REVERSE MICELLE ENCAPSULATION AND ITS USE IN EXAMINING THE INTERPLAY BETWEEN HYDRATION AND PROTEIN DYNAMICS

Bryan Stephen Marques

A DISSERTATION

in

Biochemistry and Molecular Biophysics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

____

A. Joshua Wand, Ph.D., Benjamin Rush Professor of Biochemistry and Molecular Biophysics

Graduate Group Chairperson

____

Kim A. Sharp, Ph.D., Associate Professor of Biochemistry and Biophysics

Dissertation Committee

Kim A. Sharp, Ph.D., Associate Professor of Biochemistry and Biophysics (Committee Chair)

Ronen Marmorstein, Ph.D., George W. Raiziss Professor in Biochemistry and Molecular Biophysics

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

Jeffery G. Saven, Ph.D., Professor of Chemistry

Ponzy Lu, Ph.D., Professor of Chemistry

REVERSE MICELLE ENCAPSULATION AND ITS USE IN EXAMINING THE INTERPLAY BETWEEN HYDRATION AND PROTEIN DYNAMICS

Bryan Stephen Marques

To my parents for getting me here…

…And to my wife and baby boy for keeping me going

Acknowledgements

First I must acknowledge my thesis mentor Josh Wand for all of his support for both my undergraduate and graduate careers. More than anyone else, for the past ten years Josh has constantly tried to bring out my full scientific potential by always pushing me to squeeze as much information out of data and results as possible. I'd also like to thank him for trusting me to take more responsibility in all aspects of the lab; these experiences taught me more than possibly even the scientific techniques did. In that same vein, I'd like to express my deepest gratitude to Sabrina Bedard and Nathaniel Nucci: Sabrina for teaching me essentially everything I know about Molecular Biology and sterile lab technique and Nathaniel for helping me through the extremely steep learning curve of NMR, hydration, and protein dynamics while simultaneously becoming one of my best friends and mentors (even though he sadly supports a terrible baseball team). I'd also like to thank Kathy Valentine and Matt Stetz for helping with all things NMR from sample shimming to pulse sequence development. Of all other past and current members of the Wand lab, I'd also like to specifically thank Christine Jorge for essentially being my other half in our navigation through the world of NMR and hydration dynamics. I'll also be forever grateful for all other members of the Wand and Tommos labs I've interacted with, both past and present (Cecilia, Li, Kyle, Vonni, Melissa, Sarah, Ellen, Jack, Joe, John, Vignesh, Starla, Nikki, Sravya, Shannen, Adam, Igor, and Yao).

Next I'd like to thank all my friends that have kept me sane throughout this whole process. Some of the best friends I've ever made have come from the Wand/Tommos labs (Nucci, Evan, Brian, and Bruce) and I can't imagine where I'd be without all of the fun times and great scientific conversations we've had. My other amazing friends that I met in college were also invaluable from my long-time roommates Matt and Ross to my fellow father and comic buddy Pete to all the other amazing friends I made throughout my college career (Eric, Ben, Bob, and Paul).

I’d also like to thank all the amazing faculty and staff during my eleven years at Penn. The women in the BMB office, both past and present, are some of the nicest and most helpful people in existence and their assistance during my time at Penn has been invaluable. The members of my thesis committee have also been amazing resources to bounce ideas off of and to get excellent advice from despite my admittedly low number of committee meetings (Kim Sharp, Ronen Marmorstein, Feng Gai, Jeffery Saven, and Ponzy Lu – also my undergraduate advisor).

Last, but not even close to least, I’d like to thank everyone in my family. My parents, Jacinto and Lucy, have been pillars of knowledge and responsibility that I’ve leaned on throughout my entire life. As Portuguese immigrants and without going to college they provided the best life I could imagine for my sister and me, and I couldn’t have more respect for how hard working and caring they are. My sister, Sam, has similarly always been there for me and I couldn’t imagine growing up with a better sibling (and you’re welcome for me naming you). Above all I need to thank my wife, Jamie, and son, Teddy. Through the 11+ years we’ve been dating, the nearly 4 years we’ve been married, and the year and a half we’ve had a son I’ve learned more from you both than I have from any other person I’ve thanked here. I can’t imagine my life without you both; I love you and look forward to everything to come in our lives.

ABSTRACT

REVERSE MICELLE ENCAPSULATION AND ITS USE IN EXAMINING THE INTERPLAY BETWEEN HYDRATION AND PROTEIN DYNAMICS

Bryan Stephen Marques

A. Joshua Wand, Ph.D.

As the universal solvent, water is unquestionably essential to most aspects of protein biophysics from protein folding to enzymatic activity. Much has been learned about the relationship between proteins and surrounding solvent waters; however, it is often difficult to experimentally examine these interactions in a site-specific manner without perturbing molecular structure. Furthermore, the effect of nearby hydration dynamics on protein dynamics (and, in effect, protein conformational entropy) is poorly understood at atomic resolution. With the use of a combination of Nuclear Magnetic Resonance (NMR) spectroscopy and protein reverse micelle (RM) encapsulation, it is possible to examine both the dynamic behavior of waters in the protein hydration layer as well as protein dynamics for the same sample without physically altering the protein. The goal of this work is to use these complementary techniques in order to better understand the interplay between hydration and protein dynamics. First, we demonstrate the utility of NMR spectroscopy in monitoring and controlling the pH of the aqueous interior of reverse micelle ensembles. This leads to the ability to reliably confirm sample pH and structural fidelity upon RM encapsulation which is often difficult to accomplish using other techniques. Next we propose a novel approach to collecting and analyzing NMR hydration dynamics experiments with the use of non-uniform sampling (NUS) and nuclear Overhauser effect (NOE) mixing time buildup experiments. We examine factors contributing to the reproducibility and reliability of hydration ratios. Using these NOE-based hydration experiments, we then examine the hydration dynamics of hen egg-white lysozyme (HEWL) with and without a bound inhibitor. We find minimal retardation of hydration dynamics within a partially hydrophilic binding cleft; we detect waters within an internal pocket which are relatively fast; and we inspect trapped interfacial waters upon ligand binding. Finally, we use RM encapsulation to examine the effect of changes in solvent dynamics on fast (ps-ns) protein dynamics. While retardation of hydration dynamics seems to affect dynamics of aromatic side chains, it has little to no effect on other fast protein dynamics effectively confirming that protein conformational entropy is not slaved to solvent. This work represents a large leap forward in our understanding of the relationship between proteins and their hydrating environment.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: Introduction

Protein hydration and dynamics

Through decades of study, it has become abundantly clear that the interactions between water and protein molecules are absolutely vital for maintenance of protein structure [1-4] and function [5-7]. In a global sense, much is understood about the behavior of water solvent both in the hydration/biological layer (one to two layers of water surrounding the protein surface) and in bulk (Figure 1-1).

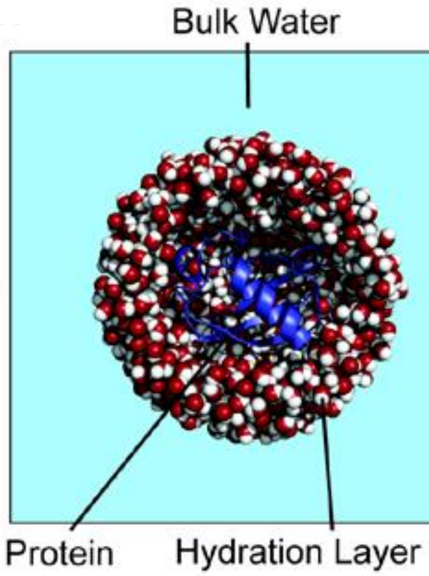

Figure 1-1. Schematic of a protein (blue cartoon structure) in aqueous solution. Two different types of waters in the context of protein hydration are depicted: the hydration layer (red and white spheres) and bulk solvent (cyan background).

With the use of techniques such as magnetic dispersion [8] and neutron/X-ray scattering [9], it has been demonstrated that the waters within the hydration layer of a protein molecule are significantly slowed relative to bulk solvent [10]; however, the amount of retardation often varies depending on the technique used [11]. While this global understanding of hydration layer water retardation is useful, a site-specific understanding would be more informative.

It is often difficult to obtain an atomistic understanding of these interactions and the relationship between protein hydration and dynamics without using techniques that physically alter the protein (such as Overhauser dynamic nuclear polarization [12] or fluorescence [13]). Historically, much of the experimental information about protein-water interactions is derived from resolved water molecules in X-ray crystal structures. While crystal structures provide an excellent scaffold to study some aspects of protein-water interactions (location of buried or interfacial waters), these structures provide a static picture of protein structures and thus cannot provide much information about the likely dynamic behavior between protein and water. Hence for an atomistic understanding of protein-water interactions we often must look to molecular dynamics simulations. While studies of this type are informative, they are generally difficult to experimentally verify and often contain contradictory interpretations [14-16] of hydration dynamics, especially involving the effect of factors such as side chain type or local protein structure. It is therefore desirable to experimentally study hydration dynamics in a site-specific manner without physically altering the protein.

Similarly, the effect of water on protein dynamics has not been experimentally quantified in a site-specific manner. The prevalent theory concerning the effect of water on protein dynamics, originally developed by Frauenfelder, is the “solvent slaving model [17-19].” This model broadly bins protein dynamics into three classes of solvent-slaved motions (Table 1-1).

Table 1-1: Solvent Slaving Motions

Class of Motion	**Solvent Slaving Dependence**	**Type of Protein Motion**
Class I (or α)	Bulk Solvent	Protein Folding Conformational Sampling
Class II (or β)	Hydration Layer	Side chain Motions (Librations and Rotations)
Class III	Independent of Solvent	Vibrational Processes

Class I (or α) motions are relatively slow (ms – μs) motions exemplifying large structural changes and conformational flexibility: these motions are slaved to the motions of bulk solvent. Class II (or β) motions are relatively fast (ns – ps) motions such as side chain librations and rotations: these motions are slaved to the motions of waters within the hydration layer. The classification of these two types of motion is based on evidence from low-resolution dielectric and scattering studies with little to no site specificity. Furthermore, a recent neutron scattering study argues that reported solvent slaving as demonstrated by previous interpretations of the technique is only due to a lack of resolution [20]. Interestingly, a third class of motions (such as vibrational processes) has recently been introduced in the model: these motions are inherently independent of solvent; however, this has been shown not to necessarily be the case with bond vibrations [21, 22]. It is clear that, much like for protein hydration dynamics, it is imperative that the effect of solvent on protein dynamics be studied in a site-specific manner without altering the protein structure. The most optimal method for studying both processes is NMR spectroscopy.

Using NMR for hydration and protein dynamics

NMR for hydration

The potential for the use of NMR spectroscopy for the detection of hydration dynamics near the surface of a protein was first developed by Wüthrich and colleagues in the early 1990s [23, 24]. The premise of using NMR to measure protein-water interactions is relatively straightforward: detect the dipolar magnetization transfer between hydrogens on the surface of the protein via the nuclear Overhauser effect (NOE). While the NOE is traditionally used to determine intramolecular distance restraints in protein structure determination, for intermolecular NOEs between protein and water hydrogens the NOE must be detected in both the laboratory and rotating (ROE) frames. The magnetization transfer rate of the two processes (σ_{NOE} and σ_{ROE}, respectively) are dependent on the spectral density function, J(ω), as follows:

$$\sigma_{NOE} = k[6J(2\omega) - J(0)] \quad (1)$$

$$\sigma_{ROE} = k[3J(\omega) + 2J(0)] \quad (2)$$

$$J(\omega) = \left[\frac{1}{r^6}\right]\left[\frac{\tau_m}{1+(\omega\tau_m)^2}\right] \quad (3)$$

Where $k = (\mu_0/(4\pi)^* \hbar^*\gamma_H^2)^2$ is a combination of fundamental constants and the proton gyromagnetic ratio (γ_H), τ_m is the rotational correlation time of the protein, r is the distance between the protein and water protons, and ω is the proton Larmor frequency. At the slow tumbling regime ($\tau_m\ \omega >> 1$), hydration dynamics are quantified by the ratio of the laboratory frame and laboratory frame magnetization transfer rates. At short NOE mixing times (τ_{mix}), the relaxation rate-corrected NOE and ROE cross-peak intensities can be used as follows:

$$\frac{I_{NOE}(\tau_{mix})}{I_{ROE}(\tau_{mix})} = \left[\frac{\sigma_{NOE} e^{-R_1 \tau_{mix}}}{\sigma_{ROE} e^{-R_{1\rho} \tau_{mix}}} \right] \tag{4}$$

Where I_{NOE} and I_{ROE} are the cross-peak intensities of the NOE and ROE, respectively and R_1 and $R_{1\rho}$ are the longitudinal relaxation rates in the laboratory and rotating frames, respectively. Because detailed interpretation of the NOE/ROE ratio can be quite complicated, the limits of the ratio are often considered. At slow hydration dynamics (long-lived protein-water interaction), the limit of the ratio is -0.5. Such an interaction is rigid within the molecular frame of the protein and thus has a "retention time" of τ_m or longer. At fast hydration dynamics (short-lived protein-water interactions) the limit tends toward 1 (although practically becomes 0 because I_{NOE} tends towards 0). Such an interaction is on the order hundreds of picoseconds. Differential hydration dynamics across the surface of a protein is thus defined as having a NOE/ROE ratio between -0.5 and 0 [25].

Unfortunately, in aqueous solution the NOE/ROE ratio between protein and water protons can become irrecoverably contaminated. Potential contamination results because of three aspects of water hydration layer waters in aqueous solution. Although hydration layer waters are slowed relative to bulk solvent by up to 2 orders of magnitude [11], they are often still too fast to detect a quantifiable protein-water NOE [26]. The NOE/ROE ratio can also become contaminated due to magnetization exchange from hydrogen-exchanged protons from nearby exchangeable protons (from nearby waters, backbone amide protons, or exchangeable side chain protons) [8, 26, 27]. Finally, the ratio can be further contaminated by contributions from waters in the bulk solvent [27]; however, this avenue of contamination has been recently disputed [28]. In effect, in order to reliably examine protein hydration dynamics with NMR, one must first eliminate contaminants: this has been accomplished by the encapsulation of proteins in reverse micelles [29] (see below).

NMR for protein dynamics

The use of NMR to study protein dynamics has been well established for a wide array of timescales (Figure 1-2) including both fast, ps-ns (Class II)[30-32], and relatively slow, ms-μs (Class I)[33, 34] motions.

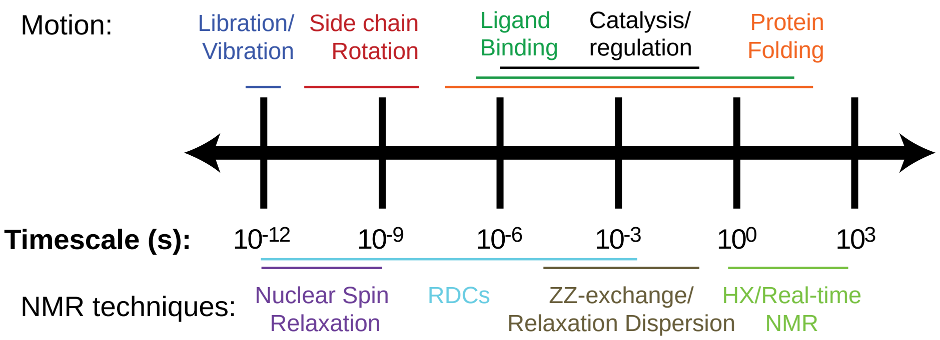

Figure 1-2. Schematic of NMR-observable timescale motions. Nuclear spin relaxation (purple) is the technique used to acquire backbone and side chain order parameters which report on fast protein dynamics (ps –ns)

Dependencies of Class I motions such as protein folding and large-scale conformational fluctuations on solvent viscosity and crowding has been demonstrated in the past [35]. In effect, we will focus mainly on the effect of changes in solvent conditions on Class II motions of protein side chains and their contributions to conformational entropy [36].

Through the use of typical spin relaxation experiments it is possible to obtain information about ps-ns timescale motions in both protein backbones [37, 38] and side chains [31, 32, 39] through the squared generalized order parameter, O^2. The order parameter is defined as the limiting value of a bond vector's position with time in the molecular frame. In other words, it is a numerical description of the degree of flexibility of a given bond vector in the molecular frame ranging from a value of 0 (completely isotropic motion) to 1 (completely rigid bond vector relative to the molecular frame). Physically the order parameter can be related to the motion of the bond vector in a cone at an angle θ: the larger the angle, the smaller the order parameter and vice versa (Figure 1-3).

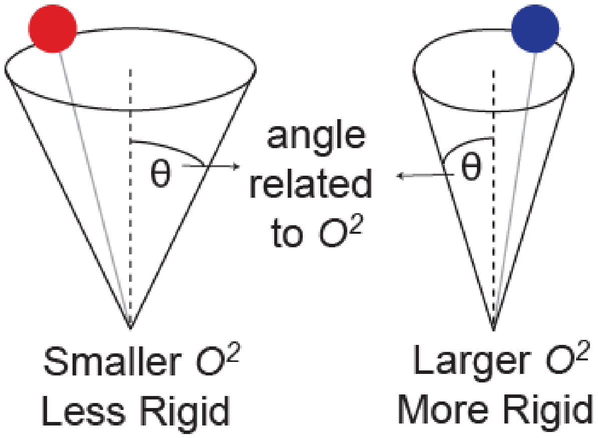

Figure 1-3. Relation of order parameter (O2) to motion of a bond vector (gray lines) in a cone at an angle θ. As θ increases, the bond vector moves more rapidly in the molecular frame (smaller O^2, left). As θ decreases, the bond vector becomes more rigid in the molecular frame (larger O^2, left).

The order parameters for backbone amide N-H vectors (O^2_{NH}) are typically rigid for globular proteins, especially in structured regions [40]. On the other hand, methyl side chain order parameters (typically defined as the order parameter with respect to the methyl C-C bond symmetry axis, O^2_{axis}) are much more heterogeneous [41] and are completely independent of backbone motions. Methyl order parameters are typically collected via well-established ^{2}H [31] or ^{13}C [32] spin relaxation methods. However, these types of experiments are either unreliable for large protein systems (^{2}H) or difficult to collect for poorly-behaved, low concentration samples with no published structure (^{13}C). In effect, we have implemented a relatively new method for the determination of methyl order parameters first developed by Kay and colleagues [42, 43] wherein O^{2}_{axis} is related to the ^{1}H-^{1}H intra-methyl cross-correlated relaxation rate. The ratio of cross-peak heights from measures of bi-exponential methyl proton single quantum decay (I_{sq}) and measures of triple quantum buildups (I_{mq}) can be fit directly to the cross-relaxation rate:

$$\left|\frac{I_{mq}}{I_{sq}}\right| = \frac{0.75 \bullet \eta \tanh(\sqrt{\eta^2+\delta^2}T)}{\sqrt{\eta^2+\delta^2} - \delta \tanh(\sqrt{\eta^2+\delta^2}T)} \tag{5}$$

Where T is the variable time delay used during experimental collection, δ is a parameter that accounts for remote dipolar contributions solely depending on the rotational correlation time of the protein (τ_m) and other constants:

$$\delta = -4\sum_{ext} \frac{1}{20} \frac{\hbar^2 \gamma_H^4 \tau_m}{r_{HHext}^6} \tag{6}$$

And η is the intra-methyl proton-proton cross-relaxation rate for a methyl group undergoing rapid rotation about its symmetry axis. For proteins tumbling slower than ~ 5 ns, the η rate can be approximated as:

$$\eta \approx \frac{9}{10}\left(\frac{\mu_0}{4\pi}\right)^2 [P_2(\cos\theta_{axis,HH})]^2 \frac{O_{axis}^2 \gamma_H^4 \hbar^2 \tau_m}{r_{HH}^6} \tag{7}$$

Of note with this method of calculating methyl order parameters is the relatively simple calculation of O^2_{axis}. Because contributions from remote protons can be fit directly from experimental data via the δ parameter, it is no longer necessary to obtain a high-resolution structure. Additionally, the relationship between the cross-relaxation rate, η, and the methyl order parameter is only dependent on constant values and the angular reorientation time of the protein which can be estimated or calculated with typical backbone dynamics experiments. Thus, unlike with the aforementioned deuterium and carbon spin relaxation experiments, the cross-correlated experiments provide extremely high sensitivity data that only require one set of observables eliminating the need of data collection at multiple fields and thus saving days or weeks of experiment time. All methyl order parameters discussed in Chapter 5 were analyzed via the cross-correlated method.

Methyl order parameters have been shown to bin into three motional classes [44, 45]: the J class exemplifies fast methyl dynamics with rapid conversion between rotameric wells; the α class exemplifies intermediate methyl dynamics with significant excursions within a rotameric well; and the ω class exemplifies slow, rigid methyl dynamics with highly restrictive motion within a rotameric well. The O^2_{axis} and its binning can vary due to factors such as protein conformational state/structure [46], temperature [47], and applied hydrostatic pressure [48]. Most importantly, the changes in average methyl order parameter upon ligand binding can be used as a proxy for changes in conformational entropy [36, 49-51].

Briefly, the binding affinity of a protein for some ligand molecule is dictated by the change in free energy upon ligand binding (ΔG) expressed as a linear combination of the change in enthalpy (ΔH) and entropy (-TΔS):

$$\Delta G = \Delta H - T\Delta S \tag{8}$$

Wherein the change in entropy upon binding can be divided into contributions from changes in solvent entropy ($\Delta S_{solvent}$), protein and ligand conformational entropy (ΔS_{conf}), rotational-translational entropy (ΔS_{RT}), and relatively uniform contributions from unrecognized sources (ΔS_{other}) as follows:

$$-T\Delta S_{Total} = -T\left(\Delta S_{conf} + \Delta S_{solvent} + \Delta S_{RT} + \Delta S_{other}\right) \tag{9}$$

The notion that there is a relationship between methyl order parameters and protein conformational entropy was first demonstrated as a linear relationship for calmodulin-peptide binding events between the change in total binding entropy measured by isothermal titration calorimetry and the change in conformational entropy calculated via the methyl order parameter [49]. Upon the incorporation of additional protein-ligand systems, it became clear that changes in conformational entropy and average methyl order parameter could be related by a relatively simple and generalized linear equation:

$$\Delta S_{Total} = s_d\left[\left(N_\chi^{protein}\left\langle \Delta O_{axis}^2\right\rangle^{protein}\right) + \left(N_\chi^{ligand}\left\langle \Delta O_{axis}^2\right\rangle^{ligand}\right)\right] + \Delta S_{RT} + \Delta S_{other} \tag{10}$$

Where N_χ is the total number of torsion angles and s_d is a constant scaling factor fit to the slope of the entropy meter (-4.8 ± 0.5 J mol^{-1} K^{-1}) [52]. This linear relationship was demonstrated for 28 protein-ligand complexes [36] with a broad range of binding affinities (Figure 1-4).

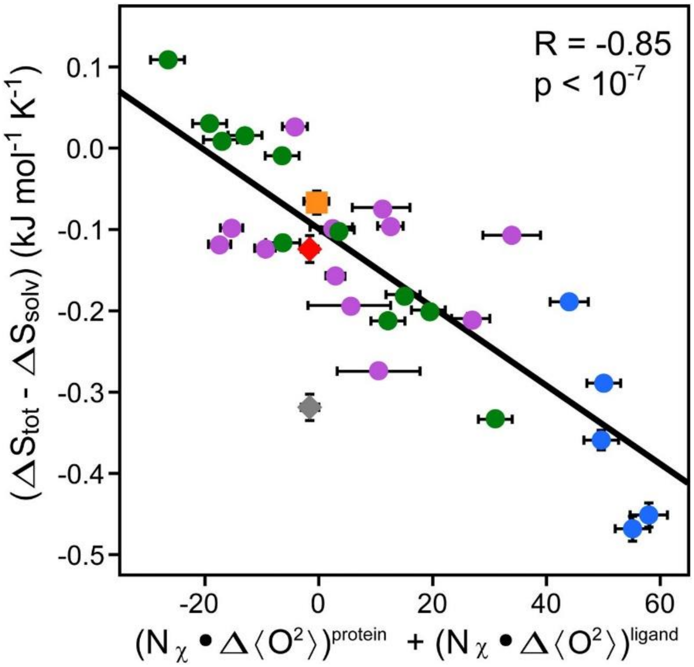

Figure 1-4. The universal entropy meter can be used to relate changes in protein conformational entropy to changes in average methyl order parameters upon ligand binding. Figure adapted from Caro et al [36].

Clearly this entropy meter is a universal phenomenon across a multitude of protein systems. It would effectively be problematic if Class II type methyl order parameter motions (and in effect protein conformational entropy) drastically changed upon introduction into a novel solvent environment. Thus in Chapter 5 we compare the backbone, methyl, and aromatic order parameters of three different proteins encapsulated in reverse micelles to examine the dependence of protein dynamics on changes in solvent environment.

Protein reverse micelle encapsulation

The spontaneous formation of reverse micelles from the proper amounts of aqueous solution (with or without protein), amphiphilic surfactant molecules, and low-viscosity nonpolar solvent (Figure 1-5) has been studied for decades [53-56].

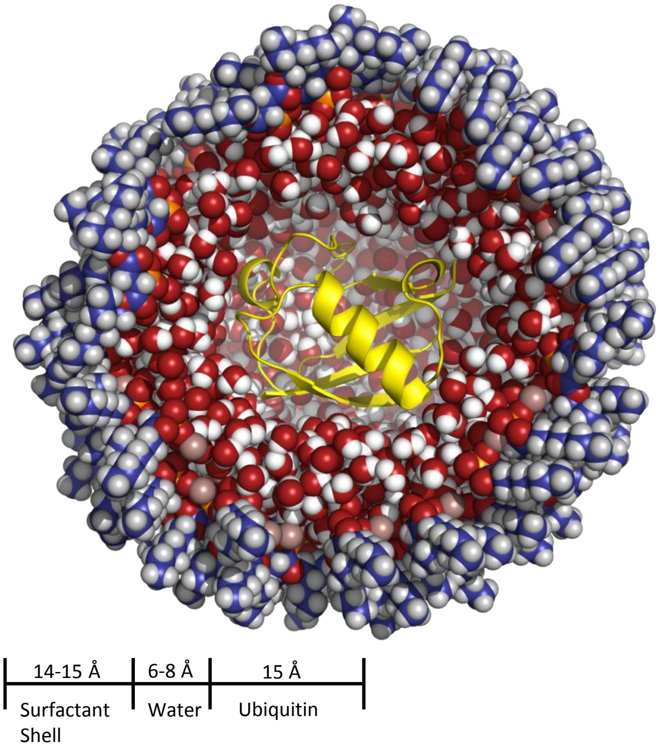

Figure 1-5. Schematic of a protein (yellow cartoon structure) encapsulated in a reverse micelle.

Originally intended to increase NMR spectral quality by allowing for faster protein tumbling [57], reverse micelle encapsulation using many different types of surfactants [58, 59] has been used to study proteins of all sizes [58, 60] including those that are typically difficult to study in aqueous solution such as metastable proteins [61], integral membrane proteins [62], and membrane anchored proteins [63]. Reverse micelle encapsulation has even been used to examine protein biophysics [64] and enzymatic catalysis [65]. In the context of the interplay between hydration and protein dynamics, the most notable characterization of reverse micelle encapsulation is its effect on the dynamics of waters within the aqueous nanopool.

Optimization of reverse micelle conditions varies depending on protein being examined with factors such as surfactant type, buffer/surfactant pH (see Chapter 2), and amount of water within the aqueous nanopool as determined by the water loading: the molar ratio of water to surfactant (W_0). Interestingly, proteins typically maintain their hydrated radii within reverse micelles [57] under all conditions in which the protein maintains its structural fidelity with excess water likely being sequestered into empty (protein-free) reverse micelles. It has also been demonstrated that waters within the aqueous core of reverse micelles are retarded relative to bulk solvent by up to an order of magnitude [66, 67] while more recently a gradient of hydration dynamics between the surface of the protein and polar surfactant head groups has been simulated and supported experimentally by small-angle X-ray and neutron scattering [68]. Finally, it has been demonstrated that the rate of hydrogen exchange of amide protons within reverse micelles is slowed by approximately two orders of magnitude [29]. All of these effects combine to allow for the quantitative determination of hydration dynamics near the protein surface using NMR [29] and provide a system in which the solvent environment is so drastically different than in aqueous solution that it provides an excellent template to site-specifically examine solvent slaving.

What to expect from previous studies and theory

As described above, reverse micelle encapsulation allows for the quantitative analysis of protein hydration dynamics using the NMR-derived NOE/ROE ratios [29]. The NMR-detected hydration surface of ubiquitin demonstrates near-global coverage of detectable hydration dynamics with evident regional clustering. While it has been suggested previously that hydration dynamics correlates with factors such as amino acid type [16], protein curvature [14], and structural rigidity [69], the hydration dynamics as determined by NMR demonstrate none of these preconceived relationships. Interestingly, areas on the protein surface that contain slow hydration dynamics typically are also areas that form large, dry protein-protein interaction surfaces and vice versa (Figure 1-6).

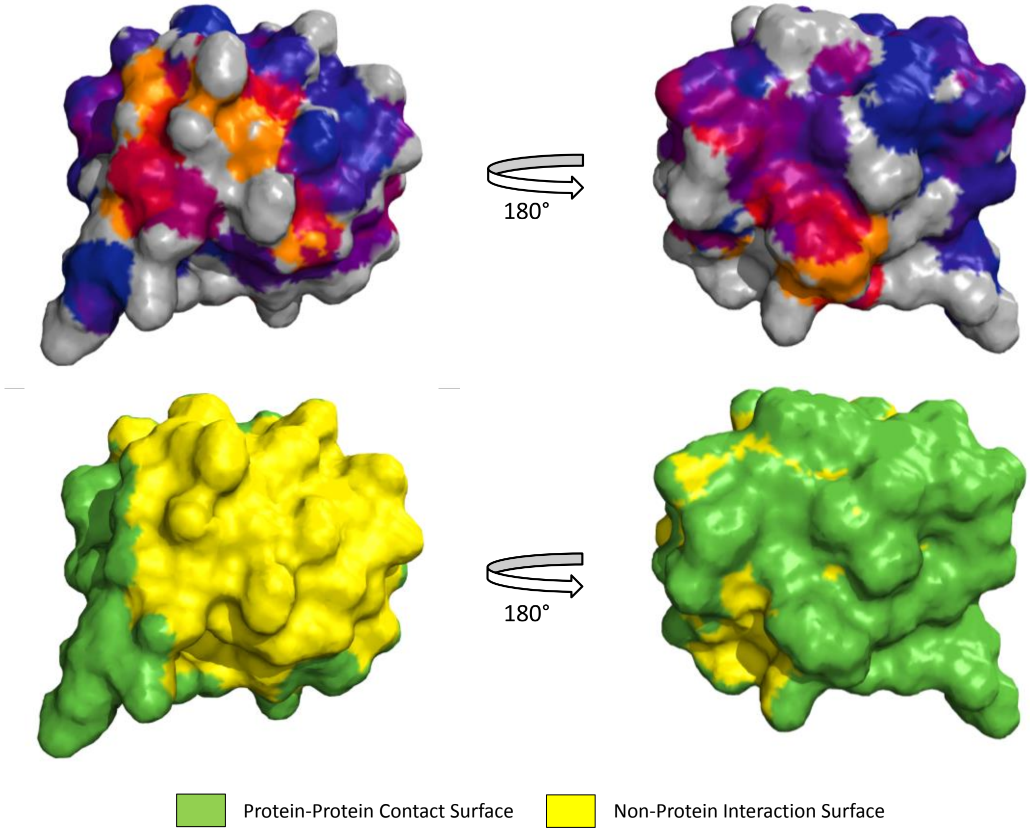

Figure 1-6. Clustering of hydration dynamics near the surface of ubiquitin clearly demonstrates a correlation with protein-protein interaction surfaces. Top: the NOE/ROE ratios are mapped to the surface of ubiquitin with a color range from blue indicating slow hydration dynamics (NOE/ROE = -0.5) to red indicating fast hydration dynamics (NOE/ROE = 0). Orange indicates points on the surface at which hydration dynamics are still too fast to detect. The clustering of slow hydration dynamics correlates well with areas of the protein that form dry protein-protein interaction surfaces (green in the bottom panel).

This correlation with protein-protein interaction surfaces suggests ubiquitin (and likely other proteins) evolved in order to maximize the contributions of solvent entropy in the thermodynamics of ligand binding. In effect, in Chapter 3 we demonstrate the optimization of collecting NOE/ROE ratios and in Chapter 4 we apply this technique to a system with a relatively wet and small protein-ligand binding interface: hen egg-white lysozyme.

In contrast to hydration dynamics, protein dynamics of reverse micelle-encapsulated proteins have not been extensively inspected. The only experimental evidence of NMR-detected protein dynamics of an encapsulated protein was performed by Simorellis and Flynn [70]. This study analyzed the changes in backbone order parameters of ubiquitin upon encapsulation in AOT reverse micelles with the findings that, although there is some localized minor retardation of backbone dynamics, overall the dynamics of the encapsulated protein are nativelike. However, while protein backbone dynamics of encapsulated protein may be enlightening, O^2_{NH} has been shown to contribute very litter to overall changes in protein conformational entropy [52]. Thus we will collect backbone, methyl, and aromatic dynamics of proteins encapsulated in reverse micelles in order to obtain a better understanding of the effect (if any) of the change in hydration dynamics on protein dynamics.

Assuming the solvent slaving model is correct, one should anticipate certain relationships of protein dynamics within reverse micelles (Figure 1-7).

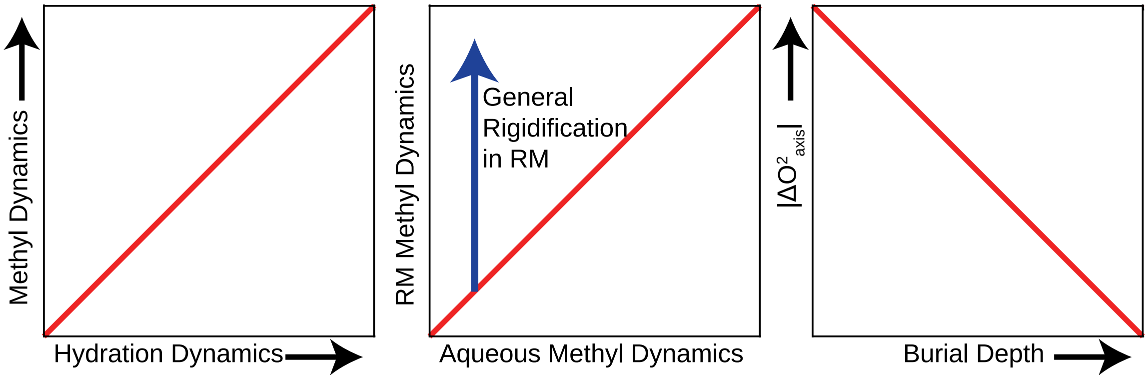

Figure 1-7. Expectations for the relationship between methyl dynamics (i.e. – conformational entropy) and hydration dynamics upon encapsulation within a revere micelle (RM) according to the solvent slaving model [18]. Methyl dynamics and hydration dynamics should demonstrate a strong, direct correlation (left). Upon encapsulation, a general rigidification of methyl dynamics should occur (middle). Finally, the changes in order parameter upon encapsulation should be indirectly correlated with the burial depth of the methyl probe (right).

Because the solvent slaving model dictates that class II-like methyl dynamics motions should be slaved to the dynamics of the hydration layer of waters, there should be a direct correlation between hydration and protein dynamics as experimentally determined in the reverse micelle sample (Figure 1-8, left). Also, since the waters within the aqueous pool of a reverse micelle are significantly slowed by up to an order of magnitude [66, 67] one would expect a general rigidification of methyl dynamics upon protein encapsulation (Figure 1-8, middle). Finally, because probes near the surface of the protein are more exposed to solvent, they should be more sensitive to changes in hydration dynamics. In effect there should be an indirect correlation of changes in methyl order parameters (if any) with the burial depth of the methyl probe (Figure 1-8, right). In Chapter 5 we examine the effect of changes in hydration dynamics on protein dynamics for three proteins: ubiquitin, maltose-binding protein (MBP), and malate-synthase G (MSG). This is accomplished by slowing the water dynamics via addition of high percentages of glycerol or the encapsulation of the proteins in reverse micelles.

Dissertation objectives

This dissertation examines how reverse micelle encapsulation is used in order to study protein hydration dynamics and their effect on protein dynamics and conformational entropy. Chapter 2 describes the optimization of reverse micelle sample preparation by demonstrating the determination and maintenance of the pH of the encapsulated aqueous nanopool. Chapter 3 presents a modified procedure in the determination of protein hydration dynamics by using non-uniformly sampled hydration experiments collected in buildup series at sequential mix times. Additionally, Appendix A compares the hydration dynamics of ubiquitin as determined in Chapter 2 to those determined by Overhauser dynamic nuclear polarization (ODNP). Chapter 4 examines the hydration dynamics of hen egg-white lysozyme (HEWL) in the apo state and with a bound inhibitor in order to determine the hydration dynamic landscape of a wet ligand-binding cleft along with changes of hydration dynamics upon ligand binding. Chapter 5 compares protein dynamics obtained under different solvation conditions (high percentages of glycerol and encapsulated in reverse micelles) in order to demonstrate that protein dynamics (i.e. – protein conformational entropy) are not slaved to solvent. This work provides the first example of NMR's ability to site-specifically examine and compare both hydration and protein dynamics in varying solvent conditions and represents a large step forward on decades' worth of examination of the interplay between hydration and protein dynamics.

CHAPTER 2: Measurement and control of pH in the aqueous interior of reverse micelles

The majority of this chapter was published in Marques et al, (2014) Journal of Physical Chemistry B. ***118(8)****:2020-31* found at https://pubs.acs.org/doi/10.1021/jp4103349
Further permissions related to the material excerpted should be directed to the ACS.

Abstract

The encapsulation of proteins and nucleic acids within the nanoscale water core of reverse micelles has been used for over three decades as a vehicle for a wide range of investigations including enzymology, the physical chemistry of confined spaces, protein and nucleic acid structural biology, and drug development and delivery. Unfortunately, the static and dynamical aspects of the distribution of water in solutions of reverse micelles complicate the measurement and interpretation of fundamental parameters such as pH. This is a severe disadvantage in the context of (bio)chemical reactions and protein structure and function, which are generally highly sensitive to pH. There is a need to more fully characterize and control the effective pH of the reverse micelle water core. The buffering effect of titratable head groups of the reverse micelle surfactants is found to often be the dominant variable defining the pH of the water core. Methods for measuring the pH of the reverse micelle aqueous interior using one-dimensional ^{1}H and two-dimensional heteronuclear NMR spectroscopy are described. Strategies for setting the effective pH of the reverse micelle water core are demonstrated. The exquisite sensitivity of encapsulated proteins to the surfactant, water content, and pH of the reverse micelle are also addressed. These results highlight the importance of assessing the structural fidelity of the encapsulated protein

using multidimensional NMR before embarking upon a detailed structural and biophysical characterization.

Introduction

Reverse micelles are nanoscale assemblies that spontaneously organize from mixtures of appropriate surfactant molecules, small volumes of polar solvent, and bulk nonpolar solvent. The polar, typically aqueous [71] interior of a reverse micelle provides a stable nanoscale confinement volume that has been used for decades in a wide variety of applications in synthetic, physical and biological chemistry. Reverse micelles have been used for studies of nanoconfinement effects on water behavior [8, 13, 72], protein structure [53, 62, 63, 73-77] and biophysics [61, 64, 78, 79], enzymatic catalysis [65], nucleic acid structure and dynamics [80] and even used as a vehicle for drug delivery [81]. In recent years, reverse micelle encapsulation of proteins and nucleic acids has emerged as a particularly powerful tool for the study of macromolecular structure, function and biophysics using high-resolution heteronuclear NMR spectroscopy [59, 82, 83].

The adaptation of reverse micelle encapsulation for high-resolution solution NMR spectroscopy of biological macromolecules was originally undertaken to overcome the deleterious effects of their slow molecular reorientation. Single protein molecules encapsulated within the aqueous core of reverse micelles dissolved in low viscosity fluids can thus be made to tumble faster than they would by themselves in aqueous solution [82]. Faster molecular reorientation leads to more optimal NMR relaxation properties and improved performance. In the context of structural biology and biophysics, the homogeneity and fidelity of encapsulation become critical and largely define the utility of this approach. In the context of high-resolution NMR of encapsulated proteins, small spherical reverse micelles containing a single protein molecule can be prepared with appropriate surfactant mixtures under water-limited conditions in short-chain alkane solvents including propane [82] and ethane [59, 83].

Over the past decade there has been a renewed interest in expanding the library of surfactant systems that can support encapsulation of proteins with a range of properties such as isoelectric point, size, oligomerization state, and the presence of bound ligands or cofactors. The classic anionic surfactant bis(2-ethylhexyl)-sulfosuccinate (AOT) has proven to be poor in this regard [59]. Appropriate mixtures of amphiphilic surfactants such as cetyltrimethylammonium bromide (CTAB), dodecyltrimethylammonium bromide (DTAB), lauryldimethylamine-oxide (LDAO), and decylmonoacyl-glycerol (10MAG) and cosurfactants such as hexanol form small homogeneous reverse micelles in the low viscosity short-chain alkanes and have proven quite successful in the encapsulation of proteins and nucleic acids with high structural fidelity [58, 59, 76, 83, 84].

A critical parameter for any chemical application but particularly in the context of protein biochemistry and biophysics is the pH of the reverse micelle aqueous core. The concept of pH in the reverse micelle water pool presents a somewhat complicated situation [85, 86]. The measurement and meaning of pH in reverse micelles is complicated by the potential for interactions between the buffers of the water pool and the surfactants and by the possibility of an inhomogeneous distribution within the water core [87, 88]. A number of studies have been undertaken to experimentally characterize pH within reverse micelle water cores including the use of oxovanadate probes with ^{51}V NMR [86, 89-92], phosphate and pyrophosphate with ^{31}P NMR [88, 93], measurements of water T_2 relaxation times with proton NMR [94], and hydroxypyrenetrisulfonate and fluorescein measurements using optical spectroscopy [85, 95].

Here we focus on the view of pH provided by solution NMR spectroscopy of reverse micelles where dynamical effects can be particularly important to consider. Unlike many other types of spectroscopy, NMR parameters such as the chemical shift can be averaged by relatively slow processes on the order of milliseconds or faster. This is an important consideration in the context of pH where the number of waters in a typical single reverse micelle core is insufficient to present, on average, even a single hydronium or hydroxide ion. As a result, the instantaneous "pH" in the core of an individual reverse micelle may vary widely. Importantly, however, reverse micelles dissolved in liquid alkane solvents collide and exchange water cores on the microsecond timescale [96-98]. These exchange events lead to averaging of ionization states on the chemical shift timescale such that a single average spectrum is generally obtained. This averaged spectrum offers an assessment of the overall or effective pH of the ensemble of reverse micelles in a particular solution.

In this study we implement in the context of the reverse micelle aqueous core a method that has been previously established for pH monitoring in bulk aqueous solution, namely observation of ^{1}H NMR signals of common, unlabeled buffer molecules [99, 100]. The approach is validated by reference to the corresponding pH-dependence of amide ^{1}H and ^{15}N resonances of an encapsulated protein. It is found that the titratable surfactant molecules can dominate the effective pH of the water core and generally overwhelm the buffering contributions of molecules in the aqueous core. Methods are described to set the effective pH of the reverse micelle water pool when a reverse micelle sample is prepared and to adjust the pH after sample preparation. Importantly, it is shown that the structural fidelity of an encapsulated protein is often exquisitely sensitive to the reverse micelle conditions, including the effective pH. It is highly recommended that the structural integrity of encapsulated proteins be directly characterized using multidimensional NMR spectroscopy rather than interrogated with less comprehensive methods such as UV/visible absorbance or fluorescence emission spectroscopy.

