frequency on the electric field exerted by the solvent molecules on the carbonyl vibrator, we carried out molecular dynamics simulations on an ester model compound in various solvents, including both protic and aprotic solvents. Our results show that the ester carbonyl stretching frequency exhibit a simple linear dependence on the local electric field and, thus, can be used to quantify, in a site-specific manner, the local electric environment of proteins. This is also supported by a subsequent study using NMR spectroscopy.[426] Since at neutral pH the ester carbonyl stretching vibrational transition is

located in the frequency range of 1700-1760 $cm^{-1}$, where no other protein vibrational bands exist, this vibrational transition can be detected and interpreted relatively easily. In comparison with the existing protein electric field IR probes, the ester IR probe described in Chapter 6 thus offer advantages.

To demonstrate the applicability of these ester-based IR probes in biophysical studies, we used them to monitor the aggregation processes of two amyloid forming peptides as well as the local hydration status of the amyloid fibrils thus formed. As shown in Chapter 7, when a model peptide aggregated into fibrils, the vibrational band of the ester transformed from a broad hydrated peak shape into sharper well-resolved peaks at higher frequencies. These peaks were assigned to different regions of the fibril therefore providing structural information. To expand on this interesting and important problem, 2D-IR spectra of this Alzheimer's peptide and another protein responsible for the neurodegenerative disease called transthyretin (TTR) amyloidosis were measured. Specifically, the TTR has a water channel inside the fibrils and an ester probe was sequestered inside the channel allowing for the direct measurement of the dynamics of trapped biological water.

## References


1. Fecko, C. J.; Loparo, J. J.; Roberts, S. T.; Tokmakoff, A., Local hydrogen bonding dynamics and collective reorganization in water: ultrafast infrared spectroscopy of HOD/$D_2O$. *J Chem Phys* **2005,** *122* (5), 54506.
2. Laage, D.; Hynes, J. T., A molecular jump mechanism of water reorientation. *Science* **2006,** *311* (5762), 832-835.
3. Kumar, R.; Schmidt, J. R.; Skinner, J. L., Hydrogen bonding definitions and dynamics in liquid water. *J. Chem. Phys.* **2007,** *126* (20), -.
4. Sharp, K. A.; Vanderkooi, J. M., Water in the Half Shell: Structure of Water, Focusing on Angular Structure and Solvation. *Acc. Chem. Res.* **2010,** *43* (2), 231-239.
5. Baskakov, I.; Wang, A. J.; Bolen, D. W., Trimethylamine-N-oxide counteracts urea effects on rabbit muscle lactate dehydrogenase function: a test of the counteraction hypothesis. *Biophys. J.* **1998,** *74* (5), 2666-2673.
6. Timasheff, S. N., Control of protein stability and reactions by weakly interacting cosolvents: the simplicity of the complicated. *Adv. Protein Chem.* **1998,** *51*, 355-432.
7. Yancey, P. H., Water Stress, Osmolytes and Proteins. *Am. Zool.* **2001,** *41* (4), 699-709.
8. Blose, J. M.; Pabit, S. A.; Meisburger, S. P.; Li, L.; Jones, C. D.; Pollack, L., Effects of a protecting osmolyte on the ion atmosphere surrounding DNA duplexes. *Biochemistry* **2011,** *50* (40), 8540-7.
9. Wang, A.; Bolen, D. W., A naturally occurring protective system in urea-rich cells: mechanism of osmolyte protection of proteins against urea denaturation. *Biochemistry* **1997,** *36* (30), 9101-8.
10. Zou, Q.; Bennion, B. J.; Daggett, V.; Murphy, K. P., The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea. *J. Am. Chem. Soc.* **2002,** *124* (7), 1192-1202.
11. Tanford, C., Protein denaturation. C. Theoretical models for the mechanism of denaturation. *Advances in protein chemistry* **1970,** *24*, 1-95.
12. Anfinsen, C. B., Principles that Govern Folding of Protein Chains. *Science* **1973,** *181* (4096), 223-230.
13. Pace, C. N., Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods in enzymology* **1986,** *131*, 266-80.
14. Alonso, D. O.; Dill, K. A., Solvent denaturation and stabilization of globular proteins. *Biochemistry* **1991,** *30* (24), 5974-85.
15. Duffy, E. M.; Kowalczyk, P. J.; Jorgensen, W. L., Do denaturants interact with aromatic hydrocarbons in water? *J. Am. Chem. Soc.* **1993,** *115* (20), 9271-9275.
16. From, N. B.; Bowler, B. E., Urea Denaturation of Staphylococcal Nuclease Monitored by Fourier Transform Infrared Spectroscopy. *Biochemistry* **1998,** *37* (6), 1623-1631.
17. Bolen, D. W.; Baskakov, I. V., The Osmophobic Effect: Natural Selection of a Thermodynamic Force in Protein Folding. *J. Mol. Biol.* **2001,** *310* (5), 955-963.
18. Timasheff, S. N., Thermodynamic binding and site occupancy in the light of the Schellman exchange concept. *Biophys. Chem.* **2002,** *101*, 99-111.
19. Schellman, J. A., Fifty years of solvent denaturation. *Biophys. Chem.* **2002,** *96* (2-3), 91-101.
20. Bennion, B. J.; Daggett, V., Counteraction of urea-induced protein denaturation by trimethylamine N-oxide: A chemical chaperone at atomic resolution. *Proc. Natl. Acad. Sci. U.S.A.* **2004,** *101* (17), 6433-6438.
21. Moglich, A.; Krieger, F.; Kiefhaber, T., Molecular basis for the effect of urea and guanidinium chloride on the dynamics of unfolded polypeptide chains. *J. Mol. Biol.* **2005,** *345* (1), 153-162.
22. Auton, M.; Bolen, D. W., Predicting the energetics of osmolyte-induced protein folding/unfolding. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (42), 15065-15068.

23. Caballero-Herrera, A.; Nordstrand, K.; Berndt, K. D.; Nilsson, L., Effect of urea on peptide conformation in water: molecular dynamics and experimental characterization. *Biophys. J.* **2005,** *89* (2), 842-57.
24. Street, T. O.; Bolen, D. W.; Rose, G. D., A molecular mechanism for osmolyte-induced protein stability. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103* (38), 13997-4002.
25. Di Michele, A.; Freda, M.; Onori, G.; Santucci, A., Hydrogen bonding of water in aqueous solutions of trimethylamine-N-oxide and tert-butyl alcohol: A near-infrared spectroscopy study. *J. Phys. Chem. A* **2004,** *108* (29), 6145-6150.
26. Paul, S.; Patey, G. N., The influence of urea and Trimethylamine-N-oxide on hydrophobic interactions. *J. Phys. Chem. B* **2007,** *111* (28), 7932-7933.
27. Athawale, M. V.; Sarupria, S.; Garde, S., Enthalpy-entropy contributions to salt and osmolyte effects on molecular-scale hydrophobic hydration and interactions. *J. Phys. Chem. B* **2008,** *112* (18), 5661-5670.
28. Zangi, R.; Zhou, R. H.; Berne, B. J., Urea's Action on Hydrophobic Interactions. *J. Am. Chem. Soc.* **2009,** *131* (4), 1535-1541.
29. Meersman, F.; Bowron, D.; Soper, A. K.; Koch, M. H., Counteraction of urea by trimethylamine N-oxide is due to direct interaction. *Biophys. J.* **2009,** *97* (9), 2559-66.
30. Kuffel, A.; Zielkiewicz, J., The hydrogen bond network structure within the hydration shell around simple osmolytes: urea, tetramethylurea, and trimethylamine-N-oxide, investigated using both a fixed charge and a polarizable water model. *J. Chem. Phys.* **2010,** *133* (3), 035102.
31. Zhang, Y.; Cremer, P. S., Chemistry of Hofmeister Anions and Osmolytes. *Annu. Rev. Phys. Chem.* **2010,** *61*, 63-83.
32. Holthauzen, L. M.; Rosgen, J.; Bolen, D. W., Hydrogen bonding progressively strengthens upon transfer of the protein urea-denatured state to water and protecting osmolytes. *Biochemistry* **2010,** *49* (6), 1310-8.
33. Hu, C. Y.; Lynch, G. C.; Kokubo, H.; Pettitt, B. M., Trimethylamine N-oxide influence on the backbone of proteins: An oligoglycine model. *Proteins: Struct., Funct., Bioinf.* **2010,** *78* (3), 695-704.
34. Auton, M.; Roesgen, J.; Sinev, M.; Holthauzen, L. M. F.; Bolen, D. W., Osmolyte Effects on Protein Stability and Solubility: A Balancing Act between Backbone and Side-Chains. *Biophys. Chem.* **2011,** *159* (1), 90-99.
35. Guinn, E. J.; Pegram, L. M.; Capp, M. W.; Pollock, M. N.; Record, M. T., Jr., Quantifying why urea is a protein denaturant, whereas glycine betaine is a protein stabilizer. *Proc. Natl. Acad. Sci. U.S.A.* **2011,** *108* (41), 16932-16937.
36. Canchi, D. R.; Garcia, A. E., Backbone and Side-Chain Contributions in Protein Denaturation by Urea. *Biophys. Chem.* **2011,** *100* (6), 1526-1533.
37. Rossky, P. J., Protein denaturation by urea: slash and bond. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105* (44), 16825-16826.
38. Yang, Z.; Xiu, P.; Shi, B.; Hua, L.; Zhou, R., Coherent microscopic picture for urea-induced denaturation of proteins. *J. Phys. Chem. B* **2012,** *116* (30), 8856-62.
39. Bennion, B. J.; Daggett, V., The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100* (9), 5142-5147.
40. Bolen, D. W.; Rose, G. D., Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annu. Rev. Biochem.* **2008,** *77*, 339-362.
41. Guo, F.; Friedman, J. M., Osmolyte-induced perturbations of hydrogen bonding between hydration layer waters: correlation with protein conformational changes. *J. Phys. Chem. B* **2009,** *113* (52), 16632-42.
42. Rezus, Y. L.; Bakker, H. J., Destabilization of the hydrogen-bond structure of water by the osmolyte trimethylamine N-oxide. *J. Phys. Chem. B* **2009,** *113* (13), 4038-44.
43. Rezus, Y. L. A.; Bakker, H. J., Effect of urea on the structural dynamics of water. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103* (49), 18417-18420.

44. Bakulin, A. A.; Pshenichnikov, M. S.; Bakker, H. J.; Petersen, C., Hydrophobic molecules slow down the hydrogen-bond dynamics of water. *J. Phys. Chem. A* **2011,** *115* (10), 1821-9.
45. Sharp, K. A.; Madan, B.; Manas, E.; Vanderkooi, J. M., Water structure changes induced by hydrophobic and polar solutes revealed by simulations and infrared spectroscopy. *J. Chem. Phys.* **2001,** *114* (4), 1791-1796.
46. Vorobyev, D. Y.; Kuo, C. H.; Kuroda, D. G.; Scott, J. N.; Vanderkooi, J. M.; Hochstrasser, R. M., Water-induced relaxation of a degenerate vibration of guanidinium using 2D IR echo spectroscopy. *J. Phys. Chem. B* **2010,** *114* (8), 2944-53.
47. Shimizu, S., The effect of urea on hydrophobic hydration: Preferential interaction and the enthalpy of transfer. *Chem. Phys. Lett.* **2011,** *517* (1-3), 76-79.
48. van der Post, S. T.; Tielrooij, K. J.; Hunger, J.; Backus, E. H. G.; Bakker, H. J., Femtosecond study of the effects of ions and hydrophobes on the dynamics of water. *Faraday Discuss* **2013,** *160*, 171-189.
49. Yancey, P. H.; Clark, M. E.; Hand, S. C.; Bowlus, R. D.; Somero, G. N., Living with Water Stress: Evolution of Osmolyte Systems. *Science* **1982,** *217* (4566), 1214-1222.
50. Uversky, V. N.; Li, J.; Fink, A. L., Trimethylamine-N-oxide-induced folding of α-synuclein. *Febs Letters* **2001,** *509* (1), 31-5.
51. Mello, C. C.; Barrick, D., Measuring the stability of partly folded proteins using TMAO. *Protein Sci.* **2003,** *12* (7), 1522-1529.
52. Auton, M.; Bolen, D. W.; Rosgen, J., Structural thermodynamics of protein preferential solvation: osmolyte solvation of proteins, aminoacids, and peptides. *Proteins* **2008,** *73* (4), 802-13.
53. Auton, M.; Bolen, D. W.; Rosgen, J., Structural Thermodynamics of Protein Preferential Solvation: Osmolyte Solvation of Proteins, Aminoacids, and Peptides. *Proteins: Struct., Funct., Bioinf.* **2008,** *73* (4), 802-813.
54. Panuszko, A.; Bruzdziak, P.; Zielkiewicz, J.; Wyrzykowski, D.; Stangret, J., Effects of Urea and Trimethylamine-N-oxide on the Properties of Water and the Secondary Structure of Hen Egg White Lysozyme. *J. Phys. Chem. B* **2009,** *113* (44), 14797-14809.
55. Ghosh, A.; Hochstrasser, R. M., A peptide's perspective of water dynamics. *Chem. Phys.* **2011,** *390* (1), 1-13.
56. Hamm, P.; Zanni, M. T., *Concepts and methods of 2d infrared spectroscopy*. Cambridge University Pres: Cambridge ; New York, 2011; p ix, 286 p.
57. 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* (12).
58. Zimmerman, S. B.; Trach, S. O., Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *J. Mol. Biol.* **1991,** *222* (3), 599-620.
59. Ellis, R. J., Macromolecular crowding: obvious but underappreciated. *Trends in Biochemical Sciences* **2001,** *26* (10), 597-604.
60. Elcock, A. H., Models of macromolecular crowding effects and the need for quantitative comparisons with experiment. *Current Opinion in Structural Biology* **2010,** *20* (2), 196-206.
61. Zhou, H.-X.; Rivas, G.; Minton, A. P., Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. In *Annual Review of Biophysics*, 2008; Vol. 37, pp 375-397.
62. Cho, S. S.; Reddy, G.; Straub, J. E.; Thirumalai, D., Entropic Stabilization of Proteins by TMAO. *J. Phys. Chem. B* **2011,** *115* (45), 13401-13407.
63. Cheung, M. S.; Klimov, D.; Thirumalai, D., Molecular crowding enhances native state stability and refolding rates of globular proteins. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (13), 4753-4758.
64. Warshel, A.; Sussman, F., Toward computer-aided site-directed mutagenesis of enzymes. *Proc Natl Acad Sci U S A* **1986,** *83* (11), 3806-10.
65. Warshel, A.; Aqvist, J., Electrostatic Energy and Macromolecular Function. *Annu. Rev. Biophys. Biophys. Chem.* **1991,** *20* (1), 267-298.

