

University of Nevada, Reno

Molecular Dynamics Simulation Studies on Interfacial Dynamics of Proteins

A dissertation submitted in partial fulfillment of the requirements for the degree
of Doctor of Philosophy in Chemical Physics

by

Yao Xu

Dr. David M. Leitner/Dissertation Advisor

August, 2013

© Copyright by Yao Xu 2013

All Rights Reserved



THE GRADUATE SCHOOL

We recommend that the dissertation
prepared under our supervision by

YAO XU

entitled

Molecular Dynamics Simulation Studies On Interfacial Dynamics Of Proteins

be accepted in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

David M. Leitner, Advisor

Peter Winkler, Committee Member

Andrei Derevianko, Committee Member

Sergey A. Varganov, Committee Member

Mark Pinsky, Graduate School Representative

Marsha H. Read, Ph. D., Dean, Graduate School

August, 2013

ABSTRACT

Clusters of water molecules embedded in proteins or at the interface between globules mediate protein dynamics, allostery and charge transfer. Confined clusters of water molecules exhibit dynamics distinct from dynamics of water in the bulk or in the hydration layer around a biomolecule. Here, we describe and compare the dynamics of water in the hydration layer around antifreeze proteins, the dynamics of a cluster of water molecules at the interface between the globules of a homodimeric hemoglobin, and the dynamics of water molecules within green fluorescent protein (GFP) and the hydration water in the vicinity of photoactive yellow protein (PYP), all studied by molecular dynamics simulations. The water dynamics in these systems span a wide range of time scales and are closely related to function, ranging from relatively fast hydration water dynamics to tightly bound water clusters that mediate protein conformational change during ligand binding. Subsequent quantum mechanical studies on energy relaxation in proteins GFP and PYP provide insights into the microenvironment of the chromophore and identify energy flow channels that may mediate the kinetics of photochemical reactions such as charge transfer and conformational isomerization.

Acknowledgements

I would like to express my deepest appreciation to my PhD research advisor Prof. David M. Leitner. He is definitely a great mentor with a passion for science, compassion for students, and commitment to excellence. His well-roundedness, passion and perseverance in research will always inspire me.

I would also like to acknowledge with much gratitude my PhD committee members, Profs. Peter Winkler, Andrei Derevianko, both of Physics Department, Sergey A. Varganov of Chemistry Department, and Mark Pinsky of Mathematics and Statistics Department for their supervision on my progress towards PhD.

Many thanks should go to Dr. Renante Yson for his enormous assistance in our research group computing cluster system administration, and former and current research group members, Drs. Ramachandran Gnanasekaran, Johnson K. Agbo for their tremendous help in my research, as well as students I worked with, Hari Datta Pandey, Marcus Sacchetti, Taylor Kaplan and Mia Swain. Special thanks should go to the staff and friends of mine at the Department of Chemistry, and the University of Nevada, Reno. I would never have been able to finish this work without their kind and constant support.

Finally, this work would have remained a dream had it not been for the everlasting love from my parents.

Table of Contents

1	Introduction.....	1
1.1	Introduction.....	1
1.2	References.....	6
2	Analysis of Water and Hydrogen Bond Dynamics at the Surface of an Antifreeze Protein.....	7
2.1	Introduction.....	8
2.2	Computational Methods.....	10
2.3	Results and Discussions.....	12
2.4	Concluding Remarks.....	16
2.5	References.....	17
3	Long-range Protein-water Dynamics in Hyperactive Insect Antifreeze Proteins.....	32
3.1	Introduction.....	34
3.2	Results.....	38
3.3	Discussion.....	43
3.4	Materials and Methods.....	45
3.5	Acknowledgements.....	48
3.6	References.....	49
4	Dynamics of Water Clusters Confined in Proteins: A Molecular Dynamics Simulation Study of Interfacial Waters in a Dimeric Hemoglobin.....	57
4.1	Introduction.....	58
4.2	Computational Details.....	61
4.3	Results and Discussion.....	64
4.4	Concluding remarks.....	72
4.5	Acknowledgements.....	73
4.6	References.....	80
5	The Dielectric Response to Photoexcitation of GFP: A Molecular Dynamics Study.....	85
5.1	Introduction.....	86
5.2	Computational Methods.....	89
5.3	Results and Discussion.....	93
5.4	Concluding Remarks.....	97
5.5	References.....	104
6	Dielectric Response in Photoactive Yellow Protein: Quantum Mechanics/Molecular Mechanics and Molecular Dynamics Simulation Studies.....	106
6.1	Introduction.....	107
6.2	Computational Methods.....	110
6.3	Results and discussion.....	112
6.4	Concluding Remarks.....	114
6.5	Acknowledgements.....	114
6.6	References.....	119
7	Communication Maps and Non-equilibrium Simulation Computed for Green Fluorescent Protein and Photoactive Yellow Protein: Studies of Energy Transport in Proteins.....	121

7.1	Introduction.....	122
7.2	Methods	124
7.3	Results and Discussion	131
7.4	Concluding remarks.....	135
7.5	Acknowledgments	136
7.6	References.....	156
8	Conclusion	161
8.1	References.....	166

List of Figures

Figure 2-1 Structure of antifreeze protein from spruce budworm <i>Choristoneura fumiferana</i> (PDB: 1LOS).....	19
Figure 2-2 Structure of wild-type AFP with original Theronine (THR) on ice-binding Plane 1.....	20
Figure 2-3 Structure of mutated AFP with Leucine (LEU) on ice-binding Plane 1.	21
Figure 2-4 Power spectra of hydrogen-bonded water around the three different planes of wild-type AFP at 200 K.	22
Figure 2-5 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of wild-type AFP at 200 K.	23
Figure 2-6 Power spectra of hydrogen-bonded water around the three different planes of wild-type AFP at 300 K.	24
Figure 2-7 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of wild-type AFP at 300 K.	25
Figure 2-8 Power spectra of hydrogen-bonded water around the three different planes of mutated AFP at 200 K.....	26
Figure 2-9 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of mutated AFP at 200 K.	27
Figure 2-10 Power spectra of hydrogen-bonded water around the three different planes of mutated AFP at 300 K.....	28
Figure 2-11 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of mutated AFP at 300 K.	29
Figure 2-12 The hydrogen-bond lifetime autocorrelation function of hydrogen-bonded water around the three different planes of wild-type AFP at 300 K.....	30
Figure 2-13 The hydrogen-bond lifetime autocorrelation function of hydrogen-bonded water around the three different planes of mutated AFP at 300 K.	31
Figure 3-1 Difference $\Delta\alpha$ in THz absorbance (integrated between 2.4 and 2.7 THz) of dissolved DAFP-1 in water relative to a water reference, plotted against the concentration. Measurements were carried out at (20 ± 0.5) and (5 ± 0.5) °C blue data points respectively. DAFP-1 shows a concentration dependent excess of THz absorbance that shifts to lower protein concentration at the lower temperature (blue data points).	52
Figure 3-2 Difference $\Delta\alpha$ in THz absorbance (integrated between 2.4 and 2.7 THz) of dissolved DAFP-1 in water relative to a water reference (black data point) and dissolved DAFP-1 in 0.5 M Na ₃ Citrate relative to a citrate buffer reference (green data points).	53
Figure 3-3 <i>Upper Panel.</i> Difference $\Delta\alpha$ in THz absorbance (integrated between 2.4 and 2.7 THz) of dissolved AFP-1 (left) and AFGP (right) relative to a water reference (black) and enhanced AFP-1 and AFGP in 0.5 M Na ₃ Citrate relative to a citrate buffer reference (green). <i>Lower Panel.</i> (left) Thermal Hysteresis of AFP-1 (black) and AFP-1 in 0.5 M citrate buffer as a function of protein concentration in water. (right) Thermal Hysteresis of AFGP (black) and AFGP in citrate buffer.	54
Figure 3-4 Computed power spectra for water molecules in the hydration layer	

- around the protein DAFP-1 and for water molecules in the bulk. Spectra were calculated at 300 K (left) and 250 K (right).55
- Figure 3-5 Hydrogen bond lifetime correlation function $C(t)$ for bonds between water molecules around the binding (red) and non-binding plane (blue) of DAFP-1 at 300 K. Dark red (binding) and dark blue (non-binding plane) indicate the mutant discussed in text. $C(t)$ is also plotted for hydrogen bonds between water molecules at the binding plane (green) and non-binding plane (dark green) of the wild type in sodium citrate solution. ..56
- Figure 4-1 Structures of deoxy- (top) and oxy-homodimeric (bottom) hemoglobin from *Scapharca inaequivalvis* (HbI). The unliganded protein and liganded protein have 17 and 11 water molecules at the interface, respectively.76
- Figure 4-2 The top two black curves are hydrogen bond time correlation functions, $C_{HB}(t)$, for hydrogen bonds between water and protein for the 11 interfacial water molecules of oxy-HbI (dashed curve) and the 17 water molecules of deoxy-HbI (solid curve). Thin solid lines through $C_{HB}(t)$ are stretched exponentials fit to the computed $C_{HB}(t)$ from 0.2 to 200 ps; the parameters for the fits are listed in Table 4-1. Gray solid and dashed curves, which are nearly identical, correspond to $C_{HB}(t)$ for hydration waters with the protein for the deoxy- and oxy- system, respectively. Thin lines through the computed $C_{HB}(t)$ are fits to stretched exponentials. The dot-dashed curve corresponds to $C_{HB}(t)$ computed for hydrogen bonds of pure water.77
- Figure 4-3 The top two black curves are rotational anisotropy decay functions ($n = 1$), $C_1(t)$, for the 11 interfacial water molecules of oxy-HbI (dashed curve) and the 17 water molecules of deoxy-HbI (solid curve). Thin solid lines through $C_1(t)$ for the interfacial water molecules are stretched exponentials fit to the computed $C_1(t)$ from 0.2 to 200 ps; the parameters for the fits are listed in Table 4-2. Gray solid and dashed curves, which are nearly identical, correspond to $C_1(t)$ for hydration waters within 3 Å of the protein surface for the deoxy- and oxy- system, respectively. The dot-dashed curve corresponds to $C_1(t)$ computed for hydrogen bonds of pure water.78
- Figure 4-4 The top two black curves are rotational anisotropy decay functions ($n = 2$), $C_2(t)$, for the 11 interfacial water molecules of oxy-HbI (dashed curve) and the 17 water molecules of deoxy-HbI (solid curve). Thin solid lines through $C_2(t)$ for the interfacial water molecules are stretched exponentials fit to the computed $C_2(t)$ from 0.2 to 200 ps; the parameters for the fits are listed in Table 4-2. Gray solid and dashed curves, which are nearly identical, correspond to $C_2(t)$ for hydration waters within 3 Å of the protein surface for the deoxy- and oxy- system, respectively. The dot-dashed curve corresponds to $C_2(t)$ computed for hydrogen bonds of pure water.79
- Figure 5-1 Illustration of green fluorescence protein from *Aequorea victoria* in its PDB structure, produced with VMD. The chromophore, surrounded by 10 confined crystallographic water molecules, is shown in atomic detail in cyan. Several additional water molecules are at times present in the β -barrel during the MD simulations. Inset (black background) shows the isolated chromophore used in MD simulations to determine the dielectric response of the water due to photoexcitation of this molecule in aqueous solution. 100

- Figure 5-2 Hydrogen bond correlation function, $C_{HB}(t)$, for hydrogen bonds between water molecules and protein atoms in the hydration layer (blue), in the β -barrel (red) and hydrogen bonds between the chromophore and water (green) at 300 K, plotted to 100 ps. $C_{HB}(t)$ for hydrogen bonds between water molecules in the bulk is shown for comparison (black). Thin black curves are fits to a stretched exponential, $\exp[-(t/\tau)^\beta]$, for which the parameters are listed in Table 5-1. Inset shows a log-log plot of the time-dependence of C_{HB} to 1 ns (see text). 101
- Figure 5-3 Normalized equilibrium time autocorrelation functions of fluctuations in the solvation energy differences at 300 K (solid curves) for the interactions between the chromophore and the rest of the system (black), the chromophore and bulk water (blue). Thin black curves are fits to the data using Eq. (5-8). The parameters for the fits are listed in Table 5-2..... 102
- Figure 5-4 (a) Dielectric response for chromophore bound to GFP (black) and its linear response decomposition, Eq. (5-6), in terms of the components present in the simulation, confined water (solid blue), and protein matrix (solid green); (b) auto-correlation functions for protein (dotted green) and water (dotted blue) and the cross-correlation function (dotted red) between chromophore-water and chromophore-protein interactions, both given by Eq. (5-7). They are shown with the dielectric response for chromophore bound to GFP (black). 103
- Figure 6-1 Ribbon diagram of PYP (left) showing all atoms of the chromophore, which is attached to Cys69, and nearby residues Glu46, Tyr42, to which it hydrogen bonds, and Thr50, which also appears close the chromophore anion during MD and QM/MM simulation. Arg52, which also lies near the chromophore, is also shown. The hydrogen bonding between the chromophore and protein residues is indicated on the right, as is the ground state charge minus the excited state charge on chromophore atoms, in units of electron charge..... 116
- Figure 6-2 The dielectric response of the protein and solvent to photoexcitation of the PYP chromophore, calculated using the interaction energy, E , between the chromophore and the rest of the system (thick solid curves, black for MD and red for QM/MM). The two results both fit well, other than some oscillations, to a tri-exponential, Eq. (6-3) (thin curves, blue for MD and green for QM/MM). 117
- Figure 6-3 Normalized equilibrium time autocorrelation functions of fluctuations in the solvation energy differences at 300 K (solid curves) for the interactions between the chromophore and bulk water (black for MD and red for QM/MM). Blue and green curves are fits to the MD and QM/MM data, respectively, using Eq. (6-3). The parameters for the fits are listed in Table 6-1..... 118
- Figure 7-1 Ribbon diagrams of crystallographic structures of GFP (left) and PYP (right). The chromophores in both proteins are highlighted in licorice representation. The GFP chromophore is surrounded by 10 confined crystallographic water molecules. 137
- Figure 7-2 Thermal average distance maps are plotted for ground state GFP at 300 K. In GFP, residue 64 represents the chromophore, and the confined water molecules are represented as residues 227 – 236. 138
- Figure 7-3 Thermal average communication maps are plotted for GFP at 300 K at ground and excited state, respectively. In GFP, residue 64 represents the chromophore, and

the confined water molecules are represented as residues 227 – 236.....	140
Figure 7-4 Thermal average communication maps are plotted for PYP at 300 K at ground and excited state, respectively. In PYP, the chromophore is represented as residue 126.....	142
Figure 7-5 Top 10 hot residues interacting with the chromophore of GFP, obtained by frequency-dependent communication maps (top: ground state; bottom: excited state).....	144
Figure 7-6 Top 10 hot residues interacting with the chromophore of PYP, obtained by frequency-dependent communication maps (top: ground state; bottom: excited state).....	146
Figure 7-7 Hot regions of ground state GFP shown from 1 to 3 ps after the excitation of the chromophore (top: 1 ps; middle 2 ps; bottom: 3 ps). The colors indicated correspond to different percent kinetic energy of the whole system contained in a residue or confined water. Residues with at least 5% of the total energy are colored in red, 4% – 5% in orange, 3% - 4% in yellow, 2% -3% in green, 1% - 2% in blue.	149
Figure 7-8 Hot regions of ground state PYP shown from 1 to 3 ps after the excitation of the chromophore (top: 1 ps; middle 2 ps; bottom: 3 ps). The colors indicated correspond to different percent kinetic energy of the whole system contained in a residue. Residues with at least 5% of the total energy are colored in red, 4% – 5% in orange, 3% - 4% in yellow, 2% -3% in green, 1% - 2% in blue.	152
Figure 7-9 Hot regions of excited state PYP shown from 1 to 3 ps after the excitation of the chromophore (top: 1 ps; middle 2 ps; bottom: 3 ps). The colors indicated correspond to different percent kinetic energy of the whole system contained in a residue. Residues with at least 5% of the total energy are colored in red, 4% – 5% in orange, 3% - 4% in yellow, 2% -3% in green, 1% - 2% in blue.	155