Results and Discussion

Monitoring pH in reverse micelles

The pH of the aqueous core encapsulated within a reverse micelle is a complex property of these systems [85, 86, 101]. Despite its complexity and importance, this property of reverse micelle samples is frequently overlooked under the assumption that the aqueous solution used to prepare a reverse micelle mixture determines the pH of the encapsulated aqueous core. However, in principle, the most abundant component that can contribute to the internal pH of the encapsulated solution in any reverse micelle mixture are titratable surfactant molecules, which are typically one to two orders of magnitude higher in concentration than the buffer or macromolecular components of the reverse micelle mixture [93].

Reverse micelles prepared in liquid alkanes and used in structural biology and biophysics have historically been comprised of various mixtures of the anionic AOT, cationic CTAB and its variants, and neutral surfactants such as hexanol and various polyethers. More recently, a new surfactant mixture based on the zwitterionic LDAO and the uncharged 10MAG has been described [58]. To examine the buffering capacity of the AOT, CTAB, LDAO, and 10MAG head groups, each was dissolved in water (with 12% ethanol for

10MAG) and titrated over a pH range from 4 to 10. As expected, CTAB and 10MAG were found to have no buffering capacity in this pH range while both AOT and LDAO have apparent pKa values in the range of 3-4. Aqueous solutions of AOT had an initial pH of 5 to 5.5, regardless of the manufacturer lot number. Aqueous LDAO solutions showed a broader initial pH range of 6.5 to 8, depending on which manufacturer batch was used. From these tests, it was determined that both AOT and LDAO should have dominant buffering capacity in the reverse micelle mixture. While measurement of pH in bulk aqueous solution is simple, measurement of the pH of the reverse micelle core is much more difficult.

Previous efforts to measure the pH in reverse micelles have often used either colorimetric pH indicators or pH-dependent fluorophores. Optical probes, where the time scale of the reporting phenomenon is quite short, can be tricky to interpret because of superpositions of spectra due to subtle pH gradients within the reverse micelle interior and the propensity for these amphiphilic molecules to partition into varying regions of the reverse micelle mixture [86]. From the point of view of NMR spectroscopy of macromolecules, relatively long time scale processes (i.e. millisecond) can result in chemical shift averaging and it is this average that captures information about the structural integrity of encapsulated proteins by reference to their spectra in bulk aqueous solution. The small size of the water core of an individual reverse micelle (approximately 10^3 water molecules) means that at neutral pH only about one in a million reverse micelles actually carries a hydronium ion. However, under most conditions, the frequent collision of reverse micelle particles results in complete averaging of ionization states of the encapsulated protein and other titratable molecules.

The pH-dependence of ^{1}H NMR resonances of buffer molecules have been shown previously to provide a rapid and precise measure of the pH in bulk aqueous solution [99, 100]. In order to be useful as a pH indicator in the context of the reverse micelle water core, buffer molecules should provide resolved reporter ^{1}H resonances that do not interfere with the assessment of other parameters of reverse micelle solutions such as the determination of water loading by integration of water and surfactant resonances. They should also not interact significantly with the reverse micelle surfactant shell. Four of the buffers used for pH determination in bulk solution were identified to satisfy these criteria: imidazole, Tris, formate, and acetate. These span a useful pH range of 4 to 10 [99, 100]. Imidazole (pK_a ~7), which provides an effective pH indicator range of approximately 5.5-8, contains three hydrogens whose chemical shifts change as a function of pH. Imidazole H2 has a chemical shift range of 7.7-8.7 ppm, while the two degenerate H4/5 hydrogens have a chemical shift range of 7.1-7.5 ppm. Tris (pK_a ~8) provides an effective pH indicator range of 6.5-10 and a chemical shift range of 3.5-3.75 ppm. Formate (pK_a ~3.8) provides an effective pH indicator range of 4.0-5.5 and a chemical shift range of 8.38-8.45 ppm. Acetate (pK_a ~4.75) has an effective pH range of 4.0-6.5 and a chemical shift range of 1.90-2.05 ppm. The main impediment to using this approach in reverse micelles is the low effective concentration of the buffer molecules in the sample relative to the surfactant molecules and solvent whose NMR signals can bleed into the signal of the buffer chemical shifts. Deuterated solvents and the usual solvent suppression techniques [102] can be used to mitigate interference from solvent signals.

Initial experiments were performed to investigate the presence of interactions between buffer molecules and the surfactants used here. ^{1}H NMR spectra of aqueous solutions containing each buffer at 25 mM with 25 mM DSS as a chemical shift reference were recorded in the presence and absence of small amounts (~1 mM) of each of the three water-soluble surfactants used (AOT, CTAB, and LDAO) over the full pH range from 4 to 10 in increments of 0.5. These experiments confirmed that with small amounts of surfactant present there were no discernible pK_a shifts for any of the buffer molecules and only minute changes (<0.05 ppm) in buffer chemical shifts. This result was interpreted as indicating the absence of strong interactions between the surfactants and buffer molecules used here. The effects of the three surfactant mixtures (AOT, CTAB/hexanol, 10MAG/LDAO) on the pH of the reverse micelle water core were examined using samples prepared with the surfactants as provided by the manufacturers (i.e. without further purification or manipulation). Thirteen samples for each surfactant mixture were prepared by injecting the buffer mixture prepared at 0.5 pH increments from pH 4 to pH 10. ^{1}H NMR spectra of these samples showed that the pH of all AOT samples were within 0.5 units of pH 5 while all 10MAG/LDAO samples showed an aqueous nanopool pH within 0.5 units of pH 7.5. In contrast, CTAB/hexanol samples showed a simple titration curve without the need for pre-equilibration (Figure 2-1), showing that CTAB has negligible buffering capacity in the reverse micelle mixture over the pH range examined here. These results confirm that the protonation state of the AOT and LDAO head groups dominate the pH of the reverse micelle when these surfactant mixtures are employed. To prepare samples in AOT or 10MAG/LDAO at a given target pH, the surfactants were pre-equilibrated to the target pH value (see Methods in Appendix B). Samples were prepared in this way for the thirteen pH values described above and examined by ^{1}H NMR. The chemical shifts of the buffer molecules as a function of reverse micelle sample pH along with example spectra obtained at pH 7.0 are shown in Figure 2-1. These data are shown based on the target pH, a view which shows some inconsistencies (RM vs aqueous) in the measured pH within the reverse micelle sample depending on which buffer was used to establish the pH (Figure 2-1). This is traced to a change in the effective pK_a and shows that some of the buffers interact with or are otherwise perturbed by the reverse micelle surfactants. It should also be noted that not all buffer molecules are appropriate for all surfactant mixtures. The pH-dependent hydrogen signal of Tris, for example, overlaps with a signal from AOT at high pH and overlaps significantly with one of the LDAO signals at all pH values (Figure 2-1f). These data indicate that there is not necessarily a particular buffer that will be ideal for monitoring pH in all reverse micelle surfactant mixtures. However, a mixture of acetate and imidazole offers general applicability over a wide range of pH values (4-8.5).

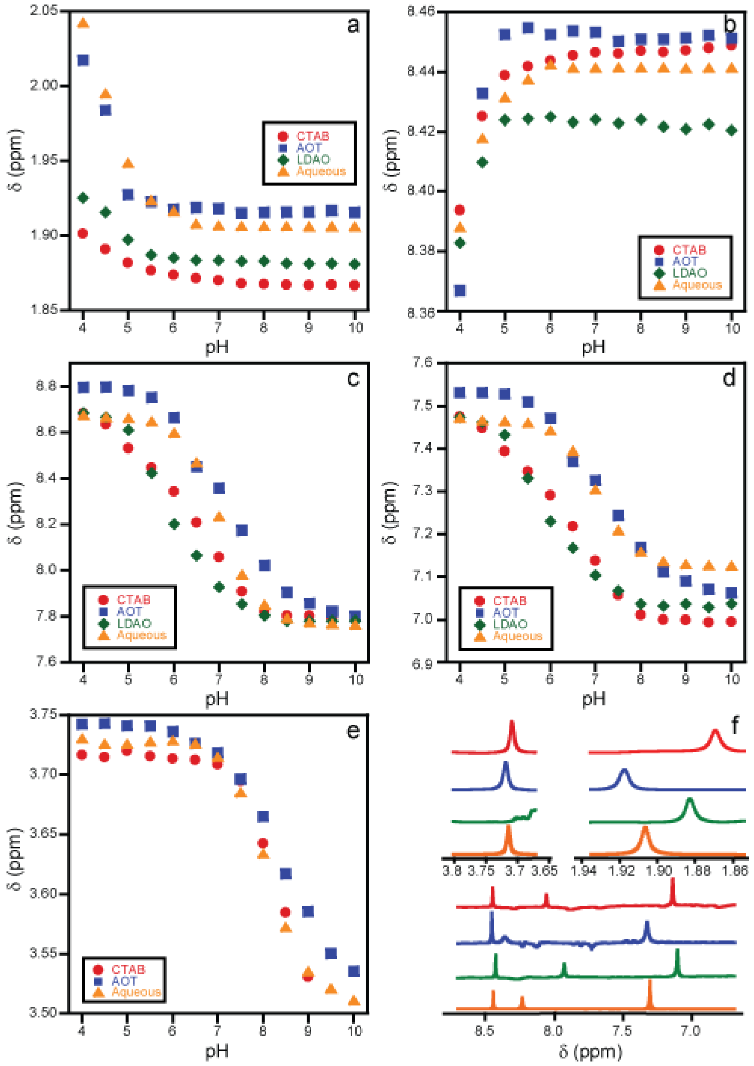

Figure 2-1: Buffer chemical shifts as a function of target pH in aqueous solution and reverse micelles composed of various surfactant mixtures. The pH-dependent chemical shifts of buffer hydrogen atoms [99, 100] are shown: **(a)** acetate; **(b)** formate; **(c and d)** imidazole H2 and H4/5, respectively; **(e)** Tris. Data are shown for aqueous buffer samples (orange triangles) and in the three reverse micelle surfactant conditions tested: AOT (blue squares), CTAB/hexanol (red circles), and 10MAG/LDAO (green diamonds). In all cases except CTAB/hexanol, the surfactants were pre-equilibrated with solutions at the target pH (see text). The signal for Tris overlaps significantly with surfactant peaks in the 10MAG/LDAO mixture and was therefore omitted from these data. Representative spectra of each buffer molecule in each environment at pH 7 are depicted in **(f)** using the same color scheme with the upper left spectra corresponding to Tris, the upper right corresponding to acetate, and the lower spectra corresponding to both formate (left peak in all spectra) and imidazole chemical shifts.

Clearly the pH indicator molecules may experience a shift in their apparent pK_a values by as much as 1 pH unit upon encapsulation in some surfactant mixtures. For example, in the imidazole plots (Fig. 2-1c and d), the chemical shifts of the AOT-encapsulated buffer appears to closely match those seen in aqueous solution while in the 10MAG/LDAO and CTAB/hexanol mixtures the chemical shifts seems to show a reduction in the pK_a from 7 to approximately 6.2. Comparison of the pH response of imidazole to that of Tris (Fig. 2-1e) in the CTAB/hexanol mixture yields a conflicting result i.e. one or more of the pH indicator molecules has experienced a shift in pK_a as a result of encapsulation.

To more quantitatively assess the apparent pK_a shifts for imidazole, the data were fit to equation 11 [100]:

$$pH = pK_a - \log\left(\frac{\delta_{obs} - \delta_{HA}}{\delta_A - \delta_{obs}}\right) \tag{11}$$

Where δ_A and δ_{HA} are the chemical shifts of the basic and acidic forms of the buffer, respectively, and δ_{obs} is the observed buffer chemical shift in a given sample. The data was fit using the target pH as the known quantity to determine an apparent pK_a for each solution condition (Table 2-1). Based on these data alone, the assumption that the sample is at the target pH is unfounded but the data from encapsulated proteins confirms the validity of this assumption (see below).

Table 2-1: Effective pK_a of imidazole in various reverse micelles.

Sample Condition	Imidazole (H2)	Imidazole (H4/5)
Aqueous	6.96	6.98
AOT (W_0 = 15)	7.29	7.30
CTAB (W_0 = 15)	6.22	6.21
10MAG/LDAO (W_0 = 12)	6.14	6.16
10MAG/LDAO (W_0 = 20)	6.72	6.79

The pK_a of imidazole varies by as much as 0.9 pH units across the various reverse micelle surfactant mixtures examined. The pK_a is also influenced by the water loading in the 10MAG/LDAO surfactant system. The sensitivity of the apparent pKa of the indicator buffers requires that the pH-dependent response must first be characterized for each surfactant mixture.

Surfactant molecules dominate the pH of reverse micelles

The chemical shifts of proteins are well known to exhibit exquisite sensitivity to pH. The amide nitrogen and amide hydrogen chemical shifts of the 8.5 kDa protein ubiquitin were used as indicators of the pH environment in the reverse micelle core. Ubiquitin was chosen as the protein indicator for these studies due to its stability over a wide pH range and its previously characterized amenability to encapsulation in many different surfactant mixtures [57, 59, 76]. Ubiquitin was prepared in aqueous solution at pH 5 or pH 7 and ^{15}N-HSQC spectra were recorded (Figure 2-2a and 2-2c, respectively). These spectra show the typical degree of pH-dependent chemical shift changes for proteins in aqueous solution. The aqueous protein sample at pH 5 was then encapsulated in 10MAG/LDAO while the aqueous sample at pH 7 was encapsulated in AOT, both without pre-adjustment of the surfactant pH. ^{15}N-HSQC spectra were collected for both of these reverse micelle samples (Figure 2-2b and d, respectively.) As described above, the chemical shifts of the buffer molecules in these samples indicated that the injected solution of ubiquitin, initially at pH 7, shifted to a pH ~5 upon encapsulation in unadjusted AOT. The protein ^{15}N HSQC spectrum agrees with this result. Indeed, this spectrum matches the previously determined assignments of encapsulated ubiquitin at pH 5 [76]. By the same token, the buffer chemical shifts of ubiquitin solution, initially set to pH 5, and then encapsulated in unadjusted 10MAG/LDAO (2-2b) indicate that the aqueous core of the reverse micelle was at pH 7. The chemical shifts of the encapsulated ubiquitin closely match those of the aqueous pH 7 spectrum (2-2a). From these data, it is clear that the surfactant dominates the pH of the aqueous nanopool for reverse micelle mixtures composed of surfactants with titratable head groups even in the presence of macromolecules.

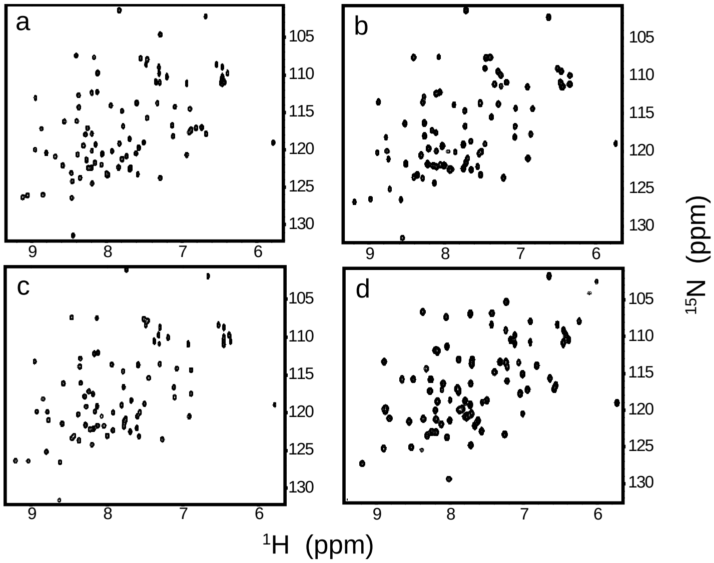

Figure 2-2: 15N-HSQC spectra of uniformily 15N-labeled ubiquitin in aqueous solution at different pH and in various reverse micelle mixtures. Aqueous ubiquitin at pH 5 **(a)** was encapsulated in 10MAG/LDAO reverse micelles without prior pH equilibration of the surfactants **(b)**. Similarly, aqueous ubiquitin at pH 7 **(c)** was encapsulated in AOT without prior pH equilibration of the surfactant **(d)**.

Calibration of the Reverse Micelle Interior for Encapsulation of Proteins

The goal of this work was to establish a method by which reverse micelles could be prepared at a target pH or adjusted to a target pH after formation of the reverse micelle mixture. Both encapsulated ubiquitin and the chemical shifts of the four buffer molecules indicate clearly that the pH of the encapsulated solution is dominated by the protonation state of the AOT or 10MAG/LDAO surfactants prior to sample preparation. The ^{15}N-HSQC spectrum of encapsulated ubiquitin was used to determine whether pre-equilibration of surfactants with an aqueous solution set at a target pH could be used to effectively set the effective pH of the reverse micelle core.

Ubiquitin was prepared in aqueous solution with a mixture of the four buffer pH indicators at three pH values: 5, 7, and 9. ^{15}N-HSQC spectra were taken for each sample to have a reference spectrum of the aqueous protein to compare to the reverse micelle-encapsulated counterparts (Figure 2-3). These protein samples were then encapsulated in 10MAG/LDAO reverse micelles. 10MAG/LDAO was chosen for these tests because of its capacity to reproduce the aqueous chemical shifts of proteins upon encapsulation and its applicability to a wide variety of proteins over a broad pH range [58]. The surfactants had been pre-equilibrated to pH 5, 7 and 9, respectively, as described in the Methods (Appendix B). The corresponding ^{1}H NMR spectra showed buffer resonance chemical shifts that closely matched those of the pre-adjusted protein-free samples (see Figure 2-1). The corresponding^{15}N-HSQC spectra (Figure 2-3) showed excellent agreement with the aqueous spectra, indicating that the effective pH of the reverse micelle water core had been set to the desired target. Thus the shift of the pK_a of imidazole is due to encapsulation in 10MAG/LDAO and the titration curves shown in Figure 2-1 can be used for calibration of the internal pH of 10MAG/LDAO reverse micelles.

To test the utility of the buffer chemical shifts for measurement and adjustment of the reverse micelle pH, the pre-equilibrated reverse micelle samples were titrated to a different target pH by direct addition of concentrated acid or base while maintaining the W_0 of the sample (see Methods in Appendix B). These titrations were monitored using only the ^{1}H NMR signals of the imidazole and acetate buffers. Once the target pH was reached, a ^{15}N-HSQC spectrum was collected (Figure 2-3, right column). The spectra of the titrated samples closely match those of both the pre-equilibrated and aqueous samples, confirming the utility of the internal buffer signals for measurement and adjustment of the pH of the reverse micelle core.

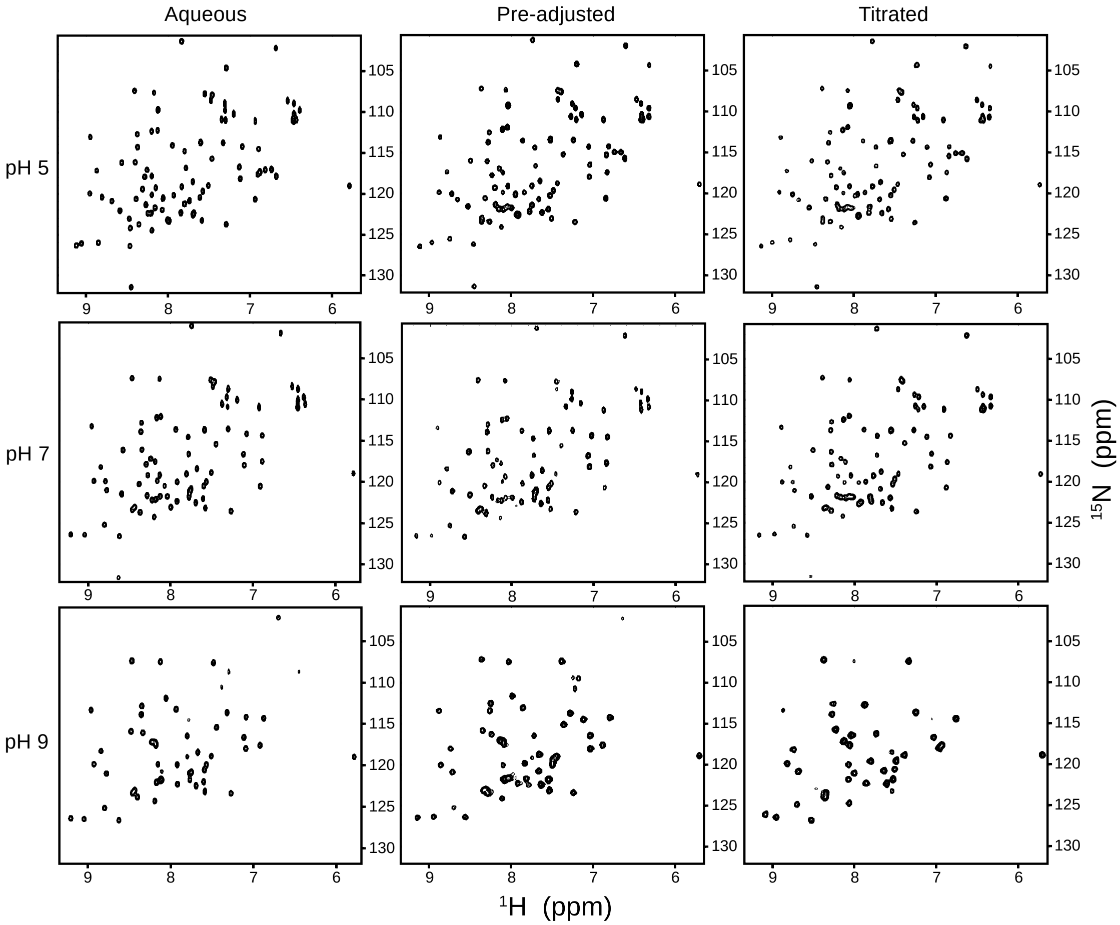

Figure 2-3: 15N-HSQC spectra of uniformily 15N-labeled ubiquitin in aqueous solution at different pH and in 10MAG/LDAO reverse micelle mixtures pre-equilibrated to a given pH and after titration. Aqueous ubiquitin samples (left column) at the indicated pH were encapsulated in 10MAG/LDAO to a W_0 of 10 that had been pre-equilibrated to the same pH (middle column). The pre-adjusted pH 9 reverse micelle sample was titrated to pH 7, then to pH 5. The pre-adjusted pH 7 reverse micelle sample was titrated to pH 9. The pH was monitored by tracking the chemical shift changes of the imidazole and acetate buffers in the solution.

Composition and pH of the reverse micelle affect protein stability

Whether or not a protein remains in its native conformation upon encapsulation is dependent on many factors including the composition of the surfactant mixture, the water loading, and the pH [57, 59, 61]. Despite the wealth of studies on proteins encapsulated in reverse micelles, efforts to discern the structural fidelity of the encapsulated protein vary widely [53, 65]. Often no detailed examination of the conformational state of the encapsulated protein is undertaken and, if considered, only low-resolution methods such as optical spectroscopy or circular dichroism have been used [53, 65, 73-75].

The vast majority of previous studies on encapsulated proteins have used AOT as the reverse micelle surfactant system. As we have demonstrated before [59] and show again here, AOT reverse micelles can frequently solubilize proteins with great efficiency, but solution NMR measurements of these samples show that ***most proteins are largely or completely unstructured in AOT reverse micelle solution***. Ubiquitin is the sole notable exception [76, 103, 104]. Figure 2-4 demonstrates the sensitivity of cytochrome c to the precise conditions of encapsulation. Though encapsulation in AOT reverse micelles solubilizes cytochrome c, a ^{15}N-HSQC of the sample reveals that the protein is highly conformationally averaged, as evidenced by the collapsed spectrum (Fig. 2-4e). The spectrum of aqueous cytochrome c, conversely, shows the typical dispersion of a properly folded protein (Fig 2-4a). Encapsulation of the protein in the charge-neutral 10MAG/LDAO surfactant mixture results in a sample with a spectrum that closely mimics that obtained in aqueous solution (Fig 2-4c). In this case, though the pH of the reverse micelle samples (as determined by the ^{1}H buffer signals) are identical to that in aqueous solution, the surfactants used dictate the structural state of the encapsulated protein. Using the buffer molecules as internal pH indicators, the unfolding of this protein as a function of pH can be reproduced in the 10MAG/LDAO reverse micelles with similar spectral results by NMR (Fig 2-4 b and d).

Evaluation of these various conditions by visible absorption spectroscopy of the Soret band yields results that are much more difficult to interpret. The Soret absorption of the aqueous and 10MAG/LDAO-encapsulated proteins at pH 5 match closely with absorbance maxima at 410 nm, but the absorption in AOT reverse micelles differs slightly. This difference has been noted previously and has been variously interpreted [105-108], but the spectral change matches closely with that induced by unfolding of cytochrome c upon binding to DOPC micelles [109]. It should be noted that the Q band of the heme at ~525 nm shows marked differences for the aqueous and 10MAG/LDAO-encapsulated protein despite the clear NMR evidence that both are in their native conformations. Though this slight shift is somewhat ambiguous, the NMR spectra clearly show the denaturation of the protein by the AOT reverse micelle solution. The response of the Soret band to pH-induced unfolding of the protein [104] varies considerably between aqueous and reverse micelle conditions. While the wavelength of maximal absorption shifts to 395 nm in aqueous solution at pH 2.5, the shift in 10MAG/LDAO reverse micelles is much more dramatic (375 nm). These data demonstrate the clarity offered by solution NMR for evaluation of encapsulated protein structural integrity as well as the utility of the present method for pH adjustment and monitoring in reverse micelle systems.

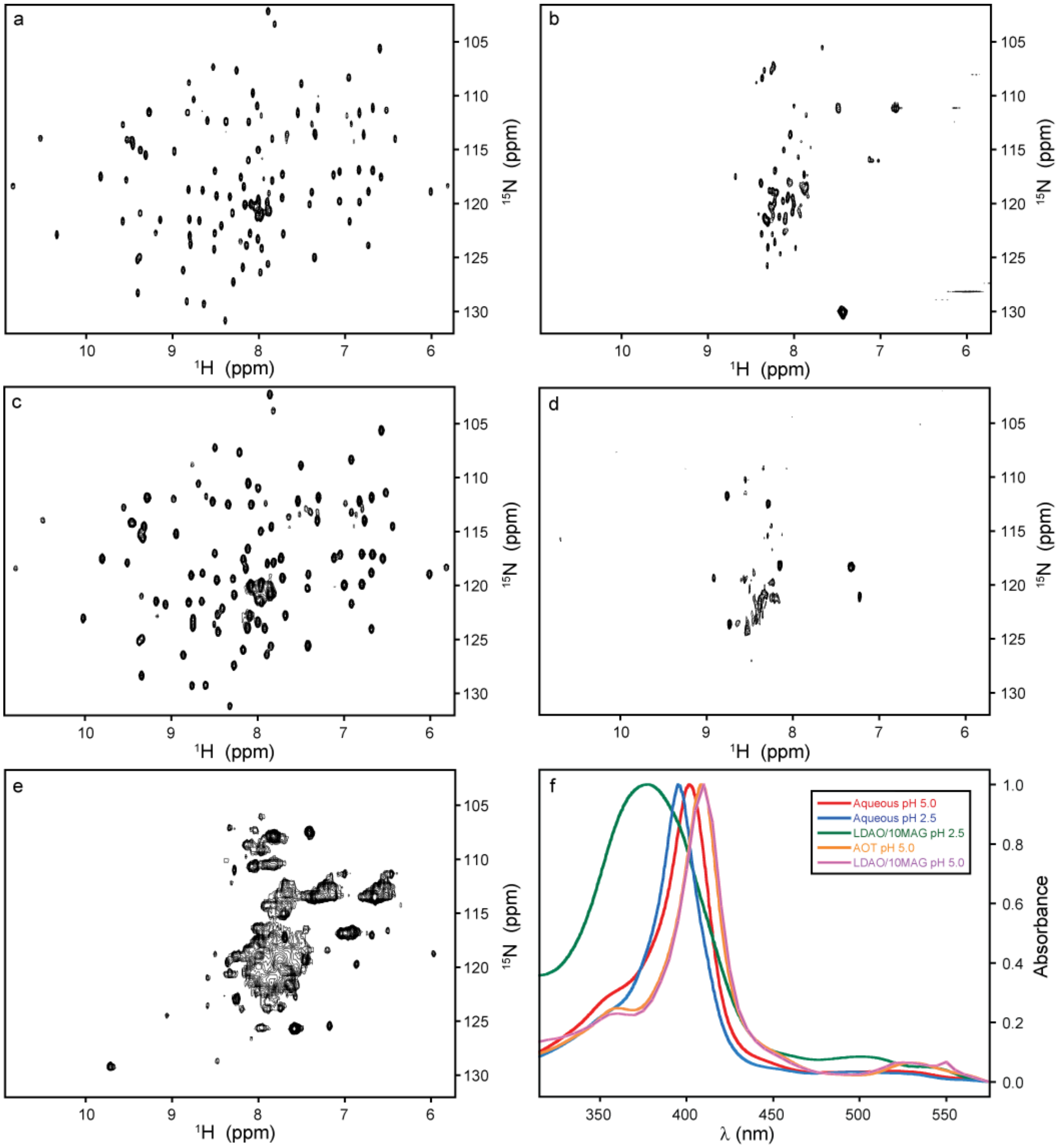

Figure 2-4: 15N-HSQC and Soret band absorption spectra of uniformily 15N-labeled cytochrome c in aqueous solution at different pH and reverse micelle mixtures to monitor protein foldedness. Native aqueous cytochrome c at pH 5 **(a)** is shown to unfold at pH 2.5 in aqueous solution **(b)**. Similarly, native RM-encapsulated cytochrome c in 10MAG/LDAO reverse micelle mixtures at pH 5 **(c)** is shown to unfold at pH ~2.5 in the same mixture **(d)**. Cytochrome c in AOT reverse micelles at pH 5, conversely is also unfolded **(e)**. All reverse micelle mixtures had a final protein concentration of 140 μM and W_0 of 15. Normalized optical absorbance spectra of all samples **(a-e)** from 315 nM to 575 nm are also shown **(f)**.

The L99A mutant of lysozyme from T4 bacteriophage is also unfolded under acidic conditions. In aqueous solution, the protein is in a natively folded state at pH 5 and becomes increasingly unfolded as the pH is lowered until it is completely unfolded at a pH of 2.5 (Figure B-1 in Appendix B). Figure 2-5 demonstrates that this is also the case for T4 lysozyme when encapsulated in the 10MAG/LDAO/DTAB reverse micelles. The encapsulated L99A mutant of T4 lysozyme is shown to maintain its native fold at a pH of 5 (2-5a). The pH was lowered by direct injection of HCl and monitored using the proton chemical shift of acetate buffer molecules. As the pH of the sample is lowered to 3.5, the ^{15}N-HSQC spectrum (2-5b) shows that the protein is beginning to unfold. This is evident due to the collapse and disappearance of multiple peaks in the pH 3.5 spectrum as highlighted by the insets in Figures 2-5a and 2-5b. Much like in aqueous solution, lowering the pH of the sample to 2.5 caused the loss and collapse of the majority of peaks in the ^{15}N-HSQC spectrum (2-5c) indicating that the T4 lysozyme is fully unfolded.

Native tryptophan fluorescence is frequently used to assess the structural state of soluble proteins both in aqueous solution and in reverse micelles [110]. Figure 2-5d shows fluorescence spectra of aqueous and encapsulated T4 lysozyme samples at pH 5, 3.5, and 2.5. Although the ^{15}N-HSQC clearly demonstrates that the lysozyme is in a partially unfolded state at pH 3.5, the fluorescence spectra are virtually identical regardless of solution condition (bulk or RM). This could lead to incorrect assumptions about the protein's conformational state in the absence of NMR data. There is only a clear shift in the peak of the fluorescence emission from 339 nm to 342 nm (2-5d blue dashed curve) once the protein is completely unfolded (2-5c) at pH 2.5. This phenomenon is replicated in aqueous solution (2-5d solid curves), although it should be noted that the shift in aqueous solution when the protein completely unfolds (from 334 nm to 348) is much larger than in the reverse micelle mixture. These data clearly indicate that solution NMR spectroscopy is the most complete tool available to unambiguously monitor the conformational state of encapsulated proteins.

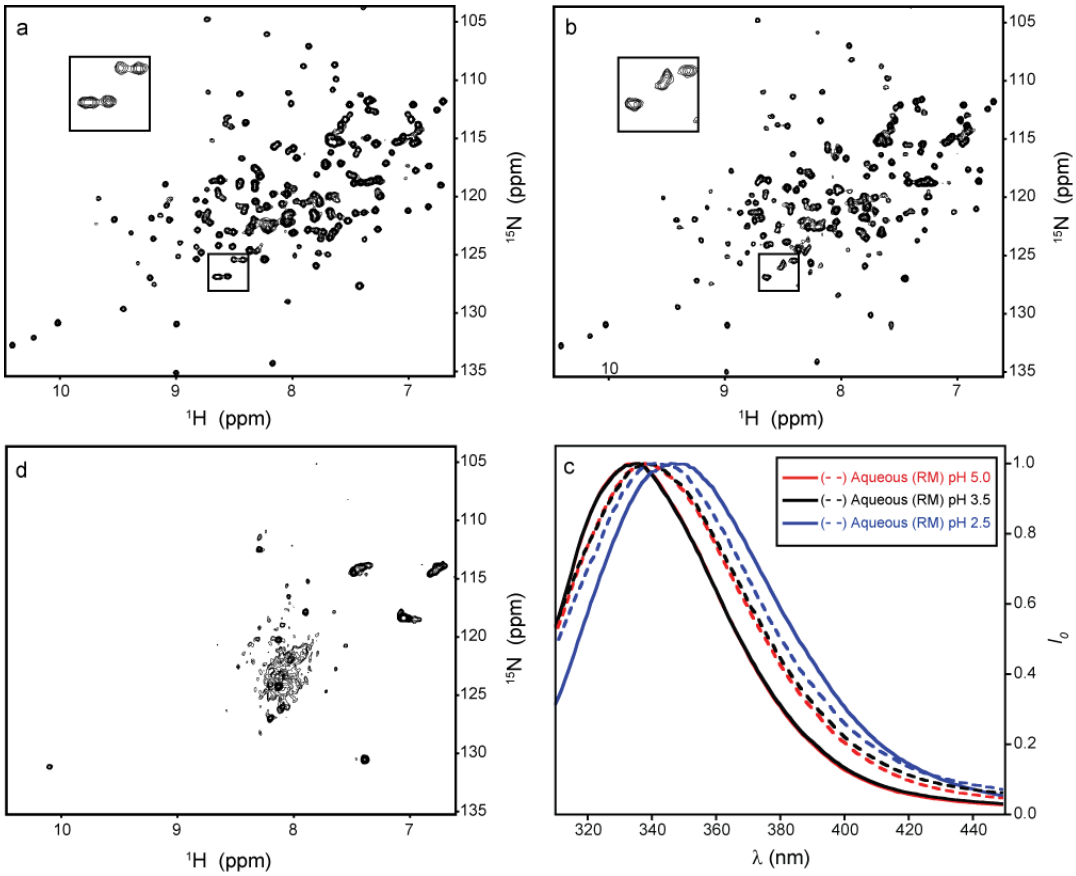

Figure 2-5: 15N-HSQC and flourenscence emission spectra of uniformily 15N-labeled T4 lysozyme L99A mutant in reverse micelle mixtures at varying pH to monitor protein foldedness. Aqueous L99A mutant of T4 lysozyme at pH 5 is encapsulated in a natively folded state in a 10MAG/LDAO/DTAB reverse micelle mixture at a water loading of 18 **(a)**. The pH of the sample was lowered to 3.5 to create a partially unfolded state **(b)** and subsequently lowered to pH 2.5 to create a fully unfolded state **(c)**. Insets in A and B demonstrate protein unfolding by showing the collapse and disappearance of peaks. All reverese micelle mixtures had a final protein concentration of 80μM. Tryptophan fluorescence emission spectra **(d)** of all three reverse micelle samples along with aqueous samples at the same pH are shown in red, black, and blue, respectively. Solid lines represent the aqueous protein while dashed lines represent the reverse micelle mixture.

In addition to the dependence of encapsulated proteins on the nature of the surfactant interface, the water content of the reverse micelle can also play an important role in maintenance of structural fidelity. Figure 2-6 shows the W_0 dependence of the L99A mutant of T4 lysozyme when encapsulated in 10MAG/LDAO/DTAB reverse micelles. The ^{15}N-HSQC spectra of T4 lysozyme L99A clearly show that, upon encapsulation of the aqueous protein (Fig. 2-6a) in 10MAG/LDAO/DTAB reverse micelles, the amount of water within the reverse micelle ensemble greatly influences the protein conformational ensemble. At low W_0 many peaks which closely match the aqueous spectrum are observed, suggesting that the encapsulated protein is largely folded. However, the collapse and disappearance of a number of peaks show that the lysozyme populates a range of partially unfolded states under this condition (Fig. 2-6b). Upon increasing the W_0 of the sample from 12 to 18, however, there is a marked improvement in spectral quality and the reappearance of the missing peaks indicating a shift toward a fully native fold (Fig. 2-6c). Native tryptophan fluorescence was used again to monitor the structural state of T4 lysozyme L99A mutant with varying W_0. Inspection of the tryptophan fluorescence spectra (Fig. 2-6d) of the aqueous protein (red) upon encapsulation yields results that would be doubly misleading in the absence of NMR data. When encapsulated at a W_0 of 12 (blue), the native tryptophan fluorescence of T4 L99A shows a clear redshift of the emission peak maximum from 333 nm to 339 nm. Such a redshift is typically interpreted as being indicative of the exposure of tryptophan side chains to solvent due to protein unfolding. In the context of the reverse micelle, exposure to solvent becomes less straightforward to interpret than in aqueous solution. The^{15}N-HSQC of T4 L99A at W_0 of 12 (Fig. 2-6b) shows clear resolution of three distinct tryptophan indole peaks (lower left of the spectrum) which closely match those in aqueous solution.

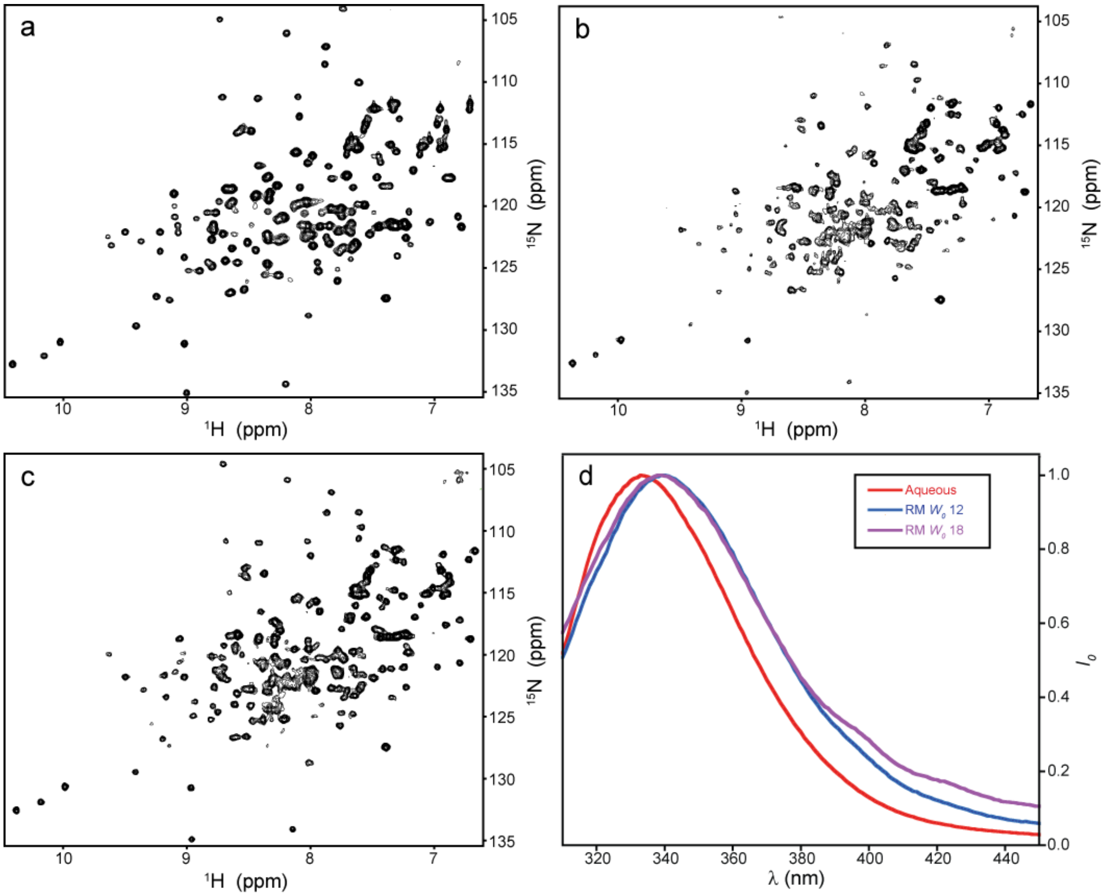

Figure 2-6: 15N-HSQC and flourenscence emission spectra of uniformily 15N-labeled T4 lysozyme L99A mutant in aqueous solution and reverse micelle mixtures at various water loadings to monitor protein foldedness.AqueousL99A mutantof T4 lysozyme at pH 5 **(a)** is encapsulated in a partially unfolded state in a 10MAG/LDAO/DTAB reverse micelle mixture at a water loading of 12 **(b)**. The water loading was increased to 18 **(c)** to allow for proper protein folding. Both reverse micelle mixtures had a final protein concentration of 80μM. Tryptophan fluorescence emission spectra **(d)** of all three protein samples **(a-c)** are shown in red, blue, and purple, respectively.

The traditional interpretation of the fluorescence data argues that the tryptophan residues are in a non-native environment, but the NMR data clearly shows that the tryptophan residues are largely in their native conformation. Upon increase of W_0 from 12 to 18, the fluorescence spectrum (purple) remains unchanged with the emission peak maximum remaining at 339 nm, despite the full structural fidelity evident in the NMR data at this condition. Here the fluorescence data alone would indicate that changing the W_0 had no impact on the encapsulated protein, but the NMR data clearly demonstrate that this is not the case. These data strongly indicate that NMR spectroscopy is optimal for both calibration of the internal pH of reverse micelle samples and confirmation of encapsulated protein structural fidelity.

Conclusions

In this study we used the observation of ^{1}H NMR signals of common buffer molecules as previously described for aqueous solution [99, 100] in order to monitor, maintain, and adjust the pH within the aqueous interior of empty reverse micelles. We then confirmed this approach by monitoring the chemical shift changes of the buffer molecules as similar pH-attributed chemical shift perturbations and ^{1}H-^{15}N cross-peak broadening were occurring in encapsulated ubiquitin at varying pH. Most importantly, we demonstrate that the pH of titratable head groups completely dominate the pH of the aqueous reverse micelle interior. In effect, we described methods allowing for the pre-adjustment of surfactant pH before the formation of reverse micelles samples. Finally, we demonstrated that protein structure within a reverse micelle is often very sensitive to the reverse micelle conditions, like the pH of the aqueous nanopool of water or the amount of water within the reverse micelle ensemble (dictated by the water loading). If possible, it is our conclusion that NMR should always be used (within reason) in order to confirm the structural fidelity of reverse micelle-encapsulated proteins before further analysis can be made.

CHAPTER 3: Optimized protocol for the determination of protein hydration dynamics using NMR spectroscopy

Most of the work presented in this chapter was done in collaboration with Christine Jorge, fellow Ph.D. candidate in the laboratory of A. Joshua Wand

Abstract

Protein-water interactions are paramount to protein structure and function; however, despite the fundamental role of water in protein function, the ability to measure these interactions in a site and time resolved manner has remained technically difficult. This has left a large hole in our understanding of protein thermodynamics. Many efforts have historically been made to measure these interactions, but have each been plagued with their own sets of caveats and technical limitations. Solution nuclear magnetic resonance (NMR) spectroscopy had been proposed as a means of measuring these interactions via the nuclear-Overhauser effect (NOE). It has previously been shown that encapsulating proteins in the core of reverse micelles serves to eliminate the issues associated with solution state measurements. Specifically, encapsulation in the reverse micelle removes bulk water while retaining the native hydration layer therefore reducing hydrogen exchange rates and slowing the motions of water. Isolating the protein and its hydration shell allows us to measure, with high reproducibility, these fundamental protein-water interactions. In this study we first demonstrate the effect of reverse micelle encapsulation on potential contamination due to bulk water according to simulation. We then describe criteria for minimum signal-to-noise of hydration data in order to achieve excellent reproducibility. Finally, we describe a new work-flow involving collecting hydration experiments with a

buildup of NOE mixing times that allows for efficient and reproducible measurement of semi-quantitative protein hydration.