66. Villa, J.; Warshel, A., Energetics and dynamics of enzymatic reactions. *J. Phys. Chem. B* **2001,** *105* (33), 7887-7907.
67. Mao, J.; Hauser, K.; Gunner, M. R., How Cytochromes with Different Folds Control Heme Redox Potentials†. *Biochemistry* **2003,** *42* (33), 9829-9840.
68. Augspurger, J. D.; Dykstra, C. E.; Oldfield, E., Correlation of carbon-13 and oxygen-17 chemical shifts and the vibrational frequency of electrically perturbed carbon monoxide: a possible model for distal ligand effects in carbonmonoxyheme proteins. *J. Am. Chem. Soc.* **1991,** *113* (7), 2447-2451.
69. Park, K. D.; Guo, K.; Adebodun, F.; Chiu, M. L.; Sligar, S. G.; Oldfield, E., Distal and proximal ligand interactions in heme proteins: correlations between carbon-oxygen and iron-carbon vibrational frequencies, oxygen-17 and carbon-13 nuclear magnetic resonance chemical shifts, and oxygen-17 nuclear quadrupole coupling constants in $C^{17}O$- and $^{13}CO$-labeled species. *Biochemistry* **1991,** *30* (9), 2333-2347.
70. Pearson, J. G.; Oldfield, E.; Lee, F. S.; Warshel, A., Chemical shifts in proteins: a shielding trajectory analysis of the fluorine nuclear magnetic resonance spectrum of the Escherichia coli galactose binding protein using a multipole shielding polarizability-local reaction field-molecular dynamics approach. *J. Am. Chem. Soc.* **1993,** *115* (15), 6851-6862.
71. Augspurger, J. D.; Dykstra, C. E., Correlation of fluorine-19 chemical shielding and chemical shift nonequivalence. *J. Am. Chem. Soc.* **1993,** *115* (25), 12016-12019.
72. Feeney, J.; McCormick, J. E.; Bauer, C. J.; Birdsall, B.; Moody, C. M.; Starkmann, B. A.; Young, D. W.; Francis, P.; Havlin, R. H.; Arnold, W. D.; Oldfield, E., $^{19}F$ Nuclear Magnetic Resonance Chemical Shifts of Fluorine Containing Aliphatic Amino Acids in Proteins: Studies on Lactobacillus casei Dihydrofolate Reductase Containing (2S,4S)-5-Fluoroleucine. *J. Am. Chem. Soc.* **1996,** *118* (36), 8700-8706.
73. Fersht, A. R.; Sternberg, M. J. E., Can a simple function for the dielectric response model electrostatic effects in globular proteins? *Protein Eng.* **1989,** *2* (7), 527-530.
74. Sharp, K. A.; Honig, B., Electrostatic Interactions in Macromolecules: Theory and Applications. *Annu. Rev. Biophys. Biophys. Chem.* **1990,** *19* (1), 301-332.
75. Chimenti, M. S.; Castañeda, C. A.; Majumdar, A.; García-Moreno E, B., Structural Origins of High Apparent Dielectric Constants Experienced by Ionizable Groups in the Hydrophobic Core of a Protein. *J. Mol. Biol.* **2011,** *405* (2), 361-377.
76. Schkolnik, G.; Utesch, T.; Salewski, J.; Tenger, K.; Millo, D.; Kranich, A.; Zebger, I.; Schulz, C.; Zimanyi, L.; Rakhely, G.; Mroginski, M. A.; Hildebrandt, P., Mapping local electric fields in proteins at biomimetic interfaces. *Chem. Commun.* **2012,** *48* (1), 70-72.
77. Schkolnik, G.; Utesch, T.; Zhao, J.; Jiang, S.; Thompson, M. K.; Mroginski, M.-A.; Hildebrandt, P.; Franzen, S., Catalytic efficiency of dehaloperoxidase A is controlled by electrostatics - application of the vibrational Stark effect to understand enzyme kinetics. *Biochem Bioph Res Co* **2013,** *430* (3), 1011-1015.
78. Andrews, S. S.; Boxer, S. G., Vibrational stark effects of nitriles I. Methods and experimental results. *J. Phys. Chem. A* **2000,** *104* (51), 11853-11863.
79. Cho, M., Vibrational solvatochromism and electrochromism: Coarse-grained models and their relationships. *J. Chem. Phys.* **2009,** *130* (9), 094505.
80. Suydam, I. T.; Boxer, S. G., Vibrational Stark effects calibrate the sensitivity of vibrational probes for electric fields in proteins. *Biochemistry* **2003,** *42* (41), 12050-12055.
81. Suydam, I. T.; Snow, C. D.; Pande, V. S.; Boxer, S. G., Electric fields at the active site of an enzyme: Direct comparison of experiment with theory. *Science* **2006,** *313* (5784), 200-204.
82. Fried, S. D.; Bagchi, S.; Boxer, S. G., Extreme electric fields power catalysis in the active site of ketosteroid isomerase. *Science* **2014,** *346* (6216), 1510-1514.
83. 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* (37), 12811-12813.

84. Choi, J.-H.; Cho, M., Vibrational solvatochromism and electrochromism of infrared probe molecules containing C equivalent to O, C equivalent to N, C=O, or C-F vibrational chromophore. *J. Chem. Phys.* **2011,** *134* (15), 54513-54525.
85. Choi, J. H.; Cho, M., Vibrational solvatochromism and electrochromism of infrared probe molecules containing C equivalent to O, C equivalent to N, C=O, or C-F vibrational chromophore. *J. Chem. Phys.* **2011,** *134* (15), 154513.
86. Klepeis, J. L.; Lindorff-Larsen, K.; Dror, R. O.; Shaw, D. E., Long-timescale molecular dynamics simulations of protein structure and function. *Curr Opin Struc Biol* **2009,** *19* (2), 120-127.
87. Dror, R. O.; Dirks, R. M.; Grossman, J. P.; Xu, H.; Shaw, D. E., Biomolecular Simulation: A Computational Microscope for Molecular Biology. *Annual Review of Biophysics* **2012,** *41*, 429-452.
88. Brunger, A. T., X-ray crystallography and NMR reveal complementary views of structure and dynamics. *Nat Struct Biol* **1997,** *4*, 862-865.
89. Maity, H.; Maity, M.; Krishna, M. M. G.; Mayne, L.; Englander, S. W., Protein folding: The stepwise assembly of foldon units. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (13), 4741-4746.
90. Lakowicz, J. R., *Principles of fluorescence spectroscopy*. 3rd ed.; Springer: New York, 2006; p xxvi, 954 p.
91. Goto, Y.; Fink, A. L., Conformational states of beta-lactamase: molten-globule states at acidic and alkaline pH with high salt. *Biochemistry* **1989,** *28* (3), 945-52.
92. Cardamone, M.; Puri, N. K., Spectrofluorimetric assessment of the surface hydrophobicity of proteins. *Biochemical Journal* **1992,** *282* (Pt 2), 589-593.
93. Acharya, P.; Rao, N. M., Stability studies on a lipase from Bacillus subtilis in guanidinium chloride. *J Protein Chem* **2003,** *22* (1), 51-60.
94. Bolen, E. J.; Holloway, P. W., Quenching of Tryptophan Fluorescence by Brominated Phospholipid. *Biochemistry* **1990,** *29* (41), 9638-9643.
95. 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* (22), 8449-8454.
96. Liu, R. H.; Siemiarczuk, A.; Sharom, F. J., Intrinsic fluorescence of the P-glycoprotein multidrug transporter: Sensitivity of tryptophan residues to binding of drugs and nucleotides. *Biochemistry* **2000,** *39* (48), 14927-14938.
97. Papadopoulou, A.; Green, R. J.; Frazier, R. A., Interaction of flavonoids with bovine serum albumin: A fluorescence quenching study. *J. Agric. Food Chem.* **2005,** *53* (1), 158-163.
98. Hu, Y. J.; Liu, Y.; Pi, Z. B.; Qu, S. S., Interaction of cromolyn sodium with human serum albumin: A fluorescence quenching study. *Bioorgan Med Chem* **2005,** *13* (24), 6609-6614.
99. Elliott, A.; Ambrose, E., Structure of synthetic polypeptides. *Nature* **1950,** *165*, 921-922.
100. Kong, J.; Yu, S., Fourier transform infrared spectroscopic analysis of protein secondary structures. *Acta Biochimica Et Biophysica Sinica* **2007,** *39* (8), 549-559.
101. Abe, Y.; Krimm, S., NORMAL VIBRATIONS OF CRYSTALLINE POLYGLYCINE-I. *Biopolymers* **1972,** *11* (9), 1817-&.
102. Krimm, S.; Bandekar, J., VIBRATIONAL SPECTROSCOPY AND CONFORMATION OF PEPTIDES, POLYPEPTIDES, AND PROTEINS. *Advances in Protein Chemistry* **1986,** *38*, 181-364.
103. Barth, A.; Zscherp, C., What vibrations tell us about proteins. *Q. Rev. Biophys.* **2002,** *35* (4), 369-430.
104. Cowie, D. B.; Cohen, G. N., Biosynthesis by Escherichia coli of active altered proteins containing selenium instead of sulfur. *Biochimica Et Biophysica Acta* **1957,** *26* (2), 252-61.
105. Brawerman, G.; Ycas, M., Incorporation of the Amino Acid Analog Tryptazan into the Protein of Escherichia-Coli. *Arch Biochem Biophys* **1957,** *68* (1), 112-117.
106. Munier, R.; Cohen, G. N., Incorporation of structural analogues of amino acid into the bacterial proteins during their synthesis in vivo. *Biochim. Biophys. Acta* **1959,** *31* (2), 378-391.
107. Heim, R.; Prasher, D. C.; Tsien, R. Y., Wavelength mutations and posttranslational autoxidation of green fluorescent protein. *Proc Natl Acad Sci U S A* **1994,** *91* (26), 12501-4.