List of Tables

Table 4-1 Listed are values of β and τ for the stretched exponential, $\exp[-(t/\tau)^\beta]$, fit to $C_{\text{HB}}(t)$ for hydrogen bonds between hydration water molecules around oxy and deoxy Hbl and the protein, and for hydrogen bonds between interfacial water molecules and the protein.....	74
Table 4-2 Listed are values of β and τ for the stretched exponential, $\exp[-(t/\tau)^\beta]$, fit to the rotational reorientation correlation function of order $n = 1$ and 2 for interfacial water molecules of oxy and deoxy Hbl.	75
Table 5-1 Parameters for the stretched exponential, $\exp[-(t/\tau)^\beta]$, used to fit the hydrogen bond correlation functions computed for hydrogen bonds between water molecules and the chromophore, the interior of the β -barrel, and the hydration layer around GFP (excluding the interior of the β -barrel). The fits are also plotted in Fig. 5-2. ..	98
Table 5-2 Parameters used to fit Eq. (5-8) to the computed dielectric response for the chromophore in solvated GFP, and the chromophore in aqueous solution.....	99
Table 6-1 Parameters used to fit Eq. (6-3) to the computed dielectric response for the chromophore in solvated PYP (Fig. 6-2), and the chromophore in bulk water (Fig. 6-3).....	115

1 Introduction

1.1 Introduction

Water is ubiquitous in biological systems and chemical processes.¹ It is usually considered essential for life, especially at microscopic level for its diverse roles in its interactions with the so-called “workhorse” molecule of life, proteins, ranging from protein architecture and folding to dynamics and function.² However, anomalous behavior of water and multi-functions of proteins make protein-water interactions notoriously complex and mysterious.³ This complexity can be reflected by the fact that water molecules in aqueous solutions of proteins have been broadly grouped into bulk-like water, the hydration water surrounding proteins and internal bound water.⁴ Bound water molecules, often identified crystallographically, occupy internal cavities and deep clefts. Such water molecules are extensively involved in the protein-water hydrogen bonding and believed to play a structural role in protein architecture and folding. Hydration water exhibits a heterogeneous dynamical behavior due to both chemical and physical interaction with and topological roughness of the solvent-exposed protein surface. And water in almost no direct contact with proteins can undergo continuous exchanging with hydration water and reveal properties similar to those of bulk water.

Perhaps the most unexpected evidence of protein-water relationship comes from the discovery that a critical level of hydration (less than 0.40 grams of water per gram of protein), even insufficient to completely cover the protein surface, is required to fully activate the dynamics and function of globular protein.⁵ Unfortunately, the mechanisms of the relationship through protein-water dynamical coupling, which is of extreme importance to understanding biological processes such as enzymatic activity, molecular

recognition and protein folding, is far from being fully understood, particularly at the protein-water interface. This interface is probably one of the most intriguing and complex systems under investigation in modern biophysics and life science, as witnessed from a growing interest over decades in constructing a unified approach to understanding the role of interfacial waters in biological and chemical processes from perspectives of theory, experimental results and computer simulations.⁶ The theoretical efforts on the interfacial dynamics can be traced back to molecular relaxation theory in classical liquid state physics. Pioneers such as Smoluchowski, Kramers, Kubo, Zwanzig built up a phenomenological picture at the hydrodynamic level for the dynamics based on principles and formalisms such as the fluctuation-dissipation theorem, linear response theory and time-correlation function.⁷ However, the dynamics and relaxation are essentially relevant to transport of mass, momentum and energy through the interface with a wide range of time scales. Therefore, it was difficult to achieve a deep understanding without constructing a microscopic picture at molecular level. The situation did not change until the extension of hydrodynamics and kinetic theory by Leutheusser, Götze, Mazenko and coworkers.⁸ They, at around the same time, developed a mode-coupling theory in supercooled liquids studies to describe the observed glassy dynamics, which was realized later on as a shared feature between proteins and glassy systems in protein physics. On the other hand, various experimental tools and techniques based on dielectric relaxation, acoustic attenuation, dynamic light scattering, NMR and ultrafast spectroscopy have been used in detecting the interfacial dynamics in different spatial and temporal resolutions.⁹ Recently, with the addition of multidimensional infrared and electronic spectroscopy and single-molecule spectroscopy, our

understanding of the dynamical processes has been greatly enhanced.⁹ Perhaps one of the most relevant experiments that shed light on interfacial dynamics throughout decades originates from solvation dynamics studies by ultrafast spectroscopies.¹⁰ The studies basically have undergone three phases, from the discovery of multiexponential solvation dynamics in the 1980s, through that of subpicosecond ultrafast dynamics in the 1990s, to solvation dynamics in complex systems after the year 2000. The observation in phase one shows that solvation times are largely dependent on solvent properties rather than probe size and shape and that the solvation time-correlation functions were found to be non-exponential. In the second phase, ultrafast subpicosecond relaxation dynamics was observed due to the available ultrashort laser pulses with single photon counting. Since the year 2000, research focus has expanded to a wide range of complex systems; some examples are micelles and reverse micelles, micro-emulsions, and hydration water of macro-biomolecules such as proteins and nucleic acids.¹¹ Water molecules are normally constrained in these systems and display sluggish solvation dynamics with amplitude and time constants varying from system to system. And the heterogeneity of the solvating environment and the multiple occupation sites of probes make the interpretation of the slow dynamics ambiguous and disputable.

To bridge the gap between theory and experiment, molecular dynamics simulation was introduced for its success in providing insights into liquid state studies at the molecular level, among others.¹² MD simulation computes system coordinates and velocities classically based on initial atomic coordinates and interaction potential between atoms. The interaction potential is typically an empirical energy function that includes both bonding and nonbonding interactions:

$$\begin{aligned}
 U(\mathbf{R}) = & \frac{1}{2} \sum_{bonds} K_b (b - b_0)^2 + \frac{1}{2} \sum_{angles} K_\theta (\theta - \theta_0)^2 + \frac{1}{2} \sum_{dihedrals} K_\chi [1 + \cos(n\chi - \delta)] \\
 & + \frac{1}{2} \sum_{impropers} K_{imp} (\phi - \phi_0)^2 + \sum_{nonbonded} \epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}
 \end{aligned} \tag{1-1}$$

where \mathbf{R} is the coordinates of atoms; K_b , K_θ , K_χ , and K_{imp} are bond, angle, dihedral angle, and improper dihedral angle force constants, respectively; b , θ , χ , and ϕ are the bond length, bond angle, dihedral angle, and improper torsion angle, respectively; the subscript zeros represent the equilibrium values for individual terms. Lennard–Jones 6–12 and Coulomb terms contribute to the external or nonbonded interactions; ϵ_{ij} is Lennard–Jones well depth, and σ_{ij} is the distance at Lennard–Jones minimum between atom i and j ; q_i is partial atomic charge, ϵ_0 is dielectric constant, and r_{ij} is the distance between atom i and j . Different sets of parameters for empirical potential energy function were developed (CHARMM, GROMOS, AMBER, etc.).¹³ The parameters are initially obtained from experimental and quantum mechanical studies of small molecules and then refined to yield correct experimental structural and spectroscopic results.

To include solvation effects, a variety of potential functions to describe intermolecular liquid water interactions were proposed and tested: for example, SPC, SPC/E, TIP3P, etc. These models differ from the number of interactions sites, geometric arrangements, and parameterization of charged sites. The macromolecule is normally solvated in a large box (cubic, rectangular, or truncated octahedron) filled with solvent molecules, and the whole system is replicated in three dimensions with periodic boundary conditions. To simulate experimental conditions directly, MD has grown into simulation approaches with the consideration of constant temperature (canonical or NVT ensemble)

or both constant pressure and temperature (isobaric–isothermal or NPT ensemble) from its initial emphasis on energy conservation only (microcanonical or NVE ensemble).

In practice, MD simulation starts from an initial set of atomic coordinates and velocities. The coordinates are commonly obtained from X-ray crystallographic or NMR structure data or by model modeling and the velocities are assigned based on the Maxwell-Boltzmann distribution at fixed temperature. Then the whole system propagates dynamically by updating coordinates and velocities according to Newton's second law. In macromolecular systems, bond vibrations are often constrained to fixed length and angle iteratively to reduce computer time and to preserve an appropriate structure.

Here, we have carried out molecular dynamics simulation studies on dynamical coupling between proteins and water molecules at interfaces. The dissertation is organized as follows: in Chapter 2 and 3, we computationally analyze water hydrogen bond dynamics near the surface of two types of antifreeze proteins and compare our results with recent terahertz spectroscopic studies; in Chapter 4, we focus on interfacial water confined between globules of invertebrate hemoglobin and relate the roles of water to protein function through both computational and theoretical analysis; in Chapter 5 and 6, intrigued by recent ultrafast spectroscopic studies, we investigate the water molecules involved in photochemical reactions in photoreceptor proteins such as green fluorescent protein and photoactive yellow protein and identify the confined water in GFP and the neighboring residues hydrogen-bonding to the PYP chromophore as mediator in charge transfer and conformational isomerization, respectively, followed by a quantum perspective on the dynamics of energy relaxation to characterize residue-based energy flow channels within the proteins in Chapter 7.

1.2 References

1. Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*. (Cambridge University Press, 2013).
2. Ball, P. Water as a Biomolecule. *ChemPhysChem* **9**, 2677–2685 (2008).
3. Mattos, C. Protein-water interactions in a dynamic world. *Trends Biochem. Sci.* **27**, 203–208 (2002).
4. Bizzarri, A. R. & Cannistraro, S. Molecular Dynamics of Water at the Protein–Solvent Interface. *J. Phys. Chem. B* **106**, 6617–6633 (2002).
5. Poole, P. L. & Finney, J. L. Sequential hydration of a dry globular protein. *Biopolymers* **22**, 255–260 (1983).
6. Bagchi, B. *Molecular Relaxation in Liquids*. (Oxford University Press, USA, 2012).
7. Hynes, J. T. Chemical Reaction Dynamics in Solution. *Annu. Rev. Phys. Chem.* **36**, 573–597 (1985).
8. Götze, W. *Complex Dynamics of Glass-Forming Liquids: A Mode-Coupling Theory*. (Oxford University Press, Oxford, 2009).
9. Serdyuk, I. N., Zaccai, N. R. & Zaccai, J. *Methods in Molecular Biophysics: Structure, Dynamics, Function*. (Cambridge University Press, 2007).
10. Maroncelli, M. & Fleming, G. R. Comparison of time-resolved fluorescence Stokes shift measurements to a molecular theory of solvation dynamics. *J. Chem. Phys.* **89**, 875–881 (1988).
11. Leitner, D. M., Gruebele, M. & Havenith, M. Solvation dynamics of biomolecules: modeling and terahertz experiments. *Hfsp J.* **2**, 314–323 (2008).
12. Allen, M. P. & Tildesley, D. J. *Computer Simulation of Liquids*. (Oxford University Press, USA, 1989).
13. Ponder, J. W. & Case, D. A. Force fields for protein simulations. *Adv. Protein Chem.* **66**, 27–85 (2003).

2 Analysis of Water and Hydrogen Bond Dynamics at the Surface of an Antifreeze Protein

Abstract

We examine dynamics of water molecules and hydrogen bonds at the water-protein interface of the wild-type antifreeze protein from spruce budworm *Choristoneura fumiferana* and a mutant that is not antifreeze active by all-atom molecular dynamics simulations. Water dynamics in the hydration layer around the protein is analyzed by calculation of velocity autocorrelation functions and their power spectra, and hydrogen bond time correlation functions are calculated for hydrogen bonds between water molecules and the protein. Both water and hydrogen bond dynamics from subpicosecond to hundred picosecond time scales are sensitive to location on the protein surface and appear correlated with protein function. In particular, hydrogen bond lifetimes are longest for water molecules hydrogen bonded to the ice-binding plane of the wild type, whereas hydrogen bond lifetimes between water and protein atoms on all three planes are similar for the mutant.

Keywords: Antifreeze Proteins, Hydration Water Dynamics, Velocity Autocorrelation Function, Power Spectra, Hydrogen Bond Lifetime.

2.1 Introduction

Widely known as biologically complex systems, proteins demonstrate complex dynamics in terms of non-Debye density of vibrational states, anomalous distribution of vibrational energy, and non-exponential decay of correlations, for example. The complexity may be to great extent due to their complex energy landscapes since proteins, as many-particle systems, possess a great number of geometrically distinct minima and saddle points on the energy surface. Perhaps a more challenging aspect would be to introduce those proteins involved in phase transition under investigation. Antifreeze proteins (AFPs), of this kind as a distinguisher between ice and liquid water, are a class of polypeptides widely distributed in certain plants, vertebrates, fungi and bacteria with protection against subzero environments.¹ They obtain a unique ability to lower the temperature at which an ice crystal will grow, as the temperature is depressed below the equilibrium melting/freezing point to a lower, hysteric freezing point.² Numerous fields³ have witnessed the applications of AFPs, ranging from cold avoidance in plant biology⁴ to cryosurgery in medicine⁵.

Ice crystals in solution normally grow only on basal and prism planes appeared as round and flat discs.⁶ However, it has been found the presence of AFPs on other preferred ice-binding planes, at least for AFP type I in winter flounder and shorthorn sculpin.⁷ The generally accepted mechanism for this function is adsorption-inhibition mechanism,^{8, 9, 10} which proposes that AFPs adsorb onto the preferred growth sites of ice surface, thereby preventing ice growth in the normal low radius of curvature and restricting growth to regions between the adsorbed AFPs in a high radius of curvature front,² and hence leading to the prevention of new ice growth until the temperature is lowered to the

hysteretic freezing point, according to the Kelvin effect.^{2,9,10,11,12}

Through studies¹³ on type I AFP, it was initially thought that ice and AFP interacted through hydrogen bonding.⁸ However, when parts of the protein that were thought to facilitate this hydrogen bonding were mutated, the hypothesized decrease in antifreeze activity was not observed.¹⁴ Recent data suggest this could be because of hydrophobic interactions.¹⁴ Since it is difficult to understand the exact mechanism of binding around the complex water-ice interface, efforts to unveil the precise details are being made through the use of theoretical calculations such as molecular dynamics (MD) simulation.⁶

The present work focused on the atomistic details of the influence of water environment on protein dynamics of the wild-type AFP (Fig. 2-1) from the spruce budworm *Choristoneura fumiferana* and its mutant (mutation from Theronine to Leucine, see Fig. 2-2 and 2-3) by employing molecular dynamics simulation, providing some phenomenological reflection of their complex dynamics and energy landscape. Following a brief introduction of the background of the system and simulation methodology, we will begin with a discussion of the protein and water dynamics by examining velocity autocorrelation functions and power spectra. Next, we will introduce the investigation of hydrogen bond lifetime correlation function to unravel protein-water interactions at different temperatures, ranging from room temperature to protein transition temperature, which is at ~200 K for soluble proteins. In the final section, we will summarize and conclude.