Introduction

The ability to measure protein-water interactions is necessary to understand the thermodynamic role solvation water plays on protein function and stability [1, 111, 112]. However, measuring these interactions has proven to be quite difficult for a number of reasons. Solution nuclear magnetic resonance (NMR) has been used to measure the dynamics of protein water interactions via the nuclear Overhauser effect [23, 26, 113]. When performed in bulk aqueous solutions these measurements have numerous artifacts that limit the ability to study protein hydration in a quantitative manner [8, 27].

We have recently shown that encapsulating proteins in the hydrophilic core of a reverse micelle (RM's) retains the native folded protein and its hydration shell while reducing the amount of water [29, 60, 76]. These factors contribute to slower water and reduced hydrogen exchange [54, 114]. Proteins encapsulated in reverse micelles allow an experimental condition that is optimal for the study of protein hydration. Nucci et al showed that the majority of ubiquitin in AOT reverse micelles had detectable hydration that spanned the entire theoretical limit [29]. He further concluded that the HX rates are indeed slowed by several orders of magnitude [29]. The combination of selective perdeuteration and reverse micelle encapsulation removes all artifacts present in aqueous solution measurements without the need for complicated pulse sequences.

However, despite this massive improvement in sample preparation one major technical problem still remains. The overall concentration of protein in a reverse micelle sample is often low (~50-300μM) [60]. Cross peaks observed in nuclear Overhauser effect spectroscopy can often be quite weak due to rapid motions or large distances between dipoles. This is especially true in the case of intermolecular interactions [115]. Thus, protein hydration experiments are often very time costly due to the need to collect a high number of transients per FID, which is necessary to obtain peaks with adequate signal-to-noise (S/N) intensity.

In order to reduce experimental time we have implemented non-uniform sampling (NUS) to the study of protein hydration. Traditional NMR experiments sample uniformly across a Cartesian grid in order to satisfy the criteria for the discrete Fourier transform (DFT). NUS takes advantage of the fact that only a small subset of sampled frequencies contain data while the rest contain noise. Therefore, when using NUS only a fraction time points need to be collected resulting in decreased experimental time per experiment. This time savings can then be applied to collecting experiments with greater signal-to-noise (S/N) or resolution in a tractable amount of time [116-118]. However, processing of NUS data with a traditional DFT results in many artifacts. A number of programs have been created to reconstruct NUS data sets that result in accurate S/N and frequency reconstruction with minimal artifacts. In this present study we use sinusoidally-weighted Poisson-gap [119] NUS schedule followed by iterative-soft thresholding (IST) reconstruction [120]. IST reconstruction belongs to a larger class of compressed sensing algorithms [121]. It has previously been shown to have accurate frequency reconstruction and high linear peak-heights relative to a uniformly sampled reference and also benefits from reduced spectral noise [116, 119, 122]. We have chosen this method due to its ease of application, broad use in the NMR community, and minimal computational requirements. However, the procedure in this present paper should be broadly applicable to a number of NUS sampling scheduling and reconstruction methods.

First we identify the minimal recommended sampling density necessary for protein hydration via the NOE. Then we assess the intra-sample and inter-sample reproducibility of hydration $\sigma_{NOE}/\sigma_{ROE}$ ratios as detected with NUS hydration. This allows us to determine a reliable minimal S/N for high data reproducibility. We then collect hydration experiments with a series of mixing times in order to improve the way in which the $\sigma_{NOE}/\sigma_{ROE}$ is calculated. The hydration measurements collected here are compared to previously published experiments and show that qualitative trends remain the same despite marked quantitative improvements. Finally, we offer recommendations of how to measure protein hydration in a quantitative way without the need to further optimize the NUS procedure for each protein system.

Results

Effect of “bulk” solvent on NOE/ROE Ratio in reverse micelle

While the use of NMR spectroscopy is a potentially powerful tool for determining quantitative information about the hydration dynamics near the surface of a protein [23, 24, 26], it has been shown in aqueous solution that the intermolecular NOE between a protein and hydration waters can become contaminated [8]. One such potential avenue of contamination is the diminishment of the distance dependence of the intermolecular NOE between protein and water protons due to the presence of bulk solvent [27]. Interestingly, we have recently noted that upon reverse micelle encapsulation proteins tend to maintain their hydrated radii with excess water likely being sequestered into empty reverse micelles [57]. In effect, bulk aqueous solvent cannot affect the NOE/ROE ratio because there is little bulk solvent in the ensemble (unless there was contamination from bulk alkane solvent – no evidence has been seen for this in any examined protein system). In effect, we used similar simulation approaches [27] in order to determine the effect of waters near the surfactant surface (~5 layers of water away from the protein) on the NOE/ROE ratios detected immediately near the protein surface.

As described in the Methods section of Appendix C, the protein-containing reverse micelle ensemble was assumed to mimic a scenario with a restricted uniform distribution of water up to a reflective boundary at the surfactant surface. While uniform distribution of hydration dynamics within the reverse micelle ensemble is likely not realistic [68], this model was selected for simplicity with the results of contamination from relatively distant solvent to be minimized further in the case of a gradient of hydration dynamics. We simulated the rank-2 spectral density function (Appendix C) for the protein-water NOE both with no effect from other solvent molecules and with the effect of all other solvent molecules within the reverse micelle ensemble (Figure 3-1):

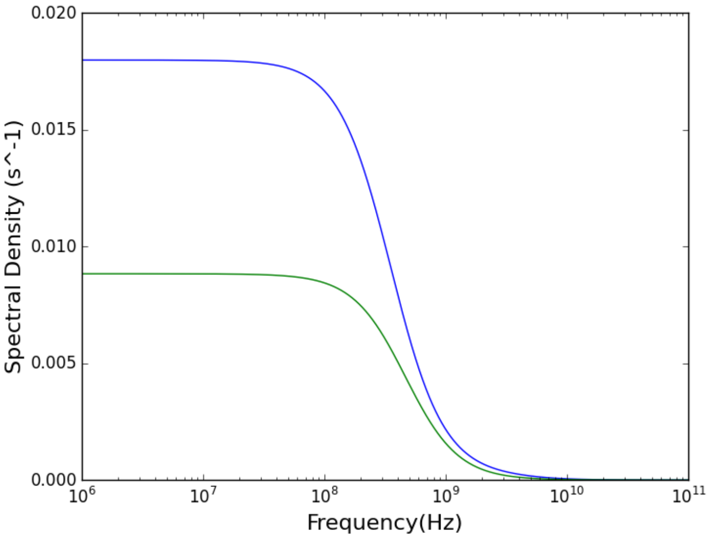

Figure 3-1: Simulations of the rank-2 spectral density function shows significant differences when including effects from other encapsulated waters. The spectral density is depicted for the protein-water NOE within 4 Å of the protein proton with no added effect from solvent molecules beyond 4 Å (green) and with the added effect of waters up to the surfactant head-group border (blue).

Clearly there is still a difference for the spectral density of the protein-water NOE immediately at the surface of the protein versus including effects from all other waters within the surfactant shell. When using the simulated spectral density functions to calculate the effective NOE/ROE ratios of both scenarios via equations 1 and 2, there is a significant difference between the hydration ratios (~30%) demonstrating an effect of distant solvent water molecules on the NOE/ROE ratio.

Recently, however, Steinhauser and colleagues have reinvestigated the effect of bulk solvent on the protein-water NOE [28]. In this study, the authors applied a similar simulation as performed by Halle [27], but this new study included an overlooked factor: the pair correlation function between the protein spin and all water spins. They find that, because the pair correlation function has a high initial peak, the dependence of protein-water NOEs on bulk water becomes entirely dependent on the burial depth of the protein probe. For protein spins at the surface of the protein, the contribution from the pair correlation function is much higher due to the higher incidence of nearby waters: this leads mathematically to less of a contribution to the protein-water NOE from bulk solvent since the contribution from the pair correlation function is so high. For protein probes away from the surface of the protein, it more closely matches the case described by Halle: for these probes the pair correlation function plays less of a roll because of the lower incidence of nearby water protons. Thus, according to recently updated simulations, it is clear that bulk solvent only affects protein-water NOEs for protein probes significantly far from water. These protein-water NOEs are, of course, not experimentally detectable because of the large distance between the protein and water and thus are only "detectable" via simulations. In effect, the authors conclude that that NOE/ROE ratio is an excellent tool for the detection of hydration dynamics near the surface of a protein without fear of contamination from bulk solvent (either in aqueous solution of encapsulated in a reverse micelle).

Choice of sampling density

In order to measure hydration quantitatively accurate peak heights in addition to the precision of peak frequencies are absolutely necessary. We set out to determine the minimal NUS sampling density necessary for a quantitative NOE. The peak height intensities of the intermolecular NOE between protein and water is low relative to intramolecular cross peaks. In the case of hydration it is often more beneficial to increase the number of transients per FID than increasing the resolution in the indirect dimensions

[116]. In order to find a minimum recommended sampling density we collected a series of ^{15}N-edited NOESY-HSQC spectra on U- [^{13}C^{15}N]-ubiquitin collected at low spectral resolution with different NUS sampling densities (Table 3-1).

Table 3-1: Determination of required NUS sampling density.

	Slope	R^2	RMSD%	% Error (NOE/ROE)
US vs. US	1.005	0.999	2.12%	2.99%
US vs. 5% NUS	0.839	0.893	32.75%	46.32%
US vs. 10% NUS	0.938	0.976	14.82%	20.96%
US vs. 15% NUS	0.952	0.989	9.82%	13.89%
US vs. 20% NUS	0.966	0.994	7.76%	10.98%
US vs. 25% NUS	0.971	0.996	5.85%	8.28%
25% vs. 25% NUS	1.01	0.998	3.30%	4.67%

US = Uniform Sampling, NUS = Non-Uniform Sampling

We collected five sampling densities: 5, 10, 15, 20, 25% NUS in both indirect dimensions as well as duplicates in the uniformly sample and 25% sampled (highest sampled density) data sets. Over 200 well-resolved cross-peaks ranging from signal-to-noise of 10-200 (~1-20% diagonal peak height) was compared across the varied densities. The R^2 converges to ~0.99 at a15% NUS density, however the <RMSD> between uniformly sampled (US) and NUS sampled data continues to decrease as the sampling density is increased. The percent RMSD in peak height between two US replicates was 2.12%, compared with two 25% NUS replicates with a percent RMSD of 3.3%. The percent RMSD error between an US and 25% NUS data set was 5.85%. Because the error of the $\sigma_{NOE}/\sigma_{ROE}$ is taken in quadrature it was important to have a final percent error between two replicates be ~5% of the total measurement.

Therefore, all experiments for testing the reproducibility of hydration were used at 25% NUS data collection yielding a total of 25% of the experimental time required for a uniformly sampled dataset. These guidelines provided a conservative estimate that was determined to be sufficient for high reproducibility independent of sampling schedule optimization, resolution, and spectral crowding.

Reproducibility of NUS-derived hydration ratios

In order to study the reproducibility of the hydration dynamics measurements we collected ^{15}N-edited NOESY-HSQC and ROESY-HSQC spectra at a single mix point on two independent samples each collected in duplicate. Each sample used >90% perdeuterated protein and was prepared in AOT reverse micelles as described in the Methods (see Appendix C). The ratio of the intensity of the NOE and ROE cross peak was then identified for each peak. This allowed us to compare both intra- and inter-sample reproducibility of hydration dynamics measurements. Both the intra- (Figure 3-2 left) and inter-sample (Figure 3-2 right) are very reproducible with several obvious outliers (see Tables C-1 and C-2, respectively for more information).

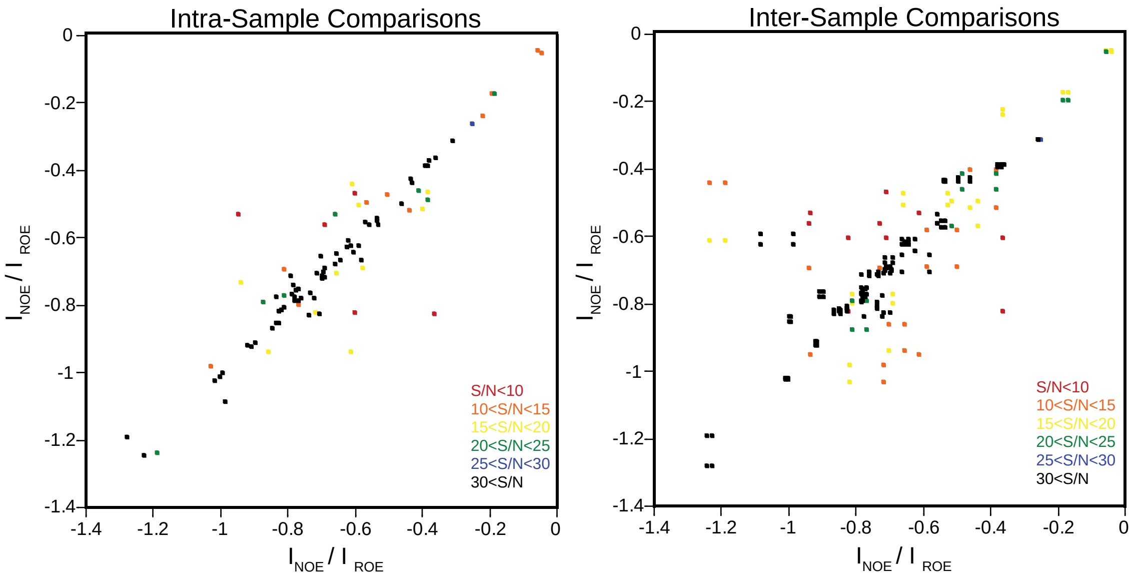

Figure 3-2: The reproducibility of the hydration experiments is dependent on signal to noise. Comparison of the uncorrected NOE/ROE ratios within the same sample **(left)** and between different samples **(right)** demonstrate excellent reproducibility with obvious outliers. The black line indicates the line of best fit. Individual sites are colored according to their minimum S/N as follows: Red= S/N<10, Orange= 10<S/N<15, Yellow=15<S/N<20, Green=20<S/N<25, Blue=25<S/N<30. (c) Shows the individual R2, RMSD, and RMSD% for both intra- and inter-sample comparisons as a function of minimum signal to noise.

The comparison of all intra sample points (Figure 3-2 left) shows a modest R^2 of 0.83 with an <RMSD%> of ~15 percent. However, this is shown to be very dependent on the minimum signal-to-noise ratio of either the NOESY or ROESY peak. As the peaks with the lowest signal to noise is decreased the statistics for reproducibility increase. The R^2 and RMSD both begin to plateau when at values of ~0.97, and ~6% of the measurement when only peaks with a minimum S/N greater than 20 is used. As the minimum S/N is increased the reproducibility does not increase dramatically.

When comparing inter-sample duplicates similar trends are observed. As peaks with the low S/N are removed from the analysis the R^2 and RMSD improve. However, unlike the case of the intra sample comparisons the overall RMSD continues to decrease as peaks with S/N as high as 30 are discarded. This suggests that slight variations in samples might contribute to slightly decreased reproducibility. However, it should be noted that this is true in the case of low S/N cross peaks, and that peaks with higher S/N remain quantitatively reproducible between samples.

It is interesting to note that even in the IST reconstruction that peaks with S/N greater than 8 are reliably re-constructed in their frequency. However, the absolute peak height is less reliably reproduced with peaks with S/N between 8 and 20. It is clear that cross-peaks of low S/N are less reproducible and hence less reliable than those of high S/N. However, these peaks can still be used to detect the presence of water in cases where precise quantification of the rate is not necessary.

Quantitative determination of hydration ratios with mixing time buildup analysis

To determine the hydration ratios for ubiquitin amides ^{15}N-edited NOESY-HSQC and ROESY-HSQC hydration experiments were collected at four different mixing times (20, 40, 60, and 80 ms) on the same sample. These data were then fit to Equation 18 (Appendix C) to obtain the effective relaxation rate ($1/T_{1\rho}$-$1/T_1$:slope) and $\sigma_{NOE}/\sigma_{ROE}$ ratios (intercept). All mix times were performed in the linear regime of the NOE to prevent any contamination from spin-diffusion. The T_1 of the NOE is negligible at the mix times used and therefore the overall relaxation rate will be referred to as just the $T_{1\rho}$, which includes contribution from both protein and water relaxation.

As demonstrated in Figure 3-3a, as the mix time of the experiment increases the $\sigma_{NOE}/\sigma_{ROE}$ gets more negative. When only the protein amide proton $T_{1\rho}$ is used to correct for the auto-relaxation $\sigma_{NOE}/\sigma_{ROE}$ approaches 0 (not shown) the ratio is not constant across all mix times. Clearly, this results in an overcorrection leading to artificially faster hydration measurements. All $T_{1\rho}$ values fit from equation 2 have rates slower than the relaxation from the amide relaxation alone owing the contribution from the slowed relaxation of water. There is no correlation between the fit $T_{1\rho}$ of the amide and water peak and the amide $T_{1\rho}$ alone. Additionally, the calculated $T_{1\rho}$ values vary substantially suggesting that the water protons interacting with the amides have unique relaxation properties.

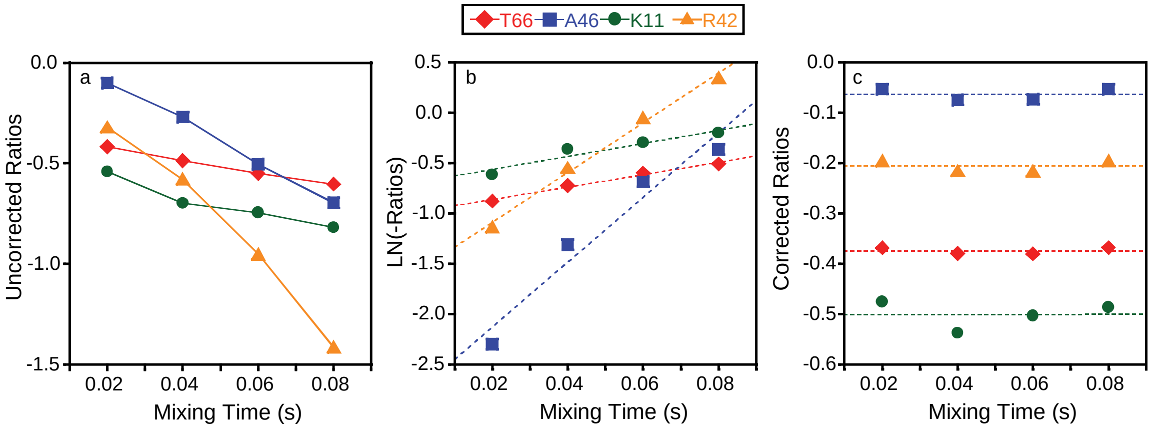

Figure 3-3: The buildup hydration ratios as a function of NOESY and ROESY mixing time (τm) demonstrate the necessity of proper T1ρ correction. **(a)** The uncorrected hydration ratios **(b)** fitting of data to equation 2 to for the effective relaxation time constant ($T_{1\rho}$) and the proper hydration ratio **(c)** Applying the fit $T_{1\rho}$ correction shows nearly constant hydration ratios at all experimental mixing times (20, 40, 60, and 80 ms). The four example residues were chosen to demonstrate the full dynamic range of the hydration ratio.

The quantitative hydration ratios and $T_{1\rho}$ time constants are listed in Table C-3 in Appendix C. The final calculated $\sigma_{NOE}/\sigma_{ROE}$ ratios remain within the theoretical limit (with the exceptions of a few outliers), and the calculated $T_{1\rho}$ values, when used to correct the raw NOE/ROE ratios yield constant $\sigma_{NOE}/\sigma_{ROE}$ ratios across all mix times (Figure 3-3c). Despite the ~10% error possible for each time point measurement, linear regression offers statistics on the precision of the obtained ratios, which would not be possible at a single mix time alone. This, in combination with the longer mix times that can be sampled, can be advantageous in the case of lower signal to noise cross peaks. In general, this fitting method appears to be highly robust and is independent of assumptions of uniform water relaxation.

Mapping of hydration ratios demonstrates clustering of hydration dynamics

In order to identify the significance of the measured hydration dynamics of ubiquitin we mapped the $\sigma_{NOE}/\sigma_{ROE}$ ratios to the protein (Figure 3-4). Over 70% of backbone amides in ubiquitin have detectable protein to water NOE's. The detectable interactions are scattered across the entire protein and includes areas of fast and slow hydration. There is a cluster of slowed hydration along the outer surface of the β-sheet, and along the loop connecting the α-helix and 3_{10} helix. There are also regions of non-detectable protein-water NOE's suggesting very rapid hydration. These regions primarily include the α-helix. The sites that have slowed hydration correlate with the major protein-protein binding interface of ubiquitin. This is very similar as to what was seen previously [29].

Previous studies used a single experimental mix time and corrected for relaxation using the amide proton $T_{1\rho}$ values measured in a separate experiment [29]. We show that both protein and water relaxation contribute to the overall relaxation of the protein-water NOE. Since protein relaxation is much faster than water relaxation this method is in general an acceptable approximation of the overall relaxation, though it may result in over-correction (therefore artificially faster hydration) as shown above. This results in slight quantitative differences in the $\sigma_{NOE}/\sigma_{ROE}$ reported here relative to previously published hydration studies. For example in the 3_{10} helix of ubiquitin previous studies identified it as having intermediate hydration dynamics but using this method exhibits very slow hydration dynamics and very large $T_{1\rho}$ time constants. Therefore, we suggest in studies that require quantitative hydration dynamics that a full series of mix times be collected. In cases where only the detection of water is important a single mix time with an amide proton specific $T_{1\rho}$ is sufficient.

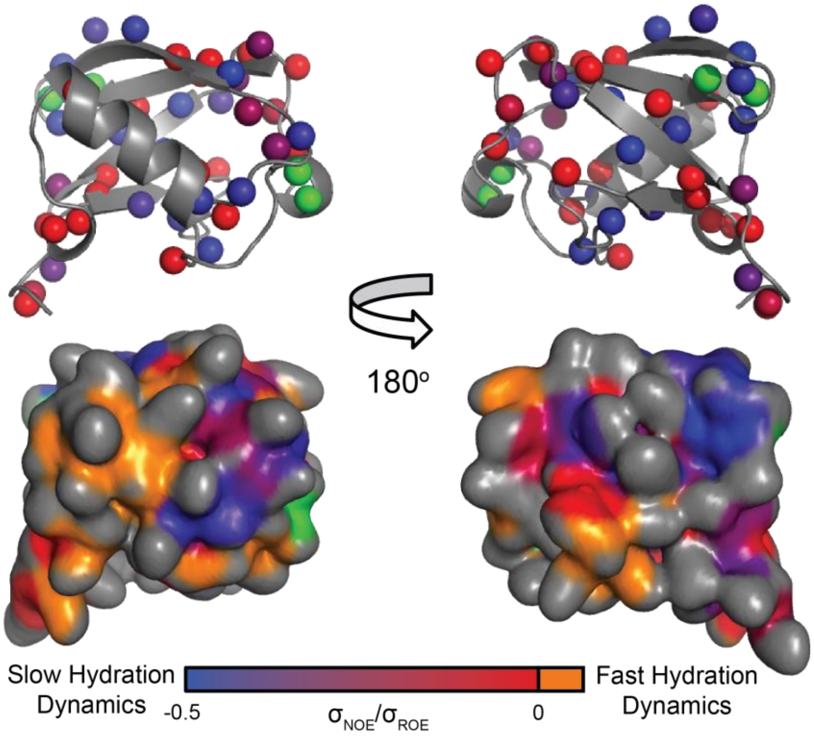

Figure 3-4: The surface of ubiquitin demonstrates a clustering of hydration dynamics. Top: colored spheres plotted on two views of the structure of ubiquitin (PDB entry 1G6J, conformer 25) represent amide hydrogens detected in the suite of buildup hydration experiments. The spheres are color-coded according to the relative hydration dynamics of waters at that probe as determined from the intercepts of the natural log linear regression (Fig 3-3b). A color gradient from blue spheres representing slow waters (high retention time) to red spheres representing fast waters (lower retention time). Green spheres represent residues whose hydration ratios are outside the theoretical limit (more negative than -0.5); most of

these residues are found in areas of high protein flexibility. Bottom: surface of ubiquitin generated with the trisrf and trigen programs [123] as described in the materials and methods. The color scheme used on the points across the generated surface is the same as used above, and in addition, orange represents points on the surface within 4 Å of amide protons that were not detected in the NOESY/ROESY experiments (the waters near these amides were too fast to detect). All protein images were generated with Pymol.

In order to show this, the amide proton $T_{1\rho}$ correction was removed from previously published values [29] and the fitted NOE/ROE $T_{1\rho}$ values as determined from the natural log plots (Fig 3-3b) were used to re-correct the published data (Figure 3-5).

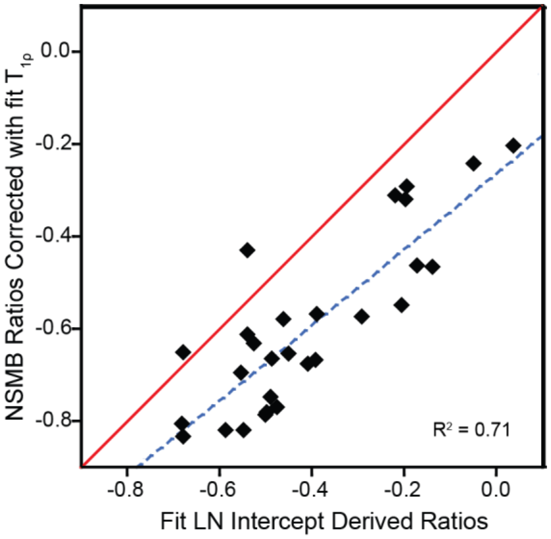

Figure 3-5: Comparison of the hydration ratios determined previously [29] with those determined using the new mixing time buildup method demonstrates a similar distribution of hydration dynamics across the surface of ubiquitin. The red line is the unity line and the blue dotted line is the best fit.

While there are obvious outliers, the data correlate very well. The red identity line is included to demonstrate that the correlation is excellent and that the slope is nearly one; however, there is an obvious offset due to differences in the effective $T_{1\rho}$ of the samples. This is due to differences in the preparation of the protein (for example, the deuterium content of the sample and the temperature at which the experiments were collected in previous publications was higher than that used in this study). However, this owes to the relative reproducibility of the method despite different means of data collection and analysis.

Discussion

Despite the importance of water in biological systems the ability to quantify water interactions has remained technically challenging. In this study we set out to improve the way the protein-water interactions are measured via NMR. It has previously been shown that optimal choices in sample preparation such as for proteins encapsulated in reverse micelles is necessary to remove experimental artifacts [29]. Here, we extend those methods to improve data collection and analysis in order to obtain reproducible and quantitative $\sigma_{NOE}/\sigma_{ROE}$ rates.

We implement sinusoidally-weighted Poisson-Gap NUS[119] schedules, and IST reconstruction to 3D ^{15}N-edited NOESY and ROESY spectra [120]. The IST method for reconstructing non-uniformly sampled data belongs to a larger family of compressed sensing (CS) techniques. One of the major limitations of CS techniques is the principle of transform sparsity [121]. In general, as the sparseness of the measured signal is decreased (i.e. more peaks in an NMR spectrum), the amount of points collected is decreased, or the sampling density is decreased the less reliable the reconstruction. Conventional 3-dimensional spectra (such as the HNCO) which are spectrally sparse are often recorded with sampling densities as low as ~10%. Unfortunately, NOESY spectra have a much larger number of peaks with a high dynamic range of peak heights. This, in combination with the low signal-to-noise of the inter-molecular NOE presents a challenge for reconstruction of NUS data [116, 117, 124].

It has been shown that in traditional NMR experiments optimal S/N is obtained when indirect dimensions are sampled to 1.26*T_2, whereas optimal resolution is present at 3.14*T_2 [125]. However, it has also been shown that with NUS demonstrates benefits from increased S/N and resolution well past 1.26*T_2 [126]. Additionally, reconstruction precision seems to be dependent on the overall number of points collected with low NUS sampling densities being more highly tolerated for high resolution data sets. There is no consensus of optimal sampling density and resolution [116, 127].

In the case of protein hydration measurements the rapid nature of water dynamics makes cross-peak intensities generally quite weak relative to the diagonal [115]. In contrast to standard NOESY experiments perdeuterated protein is necessary for hydration experiments resulting in decreased resolution requirements in the indirect proton dimension [29]. The necessary resolution in the indirect heteronuclear dimensions (^{15}N or ^{13}C) can be viewed from standard HSQC experiments. Here we recommend reducing resolution in the indirect proton dimension while collecting spectra with high number of transients/FID and maintaining adequate resolution to resolve all cross-peaks in the heteronuclear dimension.

In order to simulate an absolute worst case scenario (as may be encountered with large spectrally overlapped proteins) using these criteria we show that a sampling density of 25% results in accurate peak height reproducibility and an RMSD less than 5%. This sampling density is similar to the requirements of a standard 2D experiment [117, 127], an intuitive finding since the water-plane of a 3D NOESY-HSQC largely mimics its HSQC [29]. This conservative method does not require optimizing sampling schedules, and is tolerant to spectral overlap. Decreases in RMSD would be present for better resolved or higher resolution data collection.

In the case of hydration it is highly recommended to use NUS to increase the number of transients per FID rather than to save experimental time because increasing the number of transients improves reproducibility [116]. The implementation of NUS also allowed us to comprehensively test the reproducibility within and between samples. While thermodynamically stable, reverse micelles are spontaneously forming fluid assemblies so any slight difference in the ratio of the reverse micelle components could, in theory, affect the hydration dynamics [68, 128]. In general we show that the ratio of the $\sigma_{NOE}/\sigma_{ROE}$, which is the basis of the hydration measurements, is highly reproducible given high signal to noise both within a sample and between samples. Additionally, we show that the sites with the highest percent error of the $\sigma_{NOE}/\sigma_{ROE}$ ratio are the sites that have lowest signal to noise.

We highly recommend that a duplicate of one mix time is collected for each protein measured as a reference to determine the minimal S/N necessary for reproducibility. The conventional measure of S/N ratio in NMR spectroscopy is the peak height divided by the RMS noise of a peak-less spectral region. While this assumption holds true for fully sampled data, NUS noise is non-Gaussian and non-uniformly distributed. Several metrics for measuring S/N in NUS sampled data have been discussed, and is beyond the scope of the paper. Here, we use standard calculation of RMS noise to define our S/N cutoff. This is used as a metric to determine the S/N cutoff for reproducibility the $\sigma_{NOE}/\sigma_{ROE}$ in our hands. This should allow for a broadly applicable metric for S/N despite variations in pulse sequence, NUS sampling schedules, reconstruction methods, or S/N calculation.

Conclusions

Previous studies have measured the NOE/ROE at a fixed mixing time [23, 24, 29, 113, 129-134]. The time savings due to non-uniform sampling allowed us to collect four sets of NOESY and ROESY experiments with varied mix times. This allowed us to perform a full buildup series and fit for the desired $\sigma_{NOE}/\sigma_{ROE}$ and relaxation rates for ubiquitin to demonstrate clustering of hydration dynamics. It has previously been assumed that the amide-proton $T_{1\rho}$ is the fastest relaxing term in the hydration measurements and can therefore be used to correct for the autorelaxation in the $\sigma_{NOE}/\sigma_{ROE}$ [29]. However, we here we show that the relaxation behavior is a complex mixture of water and protein relaxation in both the laboratory and rotating frames. Performing a full buildup is the only way to obtain the coefficient of the auto-relaxation as there is no clear correlation between amide proton $T_{1\rho}$ and the $\sigma_{NOE}/\sigma_{ROE}$ relaxation term. Furthermore, linear regression of the buildup further increases the precision of the measurement. However, if only qualitative differences in hydration are required we show that a single mix point is sufficient to adequately describe hydration trends (Figure 3-4).

In general the methods presented here offer an experimentally tractable way of measuring protein hydration in a quantitative fashion via the NOE. The implementation of non-uniform sampling to 3D NOESY-HSQC and ROESY-HSQC spectra is sufficient to reduce experimental time and allow for increased signal-to-noise and greatly improved data fitting methods. The method simply requires collection of one additional spectrum at a single mix point to determine the S/N cutoff required to reproducibility. This makes the method broadly applicable so other implementations of NUS sampling and reconstruction may be used.

CHAPTER 4: Hydration dynamics of hen egg-white lysozyme

Abstract

Water, the universal solvent, is of utmost importance for virtually all biological processes, especially in its interaction with proteins and other macromolecules. By encapsulating proteins in reverse micelles, it is possible to use NMR nuclear Overhauser effect (NOE) spectroscopy in order to quantify interactions between protein surfaces and water molecules in a site-specific manner. This approach was used to investigate the surface hydration of hen egg-white lysozyme, particularly within the hydrophilic peptidoglycan binding cleft. The waters within the binding cleft of the free protein are generally fast while interfacial waters between a protein and a bound chitotriose inhibitor are relatively slow. The balance of the surface of the protein is characterized by a heterogeneous distribution of hydration dynamics. Similarly, relatively dynamic waters are also observed within a partially hydrophobic core more than 6 Å away from the surface of the protein. Finally, remote locations on the surface of the protein show an increase in hydration dynamics upon inhibitor binding.

Introduction

It has been appreciated for some time that water is critical to biological processes ranging from protein folding to enzymatic activity [111, 112]. Despite this central importance, the interaction of liquid water with protein molecules has been notoriously difficult to characterize experimentally in site-specific detail. A number of advanced spectroscopic and crystallographic strategies have been employed [29, 135-138]. Generally, it is found that the interaction of water with the surface of proteins is quite variable with respect to residence time and local dynamics. Interestingly, a recently introduced solution NMR approach, based on the favorable properties of proteins encapsulated within the protective water core of a reverse micelle, suggested a clustering of dynamics of hydration water on the surface of ubiquitin. The clustering of "slow" hydration water was loosely correlated with the surface of the protein that is involved with formation of a dry interfaces with protein binding partners [29]. This suggested that ubiquitin has evolved to maximize the hydrophobic effect for surfaces involved in molecular recognition. The generality of this observation remains to be explored.

In the recent past, the introduction of a variety of new infrared [138], fluorescence [13, 136], and EPR [137] strategies based on introduction of probes via mutations have provided valuable insights about protein hydration. Unfortunately, these methods require introduction of probes across the surface of the protein in order to achieve a comprehensive view of hydration. Molecular dynamics simulations have also been used extensively to study protein hydration but remain uncertain due to the lack of confirmatory experimental support. Importantly, additional experimental tests of emerging theories of protein-water solvation are clearly required [1, 139].

Solution NMR has been used in the past to study the interaction between protein molecules and water [26]. Approaches based on the nuclear Overhauser effect (NOE) spectroscopy in the laboratory and rotating (ROE) frame were proposed to allow for characterization of the dynamical aspects of protein hydration [23]. Though particularly useful for detection of internal "structural" water, use of the NOE for characterization of the dynamical aspects of surface hydration is constrained by a number of potential artifacts and limitations [8]. Although the one to two layers of water immediately surrounding the protein are somewhat slowed relative to bulk solvent [11], the majority of these waters are still too fast to be detected with NOE spectroscopy. Significant artifacts and limitations can arise from the action of hydrogen-mediated exchange of magnetization between the protein and solvent water [8, 23, 26, 29]. It is also noted that the exquisite localization of the NOE through the $1/r^6$ dependence associated with the intramolecular NOE is potentially greatly relaxed for intermolecular interactions [27, 140]. Fortunately, the use of reverse micelle encapsulation has been shown to largely overcome these issues [29] by significantly slowing the water comprising the hydration shell [54], quenching hydrogen exchange chemistry [29, 128] and eliminating most of the bulk water of free aqueous solutions [57].

Small, largely spherical reverse micelles (RMs) are formed in the presence of appropriate amphipathic surfactant molecules, bulk organic solvent such as the liquid alkanes under limiting water conditions [60]. With careful optimization, protein molecules can be encapsulated with high structural fidelity with their native hydration layer maintained inside of the protective core of the reverse micelle particle [60]. Here we use reverse micelle encapsulation to enable detailed study of the dynamics of hydration of hen egg-white lysozyme (HEWL), which is a relatively small (14.4 kD) protein that binds and cleaves small peptidoglycans in bacterial cell walls [141].

The average surface hydration dynamics of lysozyme have been studied extensively in the past using techniques such as terahertz (THz) spectroscopy [142], Fourier transform infrared (FTIR) spectroscopy [143], neutron scattering [144, 145], and dielectric relaxation spectroscopy [146]. These studies were designed to examine many aspects of hydration dynamics at the surface of HEWL including experiments designed to determine the effect of solvent molecules on hydration dynamics [142, 145], those designed to examine different types of surface water motions [144, 146], and those designed to study the effect of protein surfaces on hydrogen bond networks [143]. There are also some studies designed to examine relatively slow waters within internal hydrophilic pockets [147] and within the hydrated interfacial space between HEWL and an inhibitor [148] using ^{17}O-detected nuclear magnetic dispersion and triple quantum filtered NMR, respectively. None of these studies, however, present any site-specific information: all conclusions are drawn based on the average hydration dynamics within the hydration layer of HEWL (along with some that study internal pockets).

As mentioned above, it is notoriously difficult to experimentally examine protein hydration dynamics at the surface of proteins in a site-specific manor. Two-dimensional infrared (2D-IR) spectroscopy has been employed to study the surface hydration dynamics of lysozyme [138, 149]. These experiments require the covalent modification of histidine residues with ruthenium-derivatives. Molecular dynamics simulations have been used extensively to study the hydration dynamics of HEWL but have not reached a consensus and would benefit from more experimental guidance [14-16, 69]. NMR spectroscopy in conjunction with reverse micelle encapsulation is perhaps the only experimental method that can site-specifically measure hydration dynamics across the entire surface of a protein [29]. In this study, we examine the hydration dynamics of HEWL in its free and inhibitor-bound states.

Results and Discussion

Reverse micelle optimization

Lysozyme was initially encapsulated in a mixture of 75 mM cetyltrimethylammonium bromide (CTAB) and 450 mM hexanol at a pH of 4.7 and a final protein concentration of approximately 60 µM in deuterated pentane. This reverse micelle mixture had the highest protein encapsulation efficiency of all surfactant mixtures tested, maintained the structural fidelity of the encapsulated protein (Fig. D-1) showing no interactions between the protein and the reverse micelle shell, and was stable for several months. Unfortunately, under these conditions the water and hexanol hydroxyl ^{1}H resonances are unresolved (see Fig. D-2a). Separation of these resonances is required for unambiguous assignment of NOEs to water. Their separation is quite sensitive to the rate of exchange between water and the hydroxyl group of the hexanol. The rate of hydrogen exchange can be dependent on many factors including temperature and pH. Adjusting the effective pH of the reverse micelle water pool [128] to 5.4 resulted in sufficient resolution (~0.2 ppm) of the water and hexanol hydroxyl resonances (Fig. D-2b).

HEWL hydration

Hydration dynamics are characterized by the ratio of the NOE and ROE. While it was previously concluded that only $T_{1\rho}$ relaxation of the ROE needed to be corrected for the hydration ratio to be quantitative [29], it has been shown recently that this is insufficient (Chapter 3). The decay of the NOE/ROE ratio is not only dependent on the $T_{1\rho}$ of the amide protons, but on a combination of the T_1 and $T_{1\rho}$ relaxation of both the amide and water protons. Since it is not possible to determine relaxation properties of site-specific waters, the only way to detect quantitative hydration ratios is the run a linear buildup: the NOESY and ROESY spectra must be collected at a buildup of mixing times and the hydration ratios are then extrapolated as a function of mixing time. This requires significantly more experimental time, even with non-uniform sampling. While the HEWL reverse micelle sample is relatively stable (~1-2 months), the encapsulation efficiency is low enough such that it would take an inordinate amount of time to collect a full linear buildup. In effect, all the NOE/ROE ratios reported here are simply the ratio of the NOE and ROE peak intensities (Figure 4-1) at one mixing time of 35 ms with no linear buildup and no correction from relaxation. While this leads to ratios that are not quantitatively correct and outside the bound limit ratio of -0.5, the ratios are still qualitatively correct relative to other ratios in the same protein system assuming identical sample and experimental conditions.

In collecting these hydration experiments one is often concerned about incorrect assignments of the amide-water NOEs to NOEs between amide protons and aliphatic protons with chemical shifts near the water resonance. There are several approaches to evading this issue. The most direct is to simply perdeuterate the protein which eliminates the aliphatic protons while reducing the contributions of dipolar relaxation by surrounding hydrogens to the relaxation of the amide hydrogen. The amide sites are back-exchanged to hydrogen during the refolding process. Deuteration significantly improves the performance of the sample and allows for detection of several ROEs that were otherwise unobservable.

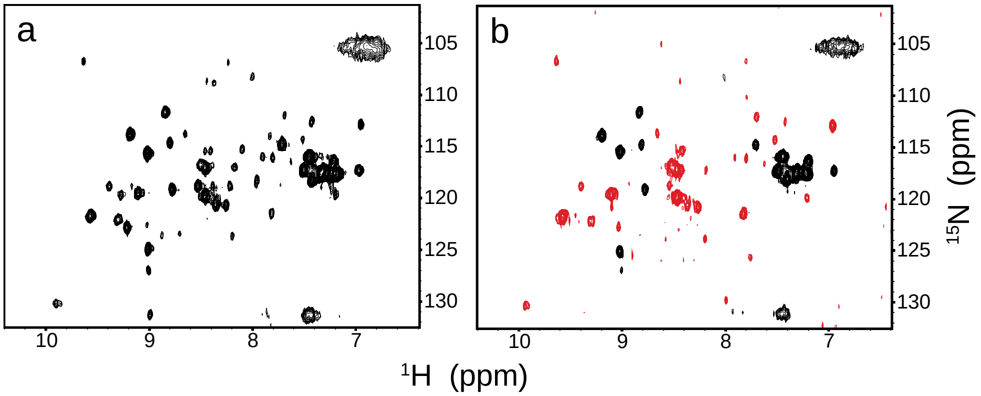

Figure 4-1: Detection of NOEs between water and encapsulated hen egg-white lysozyme. Aqueous ^{15}N-labeled, 95% deuterated HEWL was encapsulated in the described CTAB/hexanol mixture at pH 5.4 in pentane. ^{15}N-resolved NOESY **(a)** and ROESY **(b)** spectra are shown in the plane of the indirect water resonance (4.5 p.p.m.) for the encapsulated protein demonstrating clear detection of NOEs from backbone amides to waters in the hydration shell. In all cases black peaks represent positive intensity and red peaks represent negative intensity. Black, positive peaks in the ROESY **(b)** spectra represent contamination of some amide probes and side chains by exchange.

Due to the nature of the reconstruction of non-uniformly sampled data, there are inherent concerns with data precision. It has been shown (Chapter 3) that the hydration dynamics represented by the NOE/ROE ratio is reproducible to within 5% if the signal to noise (S/N) of the peaks of *both* the NOE and the ROE are at or above a value of ~18. In effect, the amide probes that did not meet this criteria were assigned a binned NOE/ROE ratio (represented by a “+” symbol in Table D-1). The binned ratio was set to -0.05 if the S/N of the ROE was less than 18 and there was no NOE peak. The binned ratio was set to -1.0 if the S/N or *either* the NOE or the ROE was below 18.