108. Furter, R., Expansion of the genetic code: Site-directed p-fluoro-phenylalanine incorporation in Escherichia coli. *Protein Sci.* **1998,** *7* (2), 419-426.
109. Ayers, B.; Blaschke, U. K.; Camarero, J. A.; Cotton, G. J.; Holford, M.; Muir, T. W., Introduction of unnatural amino acids into proteins using expressed protein ligation. *Biopolymers* **1999,** *51* (5), 343-354.
110. Cotton, G. J.; Muir, T. W., Peptide ligation and its application to protein engineering. *Chemistry & Biology* **1999,** *6* (9), R247-R256.
111. Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G., Expanding the genetic code of Escherichia coli. *Science* **2001,** *292* (5516), 498-500.
112. Connor, R. E.; Tirrell, D. A., Non-canonical amino acids in protein polymer design. *Polym. Rev.* **2007,** *47* (1), 9-28.
113. Voller, J.; Biava, H.; Koksch, B.; Hildebrandt, P.; Budisa, N., Orthogonal translation meets electron transfer: in vivo labeling of cytochrome c for probing local electric fields. *Chembiochem* **2015**.
114. Schade, M.; Moretto, A.; Crisma, M.; Toniolo, C.; Hamm, P., Vibrational energy transport in peptide helices after excitation of C-D modes in Leu-d10. *J. Phys. Chem. B* **2009,** *113* (40), 13393-7.
115. Remorino, A.; Korendovych, I. V.; Wu, Y.; DeGrado, W. F.; Hochstrasser, R. M., Residue-specific vibrational echoes yield 3D structures of a transmembrane helix dimer. *Science* **2011,** *332* (6034), 1206-9.
116. Zimmermann, J.; Thielges, M. C.; Seo, Y. J.; Dawson, P. E.; Romesberg, F. E., Cyano Groups as Probes of Protein Microenvironments and Dynamics. *Angew. Chem. Int. Ed.* **2011,** *50* (36), 8333-8337.
117. Thielges, M. C.; Chung, J. K.; Fayer, M. D., Protein dynamics in cytochrome P450 molecular recognition and substrate specificity using 2D IR vibrational echo spectroscopy. *J. Am. Chem. Soc.* **2011,** *133* (11), 3995-4004.
118. King, J. T.; Arthur, E. J.; Brooks, C. L.; Kubarych, K. J., Site-specific hydration dynamics of globular proteins and the role of constrained water in solvent exchange with amphiphilic cosolvents. *J. Phys. Chem. B* **2012,** *116* (19), 5604-11.
119. Yu, W.; Dawson, P. E.; Zimmermann, J.; Romesberg, F. E., Carbon-deuterium bonds as probes of protein thermal unfolding. *J. Phys. Chem. B* **2012,** *116* (22), 6397-403.
120. Buchanan, L. E.; Dunkelberger, E. B.; Tran, H. Q.; Cheng, P. N.; Chiu, C. C.; Cao, P.; Raleigh, D. P.; de Pablo, J. J.; Nowick, J. S.; Zanni, M. T., Mechanism of IAPP amyloid fibril formation involves an intermediate with a transient β-sheet. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110* (48), 19285-19290.
121. Park, K. H.; Jeon, J.; Park, Y.; Lee, S.; Kwon, H. J.; Joo, C.; Park, S.; Han, H.; Cho, M., Infrared Probes Based on Nitrile-Derivatized Prolines: Thermal Insulation Effect and Enhanced Dynamic Range. *J. Phys. Chem. Lett.* **2013,** *4* (13), 2105-2110.
122. Woys, A. M.; Mukherjee, S. S.; Skoff, D. R.; Moran, S. D.; Zanni, M. T., A strongly absorbing class of non-natural labels for probing protein electrostatics and solvation with FTIR and 2D IR spectroscopies. *J. Phys. Chem. B* **2013,** *117* (17), 5009-18.
123. Maekawa, H.; Sul, S.; Ge, N. H., Vibrational correlation between conjugated carbonyl and diazo modes studied by single- and dual-frequency two-dimensional infrared spectroscopy. *Chem. Phys.* **2013,** *422*, 22-30.
124. Guo, L.; Gai, F., Heterogeneous diffusion of a membrane-bound pHLIP peptide. *Biophys. J.* **2010,** *98* (12), 2914-2922.
125. Smith-Dupont, K. B.; Guo, L.; Gai, F., Diffusion as a probe of the heterogeneity of antimicrobial peptide-membrane interactions. *Biochemistry* **2010,** *49* (22), 4672-4678.
126. Rogers, J. M. G.; Lippert, L. G.; Gai, F., Non-natural amino acid fluorophores for one- and two-step fluorescence resonance energy transfer applications. *Anal. Biochem.* **2010,** *399* (2), 182-189.
127. Rogers, J. M. G.; Poishchuk, 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* (7), 3815-3821.
128. Serrano, A. L.; Bilsel, O.; Gai, F., Native State Conformational Heterogeneity of HP35 Revealed by Time-Resolved FRET. *J. Phys. Chem. B* **2012,** *116* (35), 10631-10638.

129. Guo, L.; Gai, F., Simple Method to Enhance the Photostability of the Fluorescence Reporter R6G for Prolonged Single-Molecule Studies. *J. Phys. Chem. A* **2013,** *117* (29), 6164-6170.
130. Pazos, I. M.; Roesch, R. M.; Gai, F., Quenching of p-cyanophenylalanine fluorescence by various anions. *Chem. Phys. Lett.* **2013,** *563*, 93-96.
131. Tang, J.; Yin, H.; Qiu, J.; Tucker, M. J.; DeGrado, W. F.; Gai, F., Using Two Fluorescent Probes to Dissect the Binding, Insertion, and Dimerization Kinetics of a Model Membrane Peptide. *J. Am. Chem. Soc.* **2009,** *131* (11), 3816-3817.
132. Reimers, J. R.; Hall, L. E., The solvation of acetonitrile. *J. Am. Chem. Soc.* **1999,** *121* (15), 3730-3744.
133. Mukherjee, S.; Chowdhury, P.; DeGrado, W. F.; Gai, F., Site-specific hydration status of an amphipathic peptide in AOT reverse micelles. *Langmuir* **2007,** *23* (22), 11174-11179.
134. Waegele, M. M.; Gai, F., Infrared study of the folding mechanism of a helical hairpin: porcine PYY. *Biochemistry* **2010,** *49* (35), 7659-7664.
135. 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* (49), 10354-6.
136. Urbanek, D. C.; Vorobyev, D. Y.; Serrano, A. L.; Gai, F.; Hochstrasser, R. M., The Two-Dimensional Vibrational Echo of a Nitrile Probe of the Villin HP35 Protein. *J. Phys. Chem. Lett.* **2010,** *1* (23), 3311-3315.
137. Waegele, M. M.; Culik, R. M.; Gai, F., Site-Sppecific spectroscopic reporters of the local electric field, hydration, structure, and dynamics of biomolecules. *J. Phys. Chem. Lett.* **2011,** *2* (20), 2598-2609.
138. Serrano, A. L.; Waegele, M. M.; Gai, F., Spectroscopic studies of protein folding: Linear and nonlinear methods. *Protein Sci.* **2012,** *21* (2), 157-170.
139. Pazos, I. M.; Gai, F., Solute's perspective on how trimethylamine oxide, urea, and guanidine hydrochloride affect water's hydrogen bonding ability. *J. Phys. Chem. B* **2012,** *116* (41), 12473-12478.
140. 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.
141. 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* (23), 8476-8481.
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* (1), null.
143. Tucker, M. J.; Getahun, Z.; Nanda, V.; DeGrado, W. F.; Gai, F., A new method for determining the local environment and orientation of individual side chains of membrane-binding peptides. *J. Am. Chem. Soc.* **2004,** *126* (16), 5078-5079.
144. Hu, W.; Webb, L. J., Direct Measurement of the Membrane Dipole Field in Bicelles Using Vibrational Stark Effect Spectroscopy. *J. Phys. Chem. Lett.* **2011,** *2* (15), 1925-1930.
145. Maienschein-Cline, M. G.; Londergan, C. H., The CN stretching band of aliphatic thiocyanate is sensitive to solvent dynamics and specific solvation. *J. Phys. Chem. A* **2007,** *111* (40), 10020-10025.
146. Choi, J. H.; Oh, K. I.; Lee, H.; Lee, C.; Cho, M., Nitrile and thiocyanate IR probes: quantum chemistry calculation studies and multivariate least-square fitting analysis. *J. Chem. Phys.* **2008,** *128* (13).
147. 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* (15).
148. Ghosh, A.; Remorino, A.; Tucker, M. J.; Hochstrasser, R. M., 2D IR photon echo spectroscopy reveals hydrogen bond dynamics of aromatic nitriles. *Chem. Phys. Lett.* **2009,** *469* (4-6), 325-330.
149. Lindquist, B. A.; Corcelli, S. A., Nitrile groups as vibrational probes: calculations of the C N infrared absorption line shape of acetonitrile in water and tetrahydrofuran. *J. Phys. Chem. B* **2008,** *112* (20), 6301-6303.
150. Lindquist, B. A.; Haws, R. T.; Corcelli, S. A., Optimized quantum mechanics/molecular mechanics strategies for nitrile vibrational probes: acetonitrile and para-tolunitrile in water and tetrahydrofuran. *J. Phys. Chem. B* **2008,** *112* (44), 13991-14001.

151. Weeks, C. L.; Polishchuk, A.; Getahun, Z.; DeGrado, W. F.; Spiro, T. G., Investigation of an unnatural amino acid for use as a resonance Raman probe: detection limits and solvent and temperature dependence of the $\upsilon$C≡N band of 4-cyanophenylalanine. *J. Raman Spectrosc.* **2008,** *39* (11), 1606-1613.
152. 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 β 16-22 (Ac-KLVFFAE-NH2). *Proteins-Structure Function and Bioinformatics* **2010,** *78* (10), 2306-2321.
153. 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* (43), 13984-13985.
154. Zou, H. L.; Liu, J.; Blasie, J. K., Mechanism of interaction between the general anesthetic halothane and a model ion channel protein, III: molecular dynamics simulation incorporating a cyanophenylalanine spectroscopic probe. *Biophys. J.* **2009,** *96* (10), 4188-4199.
155. Liu, J.; Strzalka, J.; Tronin, A.; Johansson, J. S.; Blasie, J. K., Mechanism of Interaction between the General Anesthetic Halothane and a Model Ion Channel Protein, II: Fluorescence and Vibrational Spectroscopy Using a Cyanophenylalanine Probe. *Biophys. J.* **2009,** *96* (10), 4176-4187.
156. Fang, C.; Bauman, J. D.; Das, K.; Remorino, A.; Arnold, E.; Hochstrasser, R. M., Two-dimensional infrared spectra reveal relaxation of the nonnucleoside inhibitor TMC278 complexed with HIV-1 reverse transcriptase. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105* (5), 1472-1477.
157. Webb, L. J.; Boxer, S. G., Electrostatic fields near the active site of human aldose reductase: 1. New inhibitors and vibrational stark effect measurements. *Biochemistry* **2008,** *47* (6), 1588-1598.
158. Silverman, L. N.; Pitzer, M. E.; Ankomah, P. O.; Boxer, S. G.; Fenlon, E. E., Vibrational stark effect probes for nucleic acids. *J. Phys. Chem. B* **2007,** *111* (40), 11611-11613.
159. Krummel, A. T.; Zanni, M. T., Evidence for coupling between nitrile groups using DNA templates: a promising new method for monitoring structures with infrared spectroscopy. *J. Phys. Chem. B* **2008,** *112* (5), 1336-1338.
160. Watson, M. D.; Gai, X. S.; Gillies, A. T.; Brewer, S. H.; Fenlon, E. E., A vibrational probe for local nucleic acid environments: 5-cyano-2 '-deoxyuridine. *J. Phys. Chem. B* **2008,** *112* (42), 13188-13192.
161. Fafarman, A. T.; Boxer, S. G., Nitrile bonds as infrared probes of electrostatics in ribonuclease S. *J. Phys. Chem. B* **2010,** *114* (42), 13536-13544.
162. Stafford, A. J.; Ensign, D. L.; Webb, L. J., Vibrational stark effect spectroscopy at the interface of Ras and Rap1A bound to the Ras binding domain of RalGDS reveals an electrostatic mechanism for protein-protein interaction. *J. Phys. Chem. B* **2010,** *114* (46), 15331-15344.
163. McMahon, H. A.; Alfieri, K. N.; Clark, C. A. A.; Londergan, C. H., Cyanylated Cysteine: A Covalently Attached Vibrational Probe of Protein-Lipid Contacts. *J. Phys. Chem. Lett.* **2010,** *1* (5), 850-855.
164. Bischak, C. G.; Longhi, S.; Snead, D. M.; Costanzo, S.; Terrer, E.; Londergan, C. H., Probing structural transitions in the intrinsically disordered C-terminal domain of the measles virus nucleoprotein by vibrational spectroscopy of cyanylated cysteines. *Biophys. J.* **2010,** *99* (5), 1676-1683.
165. 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* (41), 13356-13357.
166. Lee, H.; Choi, J. H.; Cho, M., Vibrational solvatochromism and electrochromism of cyanide, thiocyanate, and azide anions in water. *Phys. Chem. Chem. Phys.* **2010,** *12* (39), 12658-12669.
167. Ragain, C. M.; Newberry, R. W.; Ritchie, A. W.; Webb, L. J., Role of electrostatics in differential binding of RalGDS to Rap mutations E30D and K31E investigated by vibrational spectroscopy of thiocyanate probes. *J. Phys. Chem. B* **2012,** *116* (31), 9326-36.
168. Munde, M.; Poon, G. M.; Wilson, W. D., Probing the electrostatics and pharmacological modulation of sequence-specific binding by the DNA-binding domain of the ETS family transcription factor PU.1: a binding affinity and kinetics investigation. *J. Mol. Biol.* **2013,** *425* (10), 1655-69.
169. Fried, S. D.; Bagchi, S.; Boxer, S. G., Measuring electrostatic fields in both hydrogen-bonding and non-hydrogen-bonding environments using carbonyl vibrational probes. *J. Am. Chem. Soc.* **2013,** *135* (30), 11181-92.