2.2 Computational Methods

The initial coordinates of the antifreeze protein with resolution of 2.3 Å X-ray crystal structure from the spruce budworm *Choristoneura fumiferana* were taken from the Protein Data Bank file 1LOS.¹⁵ Missing residues and hydrogen atoms were built into the structure and the iodated tyrosine Y26 required for the structure determination was reverted to a standard tyrosine using the software package Swiss PDB Viewer¹⁶. For the mutant, four theronines residues on Plane 1 were mutated to leucines to explore the roles of the theronines play in the ice-binding process on three different planes of the AFP, intrigued by Davies and coworkers' ice crystal morphology method and ice-etching experiment¹⁷.

Both the wild-type and mutated structure were first minimized for 1000 steps with steepest descent algorithm under the AMBER force field,^{18,19} after its solvation in a 70 Å x 70 Å x 70 Å cubic water box of TIP5P water model, Then the systems, which individually comprised of 10539 number of water molecules, were equilibrated for 400 ps. For the first 100 ps the positions of the proteins were restrained and in the later 300 ps they were released. Constraints were applied to all bonds to hydrogen with the SHAKE algorithm²⁰ and periodic boundary conditions were applied. All the classical MD simulations were performed on the systems in canonical (NVT) ensemble by GROMACS software package²¹. Following the equilibration, trajectories of 2 ns were obtained at different temperatures of 200 K and 300 K with Nose-Hoover thermostat^{22, 23}. Non-bonded interactions were gradually brought to zero by a shift function for the electrostatics as well as a switch function for van der Waals interactions between 10 and 12 Å.^{24, 25} All the simulations were performed by integrating the Newtonian equation of

motion with a standard leapfrog algorithm²⁶ of 1 fs time steps. The system coordinates and velocities were stored at every 5 fs, and velocity autocorrelation function (VACF) was averaged over 15 ps time segments of the trajectory for the oxygen atoms that survive in the first hydration shell of thickness 5 Å and that involve in forming hydrogen bonds, respectively before Fourier transform to obtain power spectra. And the corresponding hydrogen bond lifetime autocorrelation functions were also computed at 300 K.

2.3 Results and Discussions

A. Power Spectrum

A protein molecule is known to perturb the regular water-water hydrogen bond network in bulk water with the formation of protein-water hydrogen bonds by their influence on the water dynamics at the protein surface. To investigate this, we calculated the velocity autocorrelation function and Fourier-transform it to obtain the power spectra for the water molecules, both forming hydrogen bonds with the amino acid residues of the three planes and present within a layer of thickness 5 Å from the residues of the three planes of the protein²⁷. This layer essentially corresponds to the first hydration layer around three planes. The VACF is defined as $C_v(t) = \langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle / \langle \mathbf{v}_i(0) \cdot \mathbf{v}_i(0) \rangle$, where $\mathbf{v}_i(t)$ is the velocity vector of the oxygen atom in a water molecule at time t . The angular brackets denote averaging over all atoms of the particular type present in the hydration shell and over different reference initial times. The results from a separate MD simulation of pure TIP5P water model for bulk water under identical conditions are also included for comparison purpose.

In Fig. 4-11, two broad low-frequency bands in the power spectra of water were seen at ~ 2.4 THz and ~ 8.4 THz. Although Walrafen and coworkers²⁸ interpreted the bands as the O \cdots O \cdots O bending mode from triplets of hydrogen-bonded water molecules and O \cdots O stretching mode between pairs of hydrogen-bonded water molecules respectively, there is still controversy over the microscopic origin of the bands. The results here show that for both wild-type and mutant, compared to bulk water a clear blue shift in $S_O(\omega)$ for the band corresponding to the O \cdots O \cdots O bending for water around the three planes can be seen at both 200 K and 300 K. However, only small differences were

noticed in the blue shifts of this band among the hydration layers of the three planes. This could relate to the glassy property of the first hydration layers of different planes of the protein molecule. On the other hand, the band at ~ 8.4 THz seems to remain unaffected with mutation at different temperatures. These support the notion that versatile influences of the protein on the orientation and translation on these water molecules lead to the apparently contrasting results on the two low-frequency bands in the power spectra of the water molecules near the protein surface, as an indication of the restricted translational mobility of water around the protein.

Moreover, for both wild-type (Fig. 2-4 to 2-7) and mutant (Fig. 2-8 to 2-11), we also calculated the power spectra $S_O(\omega)$ for the water molecules that are hydrogen bonded to the amino acid residues at the protein surface to investigate the effect of the protein-water hydrogen bonds in the THz spectra. There shows a considerable blue shift of the low-frequency band corresponding to the $O\cdots O\cdots O$ bending for the water molecules hydrogen bonded to the planes as compared to pure bulk water. However, the water molecules hydrogen bonded to Plane 3 have a more pronounced shift than those bonded to the other two planes at 300 K, but not at 200 K. Apart from that, a higher amplitude of the band at ~ 8.4 THz particularly for Planes 1 and 2 is also observed, as a result of more restricted longitudinal translations.

B. Hydrogen Bond Lifetime Correlation Function

Standard geometrical definition of hydrogen bond criterion was employed i.e., the distance between donor and acceptor atom is less than or equal to 0.35 nm and the bond angle between the hydrogen and the acceptor atom is less than or equal to 30°. The hydrogen bond lifetime correlation function $C_{HB}(t)$ is defined as $C_{HB}(t) = \langle h(t) * h(0) \rangle / \langle h(0) * h(0) \rangle$, where h is a binary variable equal to 1 if a hydrogen bond is present between two water molecules at time t and angle brackets indicate a time average over all pairs of water molecules. $C_{HB}(t)$ allows hydrogen bonds to be broken and reformed in the time interval 't'. Here hydrogen bond lifetime correlation function was calculated for the hydrated water molecules within each plane of the protein.

Quite similar to Nutt and Smith's work²⁹, our results in Fig. 2-12 and 2-13 also show that at 300 K within the shell, the lifetime correlation function of the water around Plane 1 decays significantly slower than that of the water around the other two planes for the wild-type, while all the decays of the three planes of the mutant become much closer over the same time of period, showing the decrease in the antifreeze activity of the AFP after the introduction of apolar leucines. Similar deactivation can also be found on the other two non-ice-binding planes. Moreover, we found that the behavior of C_{HB} has two time scales: a short time scale that is strongly dependent on the solvation shell with a characteristic time τ_{HB} of ~ 1 ps and a longer time scale that is similar for all three planes of the protein with a characteristic time τ_R on the order of ~ 10 ps. This indicates that the water dynamics here are highly complex and different processes are coupled. One explanation is from Doster and Settles' interpretation³⁰ of the role of the hydrogen bond network in the dynamical transition of the protein between a glassy state, in which the

protein side chains are hindered by rigid hydrogen bonds, and a liquid-like state in which the hydrogen bonds fluctuate on the picosecond time scale. And a further study showed that the short time scale process of hydrogen bond breaking and forming might be due to the vibration/rotation of water molecules without translational displacement, while the long time scale process may participate in the rearrangement of the hydrogen bond network via water translational displacements.³¹

2.4 Concluding Remarks

In this work, we detected blue shifts of ~ 1 THz in the low-frequency band corresponding to the O \cdots O \cdots O bending of the water molecules in the hydration layer with a more pronounced extent of the shift (> 1 THz) for the water molecules that are hydrogen bonded to the protein residues. This indicates the drastically restricted local motions by the water molecules involved in the protein-water hydrogen bonds, which are much stronger at longer lifetimes. On the other hand a somewhat prominent non-uniform blue shift of the band for protein-water hydrogen bonds was investigated than that for water molecules around the three planes. However, the low-frequency band corresponding to the O \cdots O stretching due to longitudinal translations has not been found to have any shift for the hydration layer water molecules. This implies that both the orientation and translation of the water molecules around the protein are affected non-uniformly by the influence of the protein. In addition, we have clearly seen the difference between the hydrogen bond correlations functions for Plane 1 of the wild-type and mutant, supporting the proposal that theronines on Plane 1 deactivate the antifreeze activity of the AFP at different temperatures.

From the THz spectra of water present near the surface of the protein in an aqueous solution at different temperatures by MD simulation, we attempted to interweave the complex glassy property of the hydration layers and the structural relaxation of water around different planes of the protein. It seems to us that the overall picture demonstrates that water next to the protein generally does have retardation in both translational and rotational diffusion. And the hydrated protein exhibits liquid-like motion at room temperature and complex glass-like behavior at low temperature.

2.5 References

- (1) Miller, F. P.; Vandome, A. F.; McBrewster, J. *Antifreeze Proteins*; Alphascript Publishing, **2010**
- (2) Delinger, D. L.; Lee, R. E. *Low Temperature Biology of Insects*; Cambridge University Press: New York, **2010**
- (3) <http://www.afprotein.com/ref.htm>
- (4) Moffatt, B.; Ewart, V.; Eastman, A. *Physiologia Plantarum*, **2006**, *126*, 1, 5.
- (5) Pham, L.; Dahiya, R.; Rubinsky, B. *Cryobio*. **1999**, *38*, 2, 169.
- (6) Jorov, A.; Zhorov, B. S.; Yang, D. S. *Protein Science*. **2004**, *13*, 1524.
- (7) Night, C.; Cheng, C.; DeVries, A. L. *Biophys. J.* **1991**, *59*, 409.
- (8) Raymond, J. A.; DeVries, A. L. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 2589.
- (9) Raymond, J. A.; Wilson, P. W.; DeVries, A. L. *Proc. Natl. Acad. Sci. U. S. A.* **1989**, *86*, 881.
- (10) Knight, C. A.; Cheng, C. C.; DeVries, A. L. *Biophys. J.* **1991**, *59*, 409.
- (11) Wilson, P. W. *CryoLett.* **1993**, *14*, 31.
- (12) Kristiansen, E.; Zachariassen, K. E. *Cryobio*. **2005**, *51*, 262.
- (13) Duman, J.; DeVries, A. L. *Comp. Biochem. Physiol.* **1976**, *B54*, 375.
- (14) Haymett, A.; Ward, L.; Harding, M. *FEBS Lett.* **1998**, *430*, 301.
- (15) Leinala, E. K.; Davies, P. L.; Jia, Z. *Struct.* **2002**, *10*, 619.
- (16) Guex, N.; Peitsch, M.C. *Electrophoresis* **1997**, *18*, 2714.
- (17) Graether, S. P.; Kuiper, M. J.; Gagne, S. M.; Walker, V. K.; Jia, Z.; Sykes, B. D.; Davis, P. L. *Nature*, **2000**, *406*, 325.
- (18) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, C. M.; Xiong, G.; Zhang, W.; Yang, R.;

- Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comp. Chem.* **2003**, *24*, 1999.
- (19) Sorin, E.J.; Pande, V.S. *Biophys. J.* **2005**, *88*(4), 2472.
- (20) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H.J.C. *J. Comp. Phys.* **1997**, *23*, 327.
- (21) Berendsen, H. J. C. et al. *Comp. Phys. Comm.* **1995**, *91*, 43.
- (22) Nose, S. *Mol. Phys.* **1984**, *52*, 255.
- (23) Hoover, W. G. *Phy. Rev.* **1985**, *A31*, 1695.
- (24) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- (25) Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T., Lee, H.; Pedersen, L.G. *J. Chem. Phys.* **1995**, *103*, 8577.
- (26) Hockney, R.W.; Goel, S.P.; Eastwood, J. *J. Comp. Phys.* **1974**, *14*, 148.
- (27) Bandyopadhyay, S.; Charkraborty, S.; Bagchi, B. *J. Am. Chem. Soc.* **2005**, *127*, 16660.
- (28) (a) Walrafen, G. E.; Chu, Y. C. *J. Chem. Phys.* **1995**, *99*, 11225.
- (b) Walrafen, G. E.; Chu, Y. C.; Piermarini, G. J. *J. Phys. Chem.* **1996**, *100*, 10363.
- (29) Nutt, D. R.; Smith, J. C. *J. Am. Chem. Soc.* **2008**, *130*, 13066.
- (30) Doster, W.; Settles, M. *Hydration Processes in Biology: Experimental and Theoretical Approaches*; IOS Press: Amsterdam, **1999**, 305, 177.
- (31) Tarek, M.; Tobias, D. J. *Phys. Rev. Lett.* **2002**, *88*, 138101.

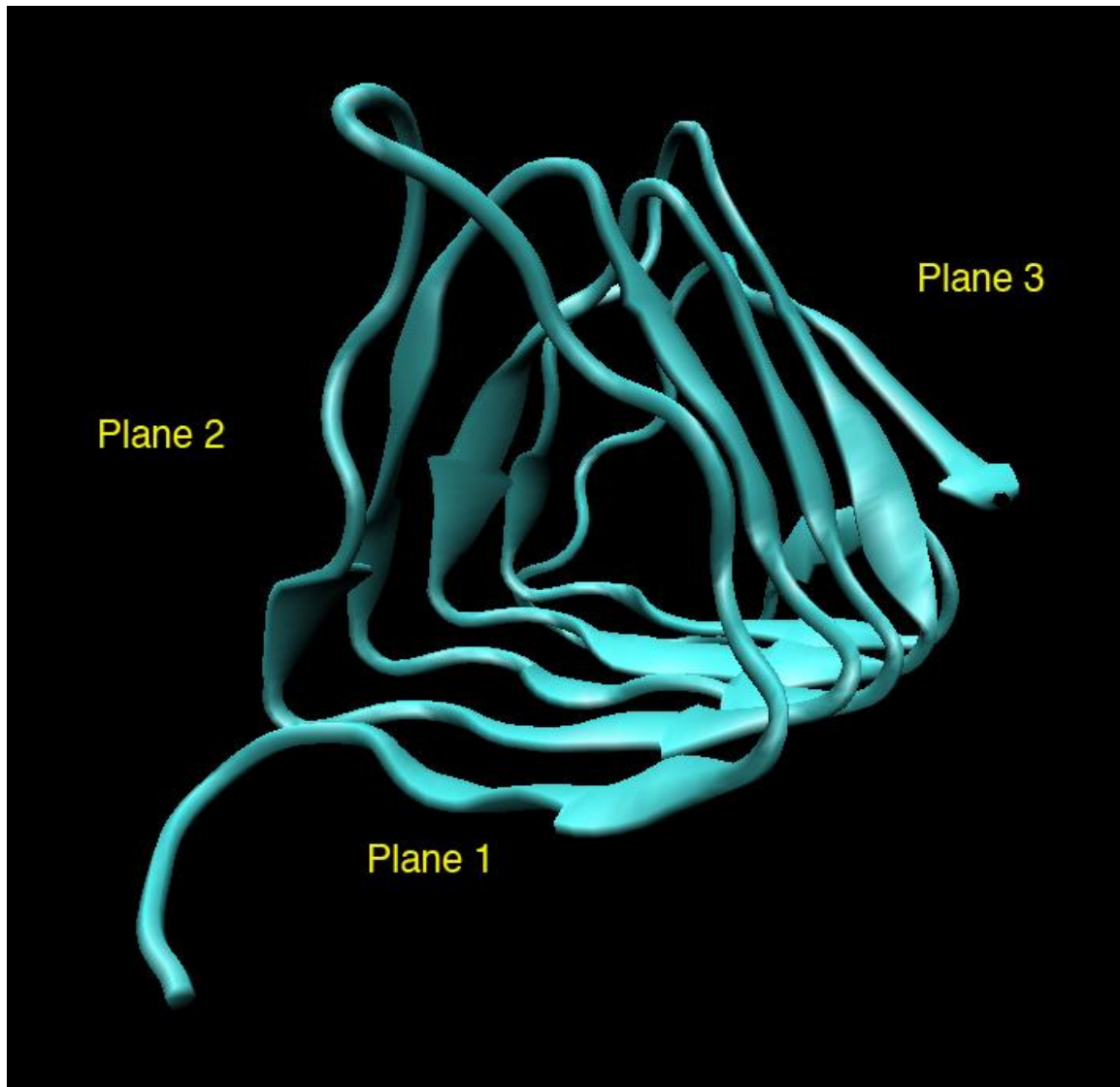


Figure 2-1 Structure of antifreeze protein from spruce budworm *Choristoneura fumiferana* (PDB: 1L0S).