Distribution of hydration dynamics across the surface of apo-HEWL

In order to analyze the hydration dynamics across the surface of the protein, the hydration dynamics ratios (Table D-1) were calculated with the NOE and ROE peak intensities. At the slow tumbling limit of the reverse micelle encapsulated HEWL, an uncorrected hydration ratio of ~ -2.0 indicates a water molecule with a long retention time [23, 26] at the surface of the protein with a correlation time on the order of or slower than the molecular tumbling time of the entire protein encapsulated within the reverse micelle particle (approximately 12 ns). The upper limit of the ratio is zero and is indicative of a water molecule with an effective correlation time of approximately 300 ps at the experimental field strength of (^{1}H) 500 MHz [26] although it is difficult to pinpoint this fast time scale due to the convolution of the internuclear distance and angle inherent in the intermolecular protein-water NOE. All ratios in between represent some range of hydration dynamics between these two effective correlation times [25, 26]; all amides at the surface of the protein that are not detected at all are interpreted as having too short of a retention time to be detected by our method. These ratios were assigned to the detected amide hydrogens and plotted as spheres on the structure of lysozyme from PDB entry 3LYZ (see Fig. 4-2 top). This representation shows a wide range of hydration dynamics across the structure of HEWL with no correlation to features such as side chain type as noted previously with ubiquitin [29]. However, with this representation it is difficult to understand the behavior of the water at the immediate surface of the protein.

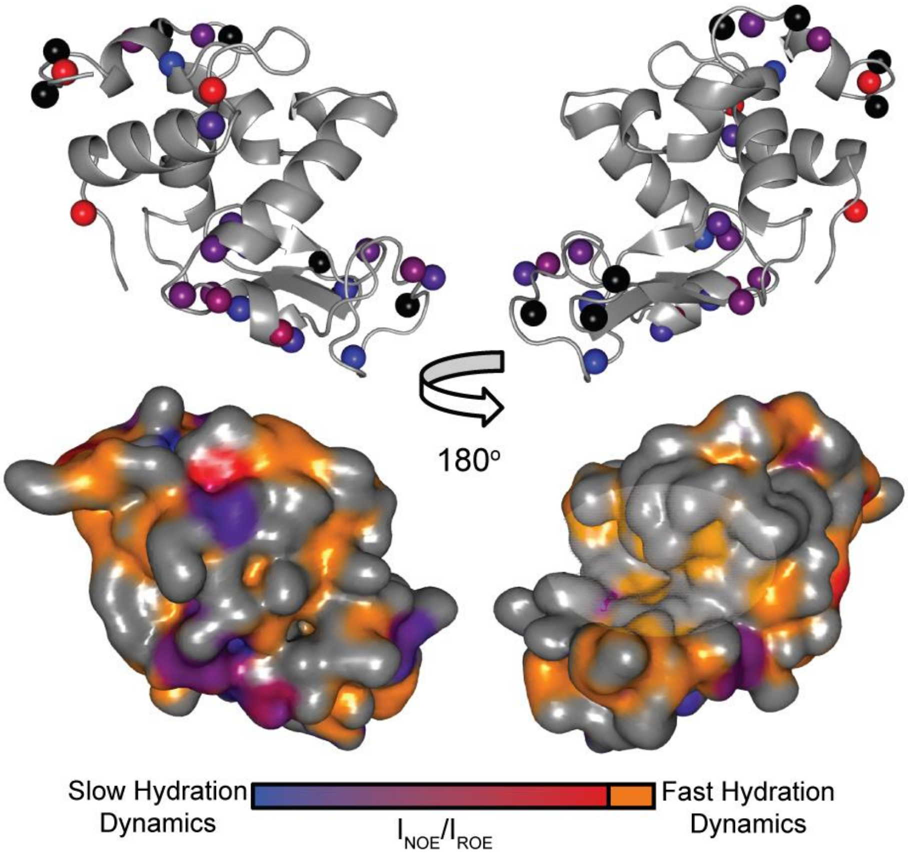

Figure 4-2: Distribution of hydration dynamics across the surface of HEWL. Top: colored spheres plotted on two views of the structure of HEWL (PDB entry 3LYZ) represent amide hydrogens detected in the suite of hydration experiments. The spheres are color coded according to the relative hydration dynamics of waters at that probe from blue spheres representing slow waters (high retention time) to red spheres representing fast waters (lower retention time). The scales ranges from ~4 ns to ~100 ps. Bottom: surface of HEWL generated with the Travel Depth program [150, 151] as described in the Methods (Appendix D). The color scheme used on the Cartesian points across the generated surface is the same as used above, and in addition, orange represents points on the surface within 4 Å of amides that were not detected in the NOESY/ROESY experiments (the waters near these amides were too fast to detect) and black represents amides that were contaminated by exchange (detected by positive ROEs). There is a wide distribution of hydration dynamics across the surface of HEWL and within the binding cleft (represented by the textured oval on the bottom right figure) there is only a very small region with significantly slowed waters. All protein images were generated with Pymol.

In order to better visualize the activity of the water directly at the surface of the protein, a surface of the protein with points in Cartesian space was generated as described in the materials and methods. If points on the surface within 4 Å of one or more amide probes, it would adopt the average NOE/ROE ratio and would be colored accordingly (Fig. 4-2 Bottom). Points outside the distance cutoff of any amides were not assigned a NOE/ROE value and colored gray. Points that were within a 1.2 Å cutoff of amide probes with positive ROEs were considered to be contaminated by exchange and colored black. It is clear there is a wide distribution of hydration dynamics across the surface of HEWL, but there are still many areas of the surface near amide probes at which waters are too fast to detect (orange in Fig 4-2).

Mostly fast waters within the binding cleft of apo-HEWL

It is clear that the majority of the slower hydration dynamics is dispersed across the surface of HEWL while within the binding cleft (textured surface area in Fig 4-32) the retained waters are either relatively fast (mostly red in color) or so fast they are undetectable (orange). This is in direct contrast to previous findings [29] which correlated protein-protein contact areas with areas on the surface of ubiquitin with slow waters. However, this is not a completely complementary comparison since the protein contact surfaces previously studied were relatively large with dry interfaces. On the other hand, the HEWL binding cleft is much smaller (only large enough to fit small peptidoglycans, approximately 1000 cubic Ångstroms) and much more hydrophilic. In fact, of the 19 water-containing structures of HEWL with a bound peptidoglycan inhibitor deposited in the PDB, 17 contain interfacial waters between the inhibitor and the protein surface (see below). The two structures that don’t (PDB entries 4DDA and 3TXJ) contain relatively few crystallographic waters in their structures. This may indicate that binding of peptidoglycans in the binding cleft (and possibly other “wet” interfaces) may need less of an entropic gain from desolvation. Further study of hydrophilic interfaces is needed in order to resolve this issue.

Comparison of HEWL hydration dynamics with other methods

Because lysozyme is such a well-studied model system, there have been many methods used to try and determine the hydration dynamics of HEWL at the surface of the protein with molecular dynamics (MD) simulations being the most prevalent. Interestingly, many of these MD simulations have led to contrary results both with each other and the data presented here. Two such studies [14, 16] examine the jump time (or the time needed to supplant one hydrogen-bond acceptor with another) of single water molecules across the surface of a solvated HEWL molecule. Here they found that the only relative retardation of the reorientation times of surface waters was found at the surface of the protein within a small region of the enzymatic binding cleft. It is proposed that the retardation of the hydration dynamics in the binding cleft is due to the abundance of hydrogen bond acceptors and its concave topology. They find little to no distribution of slower hydration dynamics across the rest of the surface. This is in contrast to the data presented above in which we find a distribution of hydration dynamics across the surface of the protein with very little slow dynamics within the binding cleft. Figures D-3 and D-4 in the Appendix Ddemonstrate virtually no correlation between the hydration dynamics measured for apo HEWL in this study and the simulated water reorientation times. This may be due to the fact that this simulation only focuses on the rotational reorientation of the water molecules while the intermolecular NOE is inherently dependent on both the rotational and translational dynamics.

There have also been studies using two-dimensional infrared (2D-IR) spectroscopy that have been used to obtain site-resolved hydration dynamics information for lysozyme [138, 149]. This technique covalently attaches surface exposed histidine residues to a ruthenium-carbonyl (RC) complex whose vibrational relaxation properties are dependent on the isotopic composition of the surrounding water [138]. The ruthenium-carbonyl probe was attached to the only surface-exposed histidine on HEWL, H15, which is in a well-structured area of the protein. They find the waters near the probe are very slow relative to bulk solution as opposed to the detected waters on a disordered loop region of a different protein which are more bulk-like [138]. Interestingly, this area of the protein is also one in which we detect a patch of slow waters (Fig. 4-2). In fact, slow hydration dynamics are detected at many of the surface-exposed amide protons within 10 Å of the ruthenium atom in the HEWL-RC complex (PDB entry 2XJW) crystal structure (Figure D-5 in the Appendix D). Relatively slow hydration dynamics is also detected in this area in a molecular dynamics simulation that focuses on residence times of water molecules near HEWL [15] as opposed to just the rotational dynamics of the water as described above. In this study, they note that the simulated waters with the longest residence times reside within hydrophilic pores leading to an internal hydrophilic cavity. The “gate” to one of these pores is near residue A90, which is within the area near H15 where slow hydration dynamics are detected using both NMR and 2D-IR. The hydrophilic cavity is the same one described below composed of residues 53-58. Also, this simulation detects high residence time water at the other end of the pore near surface residues T40, L84, and S86 all of which have detectable NOE/ROE ratios or are within 4 Å of amides with detectable hydration ratios. There is a prominent correlation between our measured hydration dynamics and both of these described studies using two completely different methodologies. It is clear that hydration dynamics measured by NMR more closely correlate with hydration dynamics measured or simulated with a contribution from translational water dynamics.

Detection of internal waters

Because hen egg-white lysozyme has been studied extensively since the 1960s using many methods, there are a large number of structures of the protein deposited in the PDB. Many of these crystal structures include crystallographic waters when the structure is solved. In fact, every crystal structure of HEWL that also includes crystallographic water molecules also contains waters within a partially hydrophobic pocket on the inside of the protein. These waters have been extensively studied [152] and have shown conservation in lysozymes from many other species [153]. Solution structures have also been solved for lysozyme [154] and internal water molecules have even been detected in other fully hydrophobic pockets using a similar method in which internal waters were detected via the NOE and occupancies were determined by filling the cavities with various gases and comparing relative peak intensities [26] in aqueous solution. This did not require the use of reverse micelle encapsulation although the reliability of the measurements is suspect due to hydrogen exchange-related effects in aqueous solution. With such a vast amount of evidence for internal waters within partially hydrophobic pockets in HEWL, it should lead to internal amide probes showing structural waters away from the surface of the protein.

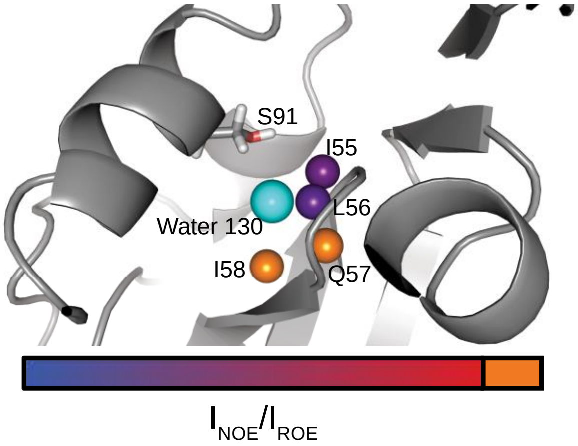

Figure 4-3: Waters detected within an internal, partially hydrophobic core. The cyan sphere is a crystallographic water that was deposited with the structure of HEWL (PBD entry 3LYZ) and is greater than 6 Å away from the protein surface. The orange colored spheres are amide protons that are far away from the surface, within 4 Å of the crystallographic water but no hydration dynamics was detected: I58 and Q57 which are 3.1 Å and 3.2 Å away from the crystallographic water, respectively. The remaining colored spheres (I55 and L56 which are 2.7 Å and 2.1 Å away from the crystallographic water, respectively) represent buried amide probes that are within 4 Å of the crystallographic water where hydration dynamics are detected hydrogens detected with a color scheme as described in Fig. 2. The only hydroxyl moiety within 4 Å of the crystallographic water is S91; however, this resides within the detection limit of only one of the internal amide probes (I55).

As shown in Fig. 4-3, there are internal amide probes (I55 and L56) that show NOEs to water. Each of these probes has a burial depth of at least 6 Å meaning that these waters can in no way be directly exposed to waters at the surface of the protein. Both of the probes are also within 4 Å of the published crystallographic water (PDB entry 3LYZ) as displayed in Figure 4-3. This structure is also just one of well over one hundred crystal structures of HEWL that include waters at multiple positions within this hydrophilic core. Although there has reportedly been evidence for multiple internal hydrophobic pockets filled with water [26], our experiments only show evidence of the one shown in Fig. 4-3. This specific internal pocket is conserved throughout multiple species of lysozyme [153]. Interestingly, the retention time of the internal waters being detected by these amide probes is not at the "bound limit" (relatively fast hydration dynamics) similar to what has been seen before [15, 26]. This indicates that either there are multiple internal waters exchanging with each other rapidly within the core, or water molecules are exchanging in and out of the core at a rapid rate. There are also two amide probes (I58 and Q57) within 4 Å of the crystallographic water but have no detected NOE/ROE ratio. This may indicate that the internal water is exchanging in and out of the core, but once inside gets coordinated by the I55 and L56 amides along with the S91 hydroxyl (the only other hydrophilic moiety in the pocket) before moving out of the pocket again. This way the water remains far enough away from the I58 and Q57 residues such that no NOEs are detected to water. These internal waters are still detected within the core of the chitotriose-bound complex at approximately the same rate (Table D-1 in Appendix D).

Hydration dynamics of inhibitor-bound HEWL and detection of interfacial waters

In order to observe the effect of an inhibitor binding on the hydration dynamics across the surface of HEWL, the suite of hydration dynamics experiments was collected on HEWL bound to chitotriose. Chitotriose is made up of three N-acetly-D-glucosamine (GlcNAc) residues that bind in half of the binding cleft of HEWL with a K_d of approximately 14 µM [155, 156]. Because HEWL has a relatively high affinity for chitotriose and can only cleave peptidoglycans between the third and fourth residues of a GlcNAc polymer, chitotriose makes for a potent competitive inhibitor [157].

To make a direct comparison between the apo- and inhibitor bound states, a mixture of HEWL and 1.2 molar equivalents of chitotriose was encapsulated in reverse micelles of the exact same composition as described above for the apo state. There were only a few chemical shift perturbations (see Fig. D-6b Appendix D) upon binding of chitotriose, as expected [158]. In this mixture, almost all of the amide residues that had a NOE to water had a corresponding ROE (Fig. D-6 c and d). The NOE/ROE ratios tabulated for the chitotriose-bound state (Table D-1) were mapped to the surface of HEWL as described above for the apo state (see Fig. D-7 in Appendix D). As described in a previous section, probes with S/N of less than 18 for either the NOE or the ROE were assigned a binned value.

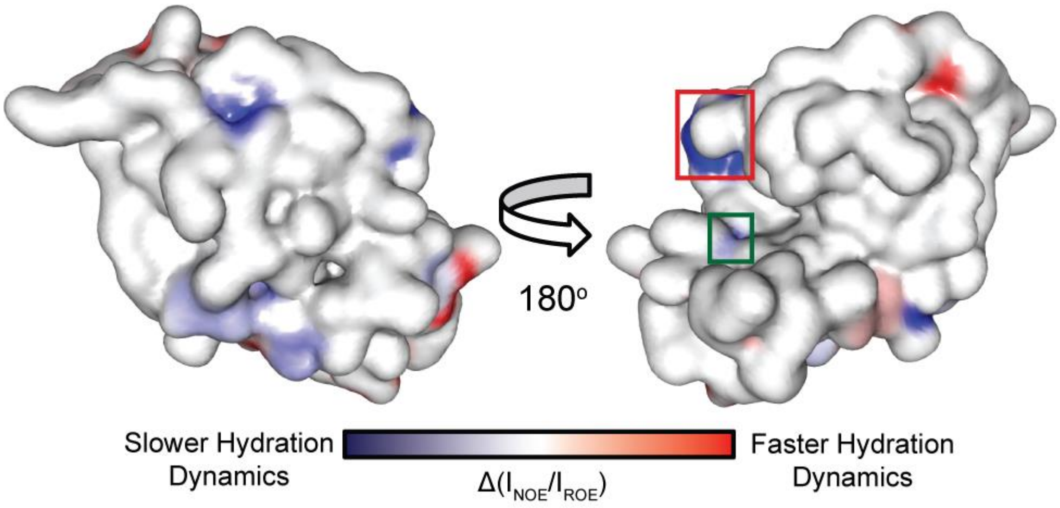

Figure 4-4: The difference in hydration dynamics between apo- and chitotriose-bound HEWL shows interfacial waters between HEWL and the inhibitor. The $\Delta(\sigma_{NOE}/\sigma_{ROE})$ represents the difference of the NOE/ROE ratio of the inhibitor-bound state and that of the apo state at every point on the rendered hydration surfaces. These differences were mapped to the solvent accessible surface of HEWL (PDB entry 3LYZ) on a blue-white-red spectrum. Blue represents points on the surface at which the water became less dynamic upon binding of inhibitor, red represents points on the surface at which the water became more dynamic upon binding of inhibitor, and white represents no change. The section boxed in green shows slowed water where there is no published crystallographic water. The section boxed in red (see Figure 4-5) represents an area where an interfacial water is

included in the crystal structure of chitotriose-bound HEWL [159]. Sections in black represent amides that were contaminated by exchange in the inhibitor-bound state and not the apo state.

The hydration surfaces for the inhibitor-bound state (Fig. D-7) and the apo state (Fig. 4-2) are very similar. To better visualize the changes in hydration dynamics between the two states, differences were taken of the NOE/ROE ratios at every point on the Cartesian Van der Waals surfaces and mapped on the same surface (Fig. 4-4). The hydration dynamics at the surface of HEWL bound to chitotriose is very similar to those at the surface of apo HEWL. Interestingly, there are two areas at the inhibitor-binding interface (red and green boxes in Fig. 4-4) which become much slower upon binding of inhibitor. Interfacial water dynamics for HEWL have been detected before [148], but there is no site-specific resolution. The area boxed in red represents an area in the HEWL-chitotriose crystal structure (PDB entry 1HEW) where a crystallographic water resides between the surface of the protein and the inhibitor (see Fig. D-8). In the chitotriose-bound hydration data, the amides corresponding to this area (Fig 4-5) have NOE/ROE ratios that are essentially bound to the protein which is much slower hydration dynamics than those found for the probes near the internal core waters described above similar to what has been seen before [148]. The interfacial water is seemingly trapped in between the protein and the ligand and is potentially crucial to the integrity or the binding interaction. As described above, all but two of the water-containing crystal structures of inhibitor-bound HEWL contain interfacial water in this area, indicating this must be a crucial interaction.

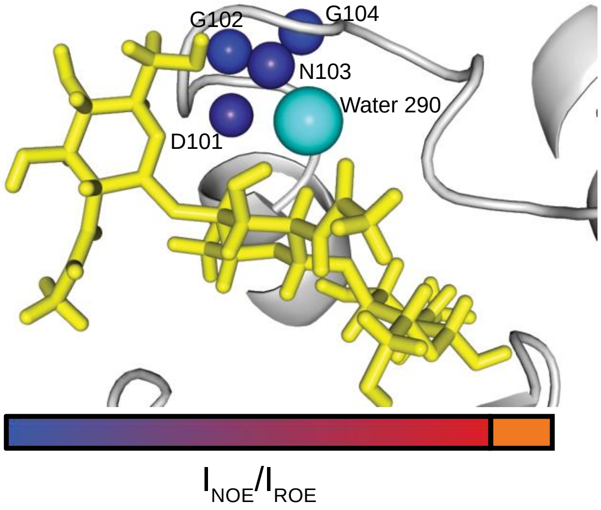

Figure 4-5: Trapped interfacial waters are detected between the inhibitor and the protein surface. The section boxed in red (see Figure 4) is a magnification of the cartoon representation of chitotriose-bound HEWL (PDB entry 1HEW) where an interfacial water is included [160]. Unlike for the water in the internal pocket (Figure 4-3), hydration dynamics is detected for all amide probes within 4 Å of the crystallographic water (cyan sphere). The hydration dynamics color scheme is the same as described in Figure 4-3 with the four amide probes (D101, G102, N103, and G104) being shown as spheres and are 3.6 Å, 3.7 Å, 2.6 Å, and 3.1 Å away from the crystallographic water, respectively. Only one of the amides (D101) is within 4 Å of potential signal contamination from hydroxyl moieties on the chitotriose (yellow stick representation).

On the other hand, the area boxed in green does not have a corresponding crystallographic water in the published structure (Fig. D-8). However, the amide residue in that area that detects slower hydration dynamics upon chitotriose binding is Trp62. Not only is the indole ring of Trp62 essential in the binding of peptidoglycans through hydrophobic ring stacking [159], but it has also been shown that the amide of Trp62 is involved in a water-mediated hydrogen bond with the OH-6 of the third GlcNAc residue of the chitotriose which is an essential interaction for enzymatic activity [158, 160]. To our knowledge, this is the first time NMR has been used to detect site-specific hydration dynamics of interfacial water between a protein and ligand (hydrogen bond-mediated or otherwise).

There are also areas on the difference surface colored in red that indicate an increase in hydration dynamics upon chitotriose binding. All of these areas are at remote locations on the surface of the protein, on the opposite face of the binding cleft and should have no direct interaction with waters in the binding pocket. These remote waters are seemingly compensating for the retardation of hydration dynamics within the binding pocket by increasing the hydration dynamics at other areas of the protein. To our knowledge, the observation that hydration shell waters at remote sites away from the binding interface becoming more dynamic upon ligand binding is a novel finding.

Conclusions

In this study we show the utility of reverse micelle encapsulation in measuring hydration dynamics at the surface of a protein with site-specific resolution. A wide range of hydration dynamics across the surface of the protein is observed with no detectable clustering of dynamics according to factors like residue type or protein dynamics. Unlike what has been previously shown with the large, hydrophobic binding interfaces of ubiquitin [29], the waters in the small, hydrophilic interface of HEWL are relatively fast. This suggests that hydrophilic binders may need to overcome less of an entropic penalty since less of the ligand needs to be desolvated. We also observe hydration dynamics of internal waters within an internal partially hydrophobic core that show intermediate hydration dynamics suggesting rapid movement of water in and out of the pocket or rapid movement of water within the pocket. Also, for the first time with NMR, we detect site-specific hydration dynamics for interfacial waters between the surface of HEWL and the chitotriose inhibitor which are much slower dynamically than the observed waters within the internal pocket. Finally, we find that, upon binding of inhibitor, remote areas of the protein show an increase in relative rate of hydration dynamics, suggesting these areas are compensating for the slow waters being trapped by the inhibitor.

Chapter 5: Protein conformational entropy is not slaved to solvent

Much of the work presented in this chapter was done in collaboration with Dr. Nathaniel V. Nucci, a former postdoctoral researcher in the Wand lab and with Dr. Matthew A. Stetz, a current postdoctoral researcher in the Wand lab.

Abstract

It is well established that the influence of water is a fundamental determinant of the structure and thermodynamic character of globular protein molecules, particularly through the hydrophobic effect that arises from changes in the entropy of solvent water as the protein folds. The impact of motion of solvent water on the internal motion of folded proteins potentially has great relevance to a number of biological functions of proteins such as catalysis and molecular recognition. The character of the surface hydration of folded protein molecules therefore naturally continues to be the subject of intense investigation. The "solvent slaving" model has commonly been employed as a descriptive framework for the coupling of solvent water and internal protein motion. Solvent slaving posits that the majority of internal protein motion is slaved to (i.e. dictated by) various structural and dynamic aspects of water. This view is seemingly at odds with the realization that fast internal protein motion is a manifestation of large reservoirs of conformational entropy that can greatly influence the thermodynamics of protein functions such as the binding of ligands. Here we show that the conformational entropy of soluble proteins is largely independent of the features of solvent pointed to by the solvent slaving model.

Introduction

The solvent slaving model proposes there are internal motions of proteins that are directly dependent on either bulk solvent motions (Class I) or on those of the layer of water nearest the protein surface, namely hydration water, which are termed Class II motions [18]. Substantial evidence indicates that large-scale, relatively long-timescale (ms - µs) collective motions such as reorientations of domains or large sections of backbone are intimately dependent on solvent viscosity and are modulated by molecular crowding[35]. Here we are mostly interested in the influence of water on the distribution of fast internal side chain motion, which reflects the rotameric entropy of proteins [36, 52, 161, 162].

Nuclear magnetic resonance (NMR) spin relaxation can be used to assess the mobility of individual bond vectors within the molecular frame [30, 163]. Recent work indicates that changes in conformational entropy of proteins can be obtained through a dynamical proxy based on NMR-relaxation in methyl-bearing side chains [36, 49, 50, 162]. Using this NMR-based "entropy meter" it has been found that conformational entropy is a highly variable and often important component of the thermodynamics governing protein-ligand interactions [36]. Thus significant effects of water on fast protein internal motion could greatly impact the fundamental thermodynamics underlying protein function.

We have used NMR relaxation methods to examine the influence of various solvent conditions on fast motion of the backbone and aromatic and methyl-bearing side chains of the proteins ubiquitin, maltose-binding protein (MBP) and malate synthase G (MSG). The fast internal motion of the protein systems in bulk aqueous solution was compared to that seen in high viscosity water/glycerol mixtures (ubiquitin only) and when encapsulated within the confined high viscosity water core of a reverse micelle. The latter was used to mimic the most extreme level of confinement a protein might experience within the cell. When encapsulation of each protein is appropriately optimized [60], the chemical shifts of backbone and methyl resonances of the encapsulated protein are statistically indistinguishable from the protein in bulk solution indicating that the native state is maintained. Indeed, the explicitly determined structure of encapsulated ubiquitin is essentially that same as that in bulk solution and in the crystal [76]. Importantly, reverse micelle encapsulation facilitates comprehensive, site-resolved measurement of hydration water dynamics [29], which enables a direct comparison of the dynamics of the protein to the mobility of the water near the protein surface site-by-site. (See Discussion).

Results and Discussion

Solvation conditions

Water/glycerol mixtures have frequently been used to investigate viscosity effects on protein systems. Such mixtures are useful because protein structure is typically maintained and because the nature of water/glycerol solvent systems are generally well-understood. The reverse micelle condition is particularly advantageous for the present study because it not only provides a level of confinement commensurate with the most extreme condition expected in the cell, but it also offers a condition wherein the protein hydration layer is satisfied, yet bulk water is replaced by the non-polar solvent. This is useful insofar as it permits direct assessment of the importance of the long-range water hydrogen bond network which has previously been implicated in direct control of protein internal motions [164].

In all conditions investigated here, the protein is provided with sufficient solvent to satisfy the hydration layer, and a direct dynamic link between the hydration layer and the solvated protein has previously been argued. We have demonstrated that encapsulation of a protein in the nanoscale water pool of reverse micelles permits site-resolved measurement of the mobility of hydrating water via measurement of the nuclear Overhauser enhancement between water and protein hydrogens [60]. Specifically, determination of the NOE/ROE ratio for each site in the protein provides a relative measure of the dynamics of nearby water molecules. The NOE/ROE value ranges from 0, indicating rapid water motion, to -0.5, indicating water that is effectively bound to the protein. This approach was used to comprehensively map the dynamics of the ubiquitin hydration layer in AOT reverse micelles, and a remarkably heterogeneous hydration surface was observed. Most importantly, a correlation was seen between areas of restricted hydration dynamics and the protein-protein interacting surface of ubiquitin [29]. The same reverse micelle condition was used to examine protein dynamics under confinement here, thereby providing the opportunity to directly compare the dynamics of the protein with the mobility of the solvating water in a site-resolved fashion for the first time.

Backbone dynamics

The mobility of the ubiquitin backbone was evaluated under each solvation condition (Figure 5-1a) using the Lipari-Szabo model-free order parameters (O^2) which ranges from 0, indicating complete isotropic flexibility, to 1, corresponding to complete rigidity within the molecular frame [165]. It is important to emphasize that this measure of flexibility is sensitive only to motions faster than overall reorientation of the protein [165], which ranges from 4 ns to 25 ns depending on context (see ED Table I). Units of secondary structure are generally rigid and differ little across the various conditions examined. The only significant change in backbone motion under confined or high-viscosity conditions is the rigidification of the reverse turn (residues 9-12). While indicative of both a viscosity and confinement effect on the fast motions of the backbone, we note that the free C-terminal tail is not generally rigidified (Figure 5-1a). The reverse turn undergoes a collective motion on the ns-μs timescale in bulk aqueous solution [166]. Because the effective rotational time of the protein is increased in the reverse micelle and glycerol samples, there is potential for an enhanced convolution of the local N-H motions of residues in the reverse turn with this collective motion. Solvent effects on collective motions have been reported in many systems [167]. This conclusion is also supported by the temperature dependence of the backbone dynamics in the presence of high glycerol (Figure E-1 in Appendix E). Interestingly, there is generally little correlation between changes in backbone motion upon encapsulation of the protein and the mobility of local hydration water (Figure 5-1b.). The sole exception is the region of slow water hydrating the reverse turn, which is consistent with the solvent-dependence of the collective motion undergone by this loop.

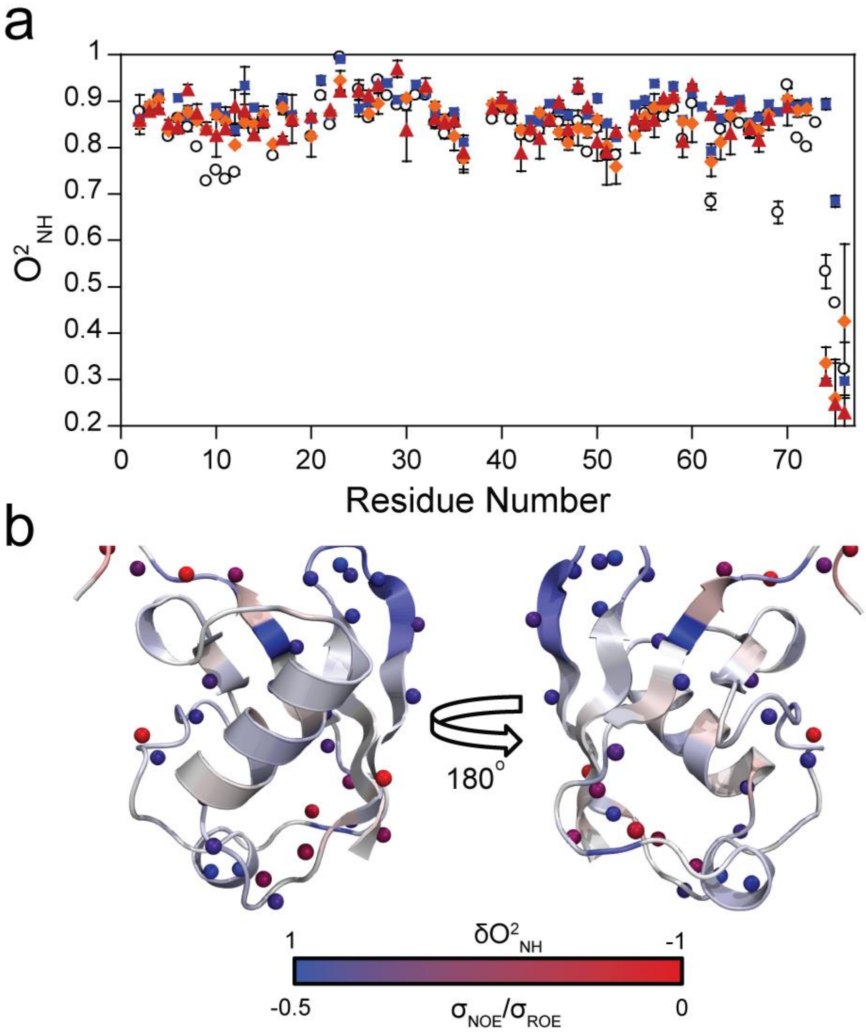

Figure 5-1: Solvent dependence of the ubiquitin backbone dynamics. **(a)**, Backbone L-S squared generalized order parameters of amide N-H bond vectors (O^2_{NH}) of ubiquitin under various solvation conditions. Solvation conditions: aqueous – open circles; AOT reverse micelles – blue squares; 30% glycerol – orange diamonds; 50% glycerol – red triangles. **(b)** The change in backbone O^2 upon encapsulation is shown against the dynamics of the hydration layer. The backbone ribbon (PDB ID 1G6J[76]) is color-coded by the difference (δO^2_{NH}) between O^2_{NH}(RM) and O^2_{NH} (aqueous)) ranging from -1 (red) to +1 (blue). The spheres represent the backbone amide hydrogens color coded by relative mobility of the hydration layer from fast (red, $\sigma^{NOE}/\sigma^{ROE}$ -> 0) to slow (blue, $\sigma^{NOE}/\sigma^{ROE}$ -> -0.5).

Encapsulation of ubiquitin in AOT reverse micelles results in a slight rigidification of the reverse turn (residues 8-11) only. This effect may be the result of confinement, but it also may be caused by the higher effective viscosity of the water within the reverse micelle core. Analysis of water mobility in AOT reverse micelles has shown that the reorientation time of water molecules in the interfacial region are restricted by a factor of approximately three to five. A simplistic interpretation of this result suggests that the effective viscosity of the water in this region is therefore three to five times higher than that of bulk water. The rigidification of the reverse turn is observed in both 30% and 50% glycerol solutions where higher bulk viscosity is present in the absence of confinement. This argues that increased local viscosity is the source of the observed effect rather than confinement.

Residual dipolar coupling analysis of ubiquitin has shown that the reverse turn undergoes a collective motion on timescales slightly longer than the rotational correlation time (τ_c), though the precise timescale of this motion is not well defined [166]. In bulk aqueous solution, τ_c is approximately 4 ns for ubiquitin, but due to the approximately three-fold higher viscosity in the 30% and six-fold higher viscosity in 50% glycerol solutions, τ_c was determined to be ~12 ns and ~25 ns, respectively. In the reverse micelle case, the bulk solvent pentane has five times lower viscosity than water, but the rotation of the encapsulated protein is determined by the total size of the reverse micelle. The resulting volume penalty results in the encapsulated ubiquitin τ_c being ~10 ns. Inherent in the Lipari-Szabo analysis is the expectation that the local bond reorientation occurs on a timescale that is substantially different than that of molecular tumbling or other motions. In the case of the reverse turn, the observed rigidification may not, therefore, be indicative of an effect on the local bond mobility, but rather on the collective supra- τ_c motion of this loop.

Relaxation dispersion measurements of motions on the ms-μs timescale have shown that these motions are dependent on viscosity, but an examination of viscosity dependence for supra- τ_c motions has not been undertaken. We argue that the observed rigidification of the reverse turn in the present data is most likely the result of a viscosity effect on the supra-τ_c motion of this loop that is manifested in this analysis as a result of the increase in τ_c. It is important to note that substantial rigidification is not observed where only a single end of the chain is tethered to secondary structure as is the case for the C-terminal tail. The high viscosity rigidification of the reverse turn is mitigated by an increase in temperature (Extended Data Figure 1b). Increased temperature not only speeds molecular tumbling considerably through direct thermal activation of rotational motion, it also reduces the solvent viscosity. As a result, the τ_c value for 30% glycerol solution at 50 °C was 5.98 ns.

The Lipari-Szabo model-free analysis also yields insight into the effective timescales of motion for backbone sites, represented by τ_e (Figure E-1a). Formally, τ_e corresponds to the total area under the autocorrelation function, but this can be taken as approximately representative of the timescale of the observed motion. The τ_e values indicate that the effective timescale of motion in the reverse turn and the C-terminal tail is increased under high viscosity conditions suggesting that the collective motions of these regions are slowing with increased viscosity. The viscosity effect on the timescales of backbone motion is also eliminated at high temperature (not shown). Importantly, the dramatically different solvation conditions examined show no measurable differences in the mobility or timescales of the secondary structural units.

Methyl side chain dynamics

Our central focus is the effect of solvent on the protein conformational entropy. Backbone contributions to the conformational entropy arise mainly from sites with $O^2 < 0.8$ [52], thus the rigidification of the reverse turn reduces the conformational entropy only slightly and maintenance of flexibility in the C-terminus indicates that the effect is not general. The majority of the conformational sampling occurs in side chain motions [36, 52, 162], thus the dynamics of the side chains are particularly informative.

Comprehensive assessment of protein conformational entropy is achieved by inventory of the methyl-bearing side chain motions, where the reporter methyl groups are sufficiently coupled to the motion of surrounding side chains to provide a quantitative measure of the residual conformational entropy [36, 45]. There is no general modification of the methyl mobility within the ubiquitin native state upon encapsulation or with increasing bulk solvent viscosity (Figure 5-2a). The solvent slaving model predicts that an increase in the bulk viscosity should result in a general rigidification of these motions [18, 168], yet none is observed. Comparison of the methyl order parameters for encapsulated ubiquitin with the dynamics of the hydration layer near these side chains (Figure 5-2b and E-2) reveals no evidence of correlation. These data collectively demonstrate that the methyl motions, and therefore the majority of the residual conformational entropy of the protein, are independent of both bulk solvent viscosity and local hydration layer mobility.

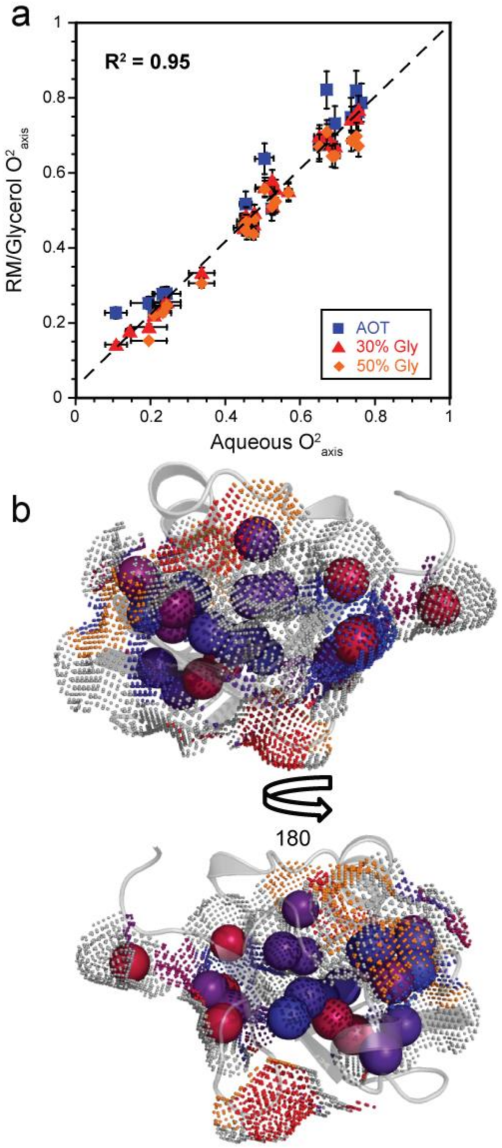

Figure 5-2: Influence of bulk solvent viscosity and confinement on fast methyl-bearing side chain motion in ubiquitin. Methyl order parameters (O^2_{axis}) are shown for ubiquitin encapsulated in reverse micelles and in high glycerol solutions versus those previously acquired in aqueous solution[48] **(a)** Solvation conditions are represented as follows: AOT reverse micelles – blue squares, 30% glycerol – orange diamonds, 50% glycerol – red triangles. These order parameters are mapped to the encapsulated ubiquitin structure **(b)** with the methyl carbons shown as spheres color-coded from their O^2_{axis} from 0 (red) to 1.0 (blue). The solvent-accessible surface is shown as dots and is colored by relative hydration dynamics as in Figure 2 with orange indicating areas on the solvent-accessible surface within NOE distance of an amide proton but with no detectable hydration dynamics.

For each solution condition of ubiquitin samples described in Figure 2, there is very little scatter about the diagonal indicating virtually no rearrangements of the distribution of side chain motions throughout the protein. The average δO^2_{axis} for the RM, 30% glycerol, and 50% glycerol conditions (-0.029, -0.013, and 0.012 respectively) are within error (±0.035, ±0.022, and ±0.020, respectively) of zero, indicating no change in conformational entropy. This shows that the degree of internal side chain motion of the protein is not dependent on solvent, but rather is the manifestation of the protein's inherent thermal motions. Many studies have shown that internal motions of the proteins require sufficient hydration to satisfy the protein hydration layer. This condition is met in all of the samples tested here. Low-temperature dynamical changes have also been heavily investigated [18, 168], but the dynamical transitions observed occur at temperatures that are not physiologically relevant and are therefore not addressed here.

Protein methyl-bearing side chains can be classified into three groups according to their dynamical behavior [45]. The most rigid class, termed the 'ω-class', represents those groups with O^2_{axis} values between ~0.7 and 1.0. These methyl groups exhibit highly restricted motion within a single rotamer well. The least rigid class, the 'J-class', represents methyl groups with $O^2_{axis} < \sim 0.5$ indicating that they rapidly interconvert between rotamers. The intermediate 'α-class' undergoes significant excursions within an individual rotamer well. As shown in Figure 2a, despite the range of motions exhibited by the methyl-bearing side chains of ubiquitin, all three classes are independent of the solvation condition, and there is no indication of significant reshuffling of methyl side chains between classes. This indicates that the viscosity of the bulk solvent and the extent of spatial confinement have no significant effect on motion within a rotamer well or on rotameric interconversion.

Further consideration of the range of solvation conditions explored here yields additional insight. The conditions employed were selected in order to examine specific aspects of the dynamic interplay between protein and solvent. Confinement within the reverse micelle provides the protein with sufficient water to satisfy its hydration layer, thereby eliminating bulk water and removing the collective, long-range motions of the water hydrogen bond network. These motions have been linked to activation of protein dynamics in many studies. Insensitivity of the methyl order parameters to reverse micelle encapsulation confirms that these internal protein motions are independent of collective water modes. The 30% glycerol condition provides a bulk viscosity approximately three times greater than that of water. The 50% glycerol mixture provides the highest viscosity glycerol/water mixture where a three-dimensional water-water hydrogen bond network is maintained [169]. Both of these mixtures, therefore, preserve the representative long-range collective modes of bulk water while increasing the bulk viscosity. Because the dynamics of the protein methyl groups are not significantly rigidified in these samples, it can be concluded that the methyl motions, and therefore the residual conformational entropy of the protein, are independent of the bulk solvent viscosity.

Figure D-2 shows a comparison of the methyl order parameters with the average NOE/ROE ratio for all solvent-accessible surface sites within NOE distance (4.5 Å) of the respective methyl carbon. It should be noted that this comparison is only shown for methyl groups that are within this distance of the solvent-accessible surface. As is evident, no correlation is observed. In contradiction of models that support a view of direct motional coupling between the protein and the hydration layer, these data indicate the absence of such correlation. This is particularly interesting because recent evidence shows that side chain methyl groups are motionally coupled to their surroundings within the protein, but the present analysis indicates that they are decoupled from, and therefore independent of, the nearby solvating water within the hydration layer.

Cross-correlated methyl dynamics for large proteins

While the range of methyl dynamics for ubiquitin is quite large, the change in depth of burial of the methyl probes is quite limited due to the protein's relatively small size. One would assume that for motions slaved to the hydration layer of water the change in order parameter near the surface of the protein should be much different than the change in order parameter buried within the interior of the protein. To that end, we encapsulated two larger proteins in reverse micelles, maltose binding protein (MBP, 41 kDa) and malate synthase G (MSG, 81 kDa), in order to determine the effect of reverse micelle encapsulation on methyl probes exposed to more varying environments. This also allowed for the study of the effect of a positively charged surfactant head group (CTAB) versus a negatively charged head group (AOT). Much like for ubiquitin, the methyl order parameters are essentially unaffected by encapsulation (Figure 5-3). Furthermore, there is no correlation to burial depth (Figure E-3). Taken together it seems clear that the so-called Class I and II motions of the solvent slaving model are absent in these three proteins at room temperature and that the conformational entropy is independent of solvent.