170. George, W. O.; Houston, T. E.; Harris, W. C., Vibrational spectra and structure of esters I. Infrared and raman-spectra of $CH_3COOCH_3$, $CH_3COOCD_3$, $CD_3COOCH_3$ and $CD_3COOCD_3$. *Spectrochim. Acta, Part A* **1974,** *A 30* (4), 1035-1057.
171. Patel, K. B.; Eaton, G.; Symons, M. C. R., Solvation of esters and dialkyl carbonates. *J. Chem. Soc., Faraday Trans.* **1985,** *81*, 2775-2786.
172. Nie, B.; Stutzman, J.; Xie, A., A vibrational spectral maker for probing the hydrogen-bonding status of protonated Asp and Glu residues. *Biophys. J.* **2005,** *88* (4), 2833-47.
173. Pazos, I. M.; Roesch, R. M.; Gai, F., Quenching of p-cyanophenylalanine fluorescence by various anions. *Chem. Phys. Lett.* **2013,** *563* (0), 93-96.
174. Pazos, I. M.; Gai, F., Solute's perspective on how trimethylamine oxide, urea, and guanidine hydrochloride affect water's hydrogen bonding ability. *J Phys Chem B* **2012,** *116* (41), 12473-8.
175. Ma, J.; 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* (23), 8476-8481.
176. 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* (24), 6080-6084.
177. 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* (2), 405-411.
178. 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* (4-6), 249-253.
179. Bloem, R.; Koziol, K.; Waldauer, S. A.; Buchli, B.; Walser, R.; Samatanga, B.; Jelesarov, I.; Hamm, P., Ligand Binding Studied by 2D IR Spectroscopy Using the Azidohomoalanine Label. *J. Phys. Chem. B* **2012,** *116* (46), 13705-13712.
180. 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* (37), 8119-8132.
181. 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* (20), 2598-2609.
182. Aprilakis, K. N.; Taskent, H.; Raleigh, D. P., Use of the novel fluorescent amino acid rho-cyanophenylalanine offers a direct probe of hydrophobic core formation during the folding of the n-terminal domain of the ribosomal protein L9 and provides evidence for two-state folding. *Biochemistry* **2007,** *46* (43), 12308-12313.
183. 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* (10), 4788-4795.
184. Tucker, M. J.; Oyola, R.; Gai, F., A novel fluorescent probe for protein binding and folding studies: p-cyano-phenylalanine. *Biopolymers* **2006,** *83* (6), 571-576.
185. Tang, J.; Gai, F., Dissecting the membrane binding and insertion kinetics of a pHLIP peptide. *Biochemistry* **2008,** *47* (32), 8250-8252.
186. Marek, P.; Gupta, R.; Raleigh, D. P., The fluorescent amino acid p-cyanophenylalanine provides an intrinsic probe of amyloid formation. *Chembiochem* **2008,** *9* (9), 1372-1374.
187. Taskent-Sezgin, H.; Marek, P.; Thomas, R.; Goldberg, D.; Chung, J.; Carrico, I.; Raleigh, D. P., Modulation of *p*-Cyanophenylalanine Fluorescence by Amino Acid Side Chains and Rational Design of Fluorescence Probes of alpha-Helix Formation. *Biochemistry* **2010,** *49* (29), 6290-6295.
188. 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* (4-6), 303-306.
189. Glasscock, J. M.; Zhu, Y.; Chowdhury, P.; Tang, J.; Gai, F., Using an amino acid fluorescence resonance energy transfer pair to probe protein unfolding: Application to the villin headpiece subdomain and the LysM domain. *Biochemistry* **2008,** *47* (42), 11070-11076.
190. Miyake-Stoner, S. J.; Miller, A. M.; Hammill, J. T.; Peeler, J. C.; Hess, K. R.; Mehl, R. A.; Brewer, S. H., Probing Protein Folding Using Site-Specifically Encoded Unnatural Amino Acids as FRET Donors with Tryptophan. *Biochemistry* **2009,** *48* (25), 5953-5962.
191. Taskent-Sezgin, H.; Chung, J.; Patsalo, V.; Miyake-Stoner, S. J.; Miller, A. M.; Brewer, S. H.; Mehl, R. A.; Green, D. F.; Raleigh, D. P.; Carrico, I., Interpretation of p-Cyanophenylalanine Fluorescence

in Proteins in Terms of Solvent Exposure and Contribution of Side-Chain Quenchers: A Combined Fluorescence, IR and Molecular Dynamics Study. *Biochemistry* **2009,** *48* (38), 9040-9046.
192. Goldberg, J. M.; Batjargal, S.; Petersson, E. J., Thioamides as Fluorescence Quenching Probes: Minimalist Chromophores To Monitor Protein Dynamics. *J. Am. Chem. Soc.* **2010,** *132* (42), 14718-14720.
193. Eftink, M. R.; Ghiron, C. A., Fluorescence Quenching Studies with Proteins. *Anal. Biochem.* **1981,** *114* (2), 199.
194. Zhang, L. Y.; Kao, Y. T.; Qiu, W. H.; Wang, L. J.; Zhong, D. P., Femtosecond studies of tryptophan fluorescence dynamics in proteins: Local solvation and electronic quenching. *J. Phys. Chem. B* **2006,** *110* (37), 18097-18103.
195. Zhang, L.; Wang, L.; Kao, Y.-T.; Qiu, W.; Yang, Y.; Okobiah, O.; Zhong, D., Mapping hydration dynamics around a protein surface. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104* (47), 18461-18466.
196. Eftink, M. R.; Ghiron, C. A., Exposure of Tryptophanyl Residues in Proteins. Quantitative Determination by Fluorescence Quenching Studies. *Biochemistry* **1976,** *15* (3), 672-680.
197. Boaz, H.; Rollefson, G. K., The Quenching of Fluorescence. Deviations from the Stern-Volmer Law. *J. Am. Chem. Soc.* **1950,** *72* (8), 3435-3443.
198. Lehrer, S. S., Solute Perturbation of Protein Fluorescence. Quenching of Tryptophyl Fluorescence of Model Compunds and Lysozyme by Iodide ion. *Biochemistry* **1971,** *10* (17), 3254.
199. Zelent, B.; Kusba, 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* (1-2).
200. Szabo, A. G.; Rayner, D. M., Fluorescence Decay of Tryptophan Conformers in Aqueous Solution. *J. Am. Chem. Soc.* **1980,** *102* (2), 554-563.
201. Kunz, W.; Lo Nostro, P.; Ninham, B. W., The present state of affairs with Hofffieister effects. *Curr. Opin. Colloid Interface Sci.* **2004,** *9* (1-2), 1-18.
202. Meloni, S.; Matsika, S., Theoretical studies of the excited states of p-cyanophenylalanine and comparisons with the natural amino acids phenylalanine and tyrosine. *Theor Chem Acc* **2014,** *133* (7), 1-13.
203. Schultz, K. C.; Supekova, L.; Ryu, Y.; Xie, J.; Perera, R.; Schultz, P. G., A Genetically Encoded Infrared Probe. *J. Am. Chem. Soc.* **2006,** *128* (43), 13984-13985.
204. Du, D.; Liu, H.; Ojha, B., Study protein folding and aggregation using nonnatural amino acid p-cyanophenylalanine as a sensitive optical probe. *Methods Mol. Biol. (N. Y.)* **2013,** *1081*, 77-89.
205. 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* (4), 878-888.
206. Tang, J.; Signarvic, R. S.; DeGrado, W. F.; Gai, F., Role of Helix Nucleation in the Kinetics of Binding of Mastoparan X to Phospholipid Bilayers. *Biochemistry* **2007,** *46* (48), 13856-13863.
207. Loll, P. J.; Upton, E. C.; Nahoum, V.; Economou, N. J.; Cocklin, S., The high resolution structure of tyrocidine A reveals an amphipathic dimer. *Biochim. Biophys. Acta, Biomembr.* **2014,** *1838* (5), 1199-1207.
208. Tucker, M. J.; Tang, J.; Gai, F., Probing the kinetics of membrane-mediated helix folding. *J. Phys. Chem. B* **2006,** *110* (15), 8105-8109.
209. Wissner, R. F.; Batjargal, S.; Fadzen, C. M.; Petersson, E. J., Labeling Proteins with Fluorophore/Thioamide Forster Resonant Energy Transfer Pairs by Combining Unnatural Amino Acid Mutagenesis and Native Chemical Ligation. *J. Am. Chem. Soc.* **2013,** *135* (17), 6529-6540.
210. Reif, M. M.; Oostenbrink, C., Molecular Dynamics Simulation of Configurational Ensembles Compatible with Experimental FRET Efficiency Data Through a Restraint on Instantaneous FRET Efficiencies. *J. Comput. Chem.* **2014,** *35* (32), 2319-2332.
211. Mintzer, M. R.; Troxler, T.; Gai, F., p-Cyanophenylalanine and selenomethionine constitute a useful fluorophore-quencher pair for short distance measurements: application to polyproline peptides. *Phys. Chem. Chem. Phys.* **2015,** *17* (12), 7881-7887.