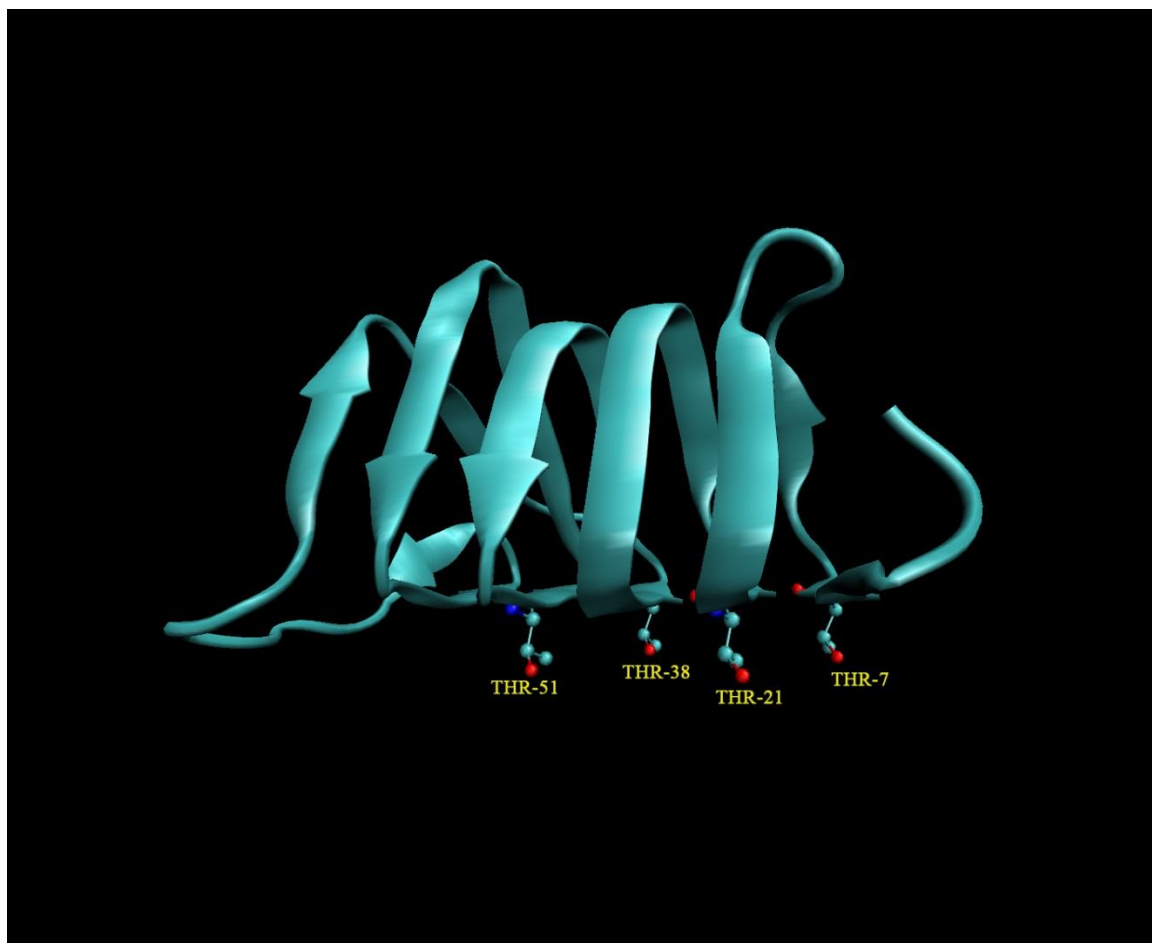


Figure 2-2 Structure of wild-type AFP with original Theronine (THR) on ice-binding Plane 1.

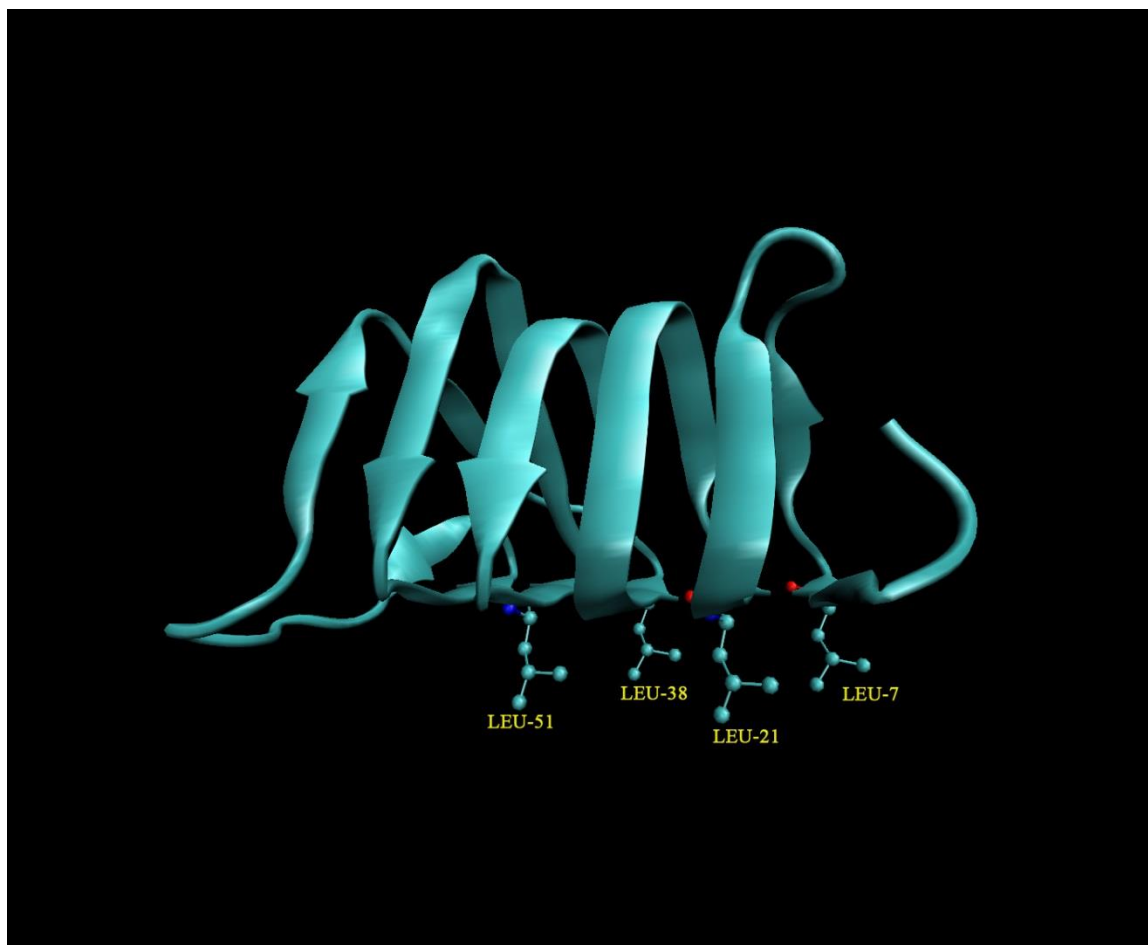


Figure 2-3 Structure of mutated AFP with Leucine (LEU) on ice-binding Plane 1.

Power spectra of AFP at 200K for 3 different planes

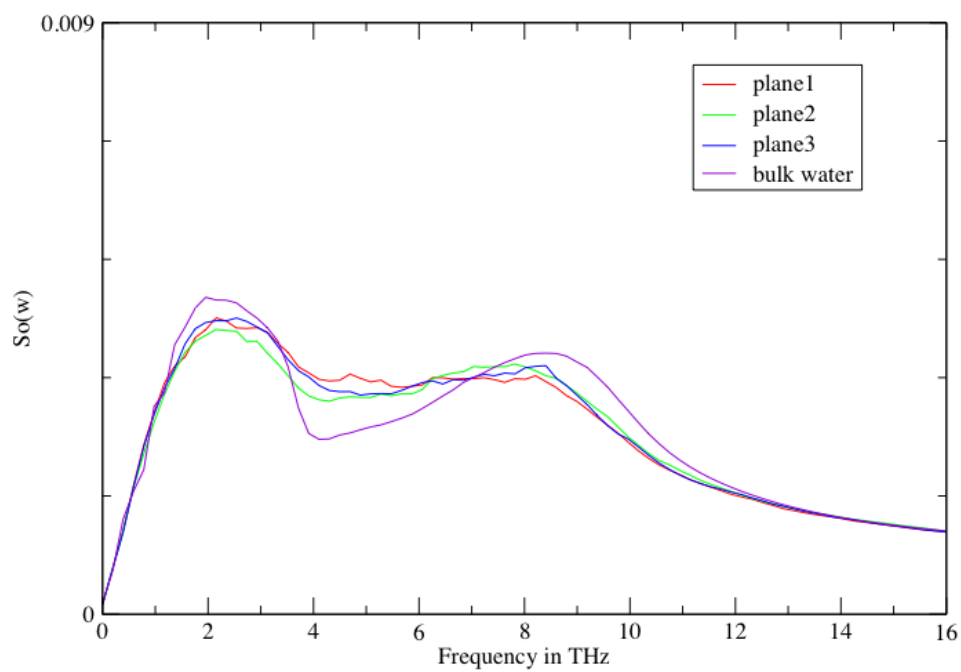


Figure 2-4 Power spectra of hydrogen-bonded water around the three different planes of wild-type AFP at 200 K.

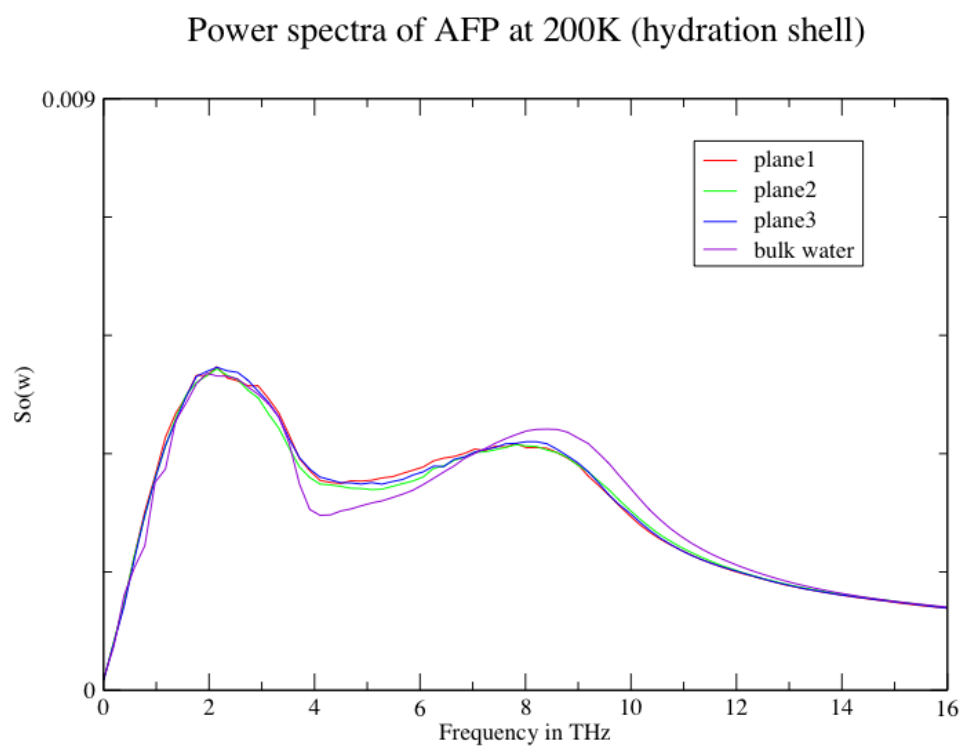


Figure 2-5 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of wild-type AFP at 200 K.

Power spectra of AFP at 300K for 3 different planes

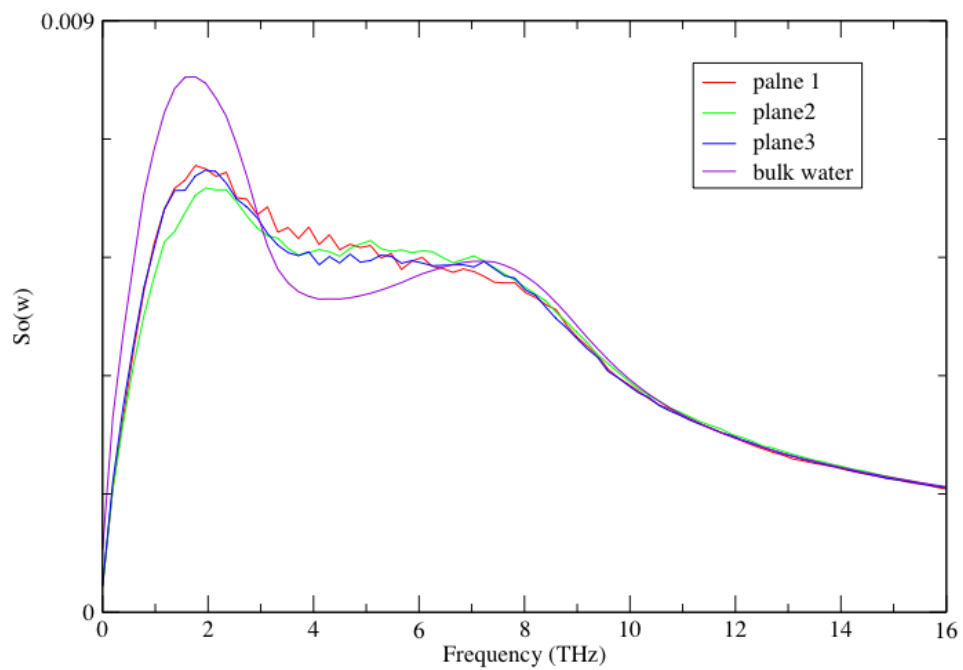


Figure 2-6 Power spectra of hydrogen-bonded water around the three different planes of wild-type AFP at 300 K

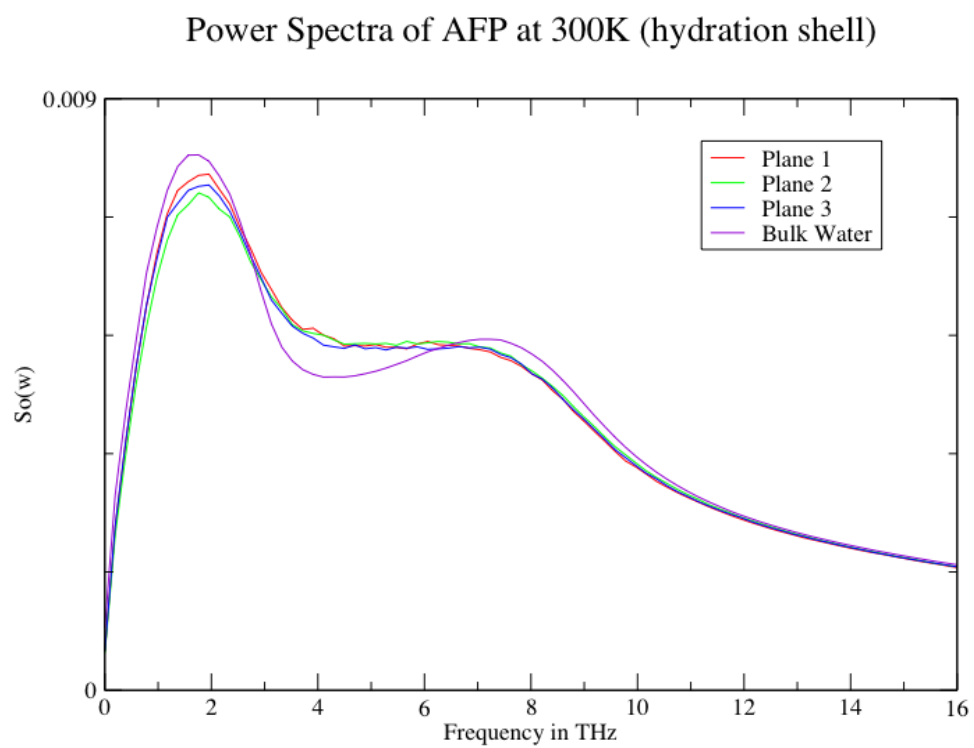


Figure 2-7 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of wild-type AFP at 300 K.