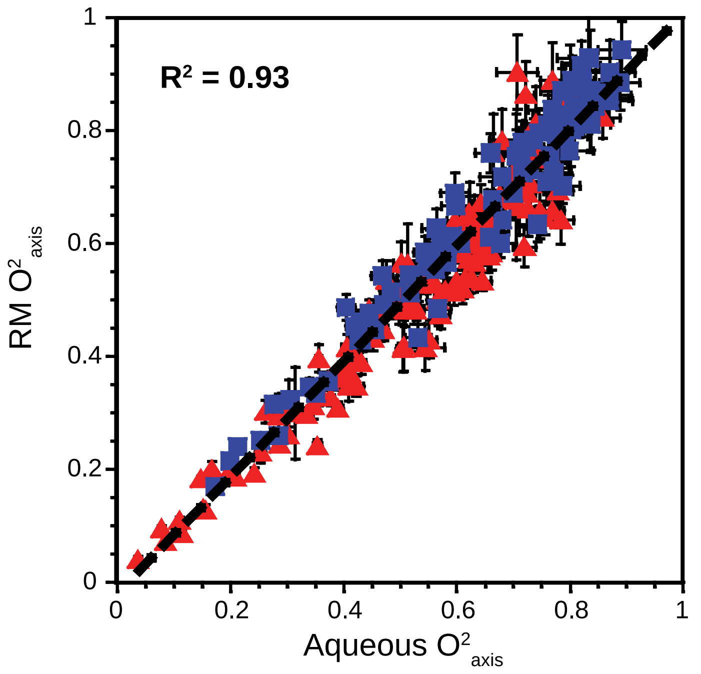

Figure 5-3: Methyl order parameters for MBP and MSG have no dependence on solvent. Correlation plots of methyl order parameters (O^2_{axis}) for MBP (blue squares) and MSG (red triangles) are displayed. For both proteins, the reverse micelle-encapsulated methyl dynamics are plotted against the aqueous methyl dynamics. The correlation of the data is excellent ($R^2 = 0.93$) indicating reverse micelle encapsulation does not affect protein methyl dynamics. In effect, methyl dynamics (and in extension conformational entropy) are not slaved to solvent.

In each of the data sets in Figure 5-3, the ratio of intensities of the triple quantum and single quantum versions of the cross-correlated experiments were fit to determine methyl order parameters as previously described[43]. The difference in order parameters for both MBP and MSG were an approximately Guassian distribution around zero (δO^2_{axis} for MBP and MSG were -0.01 and 0.02, respectively). All degenerate probes in the spectra, all probes within 40 Hz of the background streaks in the methyl spectra (Appendix E and Figure E-4), as well as all probes with fit errors in the order parameter larger than 10% were removed from statistical analyses.

While the use of deuterium relaxation experiments to obtain methyl order parameters for small proteins like ubiquitin is well documented, these experiments as well as carbon-based relaxation experiments are exceedingly difficult to collect for proteins of large size. In effect, for MBP and MSG the cross-correlated methyl dynamics experiments proposed by Kay and colleagues [43] were used to obtain methyl order parameters both in aqueous solution and encapsulated in reverse micelles (see Methods in Appendix E). These experiments eliminate the need of a high resolution structure for the order parameter determination and the need for data collection at multiple fields. Interestingly, this method also provides an excellent proxy for determining changes in molecular tumbling.

As shown previously, the ratio of the intensities of the two cross-correlated methyl dynamics experiments can be directly related to the ^{1}H-^{1}H intra-methyl cross-correlated relaxation rate, η, for a methyl group undergoing rapid rotation about its symmetry axis. As shown in the Introduction, this η rate is proportional to both the methyl order parameter (O^2_{axis}) and the rotational correlation time (τ_c):

$$\eta \propto O^2_{axis} * \tau_c$$

This indicates that, assuming there is no solvent slaving for methyl order parameters (as demonstrated in the main text) then it should be possible to determine the tumbling time of a protein in multiple environments just by determining the ratio of η values. For example, in the case of the reverse micelle:

$$\frac{\eta_{RM}}{\eta_{AQ}} = \frac{\tau_{RM}}{\tau_{AQ}}$$

The rotational correlation times for MBP and MSG were experimentally determined both in aqueous solution and in the reverse micelle using backbone T_1 and $T_{1\rho}$ experiments as described in the Methods. For MBP, the aqueous tumbling time at 25°C was 23.3 ns while the tumbling time of the RM-encapsulated MBP at the same temperature was 31.0 ns (leading to a ratio of tumbling times as described above of ~1.33). As expected, the ratios of fitted η rates (effectively the slope of the linear regression of the η rates in the RM samples versus the η rates of the aqueous sample) was ~1.35. Similarly for MSG, τ_{AQ} = 50.8 ns and τ_{RM} = 47.1 ns giving a ratio of ~0.93 while the ratio of the η rates was ~0.95. This shows the utility of using the cross-correlated experiments as a tool for detection of changes in rotational correlation times and further demonstrates the lack of a dependence of the methyl order parameter on solvent.

Aromatic side chain dynamics

Some aspects of the solvent slaving model do seem operative to some extent in the context of larger amplitude aromatic ring motion. Extensive site-resolved tryptophan fluorescence relaxation [170] has shown correlation between the mobility of tryptophan side chains and that of solvent. Ubiquitin contains three partially solvent-exposed aromatic side chains and their order parameters are clearly dependent on the viscosity of bulk solvent (Figure 5-4a). It should be noted that addition of glycerol will also perturb the hydration layer. In contrast, optimized reverse micelle encapsulation appears to preserve the natural hydration shell of protein molecules [57]. Reverse micelle encapsulation of ubiquitin does not noticeably rigidify these side chains (Table E-1) indicating that bulk solvent viscosity is the dominant parameter. Furthermore, increasing bulk solvent viscosity suppresses the thermal activation of ring-rotation by all three aromatic side chains (Table E-1). Finally, the hydration water in the regions of the solvent exposed aromatic side chains exhibits the full dynamic range of motion suggesting that the motion of hydration water and that of the aromatic side chains is not strongly coupled to solvent.

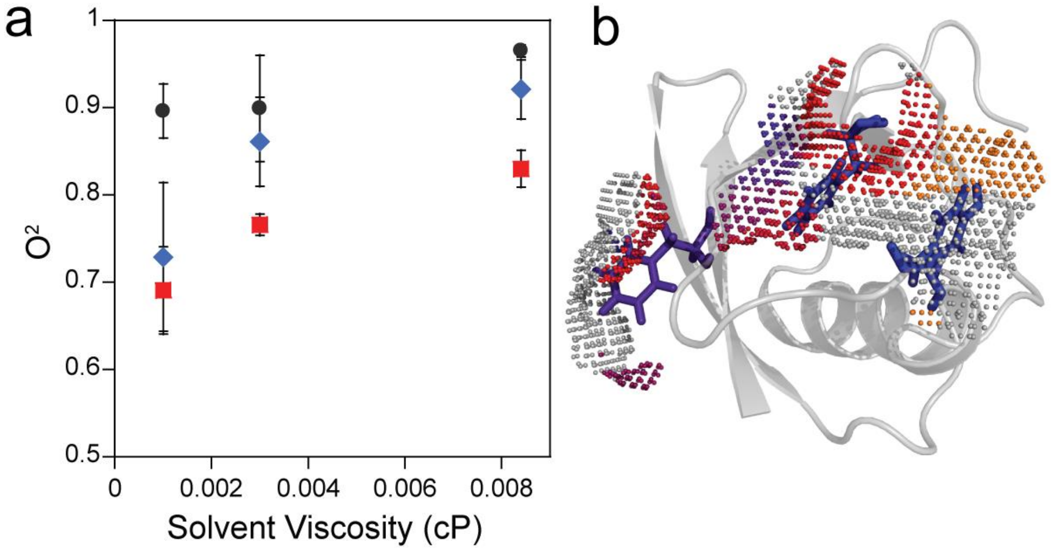

Figure 5-4: Influence of bulk solvent viscosity on aromatic side chain mobility in ubiquitin. **(a)** The order parameters (O^2) of F4 (red squares), F45 (blue triangles), and Y59 (black circles) show clear rigidification upon increase in bulk solvent viscosity (aqueous, 30% glycerol, and 50% glycerol, respectively). **(b)** The aromatic side chains are shown as sticks color-coded by their O^2 from 0 (red) to 1.0 (blue) with the solvent-accessible surface rendered as dots color-coded by relative dynamics of the hydration water from fast (red, $\sigma^{NOE}/\sigma^{ROE} = 0$) to slow (blue, $\sigma^{NOE}/\sigma^{ROE} = -0.5$) with the fastest sites (water moves too fast to be detected) color coded orange. Surface for which no data is available is colored gray.

Order parameters and τ_e values for the aromatic side chains at low temperature (20 °C) under all conditions and at high temperature (50 °C) in 30% glycerol solution are reported in Table E-1. Previous temperature dependent measurements of ubiquitin aromatic side chain dynamics showed that they undergo a dynamical transition at ~37 °C resulting from thermal activation of ring flipping. The increasing viscosity of the high glycerol solutions produces and increase in the time frame and in the rigidity of aromatic side chains at 20 °C. There is a slight rearrangement of the conformational entropy of these side chains under reverse micelle confinement, but not a general rigidification. This indicates that the solvent dependence of these side chains is more likely to be linked to the bulk solvent viscosity than the local viscosity. The starkest difference observed is in the order parameters of the aromatic side chains at high temperature in 30% glycerol. Here a considerable loss of mobility is caused by the high viscosity of the glycerol containing solution. This result is interpreted as a clear indication the ring-flipping motion is damped in the higher-viscosity glycerol condition as compared to the aqueous condition.

The aromatic rings present in ubiquitin are all partially solvent-exposed. It remains to be determined whether the viscosity effect on ring flipping observed here is the result of a direct interaction with solvent or is representative of the general viscosity dependence purported by previous studies, i.e. the result of a linkage between the effective viscosity of the protein itself and that of the bulk solvent. This topic is the purview of future studies involving larger proteins with fully buried aromatic residues.

Conclusions

These data illustrate that protein entropy is largely independent of the solvating environment. While it is well established that large-scale, slow internal motions are linked to solvent [167], the present data indicate that the mobility of bulky aromatic side chains, particularly of the ring-flipping motions, represents a reasonable cutoff in terms of the time- and length-scale for solvent dependence of internal protein motions. These findings resolve an important detail of the previously described hierarchy of protein motions [18, 168]: while slower motions that are likely linked directly to function may be effectively tuned by the protein environment, the fast motions that express the residual entropy of the protein are independent of environment. Over the past decade, several studies have indicated the potential for dynamically mediated allostery based on the potential for motional coupling within a given protein structure to provide an avenue for intramolecular signal transduction [171-173]. The environmentally ‘isoentropic’ nature of the native state revealed by our data is consistent with this view. Many proteins carry out their function across a wide range of intracellular environments. Our findings suggest that the conformational entropy of the protein is generally independent of the protein’s intracellular locale, thereby leaving the conformational entropy available for functional processes such as entropic tuning of binding free energy and conveyance of entropically-mediated allosteric responses.

Chapter 6: Conclusions and Future Directions

The goal of this work was to use NMR spectroscopy and reverse micelle encapsulation in order to better understand the relationship between hydration and protein dynamics. While this relationship has been extensively examined in a global sense, it has been difficult to study it in a site-specific manner. NMR spectroscopy allows for the study of both hydration and protein dynamics at an atomistic level. Reverse micelle encapsulation allows for the study of hydration dynamics with NMR without common concerns seen in aqueous solution, and permits for the inspection of protein dynamics under extreme confinement with significant retardation in the dynamics of hydration waters.

In Chapter 2 we used NMR spectroscopy to monitor and adjust the pH of the aqueous interior of reverse micelle ensembles by observing chemical shift changes of common buffer molecules and of encapsulated proteins. It was demonstrated that the dominating factor in the pH of the aqueous nanopool is the pH of the hydrophilic surfactant head groups indicating the absolute necessity of pH adjustment of surfactants before reverse micelle formation. Finally, we demonstrated the ability to confirm structural fidelity while monitoring the pH of encapsulated proteins – a common oversite when using alternative methods to study protein reverse micelle systems.

In Chapter 3 we present a novel, relatively fast way of using NMR to detect protein hydration dynamics via the NOE by applying non-uniform sampling techniques to typical hydration dynamics experiments. Furthermore, we examine the reproducibility and reliability of hydration dynamics NOE/ROE ratios and demonstrate the necessity of collecting a repeat spectrum in order to calibrate the minimum necessary cross-peak signal-to-noise needed to assure reliable data. Using this new approach, we were able to demonstrate clustering of hydration dynamics across the surface of ubiquitin similar to what has been seen in the past [29].

In Chapter 4 study we examined the hydration dynamics of hen egg-white lysozyme. While no slow hydration dynamics were detected in the hydrophilic substrate-binding pocket, waters were detected within a partially hydrophilic internal pocket. Interestingly, these internal waters were relatively fast demonstrating a transient nature for internal waters. Also, we detected site-specific dynamics for interfacial waters between the surface of HEWL and an inhibitor that were much slower than waters detected within the internal pocket described above. Lastly, this reduction in solvent entropy at the protein/ligand interface seems to be compensated by increases of hydration dynamics in remote areas of the protein.

Finally in Chapter 5 we studied the effect of varying solvation environments on the fast internal dynamics of three different protein systems. In conjunction with methods described in Chapters 3 and 4, we inspected the relationship between hydration and protein dynamics. While dynamics of aromatic side chains do seem to be affected by changes in solvent viscosity, those motions that most greatly contribute to protein conformational entropy (methyl side chain dynamics) exhibit no dependence on changes in solvent dynamics suggesting they are not slaved to solvent.

While this thesis provides new insight into the relationship between protein dynamics and hydration with model protein systems, it is still only the first step. With the advent of non-uniform sampling and protein dynamics techniques geared towards large protein systems (cross-correlated dynamics) it is now possible to study the protein-water relationship for much more complex systems. These techniques, in conjunction with reverse micelle surfactant systems that have been shown to encapsulate virtually every protein system attempted [58], will allow for as yet unforeseen understanding into the interplay between hydration and protein dynamics in membrane associated proteins, intrinsically disordered proteins, and large protein-protein complexes among others.

APPENDIX A: Comparison of hydration dynamics of ubiquitin using NMR and ODNP

The work in this Appendix was completed in collaboration with the laboratory of Songi Han at the University of California, Santa Barbara

Introduction

While water is very important for most aspects of protein biochemistry, dynamic information about layers of water immediately surrounding protein surfaces (the hydration layer) is traditionally difficult to obtain experimentally. By encapsulating proteins in reverse micelles, the encapsulated waters become slow enough such that NMR can be used to examine hydration dynamics of water in the hydration layer.[29] The hydration dynamics of ubiquitin was studied using this method and it was found that a large array of hydration dynamics can be detected across the surface of the protein[29] and areas of slow hydration dynamics correlated with areas of the protein that make protein-protein interaction surfaces [29]. In collaboration with the lab of Songi Han at the University of California at Santa Barbara, we compared the hydration dynamics of ubiquitin as determined by NMR to hydration dynamics determined by a different method: Overhauser effect dynamic nuclear polarization (ODNP) whose experimental hydration dynamics have been shown to be strongly dependent on the hydrophobicity of the nearby protein surface and in extension on protein-ligand binding surfaces [174].

With ODNP, the nuclear polarization of water protons is enhanced by the microwave-excited polarization of an unpaired electron in a covalently-attached spin radical. This enhancement of the water proton NMR signal is then related to local translational diffusion of water molecules 5-15 Å away from the nitroxide spin label [137]:

$$\xi = \frac{k_\sigma}{\rho} \tag{12}$$

Where k_σ is the radical concentration-dependent cross-relaxation rate between the water proton and the electron spin dependent on the enhancement of the water signal as a function of microwave power, *E(p),* the longitudinal relaxation time constant of the water protons as a function of microwave power, $T_1(p)$, the maximum attainable saturation of the unpaired electron EPR signal, S_{max} (which is 1 in the slow tumbling regime), and the gyromagnetic ratios of protons and elections, γ_H and γ_e, respectively:

$$k_\sigma = \lim_{p\to\infty}\left(\frac{1-E(p)}{T_1(p)}\left|\frac{\gamma_H}{\gamma_e}\right|\right)\left(\frac{C}{S_{\max}}\right) \tag{13}$$

ρ is the radical-induced self-relaxation rate between the proton and the electron spins dependent on the longitudinal relaxation time of the water proton with and without the presence of the spin radical (T_1 and $T_{1,0}$, respectively):

$$\rho = \frac{1}{T_1} - \frac{1}{T_{1,0}} \tag{14}$$

and ξ is a unit-less electron-hydrogen coupling factor which can be related to the time constant describing the lateral diffusion of water relative to the radical electron (τ_H) via the spectral density equation:

$$\xi(B_0;\tau_H)=\frac{6J(\gamma_e B_0-\gamma_H B_0;\tau_H)-J(\gamma_e B_0+\gamma_H B_0;\tau_H)}{6J(\gamma_e B_0-\gamma_H B_0;\tau_H)-J(\gamma_e B_0+\gamma_H B_0;\tau_H)+3J(\gamma_H B_0;\tau_H)} \quad (15)$$

Assuming the protein system and the chemically bound radical species are the same (as is the case in this study), then k_σ can be used as a proxy for the lateral diffusion time wherein k_σ increases as the average rate of the lateral diffusion of nearby water molecules increases.

In this study, we examined the ODNP-derived hydration dynamics on 16 different ubiquitin cysteine mutants at multiple different locations across the surface of the protein all exhibiting a wide array of hydration dynamics as determined by the NOE/ROE ratio (see Chapter 3). It is clear that the comparison leads to little to no correlation between the hydration dynamics determined by the two methods likely due to differences in the waters detected by each method.

Methods

Protein expression and spin radical labeling

The 16 mutants of the wild-type human ubiquitin gene were created using an Agilent QuickChange™ site-directed mutagenesis kit with suitable T7 promotor PCR primers. Purified plasmids were transformed into BL21(DE3) *E.Coli* cells and the ubiquitin mutants were expressed and purified as previously described [84]. The protein was then buffer exchanged into 50 mM Tris (pH 8.0), 50 mM NaCl and concentrated to ~ 750 μM. The protein was then incubated at room temperature overnight with ~10 times the concentration of oxidized ^{15}N isotopically labeled (1-Oxyl-2,2,5,5-tetramethyl-Δ3-pyrroline-3-methyl) Methanethiosulfonate or MTSL (Toronto Research Chemicals). In order to remove excess, unbound MTSL, the protein was then extensively buffer exchanged into NMR buffer (50 mM sodium acetate, pH 5.0, and 50 mM NaCl). The protein was then concentrated to ~500 μM for labeling efficiency determination and ODNP collection as described below.

NMR experiments for determination of labeling efficiency

For each ubiquitin mutant, ^{15}N-HSQCs were collected for the sample with the radical in the oxidized state. Half of each sample was saved for later ODNP detection. The MTSL in the other half of each sample was then reduced with 5 mM ascorbic acid. A HSQC was then collected of the reduced state and labeling efficiency was determined by comparing the intensity of peaks of the unlabeled protein species (peaks that are present in the oxidized spectrum) to the intensity of peaks of the MTSL-bound species (now visible due to the lack of the PRE effect in the reduced state). This labeling efficiency was used to determine the effective concentration of the spin radical in the ODNP experiments (and further confirmed by comparing EPR signal double integrals to standard samples of known concentrations).

EPR and ODNP experiments

Constant-wave EPR spectra were collected at 2 mW irradiation on a Bruker EMX CW EPR spectrometer for each ubiquitin mutant. In order to determine radical concentrations, the double integral of each EPR spectrum was compared to those of standard spin radicals of known concentration.

As previously described [12], the ODNP samples were then placed into quartz tubes of 0.6 mm inner diameter (VitroCom) and lowered directly into a 14.8 MHz (^{1}H) NMR probe with an additional switch to allow for constant uniform microwave power input. ODNP experiments were performed using a Bruker EMX CW EPR spectrometer and an iSpin NMR console (SpinCore Technologies). All water signal enhancement and water T_1 measurements were performed at room temperature (maintained by a slow flow of nitrogen gas to prevent sample bubbling). The sample was irradiated with up to 8 W of microwave power at the EPR frequency of the spin label at 9.8 GHz using a microwave amplifier developed by Bridge12.

The enhancement and longitudinal relaxation as a function of power were then used to calculate the cross-relaxation rate between the water proton and the electron spin as described in equation 13. These microwave power-dependent rates are then fit to the equation:

$$k_\sigma(p) = \frac{k_\sigma(S_{\max})p}{a+p} \tag{16}$$

Where p is the microwave power in Watts, a is a scaling factor set as a free parameter in the fit, and $k_\sigma(s_{\max})$ is the desired concentration-corrected cross-relaxation rate between the water proton and the electron spin at maximum saturation.

Results and Discussion

Spin-radical labeling efficiency

In order to collect ODNP experiments, the protein of interest must first be labeled with a free-radical containing species. In this case 16 dispersed ubiquitin cysteine mutants were covalently bound to ^{15}N-MTSL via a disulfide bridge. In order to determine labeling efficiency (and, in effect, spin radical concentrations) ^{15}N-HSQCs of the oxidized and reduced spin radical states were compared. This is demonstrated for two of the 16 mutants in figure A-1:

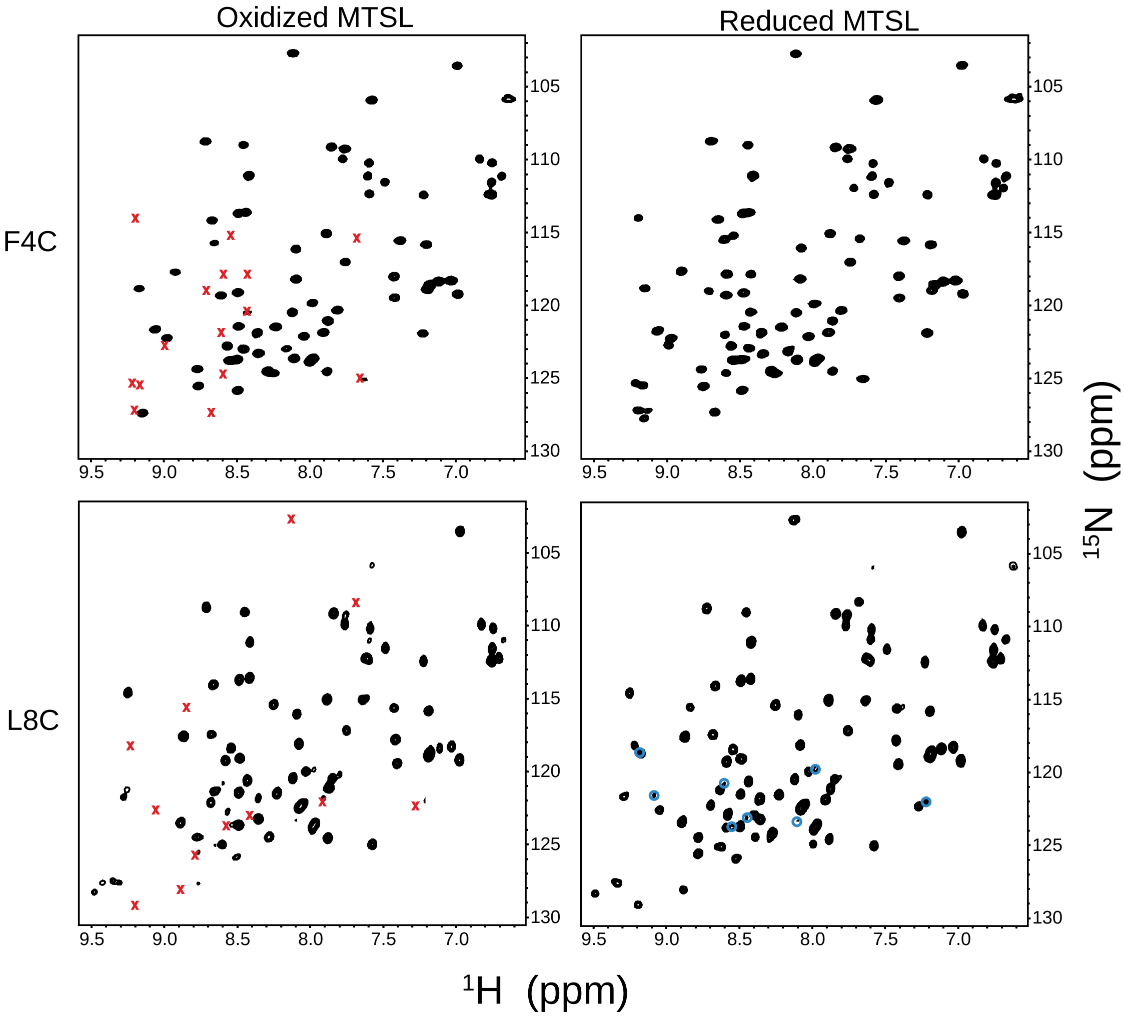

Figure A-1: 15N-HSQC spectra of two ubiquitin cysteine mutants (F4C and L8C) covalently bound to 15N-MTSL in the oxidized (left) and reduced (right) spin radical states. The red X's in the oxidized spectra indicate residues spatially near the nitroxide spin label whose peaks become broadened due to the PRE effect induced by the radical electron. The blue circles in the reduced spectra indicate peaks associated with the apo protein state. The intensity of these peaks was compared to the intensity of the un-broadened peaks in the reduced spectrum to calculate MTSL labeling efficiencies. If there are no peaks of the apo ubiquitin state evident (as is the case for the F4C mutant), the labeling efficiency is considered to be 100%.

This method was applied to all 16 ubiquitin mutants leading to a large range of labeling efficiencies:

Table A-1: Ubiquitin mutant MTSL labeling efficiencies

Mutant	Labeling Efficiency	Mutant	Labeling Efficiency
F4C	100%	G35C	<10%
L8C	66%	I44C	100%
T9C	100%	A46C	80%
T14C	40%	Q49C	30%
E18C	<10%	S57C	<10%
E24C	77%	T66C	70%
A28C	72%	H68C	74%
K33C	100%0	R72C	81%

Clearly the efficiency with which the ubiquitin cysteine mutant is labeled with MTSL is extremely variable with virtually no dependence on burial depth of the residue (data not shown) or on hydration dynamics near the residue.

ODNP data collection and analysis

As described in the methods, ODNP enhancement as a function of microwave power and longitudinal relaxation as a function of power were collected on each of the 16 ubiquitin mutants and fit to equation A-5 to obtain the concentration corrected k_σ at the maximum achievable EPR saturation. An example of the data and fitting analyses are shown in Figure A-2:

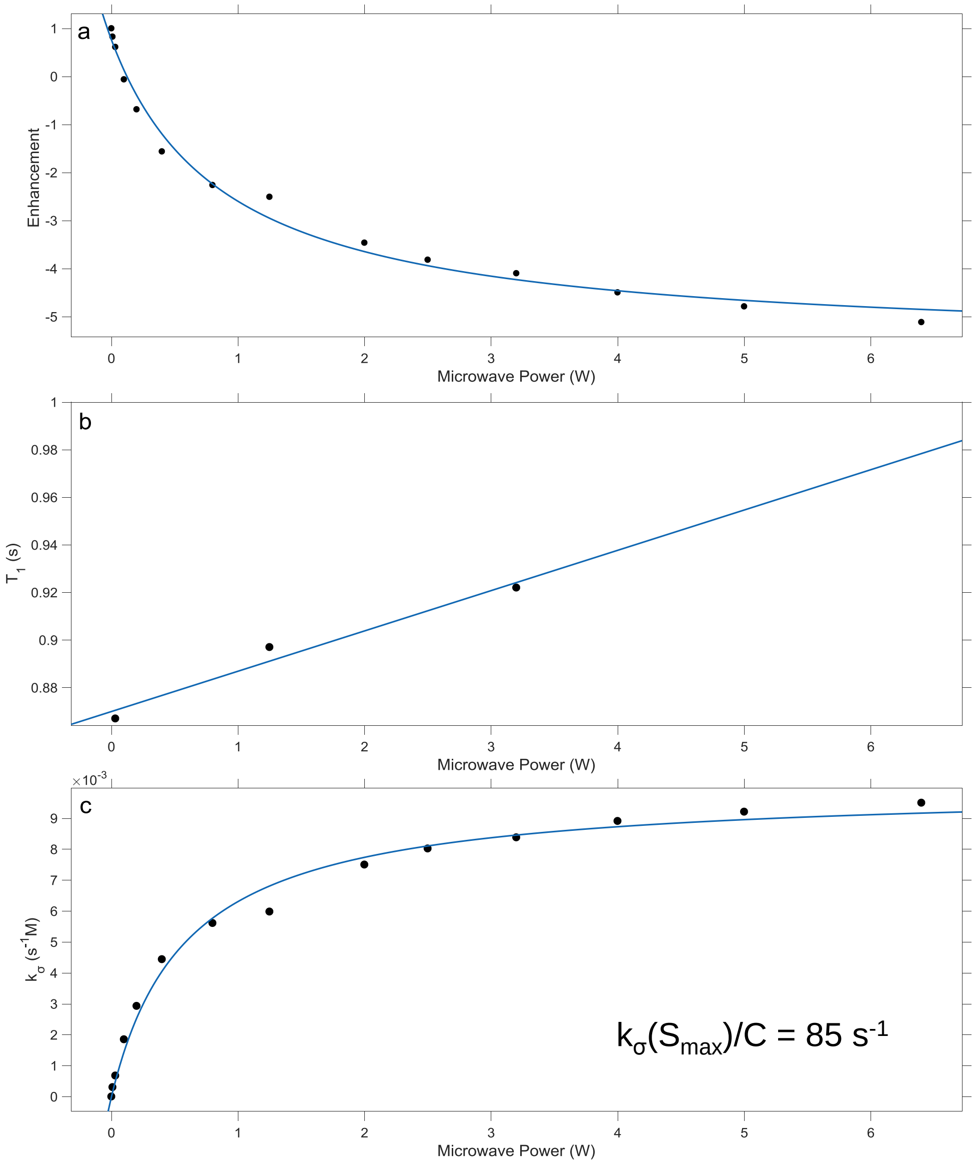

Figure A-2: Example ODNP data set for ubiquitin mutant A46C-MTSL. **(a)** Series of water signal enhancement normalized to the signal intensity in the absence of microwave power. **(b)** Series of longitudinal relaxation times as a function of microwave power. **(c)** Concentration-dependent k_σ rate constants as a function of microwave power as calculated by equation 13. The data was then fit to equation 16 in order to extrapolate the k_σ rate constant at maximum saturation.

The ODNP suite of experiments was collected for all 16 ubiquitin mutants (some mutants were collected two or three times in order to determine reproducibility) to reveal a breadth of hydration dynamics (Table A-2) across the surface of the protein much like has been seen with ODNP measurements in the past [12, 174].

Table A-2: Hydration Dynamics Derived from ODNP and NOE/ROE

Mutated Residue	k_σ 1st Repeat (s^{-1})	k_σ 2nd Repeat (s^{-1})	k_σ 3rd Repeat (s^{-1})	NOE/ROE (Chapter 3)
F4	85.3 ± 4.8	83.8 ± 2.7	83.8 ± 1.7	-0.50
L8	75.7± 2.0	67.5 ± 1.8		-0.48
T9	57.2 ± 3.1	63.8 ± 2.7	60.5 ± 1.2	-0.61
T14	65.2 ± 2.3			-0.45
E18	46.1 ± 2.0			-0.11
E24	39.4 ± 2.0			-0.37*
A28	65.0 ± 2.0	72.1 ± 1.2	70.0 ± 1.46	-0.01
K33	98.0 ± 1.4	104.9 ± 2.7		-.041
G35	65.3 ± 1.7			-0.21*
I44	66.9 ± 1.8			-0.47*
A46	84.9 ± 2.7	80.1 ± 2.6		-0.06
Q49	46.5 ± 1.3			-0.001
S57	68.0 ± 1.1			-0.53
T66	70.8 ± 4.5	65.4 ± 4.7	70.7 ±1.3	-0.37
H68	81.5 ± 2.2	76.7 ± 1.2		-0.52
R72	77.1 ± 4.2			-0.26

* Indicates residue with surface-averaged NOE/ROE ratio (no detected NOE to water at amide)

Comparison with NMR NOE/ROE hydration dynamics

By encapsulating a protein in a reverse micelle, one can experimentally determine hydration dynamics near the surface of the protein via the NOE in a site specific manner without the need for physically or chemically altering the surface of the protein. In order to reliably compare the hydration dynamics as determined by the NOE/ROE ratio to those determined by ODNP, it is necessary to choose cysteine mutation sites in multiple different environments across the surface of the protein that experience a variety of hydration environments. In effect, we chose to mutate 16 residues across the entire surface of the protein that experience the full range of hydration dynamics as determined by NMR (from slow at a NOE/ROE ratio of -0.5 to fast hydration at a ratio of 0).

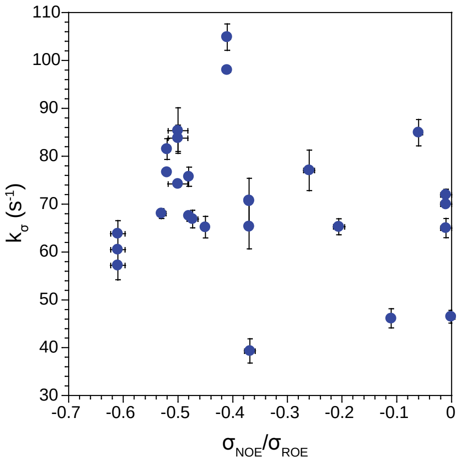

Figure A-3: There is no correlation between hydration dynamics as determined by NMR with the NOE/ROE ratio and those determined by ODNP (kσ). For all residues, ODNP-determined hydration dynamics are plotted against the amide hydration dynamics at that residue as shown in Table A-2. As explained in the footnote, all residues with no detectable NOE between the water and amide protons used a surface-averaged NOE/ROE ratio of all points on the surface within 4.5 Å of the amide proton (see Chapter 3). There is virtually no correlation ($R^2 < 0.1$) for hydration dynamics determined between the two methods.

When comparing hydration dynamics determined by NMR to those determined by ODNP (Figure A-3), it is clear there is no correlation. Even though both methods demonstrate a varying degree of hydration dynamics across the surface of the protein, there is virtually no correlation. The disparity between the two results can be clearly explained by the different information each method conveys about the nature of nearby water molecules. Unlike the NOE/ROE ratio which measures average hydration dynamics directly near the surface of the protein (within 4-5 Å of the measured probe), the ODNP method measures hydration dynamics within a much larger area (10-15 Å) and not directly at the surface of the protein since the radical electron is at the end of a

MTSL molecule which is covalently attached to a mutated cysteine residue (adding an additional ~ 9 Å distance away from the wild-type protein surface). This distance disparity between the two methods likely lead to measurements of water within and through completely different layers of hydration waters theoretically leading to average hydration measurements in water layers with very different hydration dynamics [175, 176]. In addition to this clear dissimilarity, it may be possible that the ODNP method may be susceptible to contaminations similar to those experiences by NMR in aqueous solution (hydrogen exchange and bulk solvent effects). These avenues have yet to be explored with ODNP.

Conclusions and future directions

While it is clear that both NMR and ODNP measure diverse hydration dynamics across the surface of ubiquitin, there is virtually no correlation between the hydration dynamics determined by the two methods. Although this lack of correlation was unexpected it is clearly explainable due to differences in detection distances and water layers. There also may be discrepancies due to potential contamination of the measurement itself. These avenues may be pursued in one obvious way: ODNP measurements of ubiquitin encapsulated within a reverse micelle (which will eliminate any potential bulk solvent contamination and significantly slow hydrogen exchange). In order to determine if hydrogen exchange contaminates the measurement one can also collect the experiments at different pH. These avenues will need to be further explored to better understand the lack of correlation between the NOE/ROE ratio and the ODNP k_σ.

APPENDIX B: Chapter 2

Methods

"Empty" reverse micelle sample preparation

All protein-free or "empty" reverse micelle samples were prepared by mixing appropriate amounts of surfactant in a 50/50 (v/v) mixture of pentane/*d*-pentane (d-12) (Cambridge Isotopes, Cambridge, MA) followed by injection of an appropriate volume of the desired aqueous buffer solution except samples at pH 7 which were prepared in 100% *d*-pentane for collection of spectra for Figure 2-1. Three surfactant mixtures were used: 75 mM cetyltrimethylammonium bromide (CTAB) with 450 mM hexanol as cosurfactant, 75 mM bis(2-ethylhexyl)-sulfosuccinate (AOT), and a mixture of 22.5 mM lauryldimethylamine-N-oxide (LDAO) and 52.5 mM decylmonoacyl-glycerol (10MAG). All buffers were 25 mM prepared to the stated bulk pH. The volume of buffer used for reverse micelle samples defines the molar ratio of water to total surfactant concentration (also known as "water loading" and designated as W_0). For protein-free samples, a target W_0 of 15 was used for all mixtures. 10MAG/LDAO mixtures had a final W_0 of 12, as measured by NMR integration of the ^{1}H spectra. AOT and CTAB samples had a final W_0 of 15. 10MAG/LDAO reverse micelles containing imidazole only were also tested at W_0 of 20 to examine the W_0 dependence of buffer response. All unlabeled chemicals were purchased from Sigma–Aldrich (St. Louis, MO) except LDAO (Affymetrix, Santa Clara, CA).

Characterization of surfactant pK_a values

To characterize the buffering capacity of the surfactant headgroups, each surfactant (AOT, CTAB, LDAO, and 10MAG) was individually dissolved (1-2 mM) in water (with 12% ethanol required to solubilize 10MAG) and titrated with 1 M HCl or 1 M NaOH over a pH range from 4 to 10. pH of all aqueous solutions was monitored with an Accumet AB15+ Basic pH meter and electrode (Fisher Scientific, Pittsburgh, PA).

pH and W_0 adjustment of the reverse micelle core

The pH of the water cores of reverse micelles was adjusted in two ways. In one approach, samples were prepared using surfactants as supplied without further purification or manipulation. CTAB/hexanol reverse micelles prepared with aqueous buffer cores showed pH values within 0.5 pH units of the aqueous buffer (see below for measurement methods), indicating that this surfactant mixture does not contribute appreciable buffering capacity in the pH 4-10 range. In contrast, AOT and 10MAG/LDAO reverse micelles yielded pH values of 5-5.5 and 7-7.5, respectively, *regardless of the pH of the injected buffer solution*, indicating that these surfactants have significant buffering capacity. The pH of these samples can be adjusted *post facto* by the direct addition of the appropriate amount of HCl or NaOH. After addition of a small volume of acid or base, the solution was slowly inverted, vortexed for 5-10 seconds, and allowed to equilibrate without agitation for five minutes before data collection. The pH of the aqueous nanopool was monitored with 1-dimensional ^{1}H NMR. During this pH adjustment, the W_0 often increased by four or five. In order to lower the W_0, the pentane and some of the water in the reverse micelle solution was evaporated by introducing low pressure N_2 gas. The solution was allowed to evaporate to approximately half of the total volume of the sample and then returned to the full volume with d-pentane. Depending on the total volume of the sample, this procedure lowered W_0 by approximately two or three. It should be noted that this procedure, termed the injection-evaporation method, need not have any impact on the encapsulated protein [58].

The second approach for adjustment of the pH in the reverse micelle core involved pre-adjustment of the pH of the AOT and LDAO head-groups prior to reverse micelle sample preparation as follows. AOT was dissolved in water (1 mg/mL), titrated to the target pH, and lyophilized. This procedure was repeated until the lyophilized AOT gave a consistent pH when redissolved in water. Generally three to four rounds of adjustment and lyophilization were required. It should be noted that the appearance of the dried AOT varied with pH: AOT at higher pH (8-10) had a more coarse-grained appearance than the typical pasty appearance of AOT at lower pH. The number, positions, and splitting of the AOT ^{1}H NMR signals did not change with titration, confirming that the AOT was chemically unaltered by this pH adjustment. In order to pre-adjust the pH of the LDAO head group, the appropriate amounts of LDAO and 10MAG for each sample were dissolved in a solution of 12% ethanol, titrated to the correct pH, and then lyophilized. Though LDAO is freely soluble in aqueous solution, 10MAG requires 12% ethanol to solubilize so that these surfactants could be premixed at the appropriate pH. A single round of adjustment and lyophilization was sufficient for the 10MAG/LDAO mixture at all pH values tested. These surfactants were then used for reverse micelle sample preparation as outlined above.

Protein purification and encapsulation.

Uniformly ^{15}N-labeled human ubiquitin (8.5 kDa), oxidized horse cytochrome c (11.4 kDa), and the L99A mutant of lysozyme from the bacteriophage T4 virus (18.6 kDa) were prepared as previously described ($^{15}NH_4Cl$ from Cambridge Isotopes) [84, 177-179]. Ubiquitin-containing reverse micelle samples were prepared as follows. For each sample, 2 mg (hereafter referred to as one aliquot) of dried ubiquitin was dissolved in 1 mL of water. The pH of this protein solution was adjusted to the target pH using dilute (0.1 M or 0.01 M) HCl or NaOH; the sample was then lyophilized. This pre-adjustment of the protein's pH prior to encapsulation is essential for preparation of reverse micelle samples wherein the protein structural integrity is preserved and the target pH is obtained. The dried protein was then dissolved in the appropriate volume of buffer for a target W_0 of 10. The same adjustment of the protein pH may be achieved via dialysis or buffer exchange methods followed by concentration of the protein sample to the appropriate volume for encapsulation. The same buffer mix and preparation method as described above for protein-free samples were used for ubiquitin-containing reverse micelle samples.

To illustrate the buffering capacity of the AOT and 10MAG/LDAO mixtures, one aliquot of ubiquitin at pH 5 was encapsulated in unadjusted 10MAG/LDAO, and one aliquot of ubiquitin at pH 7 was encapsulated in unadjusted AOT. In order to demonstrate the preparation of an encapsulated protein sample at a target pH, three samples were prepared with a final protein concentration of 150 µM and a W_0 of 10 using ubiquitin and 10MAG/LDAO that had been pre-adjusted to pH 5, 7, or 9. These samples were prepared with the addition of 24 mM hexanol as cosurfactant [58]. In order to demonstrate the ability to further adjust the pH of protein samples after encapsulation, the pH 5 ubiquitin sample was titrated to pH 7 and pH 9 using the methods described above for protein-free reverse micelles, and a ^{15}N-HSQC spectrum was recorded at each pH. Likewise, the pH 7 sample was titrated to pH 5 and a ^{15}N-HSQC spectrum was recorded.

Cytochrome c was used to demonstrate the degree of protein foldedness in reverse micelles detected using NMR and optical spectroscopy. Cytochrome c (6 mM aqueous solution) was encapsulated in pre-adjusted 10MAG/LDAO as described above at a W_0 of 15 and a pH of 5 to a final protein concentration of 140 μM. Cytochrome c was also encapsulated in 75 mM AOT at the same pH and protein concentration. Aqueous cytochrome c was prepared in 50 mM sodium acetate at pH 5 with 50 mM NaCl. The pH of the aqueous cytochrome c sample was adjusted to 2.5 by direct addition of HCl using a standard pH meter for measurement. The cytochrome c in 10MAG/LDAO reverse micelles was unfolded by titration of the sample through direct addition of an appropriate volume of 6 M HCl to a pH of ~2.5, as determined by the ^{1}H NMR position of the acetate peak. It should be noted that that stability of this sample was limited (~2 hours) due to the extremely low pH.