212. Hallbrink, M.; Floren, A.; Elmquist, A.; Pooga, M.; Bartfai, T.; Langel, U., Biochimica et Biophysica Acta, Biomembranes. *Biochim. Biophys. Acta, Biomembr.* **2001,** *1515* (2), 101-109.
213. Shen, D.; Liang, K.; Ye, Y.; Tetteh, E.; Achilefu, S., Modulation of nuclear internalization of Tat peptides by fluorescent dyes and receptor-avid peptides. *Febs Letters* **2007,** *581* (9), 1793-1799.
214. Jones, A. T.; Sayers, E. J., Cell entry of cell penetrating peptides: tales of tails wagging dogs. *J. Controlled Release* **2012,** *161* (2), 582-591.
215. Tseng, Y. L.; Liu, J. J.; Hong, R. L., Translocation of liposomes into cancer cells by cell-penetrating peptides penetratin and TAT: A kinetic and efficacy study. *Mol. Pharmacol.* **2002,** *62* (4), 864-872.
216. Wallis, M., p$K_a$ Values of α-Amino Groups of Peptides Derived from N-Terminus of Bovine Growth-Hormone. *Biochimica Et Biophysica Acta* **1973,** *310* (2), 388-397.
217. Chakrabartty, A.; Doig, A. J.; Baldwin, R. L., Helix capping propensities in peptides parallel those in proteins. *Proc. Natl. Acad. Sci. U.S.A.* **1993,** *90* (23), 11332-11336.
218. Doig, A. J.; Baldwin, R. L., N- and C-Capping Preferences for all 20 Amino-Acids in α-Helical Peptides. *Protein Sci.* **1995,** *4* (7), 1325-1336.
219. Wilcox, W.; Eisenberg, D., Thermodynamics of Melittin Tetramerization Determined by Circular-Dichroism and Implications for Protein Folding. *Protein Sci.* **1992,** *1* (5), 641-653.
220. Goto, Y.; Hagihara, Y., Mechanism of the Conformational Transition of Melittin. *Biochemistry* **1992,** *31* (3), 732-738.
221. Zhu, L. Y.; Kemple, M. D.; Yuan, P.; Prendergast, F. G., N-Terminus and Lysine Side-Chain p$K_a$ Values of Melittin in Aquesous-Solutions and Micellar Dispersions Measured by $^{15}$N NMR. *Biochemistry* **1995,** *34* (40), 13196-13202.
222. Brown, L. R.; Lauterwein, J.; Wuthrich, K., High Resolution H-1-NMR Studies of Self-Aggregation of Melittin in Aqueous Solution. *Biochimica Et Biophysica Acta* **1980,** *622* (2), 231-244.
223. Pace, C. N.; Grimsley, G. R.; Scholtz, J. M., Protein Ionizable Groups: pK Values and Their Contribution to Protein Stability and Solubility. *J. Biol. Chem.* **2009,** *284* (20), 13285-13289.
224. Zorko, M.; Langel, U., Cell-penetrating peptides: mechanism and kinetics of cargo delivery. *Adv. Drug Delivery Rev.* **2005,** *57* (4), 529-545.
225. Swiecicki, J.-M.; Bartsch, A.; Tailhades, J.; Di Pisa, M.; Heller, B.; Chassaing, G.; Mansuy, C.; Burlina, F.; Lavielle, S., The Efficacies of Cell-Penetrating Peptides in Accumulating in Large Unilamellar Vesicles Depend on their Ability To Form Inverted Micelles. *Chembiochem* **2014,** *15* (6), 884-891.
226. Aurora, T. S.; Li, W.; Cummins, H. Z.; Haines, T. H., Preparation and Chracterization of Monodisperse Unilamellar Phospholipid Vesicles with Selected Diameters of from 300 to 600 nm. *Biochimica Et Biophysica Acta* **1985,** *820* (2), 250-258.
227. 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* (30), 12283-12288.
228. Grimsley, G. R.; Scholtz, J. M.; Pace, C. N., A summary of the measured pK values of the ionizable groups in folded proteins. *Protein Sci.* **2009,** *18* (1), 247-251.
229. Ciobanasu, C.; Siebrasse, J. P.; Kubitscheck, U., Cell-Penetrating HIV1 TAT Peptides Can Generate Pores in Model Membranes. *Biophys. J.* **2010,** *99* (1), 153-162.
230. Matsuzaki, K.; Yoneyama, S.; Murase, O.; Miyajima, K., Transbilayer Transport of Ions and Lipids Coupled with Mastoparan X Translocation. *Biochemistry* **1996,** *35* (25), 8450-8456.
231. Suzuki, T.; Futaki, S.; Niwa, M.; Tanaka, S.; Ueda, K.; Sugiura, Y., Possible existence of common internalization mechanisms among arginine-rich peptides. *J. Biol. Chem.* **2002,** *277* (4), 2437-2443.
232. Richard, J. P.; Melikov, K.; Vives, E.; Ramos, C.; Verbeure, B.; Gait, M. J.; Chernomordik, L. V.; Lebleu, B., Cell-penetrating peptides - A reevaluation of the mechanism of cellular uptake. *J. Biol. Chem.* **2003,** *278* (1), 585-590.
233. Feigenson, G. W., Phase diagrams and lipid domains in multicomponent lipid bilayer mixtures. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2009,** *1788* (1), 47-52.

234. Marsh, D., Cholesterol-induced fluid membrane domains: A compendium of lipid-raft ternary phase diagrams. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **2009,** *1788* (10), 2114-2123.
235. Boyer, R. F., *Concepts in biochemistry*. 3rd ed.; Wiley: Hoboken, NJ, 2006; p xvi, 720 p.
236. Tanford, C., Isothermal Unfolding of Globular Proteins in Aqueous Urea Solutions. *J. Am. Chem. Soc.* **1964,** *86* (10), 2050-&.
237. Greene, R. F.; Pace, C. N., Urea and Guanidine-Hydrochloride Denaturation of Ribonuclease, Lysozyme, α–Chymotrypsin, and β-Lactoglobulin. *J. Biol. Chem.* **1974,** *249* (17), 5388-5393.
238. Alonso, D. O. V.; Dill, K. A., Solvent Denaturation and Stabilization of Globular-Proteins. *Biochemistry* **1991,** *30* (24), 5974-5985.
239. Duffy, E. M.; Kowalczyk, P. J.; Jorgensen, W. L., Do Denaturants Interact with Aromatic-Hydrocarbons in Water? *J. Am. Chem. Soc.* **1993,** *115* (20), 9271-9275.
240. Uversky, V. N.; Li, J.; Fink, A. L., Trimethylamine-N-oxide-induced folding of alpha-synuclein. *Febs Letters* **2001,** *509* (1), 31-35.
241. Athawale, M. V.; Dordick, J. S.; Garde, S., Osmolyte trimethylamine-N-oxide does not affect the strength of hydrophobic interactions: Origin of osmolyte compatibility. *Biophys. Chem.* **2005,** *89* (2), 858-866.
242. Caballero-Herrera, A.; Nordstrand, K.; Berndt, K. D.; Nilsson, L., Effect of Urea on Peptide Conformation in Water: Molecular Dynamics and Experimental Characterization. *Biophys. J.* **2005,** *89* (2), 842-857.
243. Street, T. O.; Bolen, D. W.; Rose, G. D., A Molecular Mechanism for Osmolyte-Induced Protein Stability. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103* (38), 13997-14002.
244. Auton, M.; Holthauzen, L. M. F.; Bolen, D. W., Anatomy of Energetic Changes Accompanying Urea-Induced Protein Denaturation. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104* (39), 15317-15322.
245. Di Michele, A.; Freda, M.; Onori, G.; Paolantoni, M.; Santucci, A.; Sassi, P., Modulation of hydrophobic effect by cosolutes. *J. Phys. Chem. B* **2006,** *110* (42), 21077-21085.
246. Hua, L.; Zhou, R.; Thirumalai, D.; Berne, B. J., Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105* (44), 16928-16933.
247. Paul, S.; Patey, G. N., Hydrophobic interactions in urea - Trimethylamine-N-oxide solutions. *J. Phys. Chem. B* **2008,** *112* (35), 11106-11111.
248. Meersman, F.; Bowron, D.; Soper, A. K.; Koch, M. H. J., Counteraction of Urea by Trimethylamine N-Oxide Is Due to Direct Interaction. *Biophys. J.* **2009,** *97* (9), 2559-2566.
249. Kuffel, A.; Zielkiewicz, J., The hydrogen bond network structure within the hydration shell around simple osmolytes: Urea, tetramethylurea, and trimethylamine-N-oxide, investigated using both a fixed charge and a polarizable water model. *J. Chem. Phys.* **2010,** *133* (3).
250. Canchi, D. R.; Paschek, D.; Garcia, A. E., Equilibrium Study of Protein Denaturation by Urea. *J. Am. Chem. Soc.* **2010,** *132* (7), 2338-2344.
251. Holthauzen, L. M. F.; Roesgen, J.; Bolen, D. W., Hydrogen Bonding Progressively Strengthens upon Transfer of the Protein Urea-Denatured State to Water and Protecting Osmolytes. *Biochemistry* **2010,** *49* (6), 1310-1318.
252. Yang, Z.; Xiu, P.; Shi, B.; Hua, L.; Zhou, R., Coherent microscopic picture for urea-induced denaturation of proteins. *J Phys Chem B* **2012,** *116* (30), 8856-62.
253. Rezus, Y. L. A.; Bakker, H. J., Effect of urea on the structural dynamics of water. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103* (49), 18417-18420.
254. Bakulin, A. A.; Pshenichnikov, M. S.; Bakker, H. J.; Petersen, C., Hydrophobic Molecules Slow Down the Hydrogen-Bond Dynamics of Water. *J. Phys. Chem. A* **2011,** *115* (10), 1821-1829.
255. Rezus, Y. L. A.; Bakker, H. J., Observation of immobilized water molecules around hydrophobic groups. *Phys. Rev. Lett.* **2007,** *99* (14).
256. Guo, F.; Friedman, J. M., Osmolyte-Induced Perturbations of Hydrogen Bonding between Hydration Layer Waters: Correlation with Protein Conformational Changes. *J. Phys. Chem. B* **2009,** *113* (52), 16632-16642.

257. Rezus, Y. L. A.; Bakker, H. J., Destabilization of the Hydrogen-Bond Structure of Water by the Osmolyte Trimethylamine N-Oxide. *J. Phys. Chem. B* **2009,** *113* (13), 4038-4044.
258. Weeks, C. L.; Polishchuk, A.; Getahun, Z.; DeGrado, W. F.; Spiro, T. G., Investigation of an unnatural amino acid for use as a resonance Raman probe: detection limits and solvent and temperature dependence of the nu C equivalent to N band of 4-cyanophenylalanine. *J. Raman Spectrosc.* **2008,** *39* (11), 1606-1613.
259. Boxer, S. G., Stark Realities. *J. Phys. Chem. B* **2009,** *113* (10), 2972-2983.
260. Choi, J. H.; Oh, K. I.; Lee, H.; Lee, C.; Cho, M., Nitrile and thiocyanate IR probes: Quantum chemistry calculation studies and multivariate least-square fitting analysis. *J. Chem. Phys.* **2008,** *128* (13), 8.
261. Aschaffenburg, D. J.; Moog, R. S., Probing Hydrogen Bonding Environments: Solvatochromic Effects on the CN Vibration of Benzonitrile. *J. Phys. Chem. B* **2009,** *113* (38), 12736-12743.
262. Waegele, M. M.; Gai, F., Computational Modeling of the Nitrile Stretching Vibration of 5-Cyanoindole in Water. *J. Phys. Chem. Lett.* **2010,** *1* (4), 781-786.
263. Lee, H.; Choi, J.-H.; Cho, M., Vibrational solvatochromism and electrochromism of cyanide, thiocyanate, and azide anions in water. *Phys. Chem. Chem. Phys.* **2010,** *12* (39), 12658-12669.
264. Gai, X. S.; Coutifaris, B. A.; Brewer, S. H.; Fenlon, E. E., A direct comparison of azide and nitrile vibrational probes. *Phys. Chem. Chem. Phys.* **2011,** *13* (13), 5926-5930.
265. Zhang, Z.; Guo, Y.; Lu, Z.; Velarde, L.; Wang, H.-f., Resolving Two Closely Overlapping -CN Vibrations and Structure in the Langmuir Mono layer of the Long-Chain Nonadecanenitrile by Polarization Sum Frequency Generation Vibrational Spectroscopy. *J. Phys. Chem. C* **2012,** *116* (4), 2976-2987.
266. Stafford, A. J.; Walker, D. M.; Webb, L. J., Electrostatic Effects of Mutations of Ras Glutamine 61 Measured Using Vibrational Spectroscopy of a Thiocyanate Probe. *Biochemistry* **2012,** *51* (13), 2757-2767.
267. Bagchi, S.; Fried, S. D.; Boxer, S. G., A Solvatochromic Model Calibrates Nitriles' Vibrational Frequencies to Electrostatic Fields. *J. Am. Chem. Soc.* **2012,** *134* (25), 10373-10376.
268. Huang, C. Y.; Getahun, Z.; Zhu, Y. J.; Klemke, J. W.; DeGrado, W. F.; Gai, F., Helix formation via conformation diffusion search. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99* (5), 2788-2793.
269. Purcell, K. F.; Drago, R. S., STUDIES OF BONDING IN ACETONITRILE ADDUCTS. *J. Am. Chem. Soc.* **1966,** *88* (5), 919-&.
270. Fawcett, W. R.; Liu, G. J.; Kessler, T. E., Solvent-Induced Frequency Shifts in the Infrared Spectrum of Acetonitrile in Organic Solvents. *J. Phys. Chem.* **1993,** *97* (37), 9293-9298.
271. Mountain, R. D., Microstructure and Hydrogen Bonding in Water-Acetonitrile Mixtures. *J. Phys. Chem. B* **2010,** *114* (49), 16460-16464.
272. Eaton, G.; Penanunez, A. S.; Symons, M. C. R., Solvation of Cyanoalkanesv $CH_3CN$ AND $(CH_3)_3CCN$ - An Infrared and Nuclear Magnetic-Resonance Study. *J. Chem. Soc., Faraday Trans.* **1988,** *84*, 2181-2193.
273. Huang, C. Y.; Wang, T.; Gai, F., Temperature dependence of the CN stretching vibration of a nitrile-derivatized phenylalanine in water. *Chem. Phys. Lett.* **2003,** *371* (5-6), 731-738.
274. Lee, H.; Lee, G.; Jeon, J.; Cho, M., Vibrational Spectroscopic Determination of Local Solvent Electric Field, Solute-Solvent Electrostatic Interaction Energy, and Their Fluctuation Amplitudes. *J. Phys. Chem. A* **2012,** *116* (1), 347-357.
275. Andrews, S. S.; Boxer, S. G., Vibrational Stark effects of nitriles II. Physical origins of stark effects from experiment and perturbation models. *J. Phys. Chem. A* **2002,** *106* (3), 469-477.
276. Mountain, R. D.; Thirumalai, D., Alterations in water structure induced by guanidinium and sodium ions. *J. Phys. Chem. B* **2004,** *108* (51), 19711-19716.
277. Collins, K. D., Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Methods* **2004,** *34* (3), 300-311.
278. Ham, S.; Kim, J. H.; Lee, H.; Cho, M. H., Correlation between electronic and molecular structure distortions and vibrational properties. II. Amide I modes of NMA-nD(2)O complexes. *J. Chem. Phys.* **2003,** *118* (8), 3491-3498.