Power spectra of mutated AFP at 200K (HB)

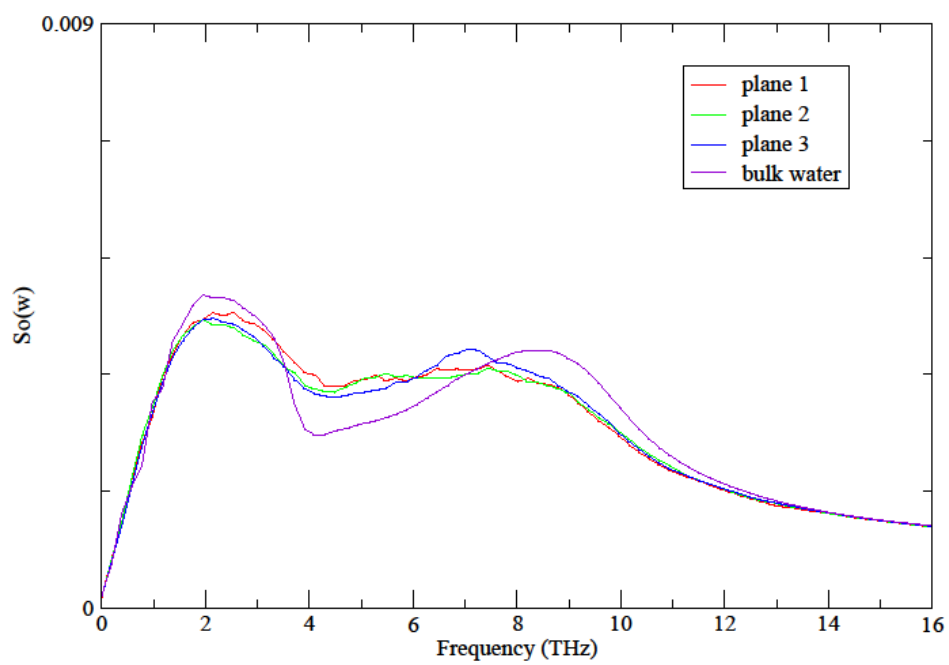


Figure 2-8 Power spectra of hydrogen-bonded water around the three different planes of mutated AFP at 200 K.

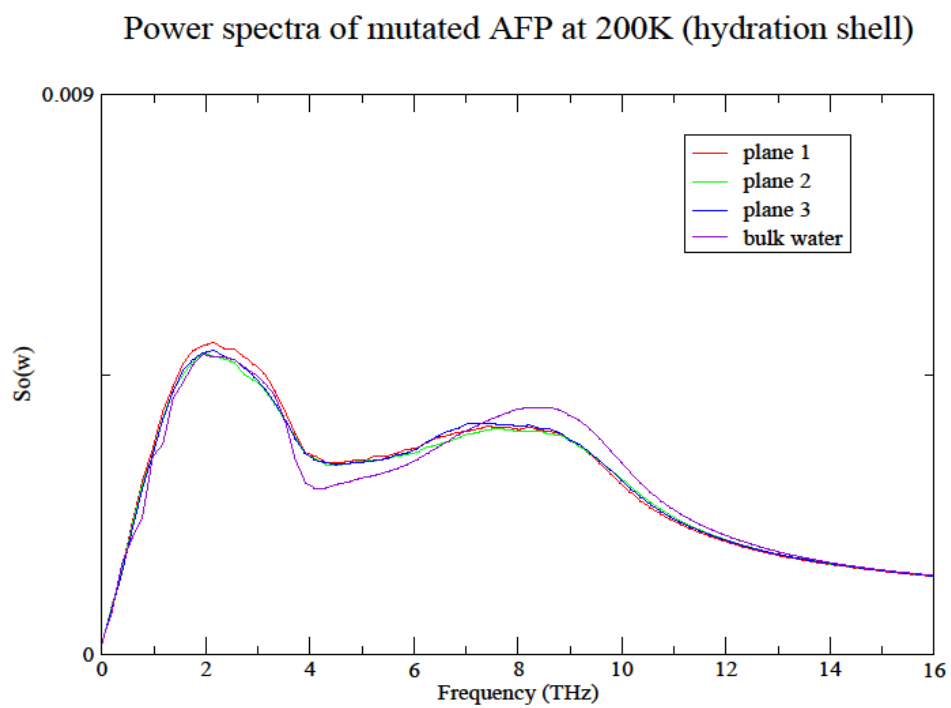


Figure 2-9 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of mutated AFP at 200 K.

Power spectra of mutated AFP at 300K (HB)

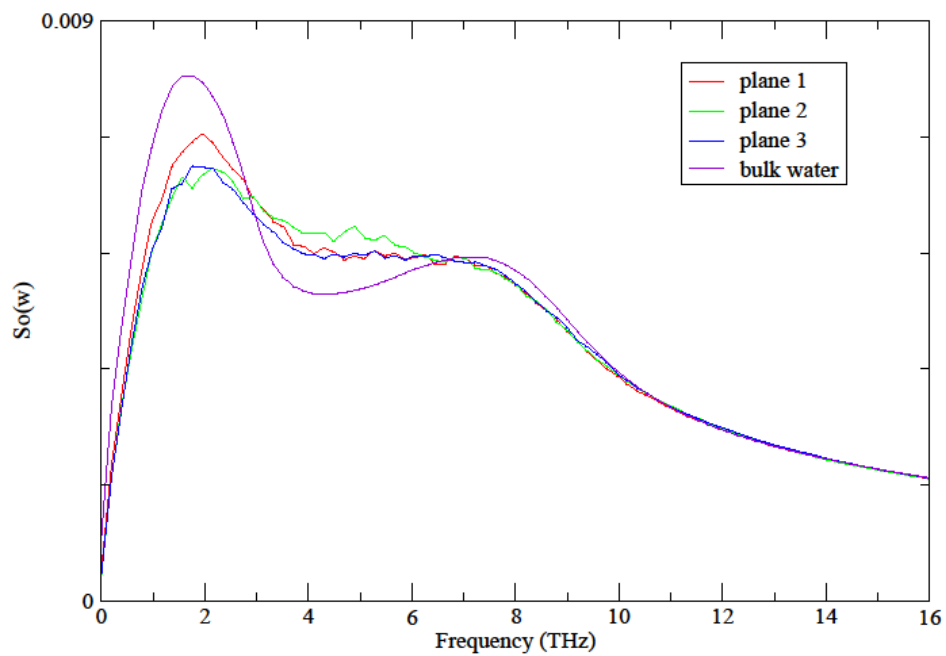


Figure 2-10 Power spectra of hydrogen-bonded water around the three different planes of mutated AFP at 300 K.

Power spectra of mutated AFP at 300K (hydration shell)

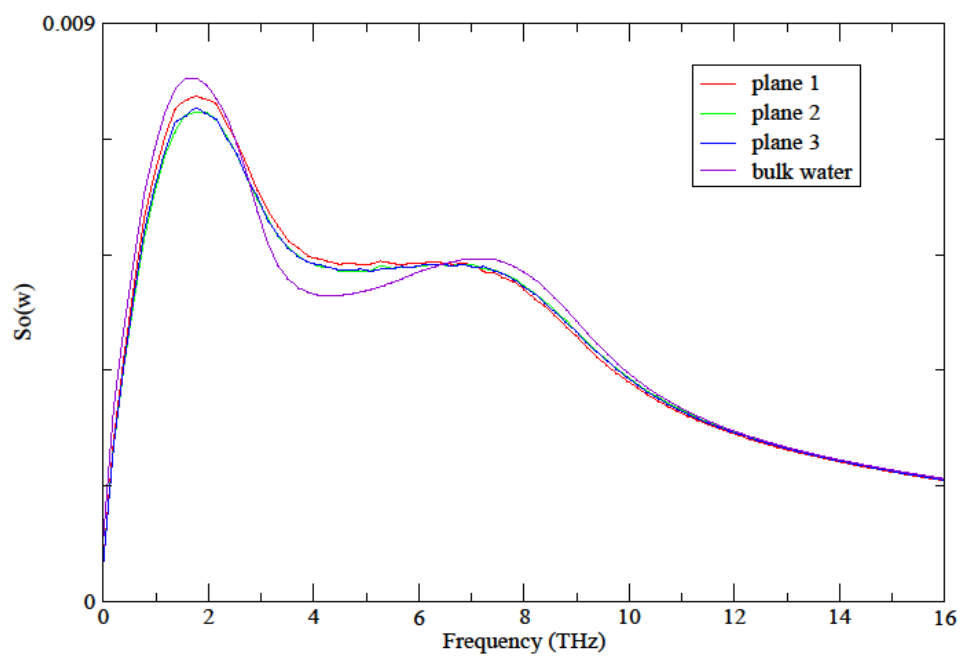


Figure 2-11 Power spectra of hydrogen-bonded water in the first hydration shell around the three different planes of mutated AFP at 300 K.

Wild AFP hydrogen bond correlation functions for three planes at 300K

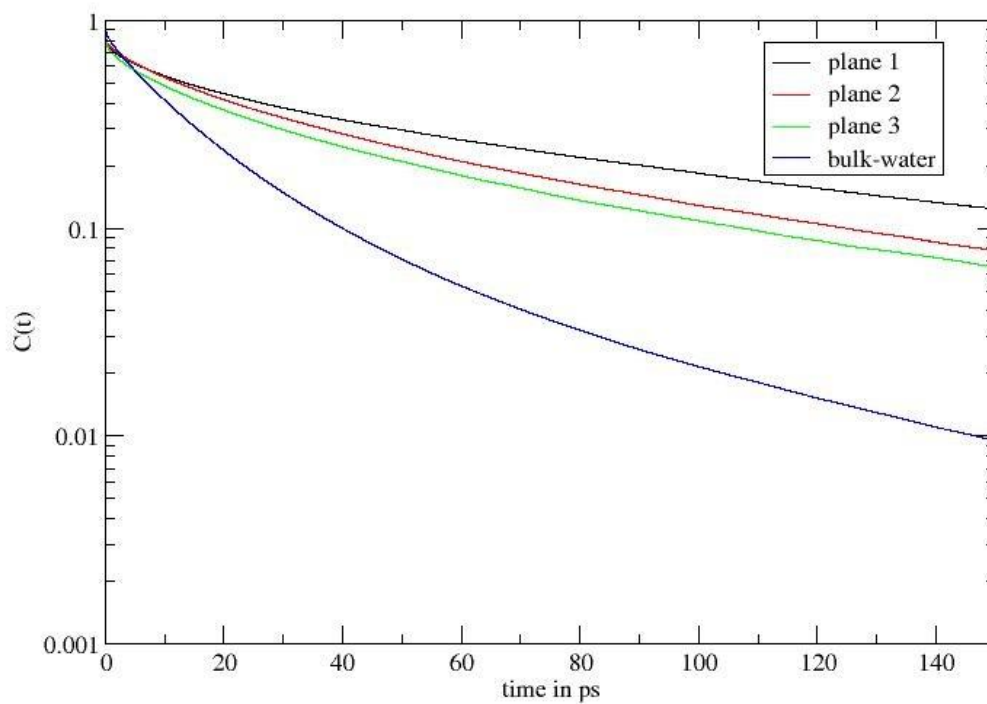


Figure 2-12 The hydrogen-bond lifetime autocorrelation function of hydrogen-bonded water around the three different planes of wild-type AFP at 300 K.

Mutated AFP hydrogen bond correlation functions for three planes at 300K

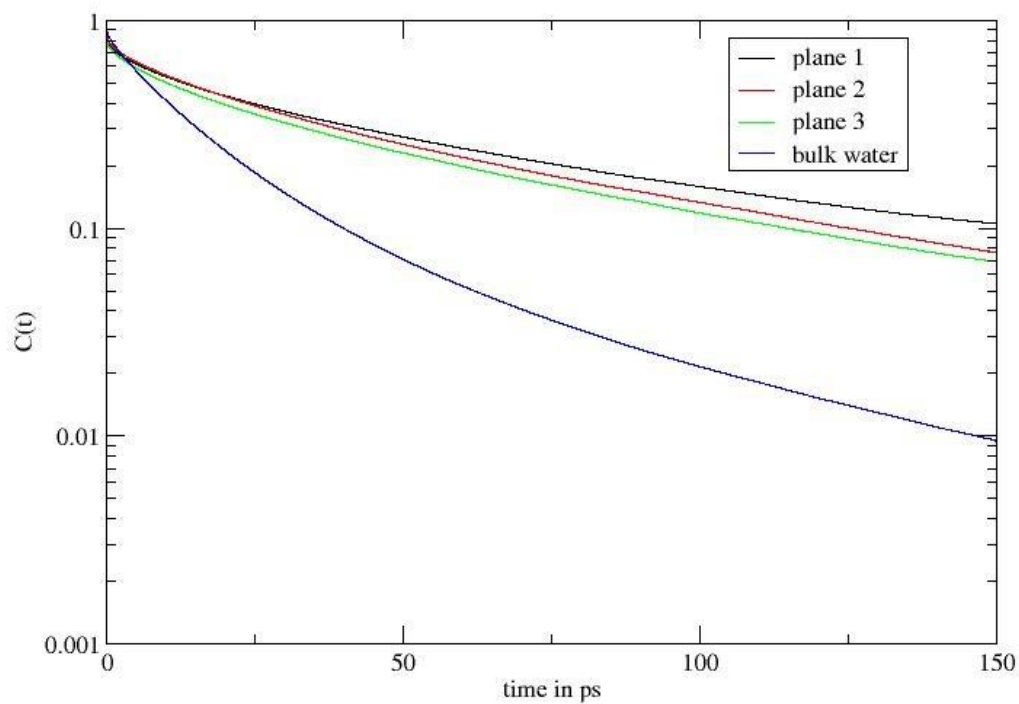


Figure 2-13 The hydrogen-bond lifetime autocorrelation function of hydrogen-bonded water around the three different planes of mutated AFP at 300 K.