The L99A mutant of T4 lysozyme was used to monitor protein foldedness in reverse micelles with changing pH and W_0 detected using NMR and tryptophan fluorescence emission spectroscopy. Aqueous T4 lysozyme (5 mM) was prepared in 50 mM sodium acetate at pH 5 with 50mM NaCl. The protein was then encapsulated in a 75 mM surfactant mixture containing 10MAG, LDAO, and DTAB in a molar ratio of 70:20:10 that had been pre-adjusted to a pH of 5. In order to monitor the effect of pH on protein foldedness, T4 lysozyme was encapsulated at pH 5 at a W_0 of 18. The pH was then adjusted to 3.5 and 2.5 by direct injection of concentrated HCl and monitored by 1-D proton NMR as described above. In order to maintain a constant W_0, the injection-evaporation method was used as described above. To monitor the effect of water loading on protein foldedness, T4 lysozyme was initially encapsulated at pH 5 with a W_0 of 12 at a final protein concentration of 80 µM. The W_0 of the sample was raised to 18 by direct injection of the proper amount of buffer followed by slow inversion and vortexing for 5-10 seconds. All mixtures were allowed to equilibrate without agitation for 30 minutes before data collection.

Optical and Fluorescence Spectroscopies

Optical spectroscopy on cytochrome c was used to observe changes in the absorption of the Soret band upon changing pH. All optical spectra were collected from 200 nm to 800 nm on a Cary 50 Bio UV-Visible Spectrophotometer (Varian/Agilent Technology, Santa Clara, CA). All aqueous spectra were baseline-corrected against buffer while all reverse micelle spectra were baseline-corrected against pentane. The normalized optical absorbance graphs were created with Kaleidagraph (Synergy Software, Reading, PA).

Fluorescence emission spectroscopy on T4 lysozyme (L99A) was used to examine native tryptophan fluorescence of the encapsulated protein as pH was lowered and water loading increased. An excitation wavelength of 297 nm was used for aqueous protein samples while a wavelength of 291 was used for reverse micelle samples. These were the wavelengths of maximum excitation for the samples, respectively. All emission spectra were collected on a Horiba Jobin Yvon (Edison, NJ) Fluorolog-3 from 310 nm to 500 nm with excitation and emission slit widths of 4 nm and 1 nm, respectively. Normalized emission spectra were created with Kaleidagraph.

NMR spectroscopy

NMR data were collected at 25ºC at 500 MHz or 600 MHz (^{1}H) on Bruker AVANCE III spectrometers equipped with TXI cryoprobes. Buffer molecule chemical shifts were determined using ^{1}H 1D NMR spectra with a selective pres-aturation pulse centered at the methyl region of the protonated pentane (~2 ppm) to suppress signal from the alkane solvent. Pre-saturation was not used for samples in 100% d-pentane. ^{1}H spectra were collected using 256 scans, which was required due to the relatively low effective concentration of buffer molecules in the reverse micelle samples (500 µM each as compared to 75 mM surfactant). ^{15}N-HSQC spectra of encapsulated ubiquitin and T4 lysozyme L99A mutant (for the W_0 titration study) were collected with 64 and 100 complex increments, respectively, at 500 MHz. Spectra of encapsulated cytochrome c and T4 lysozyme L99A mutant (for the pH titration study) were collected at 600 MHz with 48 and 100 complex increments, respectively. The cytochrome c spectra were linear-predicted to 64 complex increments. The free (bulk) aqueous protein ^{15}N-HSQC spectra for ubiquitin and cytochrome c were collected with 4 scans, while the reverse micelle spectra were collected with 8 scans. The aqueous spectra for T4 lysozyme L99A mutant were collected with 16 scans, while the reverse micelle spectra were collected with 32 scans except for the reverse micelle sample at pH 2.5 which was collected with 128 scans due to diminished signal to noise. One-dimensional ^{1}H spectra were referenced to dimethyl-silapentane-sulfonate (DSS, Sigma) [180] and processed using Topspin 3.0. All ^{15}N-HSQC spectra were processed using ALNMR [181]. Graphs of chemical shifts and fitting of pH-dependent chemical shift data were performed with Kaleidagraph.

Supplemental Figure

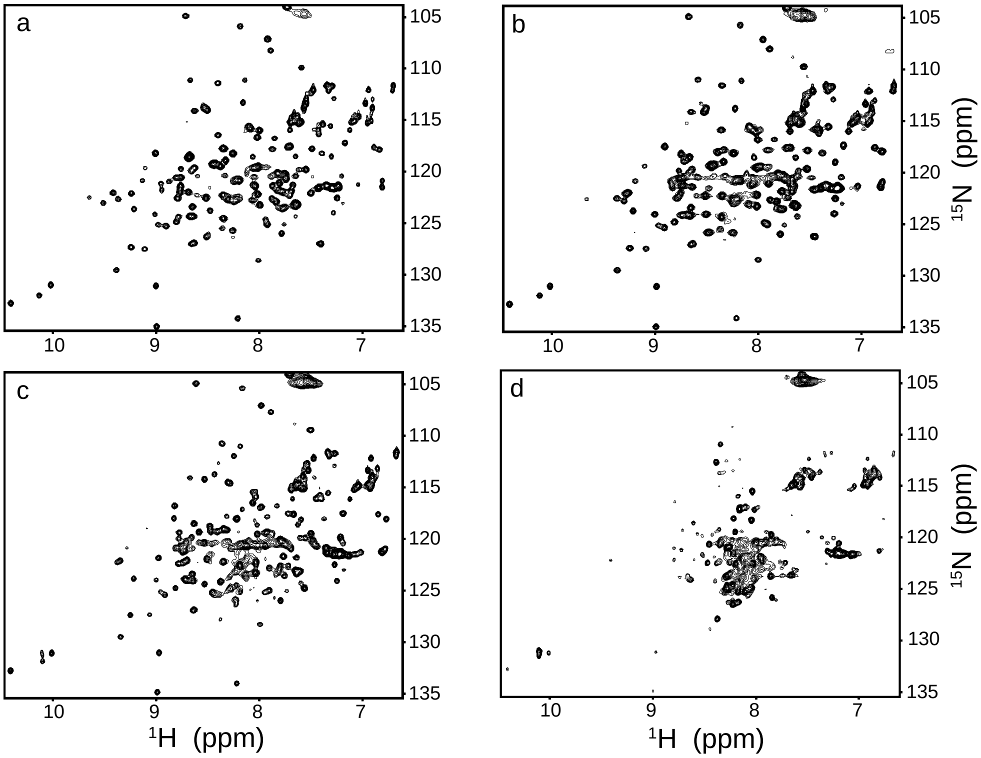

Figure B-1: 15N-HSQC spectra of uniformily 15N-labeled T4 lysozyme L99A mutant in aqueous solution at varying pH to monitor protein foldedness. Aqueous L99A mutant of T4 lysozyme at pH 4.0 (a) is titrated down to a pH of 3.5 (b), 3.0 (c), and 2.5 (d) demonstrating a partial unfolding process as the protein solution becomes more acidic until pH 2.5 at which point lysozyme is completely unfolded. All aqueous protien samples had a final protein concentration of 120µM in a buffer containing 50 mM Sodium Acetate and 50 mM NaCl.

APPENDIX C: Chapter 3

Methods

Protein purification and reverse micelle encapsulation

The gene for wild-type human ubiquitin was cloned into the pET11a expression vector (Genscript) and expressed in BL21(DE3) *E. coli* cells in M9 minimal media. Ubiquitin used for sampling density measurements was grown in 100% H_2O with ^{13}C *d*-glucose and $^{15}NH_4Cl$ as the sole carbon and nitrogen sources for uniformly labeled U- [$^{13}C^{15}N$]-ubiquitin. Ubiquitin used to determine NUS hydration reproducibility were grown in 95% D_2O and $^{15}NH_4Cl$ as the only nitrogen source to yield ~90% ^2H, U- [^{15}N] ubiquitin. The protein was then extracted from inclusion bodies, purified, and refolded as previously described [84]. After purification, the protein was extensively dialyzed against pH 5.0 adjusted water and lyophilized in 1 mg aliquots. Aqueous samples were made with U-[$^{13}C^{15}N$]-ubiquitin resuspended in buffer (50 mM sodium acetate pH 5.0 with 50 mM NaCl) to a final concentration of 1 mM. To create stable reverse micelle samples [76], each aliquot of 90% ^2H, U-[^{15}N] ubiquitin was dissolved in the proper amount of buffer (50 mM sodium acetate pH 5.0 with 50 mM NaCl) in order to create a reverse micelle using the direct injection method [60] with a molar water to surfactant ratio (water loading or W_0) of 10 using 75 mM bis(2-ethylhexyl)sulfosuccinate (AOT) as the surfactant and 99% deuterated pentane (*d-12*) as the solvent. The pH of the AOT was pre-adjusted to 5.0 [128] in order to ensure the aqueous nanopools were at an average pH of 5. All isotopically labeled materials were obtained from Cambridge Isotopes (Andover, MA) and all unlabeled chemicals were purchased from Sigma-Aldrich (St. Louis, MO). Three separate reverse micelle samples were made.

NMR spectroscopy and experimental setup

All aqueous samples were collected at 25°C on a 500 MHz (^{1}H) Bruker AVANCE III spectrometer equipped with a TXI cryoprobe. ^{15}N-resolved, sensitivity-enhanced NOESY-HSQC [182] was collected on U-[^{13}C^{15}N]-ubiquitin with an NOE mix period of 100ms. Each experiment was collected with 32 scans per free induction decay with 24 and 64 complex points in nitrogen and indirect proton dimensions, respectively. All data was collected with a 97.5 ms acquisition time and an interscan recycle delay of 1 sec. Non-uniformly sampled data was collected with both indirect dimensions collected non-uniformly. Sampling schedules were generated using the PoissonGap2 program [120] and were not tested or optimized before use. A total of 8 experiments were collected, 2 uniformely cartesian sampled (US), and 6 non-uniformly sampled (NUS) data sets at 5, 10, 15, 20, 25, 25% sampling density in the indirect dimensions, respectively.

All reverse micelle data were collected at 20°C on a 500 MHz (^{1}H) Bruker AVANCE III spectrometer equipped with a TXI cryoprobe. ^{15}N-resolved, sensitivity-enhanced NOESY-HSQC [24, 115, 182, 183] and ROESY-HSQC [184-186] spectra were collected on uniformly ^{15}N-labeled, ~90% perdeuterated ubiquitin encapsulated in AOT reverse micelles as described above. All three-dimensional (3D) NOESY-HSQC and ROESY-HSQC spectra were collected at 25% Poisson-gap sampling with non-uniformly sampled (NUS) versions of the pulse sequences [119] with 32 scans per free induction decay with 50 and 64 complex points in the nitrogen and indirect proton dimensions, respectively. The

ROESY experiments used a 8.33 kHz continuous wave spin-lock field as previously described [29], For the determination of the reproducibility of NUS-derived NOE/ROE ratios, two reverse micelle samples were used and each sample was collected at a single N(R)OE mix time in duplicate. Additionally, one sample was collected at multiple mix points (20, 40, 60, 80 ms) in order to perform a full NOE buildup curve.

All spectra were processed in NMRPipe [187] using the istHMS reconstruction algorithm for non-uniformly sampled data [120] with a threshold of 98% and 400 iterations. All water (4.6 p.p.m.) cross-peak intensities were determined using Sparky [188]. Signal-to-noise ratios were determined as the max peak intensity divided by the RMS_{NOISE} across the entire spectrum. All further analysis was completed using standard data fitting software or in house Python scripts.

Data fitting

Initial studies by Otting and Wuthrich show that ratio of the cross relaxation rates (σ) of the laboratory (σ_{NOE}) and rotating (σ_{ROE}) frame dipolar interactions can be provide an effective correlation time (τ_H) of a water protein interaction [23, 113]. In the case of a slowly tumbling molecule and assuming no effects from local motion, the limits of the $\sigma_{NOE}/\sigma_{ROE}$ range from -0.5 for long lived waters to 0 for very fast waters [26, 131].

In the linear regime of the NOESY experiment (i.e. no spin diffusion) the signal intensity (I_{NOE}) of the cross peak is proportional to the cross relaxation rates (σ_{NOE}) damped by the auto-relaxation rate (Equation 17) [115]. The auto-relaxation rate is different in the laboratory and rotating frames, and depends on the relaxation of both the protein and water protons involved in the NOE. In order to fit for the true $\sigma_{NOE}/\sigma_{ROE}$ ratio we collect NOESY-

HSQC and ROESY-HSQC spectra at series of mix times. The natural log of the I_{NOE}/I_{ROE} as a function of mix time (τ_{mix}) is fit to a line with the slope equal to the auto-relaxation rates and the intercept equal to the $\sigma_{\text{NOE}}/\sigma_{\text{ROE}}$ as shown in Equation 18.

$$\frac{I_{NOE}(\tau_{mix})}{I_{ROE}(\tau_{mix})} = \left[\frac{\sigma_{NOE} e^{-R_1 \tau_{mix}}}{\sigma_{ROE} e^{-R_1 \rho \tau_{mix}}} \right] \tag{17}$$

$$\ln\left[\frac{\sigma_{NOE} e^{-R_1 \tau_{mix}}}{\sigma_{ROE} e^{-R_1 \rho \tau_{mix}}} \right] = \ln\left[\frac{\sigma_{NOE}}{\sigma_{ROE}} \right] + (R_{1\rho} - R_1)\tau_{mix} \tag{18}$$

We use a standard Pearson R^2 to evaluate the goodness of fit for the buildups. There are several instances where the shortest mix time is missing peaks or have peaks with low signal to noise but are sufficient in longer mix times. In those cases we use a 3-point linear fit and eliminate the first point. Several cases have bad fits, or low signal to noise for all detected peaks. These sites are assigned an $\sigma_{\text{NOE}}/\sigma_{\text{ROE}}$ value of -0.01. Similarly, sites that have an ROE crosspeak but not an NOE crosspeaks are assigned values of -0.01. These sites have detectable hydration meaning they are slowed relative to other regions; however, their hydration is on the fastest end of our detection abilities making the peaks generally weak or non-quantitative.

Structural surface analysis

The trigen and trisrf algorithms [123] were implemented on the structure of ubiquitin encapsulated in AOT reverse micelles (PDB entry 1G6J, conformer 25) in order to create a Van der Waals surface (VDWS) in Cartesian points. This surface was used to map the amide hydration dynamics across the surface of ubiquitin: if a point on the rendered surface was within 4 Å of one or more amide probes, it would adopt the average $\sigma_{NOE}/\sigma_{ROE}$ ratio and would be colored accordingly (see main text for color scale). If a surface point was not within 4 Å of an amide residue, it is colored gray. All molecular images were created using PyMol (Delano Scientific).

Simulations of NOE/ROE ratios

As previously described [27], the protein-water NOE is considered for a protein of solvated radius, b, for N_s solvent spins via the real part of the rank-2 spectral density function. The case considered (case F in reference 27) assumes uniform diffusion, D_T, of water molecules between the surface of the protein and a distance, c, away from the solvated radius beyond which the diffusion of solvent waters is zero. In the scope of this study, the restrictive boundary is formed by the surfactant headgroups. All diffusion constants and distances used are those from a recent simulation [68]. The rank-2 spectral density function is calculated as follows:

$$J^2(\omega) = \frac{3N_S}{D_T b^4(\lambda^{-3}-1)}\frac{1}{\zeta^2}\left(\frac{1-\lambda^3}{3} - \frac{3(1-\lambda^5)}{\zeta^2} - \frac{9}{\zeta^3 Q_3}\times\left\{Q_4 + \lambda^6 Q_5 + \frac{2\lambda^4}{\zeta^2}\right\}\right) \qquad (19)$$

Where $\lambda=b/c$, $\zeta=(i\omega b^2/D_T)^{1/2}$, ω is the spectral frequency, and the Q functions are linear combinations of rank-2 modified spherical Bessel functions (i and k) as follows:

$$Q_3 = i_2^{'}(\zeta)k_2^{'}(\zeta / \lambda) - k_2^{'}(\zeta)i_2^{'}(\zeta / \lambda) \tag{20}$$

$$Q_4 = i_2(\zeta)k_2^{'}(\zeta / \lambda) - k_2(\zeta)i_2^{'}(\zeta / \lambda) \tag{21}$$

$$Q_5 = k_2^{'}(\zeta)i_2(\zeta / \lambda) - i_2^{'}(\zeta)k_2(\zeta / \lambda) \tag{22}$$

The real part of the spectral densities were then used to calculate the effective NOE/ROE with contributions near the protein surface and contributions from all waters within the reverse micelle.

Supplemental Tables

Table C-1: Intra-sample reproducibility dependence on S/N

Min S/N	R^2	RMSD	RMSD%
All	0.83	0.096	14.48%
10	0.92	0.066	9.96%
15	0.97	0.044	6.64%
20	0.96	0.042	6.00%
25	0.97	0.036	5.04
30	0.97	0.036	5.03%

Table C-2: Inter-sample reproducibility dependence on S/N

Min S/N	R^2	RMSD	RMSD%
All	0.54	0.159	23.80%
10	0.6	0.15	22.40%
15	0.77	0.118	17.70%
20	0.82	0.094	14.00%
25	0.78	0.097	14.00%
30	0.77	0.097	14.00%

Table C-3: Hydration dynamics of ubiquitin calculated with buildup method

Residue	NOE/ROE	Fit $T_{1\rho}$ (ms)	Residue	NOE/ROE	Fit $T_{1\rho}$ (ms)
Q2	-0.229±0.016	75.9±7.5	R42	-0.205±0.017	40.4±2.5
F4	-0.497±0.153	69.7±27.2	L43	-0.474±0.055	137.0±39.9
K6	-0.416±0.080	74.8±19.7	A46	-0.062±0.018	31.1±5.1
T7	-0.495±0.062	120.8±33.4	L50	-0.421±0.025	115.4±14.4
L8	-0.486±0.037	99.8±13.7	D52	-0.089±0.045	59.6±28.6
T9	-0.605±0.022	146.9±14.2	L56	-0.444±0.239	45.8±18.1
G10	-0.394±0.062	47.1±6.4	S57	-0.609±0.043	171.9±38.1
K11	-0.500±0.040	153.6±34.5	D58	-0.599±0.023	196.0±26.8
T12	-0.552±0.068	243.3±133.4	Q62	-0.203±0.054	71.9±21.9
T14	-0.455±0.043	235.6±95.2	E64	-0.247±0.007	38.8±0.7
L15	-0.114±0.031	38.1±7.1	S65	-0.260±0.068	62.1±18.5
V17	-0.561±0.016	285.3±42.3	T66	-0.374±0.010	162.2±13.4
E18	-0.113±0.024	52.7±10.6	L67	-0.554±0.062	102.1±21.1
S20	-0.370±0.040	91.2±16.4	H68	-0.521±0.053	111.9±23.2
D21	-0.213±0.057	37.2±6.7	L69	-0.548±0.018	141.8±12.0
T22	-0.459±0.008	173.6±9.2	V70	-0.222±0.126	42.2±16.2
I23	-0.304±0.180	38.4±14.1	L71	-0.421±0.070	146.2±64.8
N25	-0.606±0.045	224.6±67.7	R72	-0.258±0.034	76.4±14.0
K33	-0.415±0.040	72.4±9.3	L73	-0.222±0.042	49.4±8.4
E34	-0.076±0.002	29.1±0.3	R74	-0.115±0.032	37.3±7.0
Q41	-0.046±0.043	26.9±11.0	G75	-0.056±0.014	40.9±7.7

APPENDIX D: Chapter 4

Methods

Protein expression and purification

The coding gene for HEWL was cloned into pET11a expression vector (Genscript) and expressed in BL21(DE3) *E. coli* cells in M9 minimal media as previously described [155]. For all hydration experiments HEWL was expressed with $^{15}NH_4Cl$ as the sole nitrogen source and cells were grown in media containing 95% deuterium. For chemical shift assignment experiments the protein was expressed during growth on with $^{15}NH_4Cl$ and $^{13}C_6$-glucose as described previously [155]. HEWL was purified using a modified version of the previously described protocol [155]. The IPTG-induced cells were harvested and sonicated in a buffer containing 50 mM Tris, 25% Sucrose, 1 mM EDTA, pH 7.5 supplemented with 100 µg/mL deoxyribonuclease I from bovine pancreas (DNAase). Following centrifugation, the pellets were washed with the same buffer, then with 20 mM Tris, 1 % Triton X-100, 1mM EDTA, pH 7.5, then a solution of 2.5% n-Octyl-β-D-Glucopyranoside (BOG), and then twice with water. The protein was then solubilized in 8 M Guanidine HCl, 50 mM Tris, 1 mM EDTA, and 16 mM DTT, pH 8.0 and mixed gently at 37^{o}C for one hour. The protein was then rapidly diluted 8X into 1.25 M Guanidine HCl, 50 mM Tris, 1 mM EDTA, 5 mM oxidized glutathione, pH 8.0 and then further rapidly diluted 20X rapid dilution into 50 mM Tris, 20 mM NaCl, 0.8 mM KCl, 400 mM arginine, 1 mM EDTA, 2 mM reduced glutathione, 0.4 mM oxidized glutathione, pH 8.2. The refolded protein was concentrated and buffer exchanged into 50 mM Tris, 20 mM NaCl, 0.8 mM KCl, and 1 mM EDTA, pH 8.2 to remove the arginine. Non-specifically bound DNA was removed by chromatography on SP-Sepharose equilibrated with 50 mM Tris, 20 mM NaCl, 0.8 mM KCl, 1 mM EDTA, pH 8.2 and eluted with 50 mM Tris, 500 mM NaCl, 0.8 mM KCl, 1 mM EDTA, pH 8.2. The pure protein was then dialyzed against water and lyophilized in 2 mg aliquots for further use.

Protein encapsulation

Lyophilized HEWL was prepared in 20 mM sodium acetate, 10 mM NaCl, pH 5.4. For the inhibitor-bound reverse micelle samples, the lyophilized aliquots were dissolved in the same buffer with 1.2 molar equivalents of chitotriose. All pentane reverse micelle samples were made with 75 mM cetyltrimethylammonium bromide (CTAB) and 450 mM hexanol and dissolved in 99% deuterated pentane (*d-12*) at a molar ratio of water to surfactant (termed water loading or W_0) of twelve. Protein was encapsulated using the direct-injection method [60] using 6 mM stock solutions of HEWL.

NMR Spectroscopy

All ^{15}N-resolved, sensitivity-enhanced [182] water dynamics experiments: NOESY-HSQC [183, 189-191] and ROESY-HSQC [129, 185, 186] were collected on the uniformly ^{15}N-labeled, 95% deuterated HEWL on a 500 MHz (^{1}H) Bruker AVANCE III spectrometer equipped with a TXI cryoprobe. For both the apo- and inhibitor-bound reverse micelle samples, the three-dimensional NOESY and ROESY spectra were collected with the same resolution in the indirect ^{15}N and ^{1}H dimensions. The NOESY and ROESY spectra were collected with 128 scans per free induction decay. All experiments were collected at 25 ^{o}C. The ROESY experiments used a 8.33 kHz continuous wave spin-lock field as previously described [29]. A mixing time (τ_{mix}) of 35 ms was empirically determined to fall in the linear regime for the NOE. A recycle delay of 1.2 s was used throughout.

The NOESY and ROESY experiments were collected at 25% Poisson-gap sampling with non-uniform sampling (NUS) versions [119, 120] of the pulse sequences and processed in NMRPipe [187]. Peak intensities were obtained for the NOESY and ROESY spectra at the water planes in the indirect proton dimension of the apo- and inhibitor-bound states (4.5 p.p.m.) using Sparky [188]. NOE/ROE ratios were calculated for amide sites by dividing the NOE and ROE peak intensities, respectively, for both the apo and chitotriose-bound states (see Table D-1). Detected NOE/ROE probes were omitted for amide hydrogens that were within 4.0 Å of hydroxyl hydrogens of tyrosine residues. This eliminated two probes in the apo state and three probes in the inhibitor-bound state.

HNCO and HNCA [192] triple resonance spectra were collected for the free state in order to confirm amide chemical shift assignments upon reverse micelle encapsulation. Chemical shifts varied very minimally between the previously published chemical shift assignments [154]. These spectra were processed with FELIX (Accelrys). HNCA and HNCOCA [193] backbone pair experiments were collected on the inhibitor-bound state encapsulated in pentane. These experiments were collected at 10% Poisson-gap sampling with NUS versions [119, 120] of the pulse sequences and processed in NMRPipe [187].

Structural surface analysis

As previously described [29], the average burial depth of each amide probe was determined using the Travel Depth program [150, 151] on each of 51 HEWL structures within the ensemble on NMR structures represented in PDB entry 1E8L [154]. The same program was also used on PDB entry 3LYZ to construct a Van der Waals surface (VDWS) in Cartesian points that was used for all structural analyses in the apo- and inhibitor-bound states. By concurrently using the maximum detectable distances from the intramolecular NOEs determined from the NOESY spectra and the burial depth analysis of the NMR ensemble, it was determined that a distance of 4 Å was used as a distance cutoff between the points on the VDWS and the amide probes. If a point on the rendered surface was within 4 Å of one or more amide probes, it would adopt the average NOE/ROE ratio of those amides and would be colored accordingly (see Chapter 4 for color scale). Points outside the distance cutoff of any amides were not assigned a NOE/ROE value and are colored gray. Molecular images were created using PyMol (Delano Scientific).

Supplemental Figures

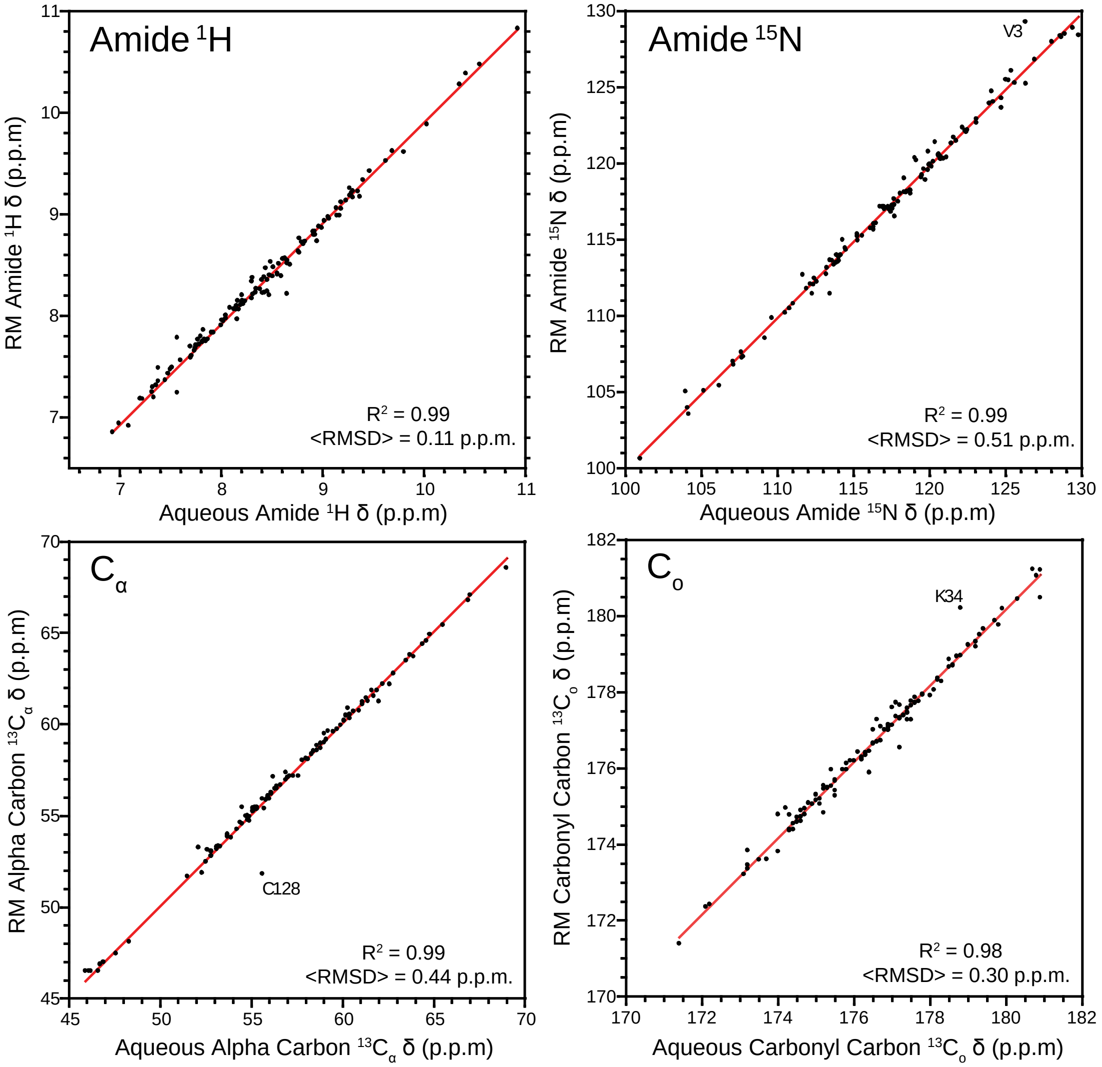

Figure D-1: Comparison of the backbone chemical shifts of aqueous and reverse micelle encapsulated HEWL at pH 4.7 demonstrates excellent structural fidelity upon reverse micelle encapsulation. Major outliers in the fit are labeled.

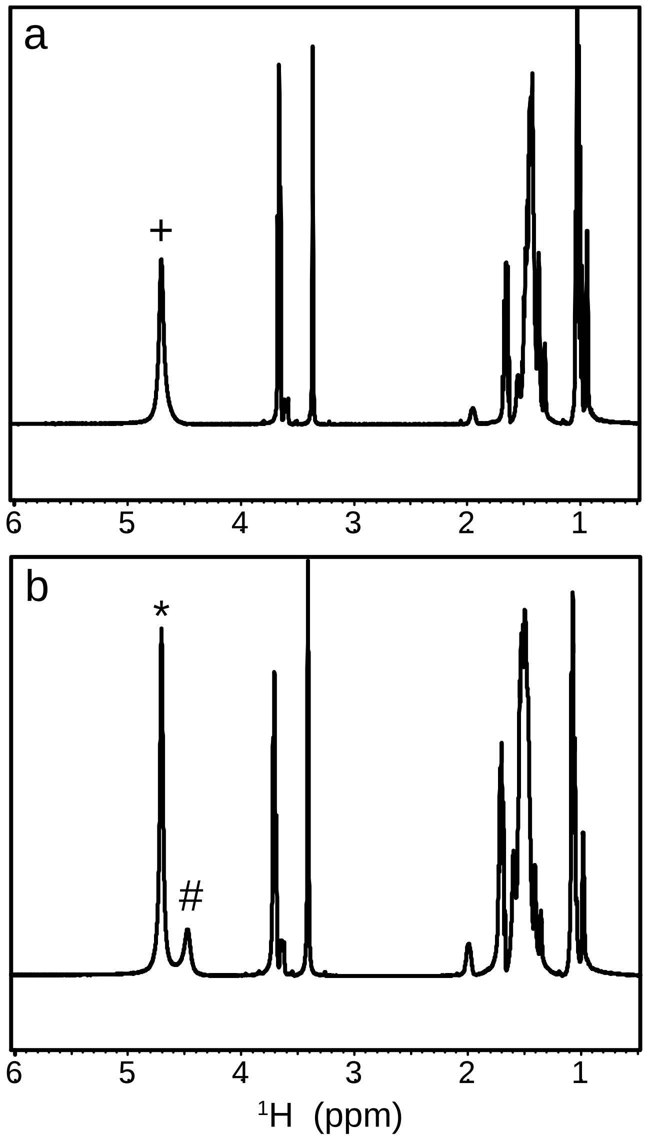

Figure D-2: 1-D 1H-spectra of hen egg-white lysozyme encapsulated in reverse micelles in pentane at an ensemble pH of 4.7 **(a)** and 5.4 **(b)**. Upon raising the pH of the reverse micelle ensemble, the hexanol and water peaks go from coalesced (+) to separate water (*) and hexanol (#) peaks. 1-D spectra were processed in NMRPipe.

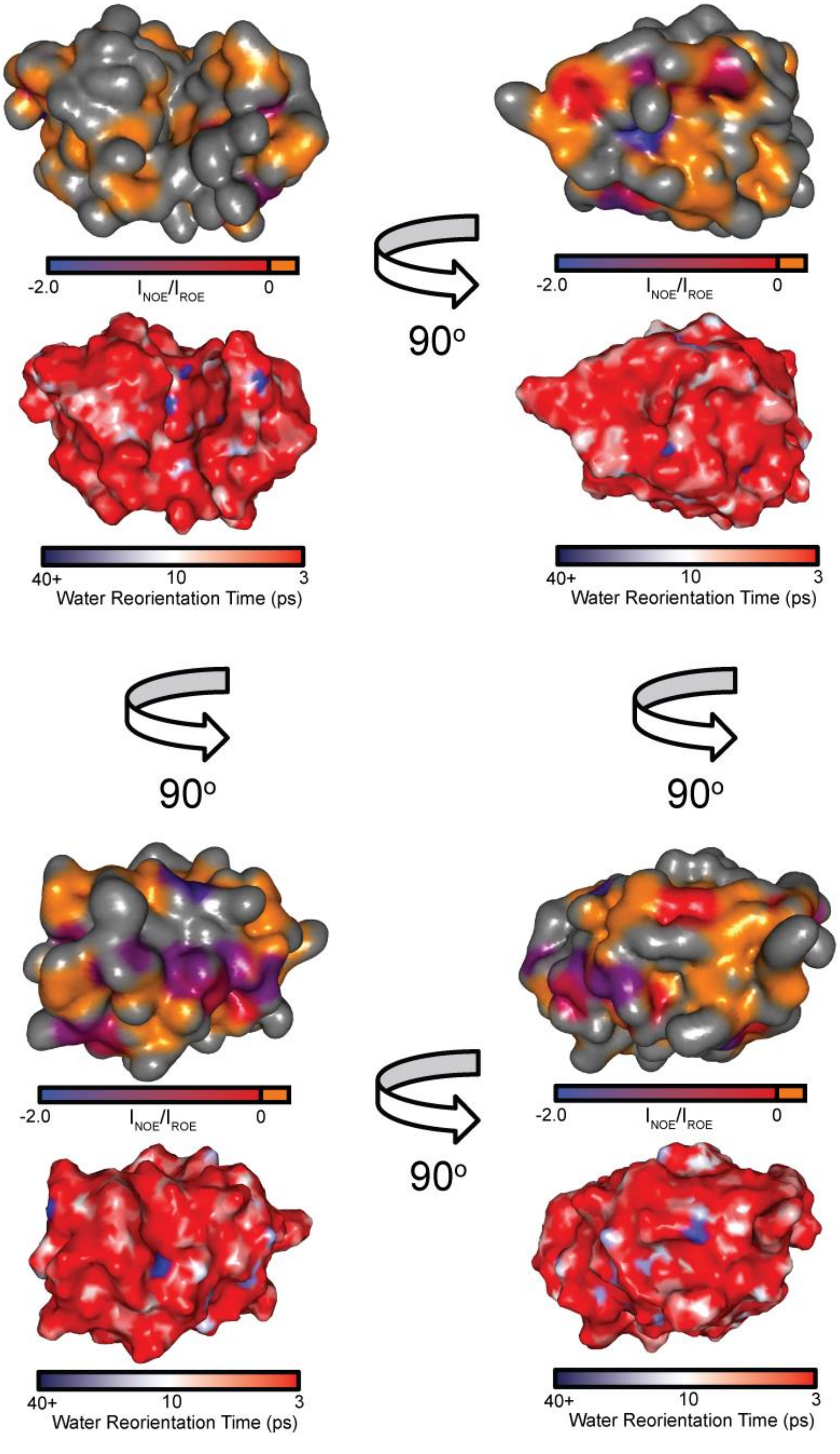

Figure D-3: Little correlation between NOE/ROE hydration dynamics measurements and water reorientation time simulations. Four views of 90o rotations of apo HEWL. The top molecule of each pair depicts the hydration dynamics measurements mapped to the surface as described in the Materials and Methods and shown in Figure 2 of the main text. The bottom molecule of each pair is a re-adaptation of the data presented in Figure 6 of the study by Fogarty and Laage [14]. The water reorientation times assigned in that study were mapped to the surface of HEWL used in this study (PDB entry 3LYZ) with red representing relatively fast hydration dynamics (~3 ps), white representing intermediate hydration dynamics (~10 ps), and blue representing slow hydration dynamics (40 ps or longer). As described in the study by Fogarty and Laage, the area with the most concentrated slow hydration dynamics in the simulation was located in the concave peptidoglycan-binding region (view of the top left pair) while there is very little slow hydration dynamics detected in the binding cleft using the NOE/ROE method described in this study. Visually, there is little correlation between the measured and simulated hydration dynamics.

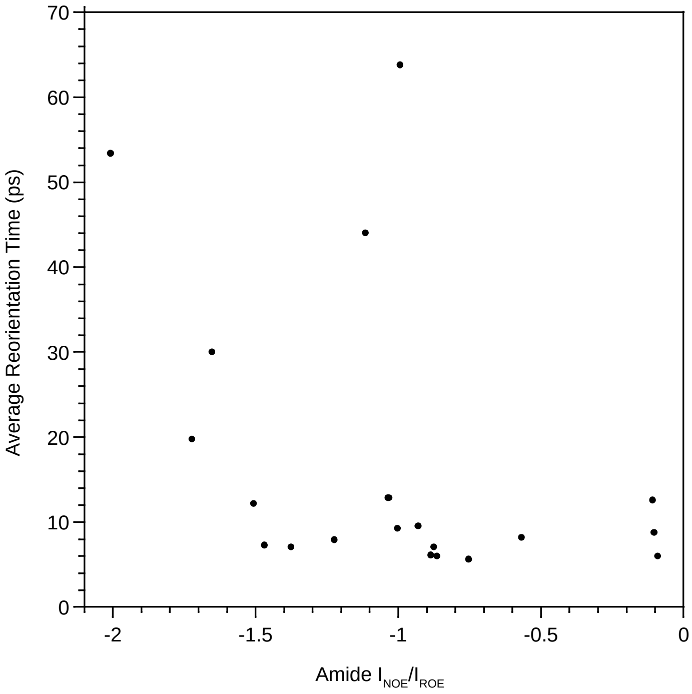

Figure D-4: Comparison of amide NOE/ROE hydration ratios to average, nearby simulated water reorientation times as described by Fogarty and Laage. For each amide hydrogen at which a NOE/ROE hydration ratio was detected (see Table S1 below), the simulated water reorientation times of atoms within 4.0 Å of the amide hydrogen were averaged together and plotted above. There is no correlation between the NMR-detected hydration dynamics and the simulated reorientation times near the amide probes.

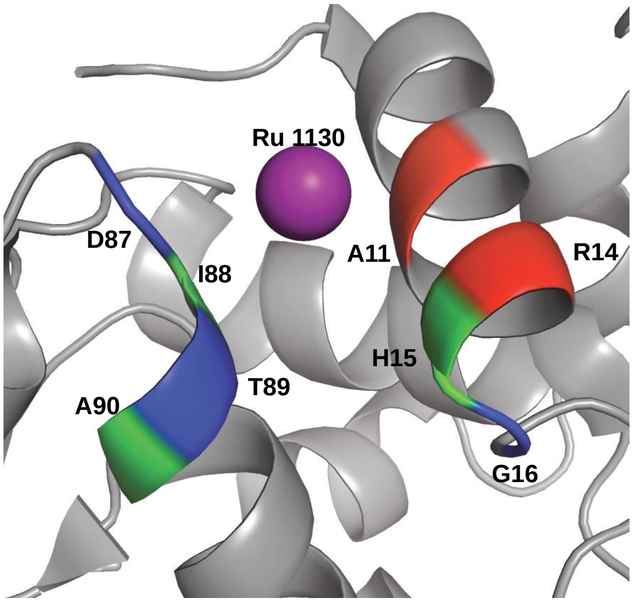

Figure D-5: Slow hydration dynamics detected at surface amides in the same area as the area described by King et al. (see Chapter 4). In purple is the ruthenium atom of the HEWL-RC complex that is attached to His15 (PDB entry 2XJW). The amides colored in blue are surface-exposed amides within 10 Å of the ruthenium atom with detectable hydration ratios: G16, D87, and T89 with NOE/ROE ratios of -1.222, -0.927, and -1.030, respectively, in the apo state. The amides colored in green are surface-exposed amides within 10 Å of the ruthenium atom, but with no detectable NOE/ROE ratios because of ambiguity due to spectral overlap in the RM-encapsulated spectra. The amides in red are those within 10 Å of the Ru atom that have no detectable hydration ratios for structural reasons: with our method it is often difficult to detect hydration ratios for amides located within α-helices.

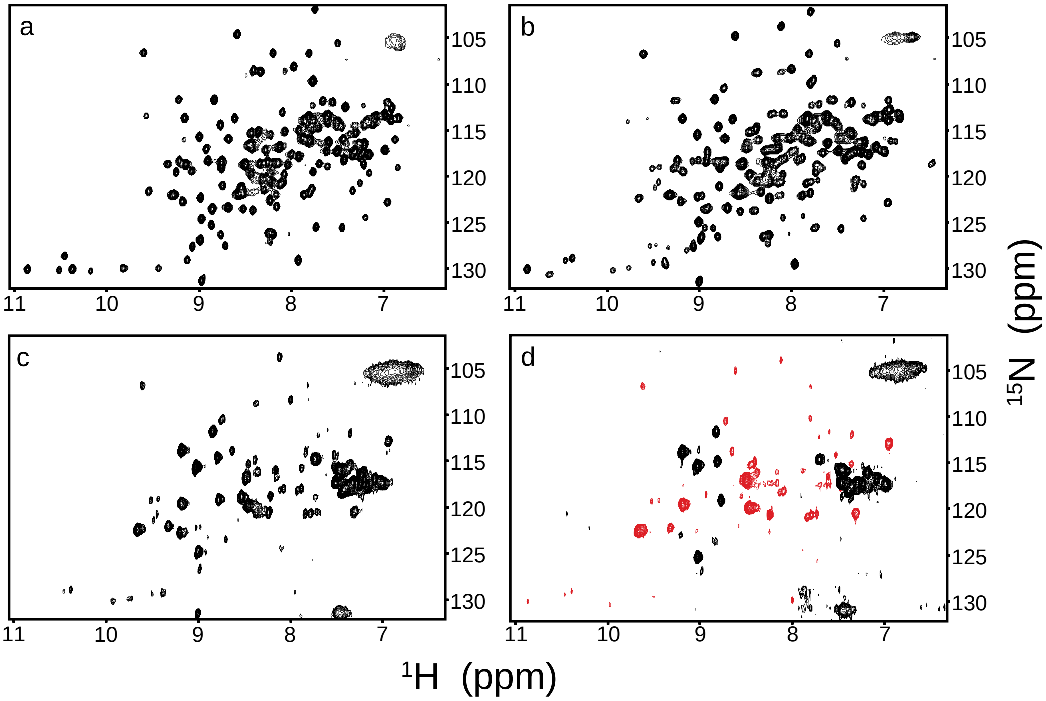

Figure D-6: Many NOE and ROE peaks are visible upon encapsulation of chitotriose-bound HEWL in reverse micelles. ^{15}N-labeled, 95% deuterated HEWL was encapsulated in the described CTAB/hexanol mixture at pH 5.4 in pentane in the apo state **(a)** and in the chitotriose-bound state **(b)**. The spectra have minimal changes in chemical shift of the amides as can be expected from the small structural changes upon binding of the inhibitor. ^{15}N-resolved NOESY **(c)** and ROESY **(d)** spectra are shown in the plane of the indirect water resonance (4.5 p.p.m.) for the chitotriose-bound state. In all cases black peaks represent positive intensity and red peaks represent negative intensity. Black, positive peaks in the ROESY **(d)** spectra represent contamination of some amide probes and side chains by exchange.