279. DeCamp, M. F.; DeFlores, L.; McCracken, J. M.; Tokmakoff, A.; Kwac, K.; Cho, M., Amide I vibrational dynamics of N-methylacetamide in polar solvents: The role of electrostatic interactions. *J. Phys. Chem. B* **2005,** *109* (21), 11016-11026.
280. Mukherjee, S.; Chowdhury, P.; Gai, F., Infrared study of the effect of hydration on the amide I band and aggregation properties of helical peptides. *J. Phys. Chem. B* **2007,** *111* (17), 4596-4602.
281. Wei, H.; Fan, Y.; Gao, Y. Q., Effects of Urea, Tetramethyl Urea, and Trimethylamine N-Oxide on Aqueous Solution Structure and Solvation of Protein Backbones: A Molecular Dynamics Simulation Study. *J. Phys. Chem. B* **2010,** *114* (1), 557-568.
282. Sagle, L. B.; Zhang, Y.; Litosh, V. A.; Chen, X.; Cho, Y.; Cremer, P. S., Investigating the Hydrogen-Bonding Model of Urea Denaturation. *J. Am. Chem. Soc.* **2009,** *131* (26), 9304-9310.
283. Hwang, S.; Shao, Q.; Williams, H.; Hilty, C.; Gao, Y. Q., Methanol Strengthens Hydrogen Bonds and Weakens Hydrophobic Interactions in Proteins - A Combined Molecular Dynamics and NMR study. *J. Phys. Chem. B* **2011,** *115* (20), 6653-6660.
284. Reddy, P. M.; Taha, M.; Venkatesu, P.; Kumar, A.; Lee, M.-J., Destruction of hydrogen bonds of poly(N-isopropylacrylamide) aqueous solution by trimethylamine N-oxide. *J. Chem. Phys.* **2012,** *136* (23), 234904.
285. Tran, H. T.; Mao, A.; Pappu, R. V., Role of backbone - Solvent interactions in determining conformational equilibria of intrinsically disordered proteins. *J. Am. Chem. Soc.* **2008,** *130* (23), 7380-7392.
286. Mason, P. E.; Neilson, G. W.; Dempsey, C. E.; Barnes, A. C.; Cruickshank, J. M., The hydration structure of guanidinium and thiocyanate ions: Implications for protein stability in aqueous solution. *Proc. Natl. Acad. Sci. U.S.A.* **2003,** *100* (8), 4557-4561.
287. Mason, P. E.; Neilson, G. W.; Enderby, J. E.; Saboungi, M. L.; Dempsey, C. E.; MacKerell, A. D.; Brady, J. W., The structure of aqueous guanidinium chloride solutions. *J. Am. Chem. Soc.* **2004,** *126* (37), 11462-11470.
288. Batchelor, J. D.; Olteanu, A.; Tripathy, A.; Pielak, G. J., Impact of protein denaturants and stabilizers on water structure. *J. Am. Chem. Soc.* **2004,** *126* (7), 1958-1961.
289. Yancey, P. H., Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* **2005,** *208* (15), 2819-2830.
290. Wang, A.; Bolen, D. W., A Naturally Occurring Protective System in Urea-Rich Cells: Mechanism of Osmolyte Protection of Proteins against Urea Denaturation†. *Biochemistry* **1997,** *36* (30), 9101-9108.
291. Koga, Y.; Westh, P.; Nishikawa, K.; Subramanian, S., Is a Methyl Group Always Hydrophobic? Hydrophilicity of Trimethylamine-N-oxide, Tetramethyl Urea and Tetramethylammonium Ion. *J. Phys. Chem. B* **2011,** *115* (12), 2995-3002.
292. Munroe, K. L.; Magers, D. H.; Hammer, N. I., Raman Spectroscopic Signatures of Noncovalent Interactions Between Trimethylamine N-oxide (TMAO) and Water. *J. Phys. Chem. B* **2011,** *115* (23), 7699-7707.
293. Sagle, L. B.; Cimatu, K.; Litosh, V. A.; Liu, Y.; Flores, S. C.; Chen, X.; Yu, B.; Cremer, P. S., Methyl Groups of Trimethylamine N-Oxide Orient Away from Hydrophobic Interfaces. *J. Am. Chem. Soc.* **2011,** *133* (46), 18707-18712.
294. Larini, L.; Shea, J.-E., Double Resolution Model for Studying TMAO/Water Effective Interactions. *J. Phys. Chem. B* **2013,** *117* (42), 13268-13277.
295. Paul, S.; Patey, G. N., Structure and interaction in aqueous urea-trimethylamine-N-oxide solutions. *J. Am. Chem. Soc.* **2007,** *129* (14), 4476-4482.
296. O'Brien, E. P.; Ziv, G.; Haran, G.; Brooks, B. R.; Thirumalai, D., Effects of denaturants and osmolytes on proteins are accurately predicted by the molecular transfer model. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105* (36), 13403-13408.
297. Rösgen, J.; Jackson-Atogi, R., Volume Exclusion and H-Bonding Dominate the Thermodynamics and Solvation of Trimethylamine-N-oxide in Aqueous Urea. *J. Am. Chem. Soc.* **2012,** *134* (7), 3590-3597.

298. Fiore, A.; Venkateshwaran, V.; Garde, S., Trimethylamine N-Oxide (TMAO) and tert-Butyl Alcohol (TBA) at Hydrophobic Interfaces: Insights from Molecular Dynamics Simulations. *Langmuir* **2013,** *29* (25), 8017-8024.
299. Macdonald, R. D.; Khajehpour, M., Effects of the osmolyte TMAO (Trimethylamine-N-oxide) on aqueous hydrophobic contact-pair interactions. *Biophys. Chem.* **2013,** *184*, 101-107.
300. Bruzdziak, P.; Panuszko, A.; Stangret, J., Influence of Osmolytes on Protein and Water Structure: A Step To Understanding the Mechanism of Protein Stabilization. *J. Phys. Chem. B* **2013,** *117* (39), 11502-11508.
301. Qu, Y. X.; Bolen, D. W., Hydrogen Exchange Kinetics of RNase A and the Urea:TMAO Paradigm. *Biochemistry* **2003,** *42* (19), 5837-5849.
302. Kim, Y. S.; Liu, L.; Axelsen, P. H.; Hochstrasser, R. M., 2D IR provides evidence for mobile water molecules in β-amyloid fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (42), 17751-17756.
303. 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* (8), 1236-1243.
304. Kuroda, D. G.; Bauman, J. D.; Challa, J. R.; Patel, D.; Troxler, T.; Das, K.; Arnold, E.; Hochstrasser, R. M., Snapshot of the equilibrium dynamics of a drug bound to HIV-1 reverse transcriptase. *Nat. Chem.* **2013,** *5* (3), 174-181.
305. Strasfeld, D. B.; Ling, Y. L.; Shim, S.-H.; Zanni, M. T., Tracking fiber formation in human islet amyloid polypeptide with automated 2D-IR Spectroscopy. *J. Am. Chem. Soc.* **2008,** *130* (21), 6698-+.
306. Strasfeld, D. B.; Ling, Y. L.; Gupta, R.; Raleigh, D. P.; Zanni, M. T., Strategies for Extracting Structural Information from 2D IR Spectroscopy of Amyloid: Application to Islet Amyloid Polypeptide. *J. Phys. Chem. B* **2009,** *113* (47), 15679-15691.
307. 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* (15), 6115-6120.
308. Chung, J. K.; Thielges, M. C.; Fayer, M. D., Dynamics of the Folded and Unfolded Villin Headpiece (HP35) Measured with Ultrafast 2D IR Vibrational Echo Spectroscopy. *Proc. Natl. Acad. Sci. U.S.A.* **2011,** *108* (9), 3578-3583.
309. Bagchi, S.; Boxer, S. G.; Fayer, M. D., Ribonuclease S Dynamics Measured Using a Nitrile Label with 2D IR Vibrational Echo Spectroscopy. *J. Phys. Chem. B* **2012,** *116* (13), 4034-4042.
310. 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* (5), 355-360.
311. Park, K.-H.; Jeon, J.; Park, Y.; Lee, S.; Kwon, H.-J.; Joo, C.; Park, S.; Han, H.; Cho, M., Infrared Probes Based on Nitrile-Derivatized Prolines: Thermal Insulation Effect and Enhanced Dynamic Range. *J. Phys. Chem. Lett.* **2013,** *4* (13), 2105-2110.
312. Mukamel, S., *Principles of nonlinear optical spectroscopy*. Oxford University Press: New York, 1995; p xviii, 543 p.
313. Cho, M., *Two-dimensional optical spectroscopy*. CRC Press: Boca Raton, 2009; p 378 p.
314. Volkov, V. V.; Palmer, D. J.; Righini, R., Distinct water species confined at the interface of a phospholipid membrane. *Phys. Rev. Lett.* **2007,** *99* (7).
315. Hochstrasser, R. M., Two-dimensional IR-spectroscopy: polarization anisotropy effects. *Chem. Phys.* **2001,** *266* (2-3), 273-284.
316. Huang, C. Y.; Getahun, Z.; Wang, T.; DeGrado, W. F.; Gai, F., Time-resolved infrared study of the helix-coil transition using C-13-labeled helical peptides. *J. Am. Chem. Soc.* **2001,** *123* (48), 12111-12112.
317. Wang, T.; Du, D. G.; Gai, F., Helix-coil kinetics of two 14-residue peptides. *Chem. Phys. Lett.* **2003,** *370* (5-6), 842-848.
318. Lin, Y. S.; Shorb, J. M.; Mukherjee, P.; Zanni, M. T.; Skinner, J. L., Empirical Amide I Vibrational Frequency Map: Application to 2D-IR Line Shapes for Isotope-Edited Membrane Peptide Bundles†. *J. Phys. Chem. B* **2008,** *113* (3), 592-602.

319. Asbury, J. B.; Steinel, T.; Stromberg, C.; Corcelli, S. A.; Lawrence, C. P.; Skinner, J. L.; Fayer, M. D., Water dynamics: Vibrational echo correlation spectroscopy and comparison to molecular dynamics simulations. *J. Phys. Chem. A* **2004,** *108* (7), 1107-1119.
320. Canchi, D. R.; Jayasimha, P.; Rau, D. C.; Makhatadze, G. I.; Garcia, A. E., Molecular Mechanism for the Preferential Exclusion of TMAO from Protein Surfaces. *J. Phys. Chem. B* **2012,** *116* (40), 12095-12104.
321. Hu, K. N.; Havlin, R. H.; Yau, W. M.; Tycko, R., Quantitative Determination of Site-Specific Conformational Distributions in an Unfolded Protein by Solid-State Nuclear Magnetic Resonance. *J. Mol. Biol.* **2009,** *392* (4), 1055-1073.
322. 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* (36), 11024-11031.
323. Chung, J. K.; Thielges, M. C.; Fayer, M. D., Conformational Dynamics and Stability of HP35 Studied with 2D IR Vibrational Echoes. *J. Am. Chem. Soc.* **2012,** *134* (29), 12118-12124.
324. Canchi, D. R.; García, A. E., Cosolvent Effects on Protein Stability. *Annu. Rev. Phys. Chem.* **2013,** *64* (1), 273-293.
325. Cannon, J. G.; Anderson, C. F.; Record, M. T., Jr., Urea-amide preferential interactions in water: Quantitative comparison of model compound data with biopolymer results using water accessible surface areas. *J. Phys. Chem. B* **2007,** *111* (32), 9675-9685.
326. Zhang, W.; Capp, M. W.; Bond, J. P.; Anderson, C. F.; Record, M. T., Thermodynamic Characterization of Interactions of Native Bovine Serum Albumin with Highly Excluded (Glycine Betaine) and Moderately Accumulated (Urea) Solutes by a Novel Application of Vapor Pressure Osmometry. *Biochemistry* **1996,** *35* (32), 10506-10516.
327. Lee, S.; Shek, Y. L.; Chalikian, T. V., Urea interactions with protein groups: A volumetric study. *Biopolymers* **2010,** *93* (10), 866-879.
328. Arakawa, T.; Timasheff, S. N., Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry* **1982,** *21* (25), 6545-6552.
329. King, J. T.; Baiz, C. R.; Kubarych, K. J., Solvent-Dependent Spectral Diffusion in a Hydrogen Bonded "Vibrational Aggregate". *J. Phys. Chem. A* **2010,** *114* (39), 10590-10604.
330. Doan-Nguyen, V.; Loria, J. P., The effects of cosolutes on protein dynamics: The reversal of denaturant-induced protein fluctuations by trimethylamine N-oxide. *Protein Sci.* **2007,** *16* (1), 20-29.
331. Qu, Y.; Bolen, D. W., Efficacy of Macromolecular Crowding in Forcing Proteins to Fold. *Biophys. Chem.* **2002,** *101–102* (0), 155-165.
332. Kohlhoff, K. J.; Shukla, D.; Lawrenz, M.; Bowman, G. R.; Konerding, D. E.; Belov, D.; Altman, R. B.; Pande, V. S., Cloud-based simulations on Google Exacycle reveal ligand modulation of GPCR activation pathways. *Nat Chem* **2014,** *6* (1), 15-21.
333. Ragain, C. M.; Newberry, R. W.; Ritchie, A. W.; Webb, L. J., Role of Electrostatics in Differential Binding of RaIGDS to Rap Mutations E30D and K31E Investigated by Vibrational Spectroscopy of Thiocyanate Probes. *J. Phys. Chem. B* **2012,** *116* (31), 9326-9336.
334. Kim, H.; Cho, M., Infrared Probes for Studying the Structure and Dynamics of Biomolecules. *Chem. Rev.* **2013,** *113*, 5817-5847.
335. Volk, M.; Kholodenko, Y.; Lu, H. S. M.; Gooding, E. A.; DeGrado, W. F.; Hochstrasser, R. M., Peptide Conformational Dynamics and Vibrational Stark Effects Following Photoinitiated Disulfide Cleavage. *J. Phys. Chem. B* **1997,** *101* (42), 8607-8616.
336. 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* (29), 12241-12252.
337. Dyer, R. B.; Gai, F.; Woodruff, W. H., Infrared studies of fast events in protein folding. *Acc. Chem. Res.* **1998,** *31* (11), 709-716.
338. Woutersen, S.; Pfister, R.; Hamm, P.; Mu, Y. G.; Kosov, D. S.; Stock, G., Peptide conformational heterogeneity revealed from nonlinear vibrational spectroscopy and molecular-dynamics simulations. *J. Chem. Phys.* **2002,** *117* (14), 6833-6840.