3 Long-range Protein-water Dynamics in Hyperactive Insect

Antifreeze Proteins

Konrad Meister^a, Simon Ebbinghaus^a, Yao Xu^b, John G. Duman^c, Arthur L. DeVries^d,
Martin Gruebele^{e,f}, David M. Leitner^b and Martina Havenith^a

^aLehrstuhl für Physikalische Chemie II, Ruhr Universität Bochum 44801 Bochum,
Germany;

^bDepartment of Chemistry, University of Nevada, Reno, NV 89557,

^cDepartment of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556;

^dDepartment of Animal Biology, University of Illinois at Urbana-Champaign, Urbana, IL
61801;

^eDepartment of Physics and

^fDepartment of Chemistry and Center of Biophysics and Computational Biology,
University of Illinois at Urbana-Champaign, Urbana, IL 61801

Abstract

Antifreeze proteins (AFPs) are specific proteins which are able to lower the freezing point compared to the melting point. Hyperactive AFPs, identified in insects, give rise to a special high ability for freezing point depression, which far exceeds that of other AFPs. In previous studies we postulated that the activity of AFP can be attributed to two distinct molecular mechanisms: a) short range direct interaction of the protein surface with the growing ice face and b) long range via changes in the water dynamics by the solute-solvent interaction exceeding up to 20 Å from the protein surface. In the present paper we combine terahertz spectroscopy and molecular simulations to prove that long-range-protein-water interactions make essential contributions to the high antifreeze activity of insect antifreeze proteins from the beetle *Dendroides canadensis*. We further support our hypothesis by studying the effect of the addition of the osmolyte sodium citrate.

3.1 Introduction

Antifreeze proteins (AFPs) and antifreeze glycoproteins (AFGPs) are classes of proteins that suppress ice nucleation and thereby enable the survival of organisms in subfreezing habitats (1). Despite its similar function many distinct structures have been identified so far. AFPs have been identified in several organisms, including polar fish (2), insects (3), bacteria (4) and plants (5). Their common characteristic is the depression of the freezing point of a solution without depressing the melting point. This nonequilibrium phenomenon leads to a difference between the freezing and melting temperature which is referred to as thermal hysteresis (TH). Thermal hysteresis is used as a characteristic measure for antifreeze activity of an AFP (6). AFGPs and AFPs, as extracted from the blood of polar fish, usually exhibit up to 2 °C thermal hysteresis activity and are termed moderately active AFPs, while insect AFPs can exhibit over 5 °C thermal hysteresis and are therefore referred to as hyperactive AFPs. Raymond and DeVries proposed a mechanism in which freezing point depression is achieved by an adsorption-inhibition mechanism in which the proteins recognize and bind quasi-irreversibly to an ice surface, thereby preventing growth of ice crystals (7, 8). The adsorption of the protein is thought to prevent macroscopic ice-growth in the hysteresis gap, but microscopic growth occurs at the interface in the form of highly curved fronts between adsorbed antifreeze molecules. This will cause a decrease of the local freezing temperature due to the Kelvin effect while leaving the melting temperature relatively unaffected (7). As recently pointed out by Sharp (9), AF activity involves one of the most difficult recognition problems in biology, the distinction between water as a liquid and ice (9). The initially proposed mechanism builds upon a local mechanism. In particular, the Thr residue was proposed to

play a decisive role: their hydroxyl groups were thought to be responsible for the high affinity required for binding of the AFP to the ice crystals. Mutations in which Thr was replaced by Ala were ambiguous in the conclusions which could be drawn: the mutation Thr→Ala was found to decrease thermal hysteresis activity for AFP I (10), whereas for AFP III the opposite mutation, Ala→Thr is found to decrease activity (11). This generally accepted molecular mechanism can explain hysteresis at fairly high protein concentrations, but a completely local mechanism cannot fully explain how some AFPs prevent freezing at very low concentrations, when water is present in great excess. Therefore the role of water molecules in the hydration shell of the AFP protein has become the focus of attention (12-15).

Evolution has produced two structurally different solutions to the antifreeze problem, the AFPs and AFGPs, which have similar THz excess and antifreeze activity on a mass-scaled basis. The latter display many saccharide OH groups in a disordered manner (16), while wf-AFP-1 displays a few threonine OH groups projecting from one side of a relatively rigid helix. In previous studies on AFGP, Havenith and coworkers proposed a distinct mechanism for their antifreeze activity: Using THz absorption spectroscopy to probe the hydration dynamics around proteins and thus the change in the collective water network motions we found a considerable long range influence on the hydration dynamics for antifreeze active intrinsically disordered AFGP. The presence of an extended dynamical hydration shell which is associated with a retardation in the H-bond network dynamics was proposed to be a mandatory for its antifreeze activity (long term hypothesis) (16). Studies of the short rigid α -helix wfAFP-1 revealed that inactive AFP mutants lacked a long range influence on the H-bond dynamics (17), but the long

range effect did not predict the strength of the activity. Instead we concluded that the occurrence of AF activity requires short range of OH group binding by wfAFP-1, which needs to be assisted by changes in the long range hydration dynamics (17). This is in accordance with the proposal of Sharp who states that the recognition event of AFP is more subtle, involving more than one kind of interaction. Only this would make it possible to combine the required high affinity (i.e. by bond using the OH-groups in AFP) with a high selectivity which could be achieved by a significant change of the hydration dynamics around ice binding site. We would like to point out here that retardation of the H-bond dynamics at the OH binding sites could reduce the entropic cost for binding of ice-crystals and thus could also further assist a local mechanism.

Hyperactive insect antifreeze proteins allow insects to survive even at temperatures of up to $-30\text{ }^{\circ}\text{C}$. Their increased hysteresis activity at a comparable binding affinity to ice (18) must be attributed to a special enhancement effect. The need for a new molecular mechanism that acts at low protein concentration is supported by the recent finding of a moderate active plant antifreeze protein that binds to both basal and primary-prism planes of ice, normally a hallmark for hyperactive insect AFPs (19). In this study we have focused on the investigation of a hyperactive antifreeze protein from the colored beetle *Dendroides Canadensis* (DAFP-1). Hyperactive insect antifreeze proteins such as DAFP or TmAFP from the mealworm *Tenebrior Melitor* are right handed β -helical proteins with nearly identical 12- or 13- amino acid repeats (20, 21). By a combination of terahertz spectroscopy and molecular dynamic simulations we can show that long range protein water interactions play an important role in explaining the hyperactive antifreeze activity of insect AFPs. Given that all AFPs, including the hyperactive insect antifreeze,

bind irreversibly to ice, we thus confirmed our previous prediction that a long range dynamical hydration shell plays an important role in insect antifreeze proteins, enhancing their activity at low concentration. AFPs show an increased activity upon the addition of osmolytes (22, 23), the molecular mechanism for this increase is still a matter of discussion. Here, we investigated the effect of the addition of sodium citrate, which increases the thermal activity by a factor of six in DAFP (24). Using THz absorption spectroscopy we found a corresponding increase in the range of retarded hydration shell, further supporting the importance of a long range mechanism that alters solvation water dynamics quite far from the surface of the hyperactive antifreeze protein.

3.2 Results

A. THz Spectroscopy

Figure 3-1 shows the measured difference, $\Delta\alpha$, in THz absorbance of DAFP-1 dissolved in water measured at 20 and 5 °C. For the hyperactive insect antifreeze protein DAFP-1 we observe a THz excess of 15 cm^{-1} , which is the largest $\Delta\alpha$ values observed for antifreeze proteins, investigated thus far. When decreasing the temperature from 20°C to 5°C, we find a shift of c_{max} from $\sim 12 \text{ mg/ml}$ to $\sim 5 \text{ mg/ml}$. This indicates an increase in the size of the dynamical hydration shell for lower temperatures from *ca.* 20 to 27 Å, i.e. from 7 to 9 hydration layers. A similar behavior has been previously reported for AFGP (16).

Previous CD- and IR spectroscopic studies revealed that changes in temperature of DAFP-1 solution have a negligible effect on secondary structure: only a slight increase in the structural order and rigidity of the β -helix was found upon lowering the temperature (25). Similar results were found for tmAFP and sbAFP using NMR techniques (26). Here, conserved side chains like the threonine residues of the ice-binding plane also revealed an increased rigidity when lowering the temperature (27).

B. Molecular Dynamics Simulation

To provide a more complete molecular-level picture, we have further explored the underlying molecular and hydrogen bond dynamics by accompanying MD simulations. In Figure 3-4 we compared the calculated power spectra, i.e. the predicted vibrational density of states, for water molecules in the hydration layer around the protein DAFP-1 and bulk water at 250 and 300 K. We observe clear differences in the spectral density: at 250 K and 300 K, we predict an increase in the frequency range from 2.5 to 3 THz and above, which increases for lower temperatures. This is in good agreement with our experimental results. To further evaluate the specific influence of the different surfaces of the protein on the hydration dynamics, site-specific hydrogen bond correlation functions were calculated. Figure 3-5 shows a remarkable difference between the water dynamics in the vicinity of the binding and the non-binding plane of DAFP-1. Both hydrogen bond correlation functions, $C(t)$, are non-exponential, with $C(t)$ decaying over longer times for water molecules near the binding plane than near the non-binding plane: $C(t)$ reaches a value of 0.2 for hydrogen bonds between water and the atoms of the ice-binding plane at 72 ps, compared with 45 ps for bonds between water and the atoms of the non-binding plane. These values can be compared to the value of 23 ps for hydrogen bonds formed between water molecules in the bulk.

We propose that the pronounced long-range retardation of H-bond dynamics as found experimentally and in the simulation is responsible for the high hysteresis activity of insect antifreeze proteins. The simulation further supports an entropic favorable docking near the binding plane.

Citrate Buffer

It is well known that the antifreeze activity can be enhanced significantly by the addition of sodium citrate, leading e.g. to a six-fold increased thermal hysteresis activity for DAFP (24). THz absorption measurements showed that upon addition of 0.5 M $\text{Na}_3\text{Citrate}$ c_{max} is decreased from 12 mg/ml to 4 mg/ml which corresponds to an increase in the size of the dynamical hydration shell to 27 Å. The addition of citrate is thus found to increase the distance over which the protein can affect hydrogen-bonding dynamics, and to simultaneously increase thermal hysteresis. Accompanying MD simulations predict increased hydrogen bond lifetimes in the presence of sodium citrate especially around the ice-binding face. As is shown in Fig. 3-4, the calculated power spectra show a significant difference between DAFP-1 solvated in water and solvated DAFP-1 in the presence of sodium citrate, with an enhancement of the THz excess when sodium citrate is present, particularly at lower temperature. The presence of sodium citrate influences the hydrogen bond lifetimes and increases the time for hydrogen bond rearrangement both on the binding and non-binding plane, as is displayed in Fig. 3-5. Nevertheless, hydrogen bond lifetimes on the binding plane remain significantly larger than on the non-bonding plane.

Finally, we investigated the effect of adding sodium citrate on other AFPs. Figure 3-3 shows the results of thermal hysteresis and THz measurements for AFGP and AFP type-1 (*wf*AFP-1). Both proteins show an increased antifreeze activity upon addition of sodium citrate. In addition, we find for AFGP a shift of c_{max} towards smaller protein concentration, which is in good agreement with the initial hypothesis that the antifreeze activity of AFGP is mainly accomplished by a change in the hydration dynamics (long

range effect). For AFP-1 we postulated another mechanism, in which the strength of the antifreeze activity depends mainly on the exact positioning of the OH groups (17). In agreement with this previous statement, we find experimentally that the size of the dynamical hydration shell remains constant upon addition of sodium citrate while the thermal hysteresis is increased. Hence a consistent picture emerges where short range and long range mechanisms make different contributions to the overall antifreeze mechanism, depending on the class of antifreeze protein we studied.

Mutation Studies

To gain more insights into the influence of DAFP-1 on the hydrogen bond lifetimes in the dynamical hydration, we address specifically the question of how the hydration dynamics is affected by mutations in the vicinity of the threonine rich ice-binding plane versus the non-binding sites. As described in Materials and Methods Section, threonine 26, 39, 41, 63 were mutated to tyrosine, a mutation (called 4TXY), which has been shown experimentally to result in 90% loss in TmAFP activity (28). The simulated hydrogen bond correlation function, plotted in Fig. 3-5, reveals that the water around the ice-binding plane of mutant 4TXY is much more mobile than in the wild-type, and the hydrogen bond dynamics over the binding plane and non-binding plane are quite similar to each other, in contrast to the wild type. Mutation of the four threonine residues to tyrosine thus appears to have a non-local, collective effect on water dynamics, in addition to any direct ice binding that may occur. We note that we found a similar trend for another antifreeze protein, from spruce budworm *Choristoneura fumiferana*, which also contains a threonine-rich ice-binding plane. In that case four point mutations, which are found experimentally to dramatically lower antifreeze activity, lead to a similar trend

in the hydrogen bond dynamics (29). Site-specific mutation had a clear non-local effect on the entire ice-binding plane while almost not affecting the non-ice binding plane. Based upon these results we conclude that in case of the insect protein the long range influence of the hydrogen bond dynamics plays a major role for the suppression of the freezing and ice nucleation.

3.3 Discussion

Our experimental and computational results give evidence that changes in the fast dynamics at the water-protein interface area play an important role in the molecular mechanism of antifreeze hyperactivity of DAFP-1. Here evolution has realized a concept in which a local binding site providing fixed OH groups for binding with high affinity is most efficiently supported by a long range effect on the retardation in the H-bond dynamics. Insect AFPs such as DAFP-1 have a well-defined surface including ice-binding and non-binding sites. The threonine-rich ice-binding site consists of regularly spaced hydroxyl groups, which can form strong bonds to water. Therefore, one might expect that antifreeze activity is dominated by short range H-bonds at the binding site. Instead we found experimentally and theoretically additional evidence for a long-range influence on the hydration dynamics. MD simulations indicate that the retardation is most pronounced in the vicinity of the ice-binding site. Therefore, we propose a long range retardation of the H-bond dynamics with a gradient towards the ice binding site. A similar gradient in the H-bond dynamics was found for enzymes near the catalytic site (30).

MD simulations on the DAFP with the addition of sodium citrate in water solution indicate a retardation of H-bond dynamics at the ice-binding face upon addition of sodium citrate which goes along with an enhancement of the DAFP-1 antifreeze activity. Based upon the experimental data we estimate an increase in the size of the dynamical hydration shell from *ca.* 20 to 27 Å. Citrate is known to interact with DAFP-1 via the guanidium group of arginine residues (21). However, these interactions are known to exist only along with hydrophobic interactions and/or hydrogen bonds with other

amino acids. The simple ion pairs between guanidinium cations and citrate anions are not very stable in aqueous solution and water molecules will compete for the formation of ion pairs by solvating individual ions.