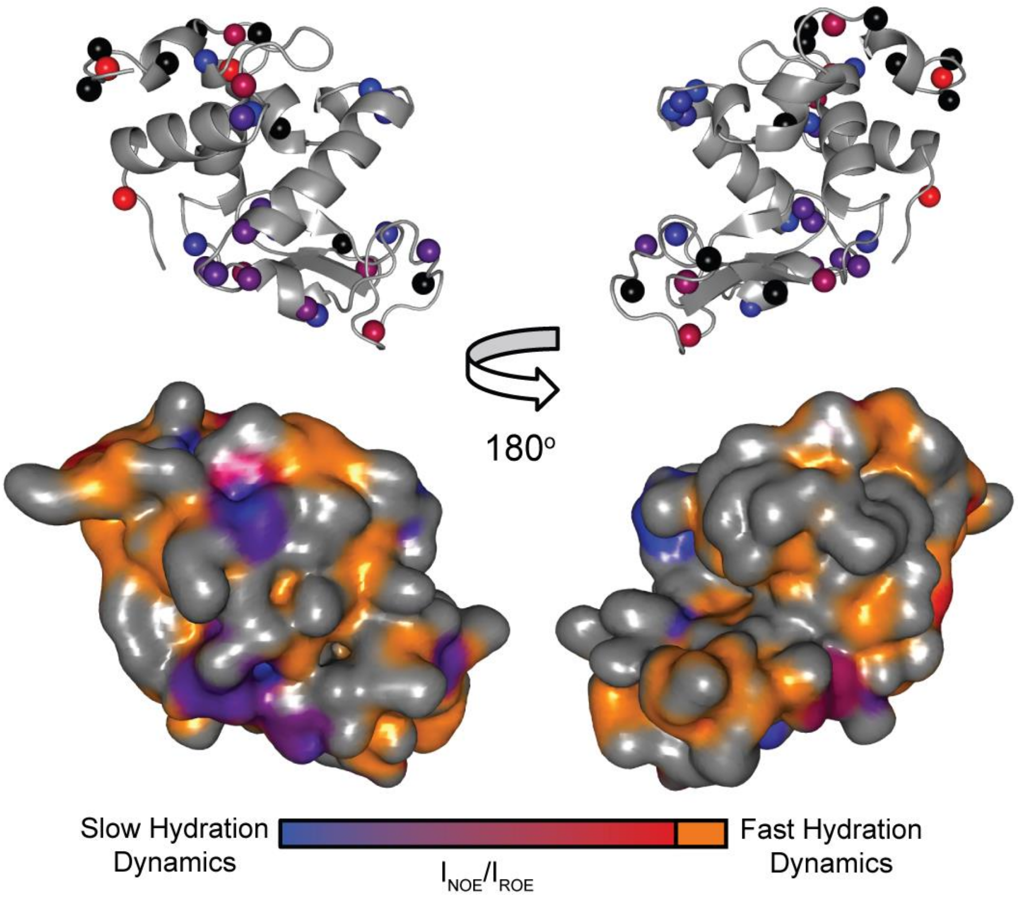

Figure D-7: There are some changes in hydration dynamics at the surface of HEWL upon binding of chitotriose. Top: colored spheres plotted on two views of the structure of HEWL (PDB entry 3LYZ) represent amide hydrogens detected in the suite of hydration experiments. The spheres are color coded according to the relative hydration dynamics of waters at that probe from blue spheres representing slow waters (high retention time) to red spheres representing fast waters (lower retention time). Bottom: surface of HEWL generated with the Travel Depth program as described in the methods. The color scheme used on the Cartesian points across the generated surface is the same as used above, and in addition, orange represents points on the surface within 4 Å of amides that were not detected in the NOESY/ROESY experiments (the waters near these amides were too fast to detect) and black represents amides that were contaminated by exchange (detected by positive ROEs). All protein images were generated with Pymol.

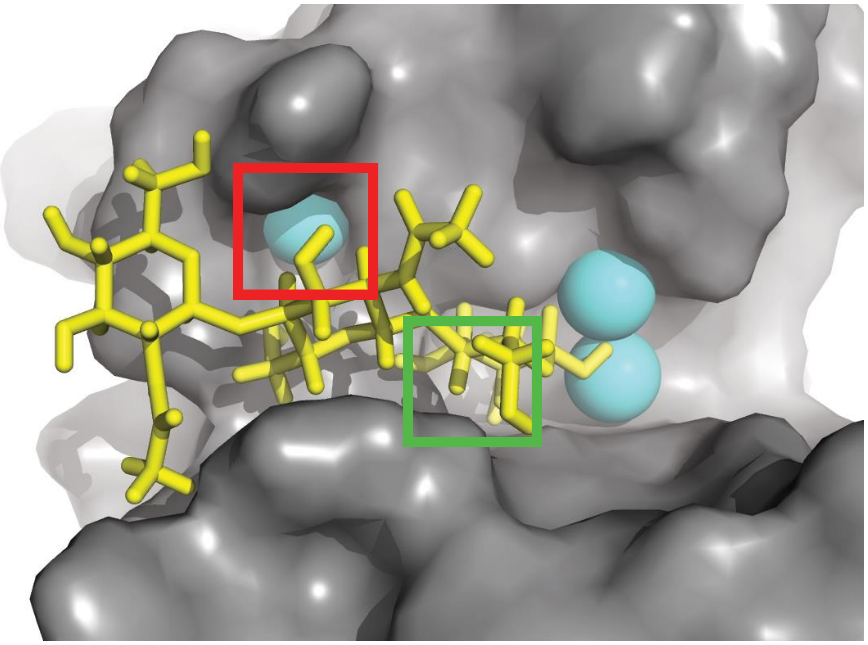

Figure D-8: Crystal structure of chitotriose-bound HEWL showing interfacial waters. The structure of chitotriose-bound HEWL (PDB entry 1HEW) with the surface of HEWL depicted in gray, the stick structure of chitotriose shown in yellow, and the published crystallographic waters shown in cyan. The red and green boxes represent areas at the surface of HEWL where hydration dynamics is slowed upon binding of inhibitor. The red and green color scheme of the boxes matches that of the scheme used in Fig. 4 of the main text. The protein image was generated with Pymol.

Supplemental Table

Table D-1: Hydration dynamics data for HEWL (apo- and chito-)

Residue	σNOE/σROE (Apo HEWL)	σNOE/σROE (Chito-HEWL)
G4	-0.010 ± 0.102	-0.010 ± 0.050
G16	-1.222 ± 0.176	-1.223 ± 0.103
L17	-	-2.439 ± 0.105
D18	-0.010 ± 0.107	-0.502 ± 0.057
S24	-	*
G26	-1.721 ± 0.132	-1.631 ± 0.086
N27	-	-0.010 ± 0.081
T40	-	-1.485 ± 0.067
E41	-	-1.073 ± 0.073
R45	-	*
T47	*	*
G49	*	-
S50	-1.504 ± 0.104	-0.710 ± 0.007
I55	-0.992 ± 0.013	-1.169 ± 0.009
L56	-1.112 ± 0.021	-1.155 ± 0.013
Q57	-	-3.568 ± 0.245
W62	-1.034 ± 0.084	-1.524 ± 0.081
R68	-1.651 ± 0.126	-0.448^{+} ± 0.069
G71	*	*
R73	-1.233 + 0.081	-
N74	-0.861 ± 0.022	-1.118 ± 0.050
S81	-1.374 ± 0.103	-1.519 ± 0.056
A82	-0.565^{+} ± 0.097	-0.927^{+} ± 0.127
L84	-1.467^{+} ± 0.179	-
S85	-0.750 ± 0.017	-0.976 ± 0.017
D87	-0.927 ± 0.009	-1.154 ± 0.017
T89	-1.030 ± 0.025	-1.054 ± 0.006
S91	-2.006 ± 0.161	-2.638 ± 0.191
D101	-	-1.450 ± 0.055
G102	-	-1.895 ± 0.092
N103	-	-1.517 ± 0.058
G104	-	-1.940 ± 0.094
W111	-	*
G117	*	*
T118	-0.999^{+} ± 0.131	-0.524^{+} ± 0.058
D119	*	*
A122	-0.874 ± 0.036	-
W123	-	*
G126	*	*
C127	-0.010 ± 0.088	-0.010 ± 0.035
R128	*	*

- Indicates there is no hydration detected for that particular amide probe.
* Indicates amide probe contaminated by exchange (positive ROEs)
+ Indicates probes whose signal to noise in the NOESY or ROESY were below 18 and were assigned a binned NOE/ROE ratio (see text)

Propagation of error in NOE/ROE ratio

In order to calculate the systematic error of the hydration ratio within the same experiment, the error propagation below was used:

$$R = \frac{NOE}{ROE} \tag{23}$$

The error in R is:

$$\Delta R = |R| \sqrt{\left(\frac{N_{NOE}}{NOE}\right)^2 + \left(\frac{N_{ROE}}{ROE}\right)^2} \tag{24}$$

Where $N_{(R)NOE}$ is the noise inherent to the (R)NOESY spectrum as determined using SPARKY [188].

APPENDIX E: Chapter 5

Methods

Sample preparation

Isotopically labeled ubiquitin was prepared as before [84] using the following labeling schemes: uniformly ^{15}N-labeled for backbone relaxation; and selective *meta*-^{13}C-H aromatic labeling for aromatic carbon relaxation [194]. Aqueous and water/glycerol samples were prepared at 1 mM total protein concentration with 50 mM sodium acetate, pH 5.0, and 50 mM sodium chloride. Samples were prepared using mixtures of appropriately labeled ubiquitin such that backbone measurements could be carried out on each sample to determine the optimal tumbling model for calculation of order parameters. The ubiquitin reverse micelle sample was prepared as before [76] in 75 mM bis-2-ethylhexylsulfosuccinate (AOT) at a w_0 (molar water:surfactant ratio) of 10 in perdeuterated hexane.

For all methyl dynamics experiments, isotopically labeled ubiquitin, β-cyclodextrin-bound MBP, and MSG were prepared as previously described[57] with selective $^{13}CH_3$ methyl labeling [195] for isoleucines (δ1 only), valines, leucines, and methionines (for MSG only) with background ^{12}C, deuterium, and ^{15}N labeling throughout. Aqueous/glycerol samples were prepared as before [57] at concentrations of 1 mM, 750 µM, and 800 µM, respectively (deuterated-*d11* tris was used as the buffer for MSG to prevent streaking in the methyl spectra). Reverse micelle samples for MBP and MSG were prepared with 150 mM deuterated cetyltrimethylammonium bromide (CTAB-*d42*) and 800 mM deuterated hexanol (*d13*) in perdeuterated pentane at w_0 of 15 and 22, respectively.

NMR experiments

For ubiquitin NMR experiments were performed at 11.7, 14.1, or 17.6 T using Bruker Avance III spectrometers equipped with cryoprobes. Backbone and aromatic relaxation data were collected at 20 °C unless otherwise noted, and all measurements were recorded with nine relaxation times (three in duplicate for error estimation) with the exception of ^{15}N T_2 measurements in 50% glycerol for which six relaxation times (two in duplicate) were measured. Backbone ^{15}N-H relaxation experiments were collected using either standard or TROSY-based [196] (50% glycerol condition only) T_1, T_2, and heteronuclear ^{1}H-^{15}N NOE pulse sequences. All aromatic relaxation measurements were performed at two magnetic field strengths using ^{13}C-H T_1 and $T_{1\rho}$ relaxation pulse sequences implemented as pseudo-2D experiments. Experiments were collected at two magnetic field strengths to provide a minimum of five observables for each backbone site and a minimum of four observables for each aromatic probe.

For all methyl relaxation data NMR experiments were collected at 14.1 and 17.6 T as described above at 25°C. Methyl order parameters were obtained via intra-methyl proton-proton cross-correlated spin relaxation experiments which were collected at one field (14.1T for ubiquitin and 17.6 T for MBP and MSG) for all samples and fit using in-house python scripts as previously described [43]. It should be noted that for the reverse micelle and glycerol samples these experiments were converted from States-TPPI to gradient-selected echo anti-echo quadrature selection to minimize streaking from the background ~2% protonated solvents and (for ubiquitin only) from the protonated AOT surfactant molecules. In order to determine molecular tumbling times, backbone TROSY-based T_1 and $T_{1\rho}$ experiments were collected at two fields per sample as described above. The fitted relaxation rates were then used to estimate tumbling times as previously described [197]. All NMR experiments were processed with NMR Pipe [187].

Data analysis

Exponential relaxation decay fitting and model-free calculations were performed using in-house software [162]. Optimal tumbling model (isotropic, axially symmetric, or anisotropic) and rotational correlation times were calculated [198] for each sample via the boundary element approach. Order parameters and τ_e values were determined using a grid search approach [199], and errors were estimated using Monte Carlo. Details of these calculations are provided elsewhere.

Methyl O^2_{axis} values are reported using an assumption of tetrahedral geometry ($O^2/0.111$). All degenerate methyl peaks were removed for all analyses. As demonstrated in Figure E-4, for the ubiquitin and MSG reverse micelle samples there was still extensive streaking from background protonated solvent (~2%) and from protonated surfactant molecules (ubiquitin only). This lead to significant outliers when comparing the methyl order parameters collected in the reverse micelle to those collected in aqueous solution. In effect, all probes within 40 Hz of the streaks in the first delay of the triple quantum cross-correlated methyl dynamics experiment (indicated by green dashed lines in Figure E-4 a and c) were removed (red circles in Figure E-4) leading to excellent correlations of methyl order parameters. MBP was not subject to these streaking restrictions (Figure E-4b) due to its extremely high encapsulation efficiency and the lack of protonated surfactant.

Backbone order parameters shown in the main text are the average of at least two replicate samples for all conditions except at 50 °C for which a single sample was measured. All methyl and all aromatic order parameters are reported from a single sample. Burial of methyl groups for comparison with hydration data and methyl order parameters was determined using Depth 2.0 [150] on reverse micelle-encapsulated ubiquitin structural ensemble 1G6J [76] (conformer 25 of 32), β-cyclodextrin bound MBP (1DMB [200]) and apo MSG (2JQX [201]). All structural images were created using PyMOL. Mapping of the hydration surface was performed by assigning each surface point a $\sigma^{NOE}/\sigma^{ROE}$ value corresponding to the average $\sigma^{NOE}/\sigma^{ROE}$ of all hydration probes (from previous measurements) within NOE distance (4.5 Å) of the surface point.

Supplemental Figures

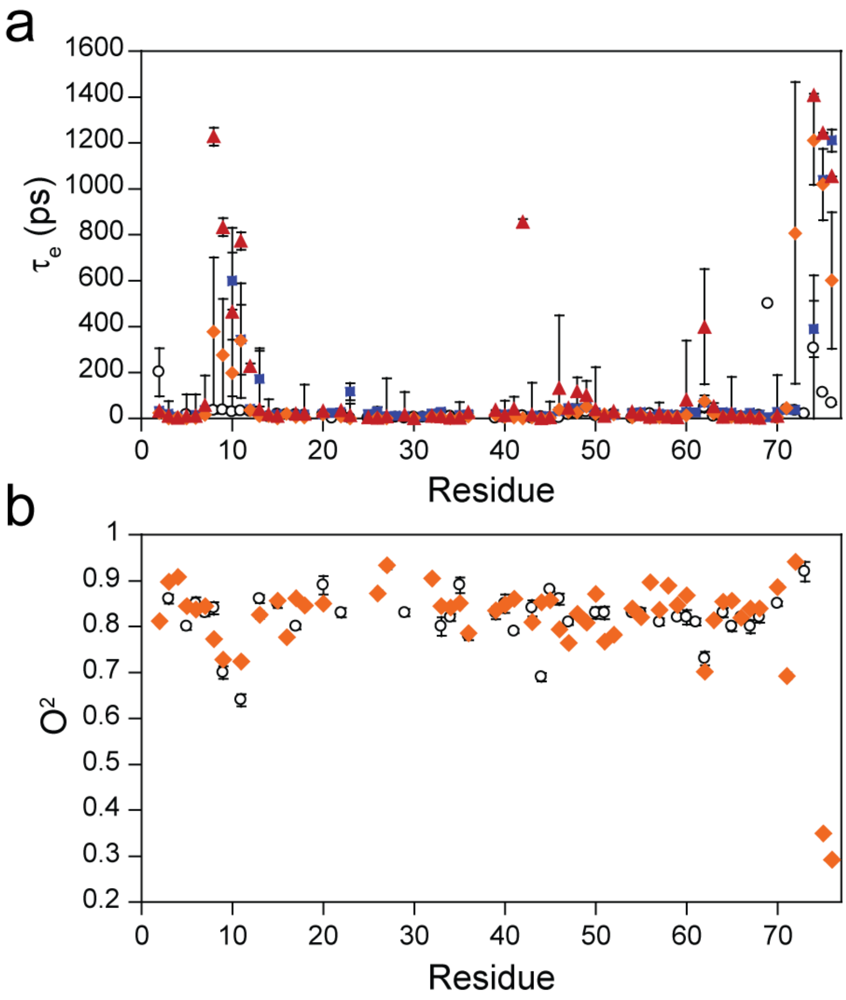

Figure E-1. Supplemental data on backbone dynamics of ubiquitin. **(a)** Correlation times (τ_e) for the backbone N-H bond vector motions of ubiquitin at 20 °C are shown under various solvation conditions represented as follows: aqueous – open circles, AOT reverse micelles – blue squares, 30% glycerol – orange diamonds, 50% glycerol – red triangles. Backbone order parameters (O^2) of ubiquitin at 50 °C are shown **(b)** for aqueous (open circles) and 30% glycerol (orange diamonds).

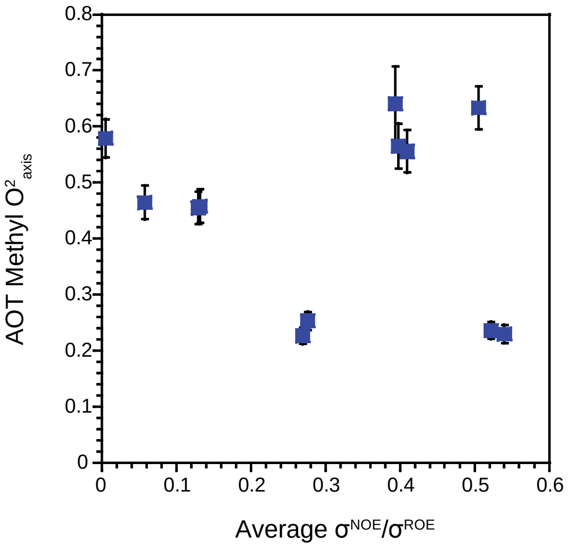

Figure E-2. Comparison of methyl dynamics with hydration dynamics. Methyl O^2_{axis} values for encapsulated ubiquitin are plotted against the average hydration dynamics ($\sigma^{NOE}/\sigma^{ROE}$) of the nearby surface. Fast hydration water corresponds to a $\sigma^{NOE}/\sigma^{ROE}$ value of 0; slow hydration water corresponds to a $\sigma^{NOE}/\sigma^{ROE}$ value of -0.5, but absolute values of $\sigma^{NOE}/\sigma^{ROE}$ were used here for simplicity. Hydration data from backbone amide-resolved measurements were mapped to the Cartesian coordinates of the solvent-accessible surface from the representative conformer of the encapsulated ubiquitin structural ensemble (1G6J) [76]. Only methyl groups within 4.5 Å (the upper limit of the hydration measurement) of the surface were included in this analysis. Each methyl group was assigned an effective $\sigma^{NOE}/\sigma^{ROE}$ value corresponding to the average $\sigma^{NOE}/\sigma^{ROE}$ of all surface points within 4.5 Å of the methyl carbon.

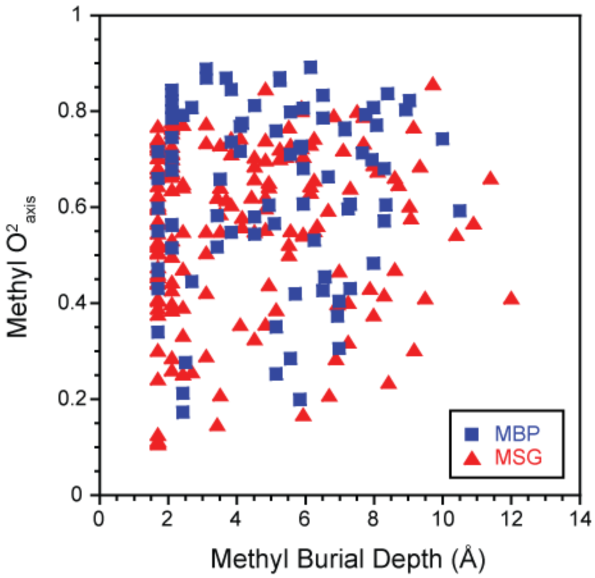

Figure E-3. Methyl order parameters are not dependent on probe burial depth. The methyl order parameters (O^2_{axis}) of aqueous MBP (blue squares) and MSG (red triangles) are plotted against the methyl burial depth as determined by the Depth 2.0 program [150]. There is clearly no correlation between methyl order parameter and probe burial depth, further demonstrating that protein conformational entropy is not slaved to solvent.

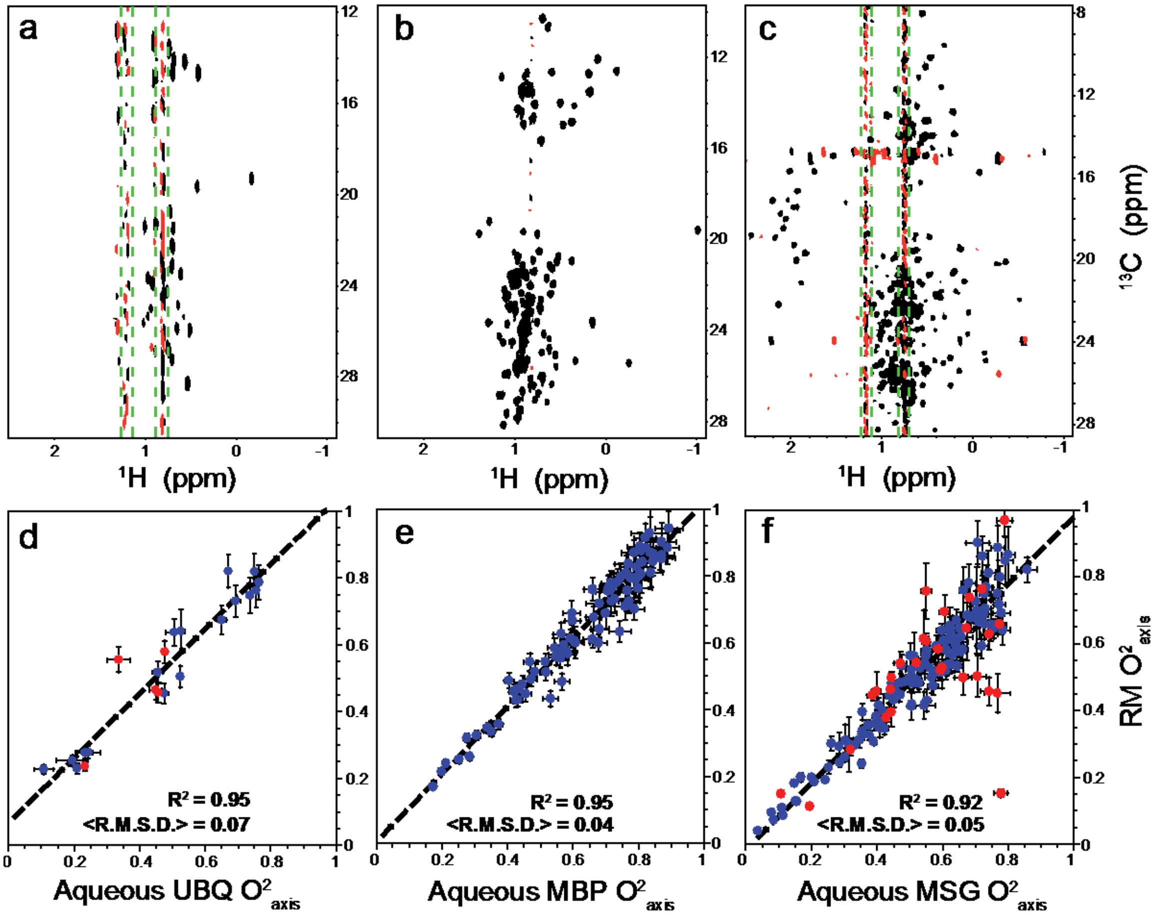

Figure E-4. Streaking in the methyl region of reverse micelle samples is caused by background solvent protonation and protonated surfactant molecules (ubiquitin only) and leads to significant outliers in methyl order parameter correlations. **(a-c).** ^{1}H-^{13}C HMQC correlation spectra of the first delay of the triple-quantum cross-correlated methyl dynamics experiment are shown for ubiquitin, MBP, and MSG, respectively. Significant streaking can be observed in the proton dimension at ~1.2 p.p.m. and ~0.8 p.p.m for the ubiquitin and MSG spectra. The stellar encapsulation efficiency and lack of protonated surfactant molecules in the MBP reverse micelle sample lead to no significant streaking (relative to methyl peak signal-to-noise) from background methyl protons. **(d-f).** Reverse micelle-encapsulated methyl dynamics are plotted against the aqueous methyl dynamics for ubiquitin, MBP, and MSG, respectively. Red circles are methyl probes within ~40 Hz of streaks in the proton dimension (green dashed lines in **a** and **c**) and are excluded statistical analyses. Significant improvements in R^2 are observed upon removal of the streak-contaminated peaks (from 0.91 to 0.95 for ubiquitin and from 0.80 to 0.92 for MSG).

Supplemental Tables

Table E-I: Order parameters and timescales of aromatic side chain motion for ubiquitin in solvation conditions of varying viscosity [202].

			O^2			τ_e (ps)		
Bulk Solvent	°C	η (Pa-s)	F4	F45	Y59	F4	F45	Y59
Water	20	0.0010	0.69 (0.05)	0.73 (0.085)	0.9 (0.031)	99 (5)	59 (20)	42 (27)
30% Glycerol	20	0.0030	0.77 (0.012)	0.86 (0.051)	0.9 (0.061)	73 (10)	220 (100)	334 (198)
50% Glycerol	20	0.0084	0.83 (0.021)	0.92 (0.034)	0.96 (0.007)	322 (68)	680 (355)	774 (42)
Pentane[2] (RM)	20	0.0002	0.68 (0.005)	0.85 (0.029)	0.93 (0.006)	100 (11)	157 (23)	86 (27)
Water[3]	50	0.0005	0.51 (0.002)	0.38 (0.015)	0.58 (0.014)	456 (4)	572 (2)	376 (4)
30% Glycerol	50	0.0014	0.87 (0.003)	0.92 (0.002)	1 (0.002)	973 (8)	980 (3)	1134 (33)

[1] Viscosities for water and pentane are as reported in the NIST webbook. For glycerol solutions, viscosities were calculated based on the volume ratio of glycerol to water[203].

[2] While pentane is the bulk solvent for the reverse micelle condition, it is important to note that the local viscosity *inside the reverse micelle* is comparable to the high viscosity glycerol solutions.

[3] Aqueous data at 50 °C [39]

Table E-2: Backbone order parameters for ubiquitin in various solvent conditions

Residue Number	AQ O^2	AOT O^2	30% Glycerol O^2	50% Glycerol O^2
2	0.877±0.037	0.861±0.003	0.852±0.026	0.860±0.008
3	0.883±0.002	0.890±0.006	0.892±0.004	0.878±0.011
4	0.906±0.010	0.915±0.002	0.904±0.002	0.883±0.001
5	0.824±0.002	-	0.836±0.004	0.852±0.008
6	0.862±0.003	0.907±0.006	0.863±0.011	0.844±0.015
7	0.844±0.008	0.867±0.017	0.877±0.015	0.925±0.011
8	0.801±0.001	-	0.858±0.009	0.874±0.019
9	0.727±0.001	-	0.838±0.009	0.841±0.013
10	0.751±0.005	0.886±0.006	0.870±0.012	0.827±0.047
11	0.731±0.004	0.869±0.019	0.856±0.024	0.841±0.008
12	0.747±0.004	0.838±0.012	0.806±0.005	0.888±0.036
13	0.866±0.003	0.933±0.041	0.855±0.010	0.879±0.036
14	0.837±0.000	0.886±0.002	0.850±0.011	0.828±0.006
15	0.863±0.004	0.870±0.007	0.871±0.022	0.858±0.016
16	0.782±0.003	-	0.807±0.002	-
17	0.896±0.021	0.902±0.002	0.885±0.016	0.819±0.005
18	0.862±0.012	0.870±0.005	0.861±0.016	0.862±0.052
20	0.823±0.004	0.861±0.007	0.825±0.051	0.868±0.006
21	0.911±0.007	0.945±0.010	-	-
22	0.849±0.002	0.869±0.006	0.875±0.005	-
23	0.995±0.003	0.991±0.004	0.944±0.024	0.921±0.003
25	0.925±0.020	0.884±0.018	0.937±0.036	0.922±0.012
26	0.864±0.001	0.898±0.002	0.873±0.017	0.913±0.010
27	0.945±0.007	0.933±0.007	0.894±0.026	0.934±0.017
28	0.912±0.003	0.939±0.006	-	-
29	0.891±0.001	0.902±0.007	0.928±0.021	0.970±0.018
30	0.890±0.010	-	0.906±0.009	0.839±0.068
31	0.912±0.007	0.936±0.004	0.929±0.019	-
32	0.912±0.002	0.917±0.003	0.926±0.003	0.933±0.017
33	0.853±0.012	0.876±0.001	0.889±0.010	0.862±0.015
34	0.829±0.009	-	0.860±0.012	0.853±0.012
35	0.867±0.007	0.877±0.004	0.825±0.036	0.859±0.014
36	0.776±0.008	0.812±0.005	0.776±0.034	0.789±0.037
39	0.860±0.004	-	0.893±0.003	0.885±0.016
40	0.889±0.003	0.899±0.006	0.892±0.016	0.908±0.011
41	0.861±0.009	0.893±0.006	0.885±0.006	0.887±0.012

42	0.827±0.009	-	0.839±0.010	0.789±0.040
43	0.822±0.002	0.860±0.004	0.837±0.009	0.842±0.003
44	0.836±0.006	0.857±0.005	0.875±0.010	0.820±0.043
45	0.879±0.005	0.895±0.002	0.862±0.017	0.862±0.007
46	0.854±0.014	0.882±0.002	0.833±0.017	0.898±0.004
47	0.812±0.006	0.870±0.011	0.809±0.016	0.840±0.010
48	0.857±0.001	0.927±0.005	0.842±0.043	0.932±0.016
49	0.790±0.002	0.870±0.004	0.836±0.016	0.887±0.001
50	0.841±0.006	0.906±0.009	0.861±0.008	0.813±0.041
51	0.783±0.010	0.852±0.002	0.804±0.017	0.790±0.070
52	0.783±0.000	0.822±0.004	0.759±0.043	0.835±0.007
54	0.840±0.010	0.892±0.007	0.827±0.044	0.863±0.028
55	0.873±0.006	0.901±0.009	0.867±0.011	0.852±0.015
56	0.880±0.005	0.938±0.004	0.886±0.034	0.859±0.037
57	0.866±0.005	0.895±0.011	0.889±0.015	0.906±0.002
58	0.883±0.005	0.931±0.009	0.906±0.008	0.910±0.002
59	0.817±0.004	0.860±0.003	0.857±0.005	0.814±0.035
60	0.895±0.005	0.917±0.009	0.852±0.046	0.934±0.003
61	-	0.889±0.006	-	-
62	0.683±0.018	0.792±0.015	0.770±0.037	0.871±0.007
63	0.839±0.005	0.863±0.000	0.811±0.042	0.906±0.004
64	0.864±0.004	0.899±0.003	0.869±0.011	0.831±0.046
65	0.852±0.001	0.903±0.005	0.885±0.007	0.892±0.010
66	0.835±0.014	0.854±0.002	0.848±0.006	0.841±0.018
67	0.857±0.007	0.865±0.011	0.838±0.014	0.816±0.025
68	0.874±0.002	0.893±0.008	0.872±0.006	0.861±0.023
69	0.660±0.024	0.877±0.004	-	-
70	0.934±0.007	0.899±0.007	0.905±0.007	0.887±0.040
71	0.820±0.002	0.889±0.012	0.880±0.016	-
72	0.801±0.007	0.897±0.006	0.883±0.017	-
73	0.853±0.002	-	-	-
74	0.533±0.036	0.893±0.011	0.336±0.039	0.299±0.002
75	0.465±0.001	0.684±0.012	0.259±0.088	0.247±0.096
76	0.322±0.004	0.297±0.031	0.426±0.192	0.229±0.151

Table E-3: Methyl order parameters for ubiquitin in various solvent conditions

Methyl	AQ [48] O^2_{axis}	50% Glycerol O^2_{axis}	30% Glycerol O^2_{axis}	AOT O^2_{axis}
i13δ1	0.475±0.012	0.437±0.015	0.463±0.021	0.455±0.029
i23δ1	0.455±0.016	0.470±0.019	0.487±0.022	0.517±0.033
i30δ1	0.685±0.004	0.645±0.028	0.676±0.030	-
i36δ1	0.477±0.003	0.478±0.019	0.494±0.021	*0.579±0.034
i3δ1	0.651±0.014	0.671±0.033	0.698±0.030	0.675±0.044
i44δ1	0.195±0.048	0.153±0.006	0.190±0.009	0.253±0.016
i61δ1	0.505±0.025	0.558±0.022	0.563±0.025	0.638±0.041
l15δ1	0.457±0.002	0.443±0.020	0.460±0.020	*0.458±0.030
l15δ2	0.449±0.018	0.462±0.018	0.471±0.022	*0.465±0.030
l43δ1	0.568±0.009	0.548±0.024	0.552±0.024	-
l43δ2	0.449±0.027	0.454±0.016	0.453±0.021	-
l50δ1	0.693±0.013	0.646±0.032	0.671±0.030	0.731±0.047
l50δ2	0.671±0.007	0.710±0.032	0.699±0.032	0.821±0.050
l56δ1	0.524±0.007	0.511±0.020	0.582±0.027	0.505±0.033
l56δ2	0.534±0.007	0.524±0.028	0.558±0.024	-
l67δ1	0.243±0.037	0.247±0.009	0.257±0.012	0.278±0.019
l67δ2	0.233±0.022	0.228±0.009	0.247±0.011	0.277±0.018
l69δ1	-	-	-	0.633±0.038
l69δ2	0.526±0.012	-	0.511±0.023	0.641±0.066
l71δ1	0.233±0.002	-	0.243±0.012	*0.237±0.015
l73δ1	0.108±0.029	-	0.143±0.007	0.227±0.015
l8δ1	0.209±0.002	0.222±0.008	0.221±0.010	0.230±0.016
l8δ2	-	0.211±0.008	0.289±0.019	-
v17γ1	0.764±0.003	-	-	0.787±0.051
v17γ2	0.750±0.003	0.698±0.036	0.746±0.034	0.819±0.053
v26γ1	0.756±0.007	0.672±0.029	0.769±0.037	0.762±0.049
v26γ2	-	-	-	0.784±0.049
v5γ2	0.736±0.010	0.686±0.028	0.743±0.034	0.748±0.052
v70γ1	-	0.343±0.013	0.368±0.016	0.565±0.040
v70γ2	0.336±0.034	0.307±0.012	0.335±0.014	*0.556±0.038

* Indicates methyl probe contaminated by surfactant/solvent streaks (see Figure E-4)

Table E-4: Methyl order parameters for MBP in aqueous solution and RM

Residue	Methyl	AQ O^2_{axis}	RM O^2_{axis}	Residue	Methyl	AQ O^2_{axis}	RM O^2_{axis}
I104	δ1	0.285± 0.010	0.260± 0.012	L280	δ2	0.837± 0.036	0.811± 0.046
I108	δ1	0.605± 0.023	0.604± 0.029	L284	δ2	0.572± 0.025	0.594± 0.028
I116	δ1	0.775± 0.044	0.794± 0.043	L284	δ1	0.532± 0.023	0.433± 0.024
I116	δ1	0.680± 0.040	0.718± 0.044	L285	δ2	0.606± 0.027	0.601± 0.046
I132	δ1	0.373± 0.014	0.357± 0.015	L290	δ1	0.727± 0.033	0.731± 0.039
I161	δ1	0.593± 0.026	0.613± 0.027	l290	δ2	0.700± 0.032	0.689± 0.044
I178	δ1	0.468± 0.019	0.543± 0.027	L304	δ2	0.792± 0.039	0.790± 0.048
I199	δ1	0.544± 0.020	0.585± 0.028	L304	δ1	0.787± 0.031	0.702± 0.032
I226	δ1	0.743± 0.033	0.634± 0.031	L311	δ1	0.658± 0.028	0.611± 0.026
I235	δ1	0.786± 0.037	0.794± 0.045	L361	δ2	0.892± 0.043	0.943± 0.050
I266	δ1	0.716± 0.032	0.780± 0.036	L7	δ2	0.426± 0.016	0.429± 0.022
I317	δ1	0.173± 0.006	0.169± 0.008	L7	δ1	0.404± 0.015	0.487± 0.023
I329	δ1	0.213± 0.008	0.240± 0.011	L75	δ1	0.808± 0.035	0.876± 0.039
I33	δ1	0.430± 0.015	0.429± 0.019	L75	δ2	0.785± 0.037	0.870± 0.039
I333	δ1	0.548± 0.021	0.559± 0.026	L76	δ1	0.864± 0.039	0.857± 0.039
I368	δ1	0.806± 0.040	0.841± 0.044	L76	δ2	0.723± 0.031	0.771± 0.035
I59	δ1	0.803± 0.039	0.890± 0.043	L89	δ2	0.253± 0.009	0.252± 0.012
I60	δ1	0.351± 0.012	0.335± 0.015	L89	δ1	0.199± 0.007	0.215± 0.009
I79	δ1	0.580± 0.023	0.576± 0.027	V110	γ2	0.710± 0.028	0.766± 0.039
I9	δ1	0.604± 0.026	0.607± 0.032	V110	γ1	0.680± 0.028	0.642± 0.030

L103	δ2	0.717±0.031	0.725±0.043	V181	γ1	0.566±0.024	0.484±0.024
L103	δ1	0.660±0.028	0.760±0.034	V183	γ1	0.420±0.016	0.454±0.023
L113	δ1	0.793±0.050	0.827±0.058	V23	γ2	0.836±0.038	0.870±0.108
L121	δ2	0.562±0.023	0.554±0.027	V23	γ1	0.820±0.034	0.916±0.043
L121	δ1	0.551±0.024	0.565±0.030	V240	γ1	0.583±0.026	0.568±0.034
L122	δ1	0.340±0.013	0.346±0.016	V244	γ2	0.812±0.035	0.843±0.046
L135	δ1	0.276±0.010	0.315±0.013	V244	γ1	0.769±0.042	0.837±0.048
L139	δ1	0.562±0.024	0.575±0.049	V259	γ2	0.845±0.041	0.867±0.041
L139	δ2	0.517±0.018	0.512±0.023	V261	γ1	0.871±0.044	0.903±0.058
L147	δ1	0.808±0.036	0.807±0.056	V261	γ2	0.870±0.039	0.863±0.046
L147	δ2	0.713±0.030	0.757±0.034	V293	γ2	0.833±0.056	0.929±0.086
L151	δ1	0.607±0.026	0.601±0.030	V302	γ2	0.765±0.041	0.756±0.068
L151	δ2	0.484±0.018	0.514±0.025	V343	γ1	0.771±0.032	0.725±0.035
L160	δ2	0.663±0.024	0.677±0.033	V347	γ1	0.799±0.042	0.764±0.042
L160	δ1	0.596±0.025	0.690±0.035	V357	γ1	0.888±0.035	0.886±0.049
L195	δ1	0.306±0.011	0.323±0.014	V357	γ2	0.802±0.036	0.801±0.065
L198	δ1	0.736±0.032	0.784±0.038	V357	γ2	0.706±0.037	0.755±0.046
L198	δ2	0.677±0.027	0.600±0.026	V37	γ1	0.869±0.042	0.853±0.039
L20	δ2	0.455±0.017	0.447±0.023	V37	γ2	0.745±0.032	0.796±0.042
L20	δ1	0.431±0.016	0.462±0.021	V50	γ1	0.598±0.025	0.666±0.029
L247	δ2	0.759±0.032	0.710±0.040	V50	γ2	0.564±0.026	0.627±0.035
L262	δ1	0.445±0.018	0.476±0.024	V8	γ1	0.843±0.040	0.844±0.042

L275	δ1	0.516± 0.021	0.544± 0.024	V97	γ1	0.822± 0.036	0.884± 0.045
L275	δ2	0.472± 0.019	0.492± 0.021	V97	γ2	0.761± 0.046	0.810± 0.053

Table E-5: Methyl order parameters for MSG in aqueous solution and RM

Residue	Methyl	AQ O^2_{axis}	RM O^2_{axis}	Residue	Methyl	AQ O^2_{axis}	RM O^2_{axis}
I105	δ1	0.547± 0.012	0.545± 0.033	L717	δ1	0.625± 0.017	0.539± 0.026
I12	δ1	0.532± 0.013	0.527± 0.029	L724	δ1	0.148± 0.004	0.183± 0.004
I147	δ1	0.528± 0.013	0.483± 0.019	L85	δ1	0.741± 0.021	0.812± 0.066
I148	δ1	0.242± 0.004	0.193± 0.009	L85	δ2	0.730± 0.018	0.707± 0.056
I200	δ1	0.319± 0.006	*0.283 ±0.015	L88	δ2	0.567± 0.020	0.500± 0.046
I229	δ1	0.356± 0.006	0.339± 0.017	L91	δ2	0.665± 0.012	0.761± 0.069
I238	δ1	0.471± 0.008	0.447± 0.019	L91	δ1	0.334± 0.006	0.297± 0.009
I242	δ1	0.210± 0.003	0.187± 0.006	M	ε	0.612± 0.036	0.583± 0.079
I248	δ1	0.445± 0.006	0.459± 0.022	M	ε	0.156± 0.004	0.128± 0.009
I256	δ1	0.640± 0.024	0.598± 0.036	M	ε	0.654± 0.040	0.644± 0.051
I256	δ1	0.428± 0.008	*0.379 ±0.024	M	ε	0.514± 0.017	0.498± 0.029
I284	δ1	0.386± 0.016	*0.446 ±0.015	M	ε	0.347± 0.008	0.313± 0.025
I309	δ1	0.108± 0.003	*0.151 ±0.010	M	ε	0.581± 0.024	0.594± 0.032
I327	δ1	0.507± 0.011	0.494± 0.018	M	ε	0.201± 0.005	0.200± 0.012
I361	δ1	0.858± 0.030	0.823± 0.037	M	ε	0.452± 0.014	0.433± 0.022
I370	δ1	0.730± 0.021	0.770± 0.021	M	ε	0.472± 0.016	0.484± 0.028
I388	δ1	0.662± 0.015	0.583± 0.030	M	ε	0.419± 0.011	0.398± 0.052

I439	δ1	0.790±0.023	*0.970±0.046		ε	0.680±0.023	0.782±0.055
I5	δ1	0.110±0.003	0.110±0.006	M	ε	0.707±0.036	0.903±0.066
I504	δ1	0.702±0.016	0.727±0.032	M	ε	0.444±0.009	*0.396±0.042
I560	δ1	0.637±0.013	0.606±0.025	M	ε	0.085±0.003	0.073±0.006
I592	δ1	0.778±0.020	*0.153±0.012	M	ε	0.508±0.015	0.482±0.029
I642	δ1	0.632±0.025	0.619±0.067	M	ε	0.631±0.028	0.612±0.033
I650	δ1	0.597±0.015	0.640±0.019	M	ε	0.623±0.025	0.651±0.059
L128	δ2	0.261±0.004	0.302±0.020	M	ε	0.078±0.003	0.095±0.005
L138	δ2	0.594±0.010	*0.521±0.013	M	ε	0.152±0.003	0.130±0.006
L142	γ2	0.506±0.017	0.523±0.048	M	ε	0.037±0.003	0.041±0.005
L178	δ1	0.357±0.006	0.329±0.016	V119	γ2	0.304±0.008	0.311±0.049
L180	δ1	0.686±0.018	0.682±0.047	V155	γ2	0.114±0.003	0.086±0.005
L180	δ2	0.645±0.016	0.616±0.050	V188	γ1	0.502±0.013	0.565±0.038
L198	δ1	0.652±0.024	0.648±0.029	V189	γ1	0.442±0.010	*0.464±0.024
L202	δ1	0.606±0.014	*0.696±0.051	V189	γ2	0.406±0.009	0.416±0.022
L202	δ2	0.616±0.014	0.613±0.023	V194	γ1	0.683±0.018	*0.738±0.049
L230	δ1	0.722±0.015	0.863±0.059	V194	γ2	0.686±0.018	0.679±0.035
L231	δ2	0.720±0.020	0.594±0.035	V24	γ1	0.287±0.009	0.245±0.013
L236	δ1	0.599±0.016	*0.529±0.063	V259	γ2	0.800±0.019	0.868±0.084
L240	δ1	0.432±0.008	0.389±0.021	V259	γ1	0.735±0.020	0.770±0.066
L240	δ2	0.411±0.007	0.370±0.015	V275	γ2	0.543±0.009	*0.615±0.087
L25	δ1	0.703±0.019	0.676±0.028	V310	γ1	0.444±0.008	*0.499±0.017