339. Demirdoven, N.; Cheatum, C. M.; Chung, H. S.; Khalil, M.; Knoester, J.; Tokmakoff, A., Two-dimensional infrared spectroscopy of antiparallel beta-sheet secondary structure. *J. Am. Chem. Soc.* **2004,** *126* (25), 7981-7990.
340. Fried, S. D.; Bagchi, S.; Boxer, S. G., Measuring Electrostatic Fields in Both Hydrogen-Bonding and Non-Hydrogen-Bonding Environments Using Carbonyl Vibrational Probes. *J. Am. Chem. Soc.* **2013,** *135* (30), 11181-11192.
341. George, W. O.; Houston, T. E.; Harris, W. C., Vibrational Spectra and Structure of Esters- I. Infrared and Raman Spectra of $CH_3COOCH_3$, $CH_3COOCD_3$, $CD_3COOCH_3$ and $CD_3COOCD_3$. *Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy* **1974,** *A 30* (4), 1035-1057.
342. Patel, K. B.; Eaton, G.; Symons, M. C. R., Solvation of Esters and Dialkyl Carbonates. *J. Chem. Soc., Faraday Trans.* **1985,** *81*, 2775-2786.
343. Nie, B.; Stutzman, J.; Xie, A., A Vibrational Spectral Marker for Probing the Hydrogen-Bonding Status of Protonated Asp and Glu Residues. *Biophys. J.* **2005,** *88* (4), 2833-2847.
344. Takimoto , J. K.; Xiang , Z.; Kang, J.-Y.; Wang, L., Esterification of an Unnatural Amino Acid Structurally Deviating from Canonical Amino Acids Promotes Its Uptake and Incorporation into Proteins in Mammalian Cells. *Chembiochem* **2010,** *11* (16), 2268-2272.
345. Kim, Y. S.; Hochstrasser, R. M., Chemical exchange 2D IR of hydrogen-bond making and breaking. *Proc. Natl. Acad. Sci. U.S.A.* **2005,** *102* (32), 11185-11190.
346. Hamm, P.; Lim, M. H.; Hochstrasser, R. M., Structure of the amide I band of peptides measured by femtosecond nonlinear-infrared spectroscopy. *J. Phys. Chem. B* **1998,** *102* (31), 6123-6138.
347. Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K., Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005,** *26* (16), 1781-1802.
348. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **1998,** *102* (18), 3586-3616.
349. Vanommeslaeghe, K.; Hatcher, E.; Acharya, C.; Kundu, S.; Zhong, S.; Shim, J.; Darian, E.; Guvench, O.; Lopes, P.; Vorobyov, I.; Mackerell, A. D., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *J. Comput. Chem.* **2010,** *31* (4), 671-690.
350. Humphrey, W.; Dalke, A.; Schulten, K., VMD: Visual molecular dynamics. *J. Mol. Graphics Modell.* **1996,** *14* (1), 33-38.
351. Kim, Y. S.; Hochstrasser, R. M., The 2D IR responses of amide and carbonyl modes in water cannot be described by Gaussian frequency fluctuations. *J. Phys. Chem. B* **2007,** *111* (33), 9697-9701.
352. Candelaresi, M.; Pagliai, M.; Lima, M.; Righini, R., Chemical Equilibrium Probed by Two-Dimensional IR Spectroscopy: Hydrogen Bond Dynamics of Methyl Acetate in Water. *J. Phys. Chem. A* **2009,** *113* (46), 12783-12790.
353. Banno, M.; Ohta, K.; Yamaguchi, S.; Hirai, S.; Tominaga, K., Vibrational Dynamics of Hydrogen-Bonded Complexes in Solutions Studied with Ultrafast Infrared Pump-Probe Spectroscopy. *Acc. Chem. Res.* **2009,** *42* (9), 1259-1269.
354. Onsager, L., Electric Moments of Molecules in Liquids. *J. Am. Chem. Soc.* **1936,** *58* (8), 1486-1493.
355. Pensack, R. D.; Banyas, K. M.; Asbury, J. B., Vibrational solvatochromism in organic photovoltaic materials: method to distinguish molecules at donor/acceptor interfaces. *Phys. Chem. Chem. Phys.* **2010,** *12* (42), 14144-14152.
356. Fried, S. D.; Wang, L.-P.; Boxer, S. G.; Ren, P.; Pande, V. S., Calculations of the Electric Fields in Liquid Solutions. *J. Phys. Chem. B* **2013,** *117* (50), 16236-16248.

357. Colletier, J.-P.; Laganowsky, A.; Landau, M.; Zhao, M.; Soriaga, A. B.; Goldschmidt, L.; Flot, D.; Cascio, D.; Sawaya, M. R.; Eisenberg, D., Molecular basis for amyloid-beta polymorphism. *Proc. Natl. Acad. Sci. U.S.A.* **2011,** *108* (41), 16938-16943.
358. 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* (15).
359. Bunagan, M. R.; Gao, J.; Kelly, J. W.; Gai, F., Probing the folding transition state structure of the villin headpiece subdomain via side chain and backbone mutagenesis. *J. Am. Chem. Soc.* **2009,** *131* (21), 7470-6.
360. Ye, S.; Zaitseva, E.; Caltabiano, G.; Schertler, G. F. X.; Sakmar, T. P.; Deupi, X.; Vogel, R., Tracking G-protein-coupled receptor activation using genetically encoded infrared probes. *Nature* **2010,** *464* (7293), 1386-9.
361. Zimmermann, J.; Thielges, M. C.; Yu, W.; Dawson, P. E.; Romesberg, F. E., Carbon-Deuterium Bonds as Site-Specific and Nonperturbative Probes for Time-Resolved Studies of Protein Dynamics and Folding. *J. Phys. Chem. Lett.* **2011,** *2* (5), 412-416.
362. Tucker, M. J.; Gai, X. S.; Fenlon, E. E.; Brewer, S. H.; Hochstrasser, R. M., 2D IR photon echo of azido-probes for biomolecular dynamics. *Phys. Chem. Chem. Phys.* **2011,** *13* (6), 2237-2241.
363. Dutta, S.; Rock, W.; Cook, R. J.; Kohen, A.; Cheatum, C. M., Two-dimensional infrared spectroscopy of azido-nicotinamide adenine dinucleotide in water. *J. Chem. Phys.* **2011,** *135* (5).
364. Dybal, J.; Stokr, J.; Schneider, B., C=O Stretching Vibrations in Raman and Infrared-Spectra of Simple Esters. *Collect. Czech. Chem. Commun.* **1982,** *47* (8), 2027-2036.
365. Nafie, L. A., Recent advances in linear and nonlinear Raman spectroscopy. Part VII. *J. Raman Spectrosc.* **2013,** *44* (12), 1629-1648.
366. Dobson, C. M., Protein misfolding, evolution and disease. *Trends in Biochemical Sciences* **1999,** *24* (9), 329-332.
367. Soto, C., Protein misfolding and disease; protein refolding and therapy. *Febs Letters* **2001,** *498* (2-3), 204-207.
368. Tarus, B.; Straub, J. E.; Thirumalai, D., Probing the initial stage of aggregation of the Aβ(10-35)-protein: assessing the propensity for peptide dimerization. *J. Mol. Biol.* **2005,** *345* (5), 1141-56.
369. Tarus, B.; Straub, J. E.; Thirumalai, D., Dynamics of Asp23-Lys28 salt-bridge formation in Aβ10-35 monomers. *J. Am. Chem. Soc.* **2006,** *128* (50), 16159-68.
370. Lashuel, H. A.; LaBrenz, S. R.; Woo, L.; Serpell, L. C.; Kelly, J. W., Protofilaments, Filaments, Ribbons, and Fibrils from Peptidomimetic Self-Assembly: Implications for Amyloid Fibril Formation and Materials Science. *J. Am. Chem. Soc.* **2000,** *122* (22), 5262-5277.
371. Haass, C.; Selkoe, D. J., Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide. *Nature Reviews Molecular Cell Biology* **2007,** *8* (2), 101-112.
372. Balbirnie, M.; Grothe, R.; Eisenberg, D. S., An amyloid-forming peptide from the yeast prion Sup35 reveals a dehydrated β-sheet structure for amyloid. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** *98* (5), 2375-2380.
373. Klimov, D. K.; Thirumalai, D., Dissecting the assembly of Aβ(16-22) amyloid peptides into antiparallel β sheets. *Structure* **2003,** *11* (3), 295-307.
374. Lipfert, J.; Franklin, J.; Wu, F.; Doniach, S., Protein misfolding and amyloid formation for the peptide GNNQQNY from yeast prion protein Sup35: simulation by reaction path annealing. *J. Mol. Biol.* **2005,** *349* (3), 648-658.
375. Zheng, J.; Ma, B. Y.; Tsai, C. J.; Nussinov, R., Structural stability and dynamics of an amyloid-forming peptide GNNQQNY from the yeast prion sup-35. *Biophys. J.* **2006,** *91* (3), 824-833.
376. Esposito, L.; Pedone, C.; Vitagliano, L., Molecular dynamics analyses of cross-β-spine steric zipper models: β-sheet twisting and aggregation. *Proc. Natl. Acad. Sci. U.S.A.* **2006,** *103* (31), 11533-11538.
377. Perutz, M., Polar Zippers - Their Role in Human-Disease. *Protein Sci.* **1994,** *3* (10), 1629-1637.
378. Nelson, R.; Sawaya, M. R.; Balbirnie, M.; Madsen, A. O.; Riekel, C.; Grothe, R.; Eisenberg, D., Structure of the cross-beta spine of amyloid-like fibrils. *Nature* **2005,** *435* (7043), 773-778.