While for DAFP-1 and AFGP we find a significant increase in the size of the dynamical hydration shell upon addition of sodium citrate, no significant change in the size of the dynamical hydration shell is found for AFP type-1. For all three AFPs the addition of sodium citrate results in an increase of thermal hysteresis by a factor of six (DAFP-1), three (AFP-1) and two (AFGP) respectively. For *wf*AFP-1 we propose that the specific binding of OH groups plays the decisive role. Thus, the strength of the antifreeze activity for AFP-1 will depend on the rigidity of the helix or proper positioning of the threonine. While for AFGP, which is intrinsically less ordered, the long range effects are more important. The most efficient AF effect for hyperactive AFP relies on the optimization of effects on both short range and long range.

In summary, we have now investigated the three major classes of antifreeze proteins, AFGP, AFP type 1 and the insect DAFP-1. Based upon the results of experimental and simulation studies we have postulated two distinct mechanisms for antifreeze activity which can assist or enhance AF activity. Nature is probably more inventive than initially thought, and makes use of short range and long range water perturbation to varying degrees in different classes of antifreeze proteins.

3.4 Materials and Methods

Materials.

DAFP-1, the predominant hemolymph AFP of the beetle *Dendroides canadensis* (including a his-tag) was expressed and purified as previously described (23). Antifreeze glycoproteins were purified from the Antarctic nothoneid toothfish *Dissostichus mawsoni*. AFP type I (HPLC6 from winter flounder *Pseudopleuronectes americanus*) was synthesized by GenScript (Piscataway, USA). Trisodium citrate (trisodium salt) was purchased from Sigma Aldrich (St. Louis, USA). Aqueous solutions were prepared using Baker HPLC analyzed water.

Terahertz measurements.

Terahertz absorption spectroscopy measurements were carried out using our p-Germanium difference laser spectrometer. THz absorption spectroscopy was used as a sensitive tool to characterize the fast collective water network motions in the dynamical hydration shell of proteins (31, 32). We recorded the change of the absorption coefficient in the spectral range from 2.1 - 2.8 THz. We have used our difference spectrometer with a double beam configuration with sample and reference cell to directly measure the change of the intermolecular water network vibrations induced by the AFP sample (16). We analyze the transmitted THz radiation through the AFP sample solution in one channel of the setup while simultaneously measuring the buffer solution as a reference under identical conditions (temperature, humidity). The humidity was controlled to be < 3%, the temperature was fixed at 293 ± 0.05 K. In general, the transmitted intensity $I(n)$, and hence the absorption of a sample, can be described by Beer's law: $I(n) = I_0 \exp(-\alpha(n)d) + c$; with I_0 , $\alpha(n)$, d , and C corresponding to the intensity of the laser source, the

absorption coefficient of the sample, the layer thickness of the sample, and the detector offset, respectively. We have recorded the changes in the THz absorption of the solvated DAFP-1, AFGP and AFP-1 in dependence of the protein concentration and the temperature. Furthermore, we have investigated the effect of adding 0.5 M sodium citrate. Assuming a two component model with protein and bulk water, an increase in protein concentration is expected to result in a decreasing absorption coefficient. This is called “THz defect” (33, 34). However, at low protein concentrations we observe for solvated proteins an initial increase in the absorption coefficient, the so called THz excess, before the THz coefficients decreases again at a characteristic protein concentration c_{\max} (35). In our previous study we could explain this general phenomenon by a shift in the vibration density of states of the hydration water in the vicinity of the protein compared to bulk water (dynamical hydration shell). The increase in the absorption coefficient could be related to a retardation of the H-bond dynamics by the protein compared to bulk water (36). Based upon c_{\max} , the size of the dynamical hydration shell can be estimated.

Thermal hysteresis measurements.

Thermal hysteresis was determined using a Clifton nanoliter osmometer as described previously (17). Resulting data were fitted to the Langmuir equation ($(\Delta T / \Delta T_{\max}) = (c/c_{1/2}) / [(c/c_{1/2}) + 1]$), where $c_{1/2}$ corresponds to the AFP concentration at half maximum.

Molecular Dynamic Simulations

The starting structure of DAFP-1 was created from a homology model of a *Tenebrio molitor* AFP, TmAFP (PDB code 1EZG), by using the 3D-JIGSAW online

structure prediction tool. Constraints were applied to all bonds to hydrogen with the SHAKE algorithm. Periodic boundary conditions were used. All MD simulations were performed in the canonical (NPT) ensemble with the GROMACS software package (37).

In order to investigate the specific role of chemical composition of the site on the water dynamics and hydrogen bond lifetimes, four threonine residues (Thr26, 39, 41, 63) on the binding plane were mutated to tyrosines using Swiss PDB Viewer (38). These mutations have been reported previously to significantly reduce the antifreeze activity of the protein (28, 39).

Both the wild-type and mutated structure were first minimized for 5000 steps with the steepest descent algorithm using the AMBER03 force field (40), after its solvation in a 60 Å cubic water box of TIP5P water model. To prepare 0.5 M sodium citrate solution of DAFP-1, 65 citrate anions and 195 sodium ions were added into the water box as co-solvents with additional sodium ions added to neutralize the system. Then the systems, each of which contained approximately 6700 water molecules, were equilibrated for 400 ps. For the first 100 ps the positions of the proteins were restrained and in the latter 300 ps they were released. Following equilibration, trajectories of 5 ns were obtained at 300 K and 250 K with a Nosé-Hoover thermostat. Non-bonded interactions were gradually brought to zero by a shift function for the electrostatics as well as a switch function for van der Waals interactions between 10 and 12 Å. All simulations were carried out by integrating Newton's equations of motion with the Verlet algorithm and 1 fs time steps. The system coordinates and velocities were stored every 10 fs. The velocity autocorrelation function (VACF) was computed as $C_v(t) = \langle V_i(t)V_i(0) \rangle / \langle V_i(0)V_i(0) \rangle$, where $V_i(t)$ is the velocity vector of the oxygen atom at time, t . The angular brackets

denote averaging over all atoms of the particular type present in the hydration shell and over different reference initial times. Averaging was done over 15 ps time segments of the trajectory for the oxygen atoms that survive in the first hydration shell of thickness 5 Å. (Criteria for hydrogen bonds are specified below.) Power spectra were obtained by Fourier transform of $C_V(t)$. The power spectra correspond to the vibrational density of the water. The vibrational density of protein molecules has been discussed in detail elsewhere (41).

Hydrogen bond time correlation functions, $C_{HB}(t)$, were computed for bonds between water molecules and the protein at 300 K. $C_{HB}(t)$ is defined as the probability that, if a hydrogen bond between donor, D, and acceptor, A, exists at $t = 0$, then it still exists at time, t , even if the bond broke at some intermediate time (42). We adopt a standard criterion for hydrogen bonds, i.e. a DA distance of 3.5 Å and a D-H-A angle greater than 150° (42, 43).

3.5 Acknowledgements

This work was funded by the VW Stiftung. Additional funding was provided by NSF grant MCB-1019958, OPP-0231006, CHE-0910669 and by the Ruhr-University Bochum. We thank E. Bründermann for the initial set-up and his advice with the THz difference spectrometer. C. Bischak is acknowledged for valuable assistance.

3.6 References

1. Duman JG (2001) Antifreeze and Ice Nucleator Proteins in Terrestrial Arthropods. *Annual Review of Physiology* 63(1):327-357.
2. DeVries A & Wohlschlag D (1969) Freezing resistance in some Antarctic fishes. *Science* 163(3871):1073.
3. Duman JG (1977) The role of macromolecular antifreeze in the darkling beetle, *Meracantha contrata*. *Journal of Comparative Physiology B: Biochemical, Systemic, and Environmental Physiology* 115(2):279-286.
4. Duman JG & Olsen TM (1993) Thermal Hysteresis Protein Activity in Bacteria, Fungi, and Phylogenetically Diverse Plants. *Cryobiology* 30(3):322-328.
5. Urrutia ME, Duman JG, & Knight CA (1992) Plant thermal hysteresis proteins. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1121(1-2):199-206.
6. DeVries A (1971) Glycoproteins as biological antifreeze agents in antarctic fishes. *Science* 172(988):1152.
7. Yeh Y & Feeney RE (1996) Antifreeze Proteins: Structures and Mechanisms of Function. *Chemical Reviews* 96(2):601-618.
8. Raymond JA & DeVries AL (1977) Adsorption inhibition as a mechanism of freezing resistance in polar fishes. *Proceedings of the National Academy of Sciences* 74(6):2589-2593.
9. Sharp KA (2011) A peek at ice binding by antifreeze proteins. *Proceedings of the National Academy of Sciences* 108(18):7281-7282.
10. Haymet ADJ, Ward LG, Harding MM, & Knight CA (1998) Valine substituted winter flounder `antifreeze': preservation of ice growth hysteresis. *FEBS Letters* 430(3):301-306.
11. DeLuca CI, Davies PL, Ye Q, & Jia Z (1998) The effects of steric mutations on the structure of type III antifreeze protein and its interaction with ice. *Journal of Molecular Biology* 275(3):515-525.
12. Modig K, Qvist J, Marshall CB, Davies PL, & Halle B (2010) High water mobility on the ice-binding surface of a hyperactive antifreeze protein. *Physical Chemistry Chemical Physics* 12(35):10189-10197.
13. Nutt DR & Smith JC (2008) Dual Function of the Hydration Layer around an Antifreeze Protein Revealed by Atomistic Molecular Dynamics Simulations. *Journal of the American Chemical Society* 130(39):13066-13073.
14. Garnham CP, Campbell RL, & Davies PL (2011) Anchored clathrate waters bind antifreeze proteins to ice. *Proceedings of the National Academy of Sciences* 108(18):7363-7367.
15. Siemer AB, Huang K-Y, & McDermott AE (2010) Protein-ice interaction of an antifreeze protein observed with solid-state NMR. *Proceedings of the National Academy of Sciences* 107(41):17580-17585.
16. Ebbinghaus S, *et al.* (2010) Antifreeze Glycoprotein Activity Correlates with Long-Range Protein-Water Dynamics. *Journal of the American Chemical Society* 132(35):12210-12211.

17. Ebbinghaus S, *et al.* (2012) Functional Importance of Short-Range Binding and Long-Range Solvent Interactions in Helical Antifreeze Peptides. *Biophysical journal* 103(2):L20-L22.
18. Marshall CB, *et al.* (2003) Partitioning of Fish and Insect Antifreeze Proteins into Ice Suggests They Bind with Comparable Affinity†. *Biochemistry* 43(1):148-154.
19. Middleton AJ, *et al.* (2012) Antifreeze Protein from Freeze-Tolerant Grass Has a Beta-Roll Fold with an Irregularly Structured Ice-Binding Site. *Journal of Molecular Biology* 416(5):713-724.
20. Liou Y-C, Tocilj A, Davies PL, & Jia Z (2000) Mimicry of ice structure by surface hydroxyls and water of a [beta]-helix antifreeze protein. *Nature* 406(6793):322-324.
21. Wang S, *et al.* (2009) Arginine, a Key Residue for the Enhancing Ability of an Antifreeze Protein of the Beetle *Dendroides canadensis*. *Biochemistry* 48(40):9696-9703.
22. Amornwittawat N, *et al.* (2009) Effects of polyhydroxy compounds on beetle antifreeze protein activity. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* 1794(2):341-346.
23. Amornwittawat N, Wang S, Duman JG, & Wen X (2008) Polycarboxylates enhance beetle antifreeze protein activity. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* 1784(12):1942-1948.
24. Li N, Andorfer C, & Duman J (1998) Enhancement of insect antifreeze protein activity by solutes of low molecular mass. *Journal of Experimental Biology* 201(15):2243.
25. Li N, Kendrick BS, Manning MC, Carpenter JF, & Duman JG (1998) Secondary Structure of Antifreeze Proteins from Overwintering Larvae of the Beetle *Dendroides canadensis*. *Archives of Biochemistry and Biophysics* 360(1):25-32.
26. Daley ME, Spyropoulos L, Jia Z, Davies PL, & Sykes BD (2002) Structure and Dynamics of a β -Helical Antifreeze Protein†,‡. *Biochemistry* 41(17):5515-5525.
27. Daley ME & Sykes BD (2004) Characterization of threonine side chain dynamics in an antifreeze protein using natural abundance ^{13}C NMR spectroscopy. *Journal of Biomolecular NMR* 29(2):139-150.
28. Marshall CB, Daley ME, Graham LA, Sykes BD, & Davies PL (2002) Identification of the ice-binding face of antifreeze protein from *Tenebrio molitor*. *FEBS Letters* 529(2-3):261-267.
29. Xu Y, Gnanasekaran R, & Leitner DM (2012) Analysis of Water and Hydrogen Bond Dynamics at the Surface of an Antifreeze Protein. *Journal of Atomic, Molecular, and Optical Physics* 2012:6.
30. Grossman M, *et al.* (2011) Correlated structural kinetics and retarded solvent dynamics at the metalloprotease active site. *Nat Struct Mol Biol* 18(10):1102-1108.
31. Bergner A, *et al.* (2005) New p-Ge THz laser spectrometer for the study of solutions: THz absorption spectroscopy of water. *Review of Scientific Instruments* 76(6):063110-063115.

32. Ebbinghaus S, *et al.* (2007) An extended dynamical hydration shell around proteins. *Proceedings of the National Academy of Sciences* 104(52):20749-20752.
33. Heyden M, Ebbinghaus S, & Havenith M (2010) Terahertz Spectroscopy as a Tool to Study Hydration Dynamics. *Encyclopedia of Analytical Chemistry*, (John Wiley & Sons, Ltd).
34. Born B, Kim SJ, Ebbinghaus S, Gruebele M, & Havenith M (2009) The terahertz dance of water with the proteins: the effect of protein flexibility on the dynamical hydration shell of ubiquitin. *Faraday Discussions* 141:161-173.
35. Heyden M, *et al.* (2010) Dissecting the THz spectrum of liquid water from first principles via correlations in time and space. *Proceedings of the National Academy of Sciences*.
36. Heyden M & Havenith M (2010) Combining THz spectroscopy and MD simulations to study protein-hydration coupling. *Methods* 52(1):74-83.
37. Berendsen HJC, van der Spoel D, & van Drunen R (1995) GROMACS: A message-passing parallel molecular dynamics implementation. *Computer Physics Communications* 91(1-3):43-56.
38. Guex N & Peitsch MC (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *ELECTROPHORESIS* 18(15):2714-2723.
39. Graether SP & Sykes BD (2004) Cold survival in freeze-intolerant insects. *European Journal of Biochemistry* 271(16):3285-3296.
40. Duan Y, *et al.* (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry* 24(16):1999-2012.
41. Leitner DM, Havenith M, & Gruebele M (2006) Biomolecule large-amplitude motion and solvation dynamics: modelling and probes from THz to X-rays. *International Reviews in Physical Chemistry* 25(4):553-582.
42. Bagchi B (2005) Water Dynamics in the Hydration Layer around Proteins and Micelles. *Chemical Reviews* 105(9):3197-3219.
43. Douglas T, Neelanjana S, & Mounir T (2009) Molecular Dynamics Simulation Studies of Coupled Protein and Water Dynamics. *Proteins, Computation in Chemistry*, (CRC Press), pp 361-386.