L25	δ2	0.550±0.010	0.565±0.037	V310	γ2	0.456±0.009	0.460±0.020
L269	δ2	0.603±0.013	0.609±0.026	V348	γ1	0.675±0.015	0.685±0.039
L269	δ1	0.663±0.028	*0.499±0.051	V377	γ1	0.302±0.007	0.261±0.014
L285	δ1	0.390±0.009	0.308±0.011	V386	γ1	0.767±0.018	0.751±0.066
L291	δ1	0.471±0.019	*0.542±0.033	V43	γ1	0.791±0.024	0.850±0.070
L30	δ1	0.550±0.011	*0.608±0.023	V490	γ1	0.774±0.019	*0.659±0.030
L329	δ2	0.769±0.020	0.889±0.066	V553	γ2	0.545±0.033	0.416±0.042
L329	δ1	0.524±0.013	0.491±0.033	V607	γ2	0.285±0.006	0.294±0.040
L334	δ1	0.721±0.012	*0.765±0.030	V608	γ2	0.706±0.027	*0.503±0.062
L343	δ1	0.392±0.008	0.360±0.020	V626	γ1	0.673±0.029	*0.646±0.094
L420	δ1	0.657±0.034	0.621±0.045	V666	γ2	0.743±0.027	*0.628±0.045
L420	δ2	0.578±0.013	0.515±0.048	V701	γ2	0.409±0.007	0.348±0.028
L43	γ2	0.552±0.012	0.430±0.034	V701	γ1	0.422±0.014	0.347±0.018
L433	δ1	0.622±0.015	0.655±0.028	V92	γ1	0.520±0.016	*0.544±0.019
L433	δ1	0.253±0.009	0.231±0.020	V92	γ2	0.507±0.009	0.419±0.045
L454	δ1	0.773±0.018	0.719±0.064	V98	γ1	0.587±0.019	*0.585±0.040
L454	δ2	0.727±0.013	0.661±0.048	?	?	0.776±0.016	0.800±0.033
L46	δ1	0.498±0.007	0.495±0.039	?	?	0.779±0.026	0.693±0.065
L471	δ1	0.645±0.016	0.534±0.017	?	?	0.707±0.013	0.692±0.043
L471	δ2	0.572±0.015	0.474±0.026	?	?	0.725±0.014	0.689±0.063
L494	δ1	0.579±0.011	0.516±0.023	?	?	0.711±0.016	0.666±0.051
L498	δ2	0.399±0.013	*0.458±0.059	?	?	0.769±0.020	0.659±0.027

L514	δ1	0.600±0.014	0.533±0.020	?	?	0.747±0.021	0.658±0.049
L526	δ2	0.552±0.018	0.581±0.041	?	?	0.783±0.023	0.642±0.043
L526	δ2	0.621±0.019	0.572±0.060	?	?	0.626±0.014	0.635±0.029
L526	δ1	0.439±0.013	0.445±0.036	?	?	0.656±0.016	0.577±0.028
L53	δ1	0.642±0.016	0.672±0.032	?	?	0.631±0.020	0.567±0.020
L54	δ1	0.742±0.026	*0.458±0.042	?	?	0.513±0.009	0.564±0.071
L543	δ2	0.705±0.022	0.700±0.129	?	?	0.615±0.013	0.538±0.027
L546	δ1	0.550±0.017	*0.759±0.084	?	?	0.610±0.026	0.519±0.025
L546	δ2	0.562±0.015	0.528±0.031	?	?	0.595±0.016	0.515±0.024
L572	δ1	0.727±0.015	0.750±0.051	?	?	0.486±0.012	0.487±0.019
L572	δ2	0.550±0.014	0.572±0.035	?	?	0.445±0.009	0.483±0.018
L573	δ2	0.377±0.007	0.327±0.014	?	?	0.356±0.006	0.396±0.025
L576	δ1	0.768±0.039	*0.452±0.058	?	?	0.399±0.008	0.384±0.018
L628	δ2	0.476±0.011	0.535±0.035	?	?	0.373±0.010	0.356±0.016
L660	δ1	0.756±0.016	0.645±0.030	?	?	0.314±0.009	0.299±0.081
L660	δ2	0.642±0.036	0.606±0.066	?	?	0.354±0.008	0.241±0.012
L699	δ1	0.642±0.015	0.582±0.021	?	?	0.167±0.004	0.201±0.014
L711	δ2	0.505±0.012	0.414±0.041	?	?	0.195±0.005	*0.113±0.008

* Indicates methyl probe contaminated by surfactant/solvent streaks (see Figure E-4)
M Indicates an unassigned methionine residue
? Indicates of methyl probe whose assignment was difficult/impossible to map

REFERENCES

1. Chandler, D., *Interfaces and the driving force of hydrophobic assembly.* Nature, 2005. **437**(7059): p. 640-7.
2. Dill, K.A., *Dominant forces in protein folding.* Biochemistry, 1990. **29**(31): p. 7133-55.
3. Smith, J.C., et al., *Structure, dynamics and reactions of protein hydration water.* Philos Trans R Soc Lond B Biol Sci, 2004. **359**(1448): p. 1181-9; discussion 1189-90.
4. Sorenson, J.M., et al., *Determining the role of hydration forces in protein folding.* J Phys Chem B, 1999. **103**(26): p. 5413-5426.
5. Levy, Y. and J.N. Onuchic, *Water mediation in protein folding and molecular recognition.* Annu Rev Biophys Biomol Struct, 2006. **35**: p. 389-415.
6. Timasheff, S.N., *Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components.* Proc Natl Acad Sci U S A, 2002. **99**(15): p. 9721-6.
7. Zaccai, G., *The effect of water on protein dynamics.* Philos Trans R Soc Lond B Biol Sci, 2004. **359**(1448): p. 1269-75; discussion 1275, 1323-8.
8. Halle, B., *Protein hydration dynamics in solution: a critical survey.* Philos Trans R Soc Lond B Biol Sci, 2004. **359**(1448): p. 1207-23; discussion 1223-4, 1323-8.
9. Michalarias, I., et al., *Recent progress on our understanding of water around biomolecules.* J Mol Liq, 2005. **117**(1-3): p. 107-116.
10. Bagchi, B., *Water dynamics in the hydration layer around proteins and micelles.* Chem Rev, 2005. **105**(9): p. 3197-219.
11. Raschke, T.M., *Water structure and interactions with protein surfaces.* Curr Opin Struct Biol, 2006. **16**(2): p. 152-9.
12. Kaminker, I., R. Barnes, and S. Han, *Overhauser dynamic nuclear polarization studies on local water dynamics.* Methods Enzymol, 2015. **564**: p. 457-83.
13. Pal, S.K. and A.H. Zewail, *Dynamics of water in biological recognition.* Chem Rev, 2004. **104**(4): p. 2099-123.
14. Fogarty, A.C. and D. Laage, *Water dynamics in protein hydration shells: the molecular origins of the dynamical perturbation.* J Phys Chem B, 2014. **118**(28): p. 7715-29.
15. Sterpone, F., M. Ceccarelli, and M. Marchi, *Dynamics of hydration in hen egg white lysozyme.* J Mol Biol, 2001. **311**(2): p. 409-19.
16. Sterpone, F., G. Stirnemann, and D. Laage, *Magnitude and molecular origin of water slowdown next to a protein.* J Am Chem Soc, 2012. **134**(9): p. 4116-9.
17. Fenimore, P.W., et al., *Slaving: solvent fluctuations dominate protein dynamics and functions.* Proc Natl Acad Sci U S A, 2002. **99**(25): p. 16047-51.
18. Frauenfelder, H., et al., *A unified model of protein dynamics.* Proc Natl Acad Sci U S A, 2009. **106**(13): p. 5129-34.
19. Frauenfelder, H., et al., *Protein folding is slaved to solvent motions.* Proc Natl Acad Sci U S A, 2006. **103**(42): p. 15469-72.
20. Benedetto, A., *Low-temperature decoupling of water and protein dynamics measured by neutron scattering.* J Phys Chem Lett, 2017. **8**(19): p. 4883-4886.
21. Pazos, I.M., et al., *Ester carbonyl vibration as a sensitive probe of protein local electric field.* Angew Chem Int Ed Engl, 2014. **53**(24): p. 6080-4.
22. Zhang, W., et al., *C≡N stretching vibration of 5-cyanotryptophan as an infrared probe of protein local environment: what determines its frequency?* Phys Chem Chem Phys, 2016. **18**(10): p. 7027-34.

23. Otting, G., E. Liepinsh, and K. Wuthrich, *Protein hydration in aqueous solution.* Science, 1991. **254**(5034): p. 974-80.
24. Otting, G. and K. Wuthrich, *Studies of protein hydration in aqueous solution by direct NMR observation of individual protein-bound water molecules.* J Am Chem Soc, 1989. **111**(5): p. 1871-1875.
25. Bruschweiler, R. and P.E. Wright, *Water self-diffusion model for protein-water NMR cross relaxation.* Chem Phys Lett, 1994. **229**(1-2): p. 75-81.
26. Otting, G., *NMR studies of water bound to biological molecules.* Prog Nucl Magn Reson Spectrosc, 1997. **31**(2-3): p. 259-285.
27. Halle, B., *Cross-relaxation between macromolecular and solvent spins: The role of longrange dipole couplings.* J Chem Phys, 2003. **119**(23): p. 12372-12385.
28. Braun, D., M. Schmollngruber, and O. Steinhauser, *Revival of the intermolecular nuclear overhauser effect for mapping local protein hydration dynamics.* J Phys Chem Lett, 2017. **8**(14): p. 3421-3426.
29. Nucci, N.V., M.S. Pometun, and A.J. Wand, *Mapping the hydration dynamics of ubiquitin.* J Am Chem Soc, 2011. **133**(32): p. 12326-9.
30. Igumenova, T.I., K.K. Frederick, and A.J. Wand, *Characterization of the fast dynamics of protein amino acid side chains using NMR relaxation in solution.* Chem Rev, 2006. **106**(5): p. 1672-99.
31. Muhandiram, D.R., et al., *Measurement of 2H T1 and T1rho Relaxation Times in Uniformly 13C-Labeled and Fractionally 2H-Labeled Proteins in Solution.* J Am Chem Soc, 1995. **117**(46): p. 11536-11544.
32. Tugarinov, V. and L.E. Kay, *Quantitative 13C and 2H NMR relaxation studies of the 723-residue enzyme malate synthase G reveal a dynamic binding interface.* Biochemistry, 2005. **44**(49): p. 15970-7.
33. Palmer, A.G., 3rd, *NMR characterization of the dynamics of biomacromolecules.* Chem Rev, 2004. **104**(8): p. 3623-40.
34. Palmer, A.G., 3rd and F. Massi, *Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy.* Chem Rev, 2006. **106**(5): p. 1700-19.
35. Gnutt, D. and S. Ebbinghaus, *The macromolecular crowding effect--from in vitro into the cell.* Biol Chem, 2016. **397**(1): p. 37-44.
36. Caro, J.A., et al., *Entropy in molecular recognition by proteins.* Proc Natl Acad Sci U S A, 2017. **114**(25): p. 6563-6568.
37. Igumenova, T.I., A.L. Lee, and A.J. Wand, *Backbone and side chain dynamics of mutant calmodulin-peptide complexes.* Biochemistry, 2005. **44**(38): p. 12627-39.
38. Schneider, D.M., M.J. Dellwo, and A.J. Wand, *Fast internal main-chain dynamics of human ubiquitin.* Biochemistry, 1992. **31**(14): p. 3645-52.
39. Kasinath, V., et al., *A sharp thermal transition of fast aromatic-ring dynamics in ubiquitin.* Angew Chem Int Ed Engl, 2015. **54**(1): p. 102-7.
40. Akke, M., R. Bruschweiler, and A.G. Palmer, 3rd, *NMR order parameters and free energy: an analytical approach and its application to cooperative calcium(2+) binding by calbindin D9k.* J Am Chem Soc, 1993. **115**(21): p. 9832-9833.
41. Lee, A.L. and A.J. Wand, *Microscopic origins of entropy, heat capacity and the glass transition in proteins.* Nature, 2001. **411**(6836): p. 501-4.
42. Sun, H., L.E. Kay, and V. Tugarinov, *An optimized relaxation-based coherence transfer NMR experiment for the measurement of side-chain order in methyl-protonated, highly deuterated proteins.* J Phys Chem B, 2011. **115**(49): p. 14878-84.

43. Tugarinov, V., R. Sprangers, and L.E. Kay, *Probing side-chain dynamics in the proteasome by relaxation violated coherence transfer NMR spectroscopy.* J Am Chem Soc, 2007. **129**(6): p. 1743-50.
44. Wand, A.J., *Dynamic activation of protein function: a view emerging from NMR spectroscopy.* Nat Struct Biol, 2001. **8**(11): p. 926-31.
45. Wand, A.J., *The dark energy of proteins comes to light: conformational entropy and its role in protein function revealed by NMR relaxation.* Curr Opin Struct Biol, 2013. **23**(1): p. 75-81.
46. Sharp, K.A., V. Kasinath, and A.J. Wand, *Banding of NMR-derived methyl order parameters: implications for protein dynamics.* Proteins, 2014. **82**(9): p. 2106-17.
47. Lee, A.L., et al., *Temperature dependence of the internal dynamics of a calmodulin-peptide complex.* Biochemistry, 2002. **41**(46): p. 13814-25.
48. Fu, Y., et al., *Coupled motion in proteins revealed by pressure perturbation.* J Am Chem Soc, 2012. **134**(20): p. 8543-50.
49. Frederick, K.K., et al., *Conformational entropy in molecular recognition by proteins.* Nature, 2007. **448**(7151): p. 325-9.
50. Marlow, M.S., et al., *The role of conformational entropy in molecular recognition by calmodulin.* Nat Chem Biol, 2010. **6**(5): p. 352-8.
51. Wand, A.J. and K.A. Sharp, *Measuring Entropy in Molecular Recognition by Proteins.* Annu Rev Biophys, 2018.
52. Sharp, K.A., et al., *On the relationship between NMR-derived amide order parameters and protein backbone entropy changes.* Proteins, 2015. **83**(5): p. 922-30.
53. De, T.K. and A. Maitra, *Solution behavior of aerosol Ot in nonpolar-solvents.* Adv Colloid Interface Sci, 1995. **59**: p. 95-193.
54. Fayer, M.D. and N.E. Levinger, *Analysis of water in confined geometries and at interfaces.* Annu Rev Anal Chem, 2010. **3**: p. 89-107.
55. Levinger, N.E. and L.A. Swafford, *Ultrafast dynamics in reverse micelles.* Annu Rev Phys Chem, 2009. **60**: p. 385-406.
56. Moilanen, D.E., et al., *Confinement or the nature of the interface? Dynamics of nanoscopic water.* J Am Chem Soc, 2007. **129**(46): p. 14311-8.
57. Nucci, N.V., et al., *Optimization of NMR spectroscopy of encapsulated proteins dissolved in low viscosity fluids.* J Biomol NMR, 2011. **50**(4): p. 421-30.
58. Dodevski, I., et al., *Optimized reverse micelle surfactant system for high-resolution NMR spectroscopy of encapsulated proteins and nucleic acids dissolved in low viscosity fluids.* J Am Chem Soc, 2014. **136**(9): p. 3465-74.
59. Peterson, R.W., et al., *Novel surfactant mixtures for NMR spectroscopy of encapsulated proteins dissolved in low-viscosity fluids.* Protein Sci, 2005. **14**(11): p. 2919-21.
60. Nucci, N.V., K.G. Valentine, and A.J. Wand, *High-resolution NMR spectroscopy of encapsulated proteins dissolved in low-viscosity fluids.* J Magn Reson, 2014. **241**: p. 137-47.
61. Peterson, R.W., et al., *Forced folding and structural analysis of metastable proteins.* J Am Chem Soc, 2004. **126**(31): p. 9498-9.
62. Kielec, J.M., et al., *Reverse micelles in integral membrane protein structural biology by solution NMR spectroscopy.* Structure, 2009. **17**(3): p. 345-51.
63. Valentine, K.G., et al., *Reverse micelle encapsulation of membrane-anchored proteins for solution NMR studies.* Structure, 2010. **18**(1): p. 9-16.
64. Pometun, M.S., et al., *Cold denaturation of encapsulated ubiquitin.* J Am Chem Soc, 2006. **128**(33): p. 10652-3.

65. Luisi, P.L., et al., *Reverse micelles as hosts for proteins and small molecules.* Biochim Biophys Acta, 1988. **947**(1): p. 209-46.
66. Fenn, E.E., D.B. Wong, and M.D. Fayer, *Water dynamics at neutral and ionic interfaces.* Proc Natl Acad Sci U S A, 2009. **106**(36): p. 15243-8.
67. Park, S., D.E. Moilanen, and M.D. Fayer, *Water dynamics--the effects of ions and nanoconfinement.* J Phys Chem B, 2008. **112**(17): p. 5279-90.
68. Fuglestad, B., et al., *Characterization of Cetyltrimethylammonium Bromide/Hexanol Reverse Micelles by Experimentally Benchmarked Molecular Dynamics Simulations.* Langmuir, 2016. **32**(7): p. 1674-84.
69. Sinha, S.K. and S. Bandyopadhyay, *Local heterogeneous dynamics of water around lysozyme: a computer simulation study.* Phys Chem Chem Phys, 2012. **14**(2): p. 899-913.
70. Simorellis, A.K. and P.F. Flynn, *Fast local backbone dynamics of encapsulated ubiquitin.* J Am Chem Soc, 2006. **128**(30): p. 9580-1.
71. Correa, N.M., et al., *Nonaqueous polar solvents in reverse micelle systems.* Chem Rev, 2012. **112**(8): p. 4569-602.
72. Mattea, C., J. Qvist, and B. Halle, *Dynamics at the protein-water interface from O-17 spin relaxation in deeply supercooled solutions.* Biophys J, 2008. **95**(6): p. 2951-2963.
73. Anarbaev, R.O., A.L. Rogozina, and O.I. Lavrik, *DNA polymerase β reveals enhanced activity and processivity in reverse micelles.* Biophys Chem, 2009. **141**(1): p. 11-20.
74. Malik, A., et al., *Myoglobin unfolding in crowding and confinement.* J Phys Chem B, 2012. **116**(43): p. 12895-12904.
75. Saha, R., et al., *Protein-cofactor binding and ultrafast electron transfer in riboflavin binding protein under the spatial confinement of nanoscopic reverse micelles.* J Mol Recognit, 2013. **26**(2): p. 59-66.
76. Babu, C.R., P.F. Flynn, and A.J. Wand, *Validation of protein structure from preparations of encapsulated proteins dissolved in low viscosity fluids.* J Am Chem Soc, 2001. **123**(11): p. 2691-2692.
77. Van Horn, W.D., M.E. Ogilvie, and P.F. Flynn, *Use of reverse micelles in membrane protein structural biology.* J Biomol NMR, 2008. **40**(3): p. 203-211.
78. Babu, C.R., V.J. Hilser, and A.J. Wand, *Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation.* Nat Struct Mol Biol, 2004. **11**(4): p. 352-357.
79. Van Horn, W.D., M.E. Ogilvie, and P.F. Flynn, *Reverse Micelle Encapsulation as a Model for Intracellular Crowding.* J Am Chem Soc, 2009. **131**(23): p. 8030-8039.
80. Workman, H. and P.F. Flynn, *Stabilization of RNA Oligomers through Reverse Micelle Encapsulation.* J Am Chem Soc, 2009. **131**(11): p. 3806-+.
81. Narang, A.S., D. Delmarre, and D. Gao, *Stable drug encapsulation in micelles and microemulsions.* Int J Pharm, 2007. **345**(1-2): p. 9-25.
82. Wand, A.J., M.R. Ehrhardt, and P.F. Flynn, *High-resolution NMR of encapsulated proteins dissolved in low-viscosity fluids.* Proc Natl Acad Sci U S A, 1998. **95**(26): p. 15299-15302.
83. Lefebvre, B.G., et al., *NMR spectroscopy of proteins encapsulated in a positively charged surfactant.* J Magn Reson, 2005. **175**(1): p. 158-62.
84. Wand, A.J., et al., *Internal dynamics of human ubiquitin revealed by 13C-relaxation studies of randomly fractionally labeled protein.* Biochemistry, 1996. **35**(19): p. 6116-25.
85. Hasegawa, M., *Buffer-like action in water pool of aerosol OT reverse micelles.* Langmuir, 2001. **17**(5): p. 1426-1431.
86. Crans, D.C. and N.E. Levinger, *The conundrum of pH in water nanodroplets: sensing pH in reverse micelle water pools.* Accounts of Chemical Research, 2012. **45**(10): p. 1637-1645.

87. Elseoud, O.A., A.M. Chinelatto, and M.R. Shimizu, *Acid-Base Indicator Equilibria in the Presence of Aerosol-Ot Aggregates in Heptane - Ion-Exchange in Reversed Micelles.* J Colloid Interface Sci, 1982. **88**(2): p. 420-427.
88. Smith, R.E. and P.L. Luisi, *Micellar Solubilization of Bio-Polymers in Hydrocarbon Solvents .3. Empirical Definition of an Acidity Scale in Reverse Micelles.* Helv Chim Acta, 1980. **63**(8): p. 2302-2311.
89. Baruah, B., et al., *When is water not water? Exploring water confined in large reverse micelles using a highly charged inorganic molecular probe.* J Am Chem Soc, 2006. **128**(39): p. 12758-12765.
90. Baruah, B., D.C. Crans, and N.E. Levinger, *Simple oxovanadates as multiparameter probes of reverse micelles.* Langmuir, 2007. **23**(12): p. 6510-6518.
91. Baruah, B., et al., *Do probe molecules influence water in confinement?* J Phys Chem B, 2008. **112**(33): p. 10158-10164.
92. Sedgwick, M.A., D.C. Crans, and N.E. Levinger, *What Is Inside a Nonionic Reverse Micelle? Probing the Interior of Igepal Reverse Micelles Using Decavanadate.* Langmuir, 2009. **25**(10): p. 5496-5503.
93. Fujii, H., T. Kawai, and H. Nishikawa, *Determination of Ph in Reversed Micelles.* Bull Chem Soc Jpn, 1979. **52**(7): p. 2051-2055.
94. Halliday, N.A., A.C. Peet, and M.M. Britton, *Detection of pH in microemulsions, without a probe molecule, using magnetic resonance.* J Phys Chem B, 2010. **114**(43): p. 13745-13751.
95. Vodolazkaya, N.A., et al., *Molecular spectroscopy studies of solvent properties of dispersed 'water pools' Fluorescein and 2,7-dichlorofluorescein in reversed AOT-based microemulsions.* J Mol Liq, 2010. **157**(2-3): p. 105-112.
96. Fletcher, P.D.I., A.M. Howe, and B.H. Robinson, *The kinetics of solubilisate exchange between water droplets of a water-in-oil microemulsion.* J Chem Soc, Faraday Trans 1, 1987. **83**: p. 985-1006.
97. Natarajan, U., et al., *Ultrafine metal particle formation in reverse micellar systems: Effects of intermicellar exchange on the formation of particles.* Langmuir, 1996. **12**(11): p. 2670-2678.
98. Oldfield, C., *Exchange Concept in Water-in-Oil Microemulsions - Consequences for Slow Chemical-Reactions.* J Chem Soc, Faraday Trans, 1991. **87**(16): p. 2607-2612.
99. Tynkkynen, T., et al., *From proton nuclear magnetic resonance spectra to pH. Assessment of 1H NMR pH indicator compound set for deuterium oxide solutions.* Anal Chim Acta, 2009. **648**(1): p. 105-12.
100. Baryshnikova, O.K., T.C. Williams, and B.D. Sykes, *Internal pH indicators for biomolecular NMR.* J Biomol NMR, 2008. **41**(1): p. 5-7.
101. Niemeyer, E.D. and F.V. Bright, *The pH within PFPE reverse micelles formed in supercritical CO2.* J Phys Chem B, 1998. **102**(8): p. 1474-1478.
102. Cavanagh, J., *Protein NMR spectroscopy : principles and practice*. 2nd ed. 2007, Amsterdam ; Boston: Academic Press. xxv, 885 p.
103. Ehrhardt, M.R., P.F. Flynn, and A.J. Wand, *Preparation of encapsulated proteins dissolved in low viscosity fluids.* J Biomol NMR, 1999. **14**(1): p. 75-78.
104. Goto, Y., L.J. Calciano, and A.L. Fink, *Acid-Induced Folding of Proteins.* Proc Natl Acad Sci U S A, 1990. **87**(2): p. 573-577.
105. Chen, J., et al., *Fluorescence studies on the microenvironments of proteins in CO2-expanded reverse micellar solutions.* J Supercrit Fluids, 2006. **38**(1): p. 103-110.
106. Gebicka, L., *Peroxidase-like activity of cytochrome c in the presence of anionic surfactants.* Res Chem Intermed, 2001. **27**(7-8): p. 717-723.

107. Ono, T. and M. Goto, *Peroxidative catalytic behavior of cytochrome c solubilized in reverse micelles.* Biochem Eng J, 2006. **28**(2): p. 156-160.
108. Valdez, D., et al., *Hydration and protein folding in water and in reverse micelles: Compressibility and volume changes.* Biophys J, 2001. **80**(6): p. 2751-2760.
109. Pinheiro, T.J., et al., *Structural and kinetic description of cytochrome c unfolding induced by the interaction with lipid vesicles.* Biochemistry, 1997. **36**(42): p. 13122-32.
110. Lakowicz, J.R., *Principles of fluorescence spectroscopy*. 3rd ed. 2006, New York: Springer. xxvi, 954 p.
111. Tanford, C., *Interfacial free energy and the hydrophobic effect.* Proc Natl Acad Sci U S A, 1979. **76**(9): p. 4175-6.
112. Kauzmann, W., *Some factors in the interpretation of protein denaturation.* Adv Protein Chem, 1959. **14**: p. 1-63.
113. Wüthrich, K., G. Otting, and E. Liepinsh, *Protein hydration in aqueous solution.* Faraday Discuss, 1992. **93**: p. 35-45.
114. Wiebenga-Sanford, B.P., et al., *Nanoconfinement's dramatic impact on proton exchange between glucose and water.* J Phys Chem Lett, 2016. **7**(22): p. 4597-4601.
115. Macura, S. and R.R. Ernst, *Elucidation of cross relaxation in liquids by two-dimensional NMR spectroscopy.* Mol Phys, 1980. **41**(1): p. 95-117.
116. Hyberts, S.G., S.A. Robson, and G. Wagner, *Exploring signal-to-noise ratio and sensitivity in non-uniformly sampled multi-dimensional NMR spectra.* J Biomol NMR, 2013. **55**(2): p. 167-78.
117. Hyberts, S.G., et al., *Perspectives in magnetic resonance: NMR in the post-FFT era.* J Magn Reson, 2014. **241**: p. 60-73.
118. Hyberts, S.G., S.A. Robson, and G. Wagner, *Interpolating and extrapolating with hmsIST: seeking a tmax for optimal sensitivity, resolution and frequency accuracy.* J Biomol NMR, 2017.
119. Hyberts, S.G., K. Takeuchi, and G. Wagner, *Poisson-gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data.* J Am Chem Soc, 2010. **132**(7): p. 2145-7.
120. Hyberts, S.G., et al., *Application of iterative soft thresholding for fast reconstruction of NMR data non-uniformly sampled with multidimensional Poisson Gap scheduling.* J Biomol NMR, 2012. **52**(4): p. 315-27.
121. Donoho, D.L., *Compressed sensing.* IEEE Trans Inf Theory, 2006. **52**(4): p. 1289-1306.
122. Donoho, D.L., *De-noising by soft-thresholding.* IEEE Trans Inf Theory, 1995. **41**(3): p. 613-627.
123. Rahi, S.J. and K. Sharp, *Mapping Complicated Surfaces onto a Sphere.* Int J Comput Geom Appl, 2006. **17**: p. 305-329.
124. Shchukina, A., et al., *Pitfalls in compressed sensing reconstruction and how to avoid them.* J Biomol NMR, 2016: p. 1-20.
125. Rovnyak, D., et al., *Resolution and sensitivity of high field nuclear magnetic resonance spectroscopy.* J Biomol NMR, 2004. **30**(1): p. 1-10.
126. Palmer, M.R., et al., *Sensitivity of nonuniform sampling NMR.* J Phys Chem B, 2015. **119**(22): p. 6502-15.
127. Stetz, M.A. and A.J. Wand, *Accurate determination of rates from non-uniformly sampled relaxation data.* J Biomol NMR, 2016. **65**(3-4): p. 157-170.
128. Marques, B.S., et al., *Measurement and control of pH in the aqueous interior of reverse micelles.* J Phys Chem B, 2014. **118**(8): p. 2020-31.
129. Clore, G.M., et al., *Identification and localization of bound internal water in the solution structure of interleukin 1 beta by heteronuclear three-dimensional 1H rotating-frame*

Overhauser 15N-1H multiple quantum coherence NMR spectroscopy. Biochemistry, 1990. **29**(24): p. 5671-6.

130. Grzesiek, S. and A. Bax, *Measurement of amide proton exchange rates and NOEs with water in 13C/15N-enriched calcineurin B.* J Biomol NMR, 1993. **3**(6): p. 627-638.
131. Denisov, V.P., et al., *Kinetics of DNA hydration.* J Mol Biol, 1997. **268**(1): p. 118-136.
132. Melacini, G., R. Boelens, and R. Kaptein, *Band-selective editing of exchange-relay in protein-water NOE experiments.* J Biomol NMR, 1999. **13**(1): p. 67-71.
133. Melacini, G., R. Kaptein, and R. Boelens, *Editing of chemical exchange-relayed NOEs in NMR experiments for the observation of protein–water interactions.* J Magn Reson, 1999. **136**(2): p. 214-218.
134. Cai, S., et al., *Solvent interaction of a Hsp70 chaperone substrate-binding domain investigated with water-NOE NMR experiments.* Biochemistry, 2003. **42**(38): p. 11100-11108.
135. Qiu, W., et al., *Protein surface hydration mapped by site-specific mutations.* Proc Natl Acad Sci U S A, 2006. **103**(38): p. 13979-84.
136. Zhang, L., et al., *Protein hydration dynamics and molecular mechanism of coupled water-protein fluctuations.* J Am Chem Soc, 2009. **131**(30): p. 10677-91.
137. Armstrong, B.D. and S. Han, *Overhauser dynamic nuclear polarization to study local water dynamics.* J Am Chem Soc, 2009. **131**(13): p. 4641-7.
138. King, J.T. and K.J. Kubarych, *Site-specific coupling of hydration water and protein flexibility studied in solution with ultrafast 2D-IR spectroscopy.* J Am Chem Soc, 2012. **134**(45): p. 18705-12.
139. Sharp, K.A., et al., *Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects.* Science, 1991. **252**(5002): p. 106-9.
140. Gabl, S., O. Steinhauser, and H. Weingartner, *From short-range to long-range intermolecular NOEs in ionic liquids: frequency does matter.* Angew Chem Int Ed Engl, 2013. **52**(35): p. 9242-6.
141. Berger, L.R. and R.S. Weiser, *The beta-glucosaminidase activity of egg-white lysozyme.* Biochim Biophys Acta, 1957. **26**(3): p. 517-21.
142. Das, D.K., A. Patra, and R.K. Mitra, *Preferential solvation of lysozyme in dimethyl sulfoxide/water binary mixture probed by terahertz spectroscopy.* Biophys Chem, 2016. **216**: p. 31-6.
143. Panuszko, A., et al., *Characteristics of hydration water around hen egg lysozyme as the protein model in aqueous solution. FTIR spectroscopy and molecular dynamics simulation.* Phys Chem Chem Phys, 2012. **14**(45): p. 15765-73.
144. Mamontov, E., et al., *Common features in the microscopic dynamics of hydration water on organic and inorganic surfaces.* J Phys Condens Matter, 2012. **24**(6): p. 064104.
145. Marconi, M., et al., *Hydration and temperature dependent dynamics of lysozyme in glucose–water matrices. A neutron scattering study.* Chem Phys, 2005. **317**: p. 274-281.
146. Cametti, C., et al., *Dielectric relaxation spectroscopy of lysozyme aqueous solutions: analysis of the delta-dispersion and the contribution of the hydration water.* J Phys Chem B, 2011. **115**(21): p. 7144-53.
147. Denisov, V.P. and B. Halle, *Protein hydration dynamics in aqueous solution.* Faraday Discuss, 1996(103): p. 227-44.
148. Baguet, E. and N. Hennebert, *Characterisation by triple-quantum filtered 17O-NMR of water molecules buried in lysozyme and trapped in a lysozyme-inhibitor complex.* Biophys Chem, 1999. **77**(2-3): p. 111-21.
149. King, J.T., et al., *Crowding induced collective hydration of biological macromolecules over extended distances.* J Am Chem Soc, 2014. **136**(1): p. 188-94.

150. Coleman, R.G. and K.A. Sharp, *Travel depth, a new shape descriptor for macromolecules: application to ligand binding.* J Mol Biol, 2006. **362**(3): p. 441-58.
151. Coleman, R.G. and K.A. Sharp, *Shape and evolution of thermostable protein structure.* Proteins, 2010. **78**(2): p. 420-33.
152. Shih, P., D.R. Holland, and J.F. Kirsch, *Thermal stability determinants of chicken egg-white lysozyme core mutants: hydrophobicity, packing volume, and conserved buried water molecules.* Protein Sci, 1995. **4**(10): p. 2050-62.
153. Malcolm, B.A., et al., *Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing.* Nature, 1990. **345**(6270): p. 86-9.
154. Schwalbe, H., et al., *A refined solution structure of hen lysozyme determined using residual dipolar coupling data.* Protein Sci, 2001. **10**(4): p. 677-88.
155. Moorman, V.R., K.G. Valentine, and A.J. Wand, *The dynamical response of hen egg white lysozyme to the binding of a carbohydrate ligand.* Protein Sci, 2012. **21**(7): p. 1066-73.
156. Garcia-Hernandez, E., et al., *Structural energetics of protein-carbohydrate interactions: Insights derived from the study of lysozyme binding to its natural saccharide inhibitors.* Protein Sci, 2003. **12**(1): p. 135-42.
157. Laible, N.J. and G.R. Germaine, *Bactericidal activity of human lysozyme, muramidase-inactive lysozyme, and cationic polypeptides against Streptococcus sanguis and Streptococcus faecalis: inhibition by chitin oligosaccharides.* Infect Immun, 1985. **48**(3): p. 720-8.
158. Maenaka, K., et al., *Dissection of protein-carbohydrate interactions in mutant hen egg-white lysozyme complexes and their hydrolytic activity.* J Mol Biol, 1995. **247**(2): p. 281-93.
159. Cheetham, J.C., P.J. Artymiuk, and D.C. Phillips, *Refinement of an enzyme complex with inhibitor bound at partial occupancy. Hen egg-white lysozyme and tri-N-acetylchitotriose at 1.75 A resolution.* J Mol Biol, 1992. **224**(3): p. 613-28.
160. Vyas, N., *Atomic features of protein-carbohydrate interactions.* Curr Opin Struct Biol 1991. **1**: p. 732-740.
161. Karplus, M., T. Ichiye, and B.M. Pettitt, *Configurational entropy of native proteins.* Biophys J, 1987. **52**(6): p. 1083-1085.
162. Kasinath, V., K.A. Sharp, and A.J. Wand, *Microscopic insights into the NMR relaxation-based protein conformational entropy meter.* J Am Chem Soc, 2013. **135**(40): p. 15092-100.
163. Mittermaier, A. and L.E. Kay, *Review - New tools provide new insights in NMR studies of protein dynamics.* Science, 2006. **312**(5771): p. 224-228.
164. Bellissent-Funel, M.C., et al., *Water determines the structure and dynamics of proteins.* Chem Rev, 2016. **116**(13): p. 7673-7697.
165. Lipari, G., Szabo, A. , *Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity.* J Am Chem Soc, 1982. **104**: p. 4546-4559.
166. Lange, O.F., et al., *Recognition dynamics up to microseconds revealed from an RDC-derived ubiquitin ensemble in solution.* Science, 2008. **320**(5882): p. 1471-5.
167. Wei, G.H., et al., *Protein ensembles: How does Nature harness thermodynamic fluctuations for life? The diverse functional roles of conformational ensembles in the cell.* Chem Rev, 2016. **116**(11): p. 6516-6551.
168. Lewandowski, J.R., et al., *Direct observation of hierarchical protein dynamics.* Science, 2015. **348**(6234): p. 578-81.
169. Dashnau, J.L., et al., *Hydrogen bonding and the cryoprotective properties of glycerol/water mixtures.* J Phys Chem B, 2006. **110**(27): p. 13670-7.

170. Qin, Y., et al., *Dynamics of hydration water and coupled protein sidechains around a polymerase protein surface.* Chem Phys Lett, 2017. **683**: p. 658-65.
171. DuBay, K.H., G.R. Bowman, and P.L. Geissler, *Fluctuations within folded proteins: implications for thermodynamic and allosteric regulation.* Acc Chem Res, 2015. **48**(4): p. 1098-105.
172. Lisi, G.P. and J.P. Loria, *Solution NMR spectroscopy for the study of enzyme allostery.* Chem Rev, 2016. **116**(11): p. 6323-69.
173. Motlagh, H.N., et al., *Disordered allostery: lessons from glucocorticoid receptor.* Biophys Rev, 2015. **7**(2): p. 257-265.
174. Barnes, R., et al., *Spatially heterogeneous surface water diffusivity around structured protein surfaces at equilibrium.* J Am Chem Soc, 2017. **139**(49): p. 17890-17901.
175. Nguyen, C.N., et al., *Thermodynamics of water in an enzyme active site: grid-based hydration analysis of coagulation factor Xa.* J Chem Theory Comput, 2014. **10**(7): p. 2769-2780.
176. Nguyen, C.N., T.K. Young, and M.K. Gilson, *Grid inhomogeneous solvation theory: hydration structure and thermodynamics of the miniature receptor cucurbit[7]uril.* J Chem Phys, 2012. **137**(4): p. 044101.
177. Alber, T. and B.W. Matthews, *Structure and thermal-stability of phage-T4 lysozyme.* Methods Enzymol, 1987. **154**: p. 511-533.
178. Muchmore, D.C., et al., *Expression and N-15 Labeling of Proteins for Proton and N-15 Nuclear-Magnetic-Resonance.* Methods Enzymol, 1989. **177**: p. 44-73.
179. Rumbley, J.N., L. Hoang, and S.W. Englander, *Recombinant equine cytochrome c in Escherichia coli: high-level expression, characterization, and folding and assembly mutants.* Biochemistry, 2002. **41**(47): p. 13894-901.
180. Morcombe, C.R. and K.W. Zilm, *Chemical shift referencing in MAS solid state NMR.* J Magn Reson, 2003. **162**(2): p. 479-86.
181. Gledhill, J.M. and A.J. Wand, *Al NMR: a novel NMR data processing program optimized for sparse sampling.* J Biomol NMR, 2012. **52**(1): p. 79-89.
182. Palmer, A.G.r., et al., *Sensitivity improvement in proton-detected two-dimensional heteronuclear correlation NMR spectroscopy.* J Magn Reson, 1991. **93**(1): p. 151-170.
183. Marion, D., et al., *Three-dimensional heteronuclear NMR of nitrogen-15 labeled proteins.* J Am Chem Soc, 1989. **111**(4): p. 1515-1517.
184. Jeener, J., et al., *Investigation of exchange processes by two-dimensional NMR spectroscopy.* J Chem Phys, 1979. **71**(11): p. 4546-4553.
185. Bothner-By, A.A., et al., *Structure determination of a tetrasaccharide: transient nuclear Overhauser effects in the rotating frame.* J Am Chem Soc, 1984. **103**(3): p. 811-813.
186. Bax, A. and D.G. Davis, *Practical Aspects of Two-Dimensional Transverse NOE Spectroscopy.* J Magn Reson, 1985. **63**(1): p. 207-213.
187. Delaglio, F., et al., *NMRPipe: a multidimensional spectral processing system based on UNIX pipes.* J Biomol NMR, 1995. **6**(3): p. 277-93.
188. Goddard, T. and D. Kneller, in *SPARKY 3.0*. 2014: University of California, San Francisco.
189. Norwood, T.J., et al., *Comparison of techniques for 1H-detected heteronuclear 1H□15N Spectroscopy.* J Magn Reson, 1990. **87**(3): p. 488-501.
190. Bax, A., et al., *Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins.* J Magn Reson, 1990. **86**(2): p. 304-318.
191. Zuiderweg, E.R. and S.W. Fesik, *Heteronuclear three-dimensional NMR spectroscopy of the inflammatory protein C5a.* Biochemistry, 1989. **28**(6): p. 2387-91.
192. Kay, L.E., et al., *Three-dimensional triple-resonance NMR Spectroscopy of isotopically enriched proteins.* J Magn Reson, 2011. **213**(2): p. 423-441.

193. Bax, A. and M. Ikura, *An efficient 3D NMR technique for correlating the proton and 15N backbone amide resonances with the alpha-carbon of the preceding residue in uniformly 15N/13C enriched proteins.* J Biomol NMR, 1991. **1**(1): p. 99-104.
194. Kasinath, V., K.G. Valentine, and A.J. Wand, *A 13C labeling strategy reveals a range of aromatic side chain motion in calmodulin.* J Am Chem Soc, 2013. **135**(26): p. 9560-3.
195. Tugarinov, V. and L.E. Kay, *An isotope labeling strategy for methyl TROSY spectroscopy.* J Biomol NMR, 2004. **28**(2): p. 165-172.
196. Lakomek, N.A., J. Ying, and A. Bax, *Measurement of (1)(5)N relaxation rates in perdeuterated proteins by TROSY-based methods.* J Biomol NMR, 2012. **53**(3): p. 209-21.
197. Kay, L.E., D.A. Torchia, and A. Bax, *Backbone dynamics of proteins as studied by 15N inverse detected heteronuclear NMR spectroscopy: application to staphylococcal nuclease.* Biochemistry, 1989. **28**(23): p. 8972-9.
198. Tjandra, N., et al., *Rotational diffusion anisotropy of human ubiquitin from 15N NMR relaxation.* J Am Chem Soc, 1995. **117**(50): p. 12562-6.
199. Dellwo, M.J. and A.J. Wand, *Model-independent and model-dependent analysis of the global and internal dynamics of cyclosporin A.* J Am Chem Soc, 1989. **111**(13): p. 4571-8.
200. Sharff, A.J., L.E. Rodseth, and F.A. Quiocho, *Refined 1.8-A structure reveals the mode of binding of beta-cyclodextrin to the maltodextrin binding protein.* Biochemistry, 1993. **32**(40): p. 10553-9.
201. Grishaev, A., et al., *Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints.* J Biomol NMR, 2008. **40**(2): p. 95-106.
202. Baldwin, R.L. and G.D. Rose, *How the hydrophobic factor drives protein folding.* Proc Natl Acad Sci U S A, 2016. **113**(44): p. 12462-6.
203. Cheng, N., *Formula for the viscosity of a glycerol−water mixture.* Ind Eng Chem Res, 2008. **47**(9): p. 3285-88.