379. Kelly, J. W., The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr Opin Struc Biol* **1998,** *8* (1), 101-106.
380. Thirumalai, D.; Klimov, D. K.; Dima, R. I., Emerging ideas on the molecular basis of protein and peptide aggregation. *Curr Opin Struc Biol* **2003,** *13* (2), 146-159.
381. Tycko, R., Insights into the amyloid folding problem from solid-state NMR. *Biochemistry* **2003,** *42* (11), 3151-3159.
382. Chiti, F.; Dobson, C. M., Protein misfolding, functional amyloid, and human disease. *Annual Review of Biochemistry* **2006,** *75*, 333-366.
383. Teplow, D. B.; Lazo, N. D.; Bitan, G.; Bernstein, S.; Wyttenbach, T.; Bowers, M. T.; Baumketner, A.; Shea, J. E.; Urbanc, B.; Cruz, L.; Borreguero, J.; Stanley, H. E., Elucidating amyloid β-protein folding and assembly: a multidisciplinary approach. *Acc. Chem. Res.* **2006,** *39* (9), 635-645.
384. Kojetin, D. J.; Venters, R. A.; Kordys, D. R.; Thompson, R. J.; Kumar, R.; Cavanagh, J., Structure, binding interface and hydrophobic transitions of Ca2+-loaded calbindin-D-28K. *Nat Struct Mol Biol* **2006,** *13* (7), 641-647.
385. Chen, M.; Margittai, M.; Chen, J.; Langen, R., Investigation of α-synuclein fibril structure by site-directed spin labeling. *J. Biol. Chem.* **2007,** *282* (34), 24970-24979.
386. Mukhopadhyay, S.; Krishnan, R.; Lemke, E. A.; Lindquist, S.; Deniz, A. A., A natively unfolded yeast prion monomer adopts an ensemble of collapsed and rapidly fluctuating structures. *Proc. Natl. Acad. Sci. U.S.A.* **2007,** *104* (8), 2649-2654.
387. Bieschke, J.; Siegel, S. J.; Fu, Y. W.; Kelly, J. W., Alzheimer's Aβ peptides containing an isostructural backbone mutation afford distinct aggregate morphologies but analogous cytotoxicity. Evidence for a common low-abundance toxic Structure(s)? *Biochemistry* **2008,** *47* (1), 50-59.
388. Kim, Y. S.; Liu, L.; Axelsen, P. H.; Hochstrasser, R. M., Two-dimensional infrared spectra of isotopically diluted amyloid fibrils from Aβ40. *Proc. Natl. Acad. Sci. U.S.A.* **2008,** *105* (22), 7720-7725.
389. Shashilov, V. A.; Lednev, I. K., 2D correlation deep UV resonance Raman spectroscopy of early events of lysozyme fibrillation: kinetic mechanism and potential interpretation pitfalls. *J. Am. Chem. Soc.* **2008,** *130* (1), 309-317.
390. Strasfeld, D. B.; Ling, Y. L.; Shim, S. H.; Zanni, M. T., Tracking fiber formation in human islet amyloid polypeptide with automated 2D-IR spectroscopy. *J. Am. Chem. Soc.* **2008,** *130* (21), 6698-9.
391. Lu, J. X.; Qiang, W.; Yau, W. M.; Schwieters, C. D.; Meredith, S. C.; Tycko, R., Molecular structure of β-amyloid fibrils in Alzheimer's disease brain tissue. *Cell* **2013,** *154* (6), 1257-68.
392. Shen, C. L.; Murphy, R. M., Solvent effects on self-assembly of β-amyloid peptide. *Biophys. J.* **1995,** *69* (2), 640-651.
393. Wood, S. J.; Maleeff, B.; Hart, T.; Wetzel, R., Physical, morphological and functional differences between pH 5.8 and 7.4 aggregates of the Alzheimer's amyloid peptide AP. *J. Mol. Biol.* **1996,** *256* (5), 870-877.
394. Harper, J. D.; Lansbury, P. T., Models of amyloid seeding in Alzheimier's disease and scrapie: mechanistic truths and physiological consequences of the time-dependent solubility of amyloid proteins. *Annual Review of Biochemistry* **1997,** *66*, 385-407.
395. Nielsen, L.; Khurana, R.; Coats, A.; Frokjaer, S.; Brange, J.; Vyas, S.; Uversky, V. N.; Fink, A. L., Effect of environmental factors on the kinetics of insulin fibril formation: elucidation of the molecular mechanism. *Biochemistry* **2001,** *40* (20), 6036-6046.
396. Gorman, P. M.; Yip, C. M.; Fraser, P. E.; Chakrabartty, A., Alternate aggregation pathways of the Alzheimer β-amyloid peptide: Aβ association kinetics at endosomal pH. *J. Mol. Biol.* **2003,** *325* (4), 743-757.
397. Pedersen, J. S.; Dikov, D.; Flink, J. L.; Hjuler, H. A.; Christiansen, G.; Otzen, D. E., The changing face of glucagon fibrillation: structural polymorphism and conformational imprinting. *J. Mol. Biol.* **2006,** *355* (3), 501-523.
398. Marek, P.; Abedini, A.; Song, B. B.; Kanungo, M.; Johnson, M. E.; Gupta, R.; Zaman, W.; Wong, S. S.; Raleigh, D. P., Aromatic interactions are not required for amyloid fibril formation by islet amyloid

polypeptide but do influence the rate of fibril formation and fibril morphology. *Biochemistry* **2007,** *46* (11), 3255-3261.
399. Wolf, M. G.; Jongejan, J. A.; Laman, J. D.; de Leeuw, S. W., Quantitative prediction of amyloid fibril growth of short peptides from simulations: calculating association constants to dissect side chain importance. *J. Am. Chem. Soc.* **2008,** *130* (47), 15772-3.
400. Balbach, J. J.; Ishii, Y.; Antzutkin, O. N.; Leapman, R. D.; Rizzo, N. W.; Dyda, F.; Reed, J.; Tycko, R., Amyloid fibril formation by Aβ(16-22), a seven-residue fragment of the Alzheimer's β-amyloid peptide, and structural characterization by solid state NMR. *Biochemistry* **2000,** *39* (45), 13748-13759.
401. Bevivino, A. E.; Loll, P. J., An expanded glutamine repeat destabilizes native ataxin-3 structure and mediates parallel β-fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2001,** *98* (21), 11955-11960.
402. Petkova, A. T.; Yau, W. M.; Tycko, R., Experimental constraints on quaternary structure in Alzheimer's beta-amyloid fibrils. *Biochemistry* **2006,** *45* (2), 498-512.
403. Paravastu, A. K.; Petkova, A. T.; Tycko, R., Polymorphic fibril formation by residues 10-40 of the Alzheimer's beta-amyloid peptide. *Biophys. J.* **2006,** *90* (12), 4618-4629.
404. Wu, Y. N.; Wang, K.; Buschle-Diller, G.; Liles, M. R., Fiber formation by dehydration-induced aggregation of albumin. *J Appl Polym Sci* **2013,** *129* (6), 3591-3600.
405. Zandomeneghi, G.; Krebs, M. R. H.; Mccammon, M. G.; Fandrich, M., FTIR reveals structural differences between native β-sheet proteins and amyloid fibrils. *Protein Sci.* **2004,** *13* (12), 3314-3321.
406. Markiewicz, B. N.; Oyola, R.; Du, D.; Gai, F., Aggregation gatekeeper and controlled assembly of Trpzip β-hairpins. *Biochemistry* **2014,** *53* (7), 1146-1154.
407. Reddy, G.; Straub, J. E.; Thirumalai, D., Dry amyloid fibril assembly in a yeast prion peptide is mediated by long-lived structures containing water wires. *Proc. Natl. Acad. Sci. U.S.A.* **2010,** *107* (50), 21459-21464.
408. Thirumalai, D.; Reddy, G.; Straub, J. E., Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2011,** *45* (1), 83-92.
409. Measey, T. J.; Gai, F., Light-Triggered Disassembly of Amyloid Fibrils. *Langmuir* **2012,** *28* (34), 12588-12592.
410. Markiewicz, B. N.; Oyola, R.; Du, D.; Gai, F., Aggregation Gatekeeper and Controlled Assembly of Trpzip beta-Hairpins. *Biochemistry* **2014,** *53* (7), 1146-1154.
411. Ma, B.; Nussinov, R., Stabilities and conformations of Alzheimer's β-amyloid peptide oligomers (Aβ16–22, Aβ16–35, and Aβ10–35): Sequence effects. *Proc. Natl. Acad. Sci. U.S.A.* **2002,** *99* (22), 14126-14131.
412. Fitzpatrick, A. W. P.; Debelouchina, G. T.; Bayro, M. J.; Clare, D. K.; Caporini, M. A.; Bajaj, V. S.; Jaroniec, C. P.; Wang, L. C.; Ladizhansky, V.; Muller, S. A.; MacPhee, C. E.; Waudby, C. A.; Mott, H. R.; De Simone, A.; Knowles, T. P. J.; Saibil, H. R.; Vendruscolo, M.; Orlova, E. V.; Griffin, R. G.; Dobson, C. M., Atomic structure and hierarchical assembly of a cross-beta amyloid fibril. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110* (14), 5468-5473.
413. Kim, Y. S.; Liu, L.; Axelsen, P. H.; Hochstrasser, R. M., 2D IR provides evidence for mobile water molecules in β-amyloid fibrils. *Proc. Natl. Acad. Sci. U.S.A.* **2009,** *106* (42), 17751-17756.
414. Lee, Y. H.; Chatani, E.; Sasahara, K.; Naiki, H.; Goto, Y., A comprehensive model for packing and hydration for amyloid fibrils of β2-microglobulin. *J. Biol. Chem.* **2009,** *284* (4), 2169-75.
415. Van Melckebeke, H.; Schanda, P.; Gath, J.; Wasmer, C.; Verel, R.; Lange, A.; Meier, B. H.; Bockmann, A., Probing water accessibility in HET-s(218-289) amyloid fibrils by solid-state NMR. *J. Mol. Biol.* **2011,** *405* (3), 765-72.
416. 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.
417. Fitzpatrick, A. W. P.; Debelouchina, G. T.; Bayro, M. J.; Clare, D. K.; Caporini, M. A.; Bajaj, V. S.; Jaroniec, C. P.; Wang, L.; Ladizhansky, V.; Mueller, S. A.; MacPhee, C. E.; Waudby, C. A.; Mott, H. R.; De Simone, A.; Knowles, T. P. J.; Saibil, H. R.; Vendruscolo, M.; Orlova, E. V.; Griffin, R. G.; Dobson, C. M., Atomic structure and hierarchical assembly of a cross-beta amyloid fibril. *Proc. Natl. Acad. Sci. U.S.A.* **2013,** *110* (14), 5468-5473.

418. Hunt, J. F.; Rath, P.; Rothschild, K. J.; Engelman, D. M., Spontaneous, pH-dependent membrane insertion of a transbilayer alpha-helix. *Biochemistry* **1997,** *36* (49), 15177-15192.
419. Wyttenbach, T.; Liu, D.; Bowers, M. T., Interactions of the Hormone Oxytocin with Divalent Metal Ions. *J. Am. Chem. Soc.* **2008,** *130* (18), 5993-6000.
420. Syme, C. D.; Nadal, R. C.; Rigby, S. E. J.; Viles, J. H., Copper binding to the amyloid-beta (A beta) peptide associated with Alzheimer's disease - Folding, coordination geometry, pH dependence, stoichiometry, and affinity of A beta-(1-28): Insights from a range of complementary spectroscopic techniques. *J. Biol. Chem.* **2004,** *279* (18), 18169-18177.
421. Ma, Q.-F.; Hu, J.; Wu, W.-H.; Liu, H.-D.; Du, J.-T.; Fu, Y.; Wu, Y.-W.; Lei, P.; Zhao, Y.-F.; Li, Y.-M., Characterization of copper binding to the peptide amyloid-beta(1-16) associated with Alzheimer's disease. *Biopolymers* **2006,** *83* (1), 20-31.
422. Massari, A. M.; Finkelstein, I. J.; Fayer, M. D., Dynamics of Proteins Encapsulated in Silica Sol−Gel Glasses Studied with IR Vibrational Echo Spectroscopy. *J. Am. Chem. Soc.* **2006,** *128* (12), 3990-3997.
423. Bakulin, A. A.; Cringus, D.; Pieniazek, P. A.; Skinner, J. L.; Jansen, T. L. C.; Pshenichnikov, M. S., Dynamics of Water Confined in Reversed Micelles: Multidimensional Vibrational Spectroscopy Study. *J. Phys. Chem. B* **2013,** *117* (49), 15545-15558.
424. Moilanen, D. E.; Levinger, N. E.; Spry, D. B.; Fayer, M. D., Confinement or the Nature of the Interface? Dynamics of Nanoscopic Water. *J. Am. Chem. Soc.* **2007,** *129* (46), 14311-14318.
425. Fayer, M. D.; Levinger, N. E., Analysis of Water in Confined Geometries and at Interfaces. *Annual Review of Analytical Chemistry* **2010,** *3* (1), 89-107.
426. Kashid, S. M.; Bagchi, S., Experimental Determination of the Electrostatic Nature of Carbonyl Hydrogen-Bonding Interactions Using IR-NMR Correlations. *J. Phys. Chem. Lett.* **2014,** *5* (18), 3211-3215.