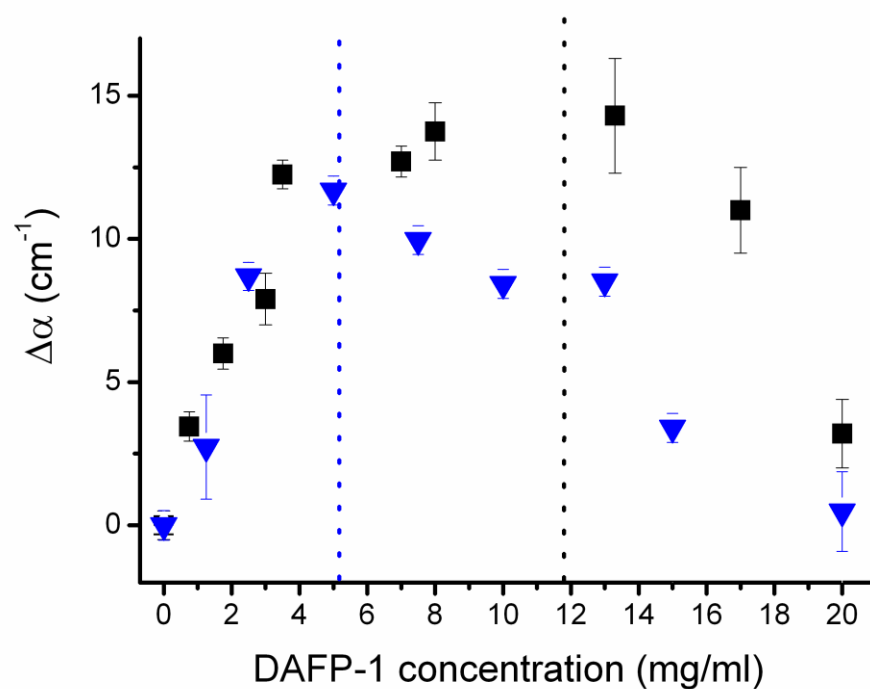


Figure 3-1 Difference $\Delta\alpha$ in THz absorbance (integrated between 2.4 and 2.7 THz) of dissolved DAFP-1 in water relative to a water reference, plotted against the concentration. Measurements were carried out at (20 ± 0.5) and (5 ± 0.5) °C blue data points respectively. DAFP-1 shows a concentration dependent excess of THz absorbance that shifts to lower protein concentration at the lower temperature (blue data points).

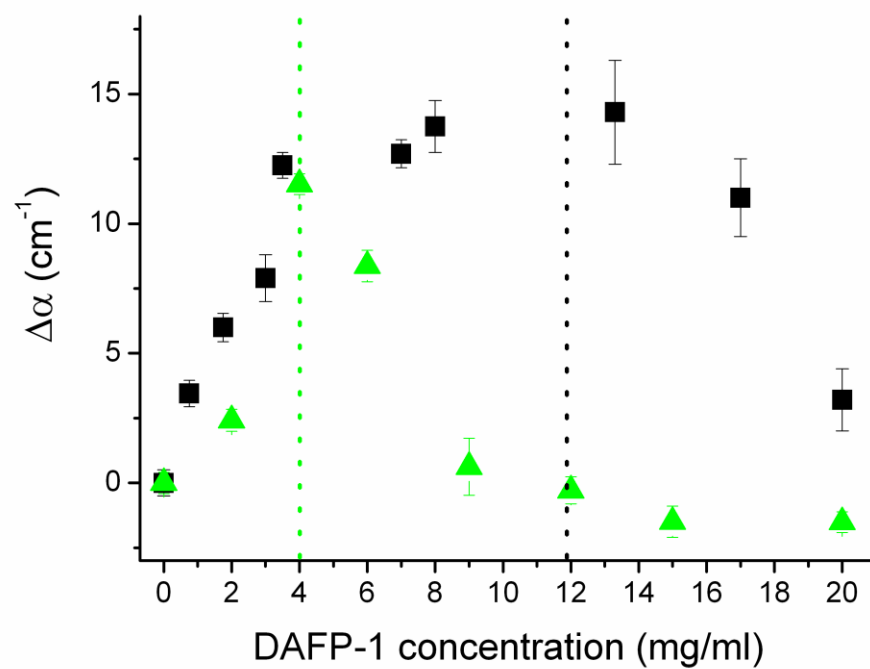


Figure 3-2 Difference $\Delta\alpha$ in THz absorbance (integrated between 2.4 and 2.7 THz) of dissolved DAFP-1 in water relative to a water reference (black data point) and dissolved DAFP-1 in 0.5 M $\text{Na}_3\text{Citrate}$ relative to a citrate buffer reference (green data points).

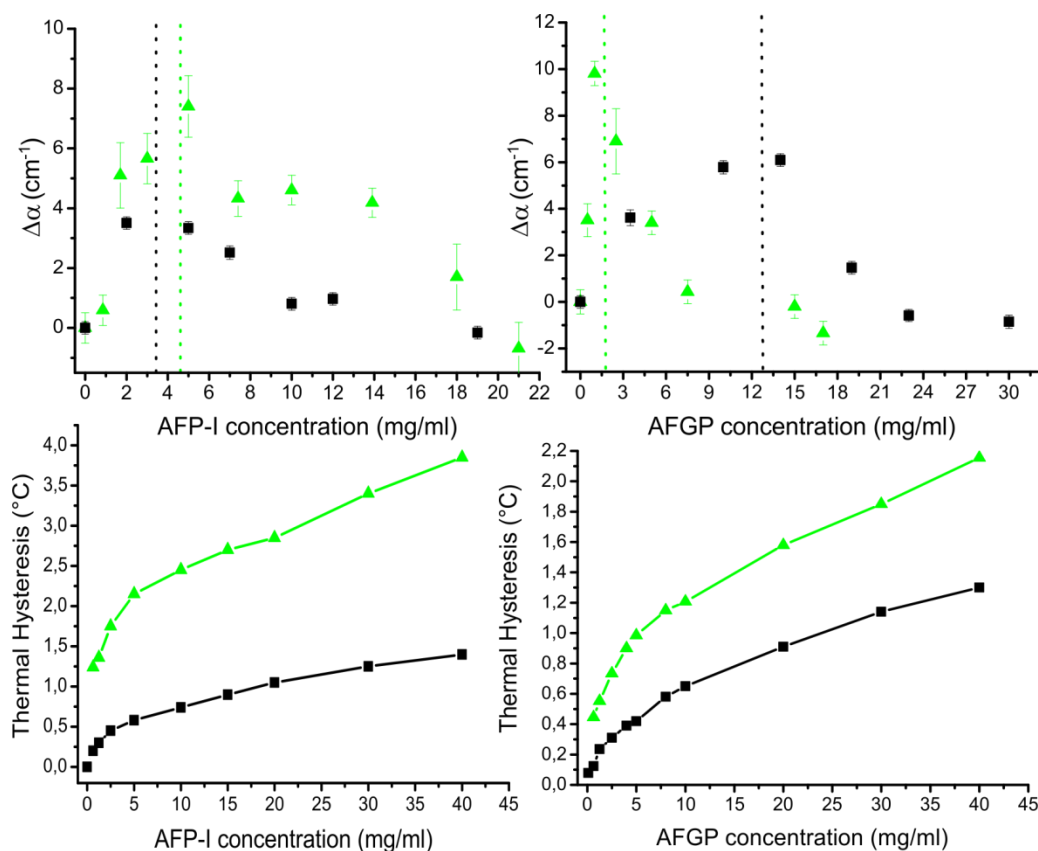


Figure 3-3 Upper Panel. Difference $\Delta\alpha$ in THz absorbance (integrated between 2.4 and 2.7 THz) of dissolved AFP-1 (left) and AFGP (right) relative to a water reference (black) and enhanced AFP-1 and AFGP in 0.5 M $\text{Na}_3\text{Citrate}$ relative to a citrate buffer reference (green). **Lower Panel.** (left) Thermal Hysteresis of AFP-1 (black) and AFP-1 in 0.5 M citrate buffer as a function of protein concentration in water. (right) Thermal Hysteresis of AFGP (black) and AFGP in citrate buffer.

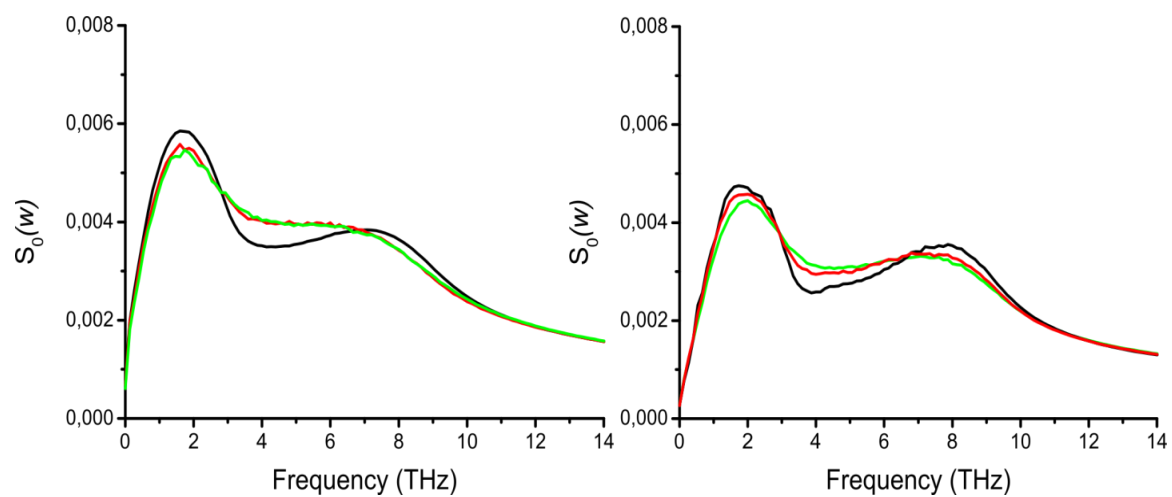


Figure 3-4 Computed power spectra for water molecules in the hydration layer around the protein DAFP-1 and for water molecules in the bulk. Spectra were calculated at 300 K (left) and 250 K (right).

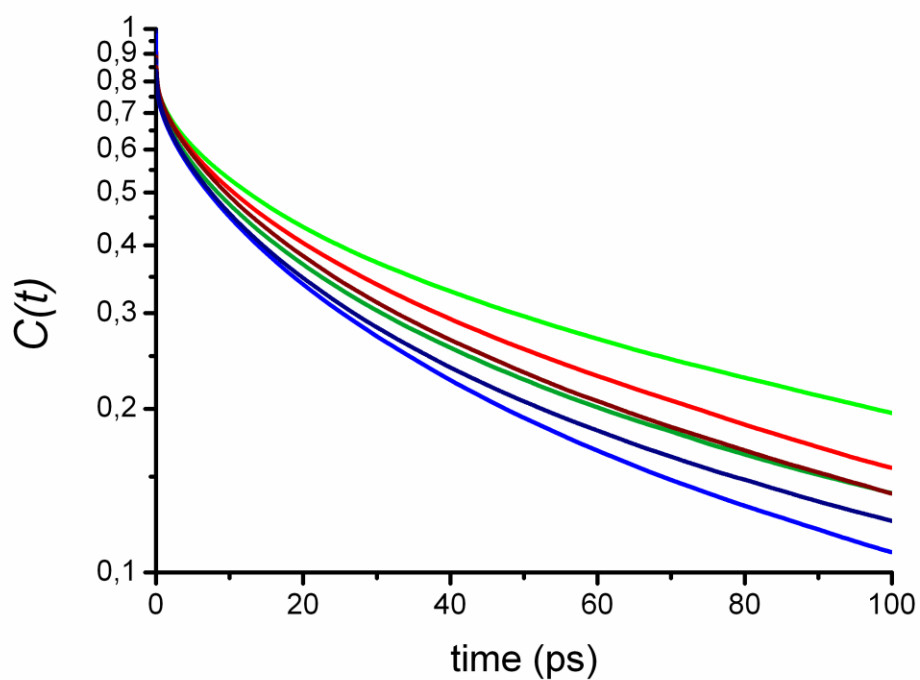


Figure 3-5 Hydrogen bond lifetime correlation function $C(t)$ for bonds between water molecules around the binding (red) and non-binding plane (blue) of DAFP-1 at 300 K. Dark red (binding) and dark blue (non-binding plane) indicate the mutant discussed in text. $C(t)$ is also plotted for hydrogen bonds between water molecules at the binding plane (green) and non-binding plane (dark green) of the wild type in sodium citrate solution.

4 Dynamics of Water Clusters Confined in Proteins: A Molecular Dynamics Simulation Study of Interfacial Waters in a Dimeric Hemoglobin

Abstract

Water confined in proteins exhibits dynamics distinct from dynamics of water in the bulk or near the surface of a biomolecule. We examine the water dynamics at the interface of the two globules of the homodimeric hemoglobin from *Scapharca inaequivalvis* (HbI) by molecular dynamics (MD) simulations, with focus on water-protein hydrogen bond lifetimes and rotational anisotropy of the interfacial waters. We find that relaxation of the waters at the interface of both deoxy- and oxy-HbI, which contain a cluster of 17 and 11 interfacial waters, respectively, is well described by stretched exponentials with exponents from 0.1 to 0.6 and relaxation times of 10s to 1000s of picoseconds. The interfacial water molecules of oxy-HbI exhibit slower rotational relaxation and hydrogen bond rearrangement than those of deoxy-HbI, consistent with an allosteric transition from unliganded to liganded conformers involving the expulsion of several water molecules from the interface. Though the interfacial waters are translationally and rotationally static on the picosecond time scale, they contribute to fast communication between the globules via vibrations.

Keywords: Hemoglobin, Interfacial Water Dynamics, Rotational Anisotropy.

Reprinted with permission from *J. Phys. Chem. B*, 2010, 114 (50), pp 16989–16996.

Copyright 2013 American Chemical Society

4.1 Introduction

Water molecules in living cells exhibit dynamics over many time scales depending on environment.¹ Properties of hydration water surrounding biomolecules have been extensively studied both experimentally and computationally in recent years,² motivated by the key role water plays in biomolecule structure, dynamics and function. These studies have detailed the distinction between properties of water in the hydration layer around biomolecules and bulk water.³⁻²⁵ Neutron scattering,²⁶ terahertz (THz) spectroscopy^{11,27,28} and MD simulations^{5,9,29} provide information about the typically sluggish dynamics of water at the surface of proteins and dynamical coupling between proteins and water.^{11,30,31} The THz absorption coefficient of biological water surrounding saccharides^{32,33}, salts³⁴ and proteins³⁵⁻³⁸ is different from the absorption coefficient of bulk water, allowing measurement of the size of the hydration layer. Water dynamics and spectroscopy depend on a variety of factors, including hydrophobicity of the surface in contact with the water,³⁹⁻⁴¹ chemical heterogeneity,⁴² surface topography,⁴³⁻⁴⁶ and partial confinement. For example, MD simulations on dynamics of water between two domains of the BphC enzyme, with inter-domain distances of one to several water layers, reveal how partially confined water molecules between the protein domains translate and rotate significantly slower than water around other regions of the protein. In this chapter, we explore distinctions between dynamics of clusters of water molecules confined inside proteins and hydration water around them, how this difference is connected to function, and the contribution of confined water to energy flow in the protein. We address the dynamics of 10 – 20 tightly bound water molecules confined to the interface between the

two subunits of the homodimeric hemoglobin (HbI) from the clam *Scapharca inaequalvis*, which mediate cooperative binding of ligands to the two hemes.^{47,48}

Water in nanoconfined environments has generally been observed to relax far more slowly than water in the bulk. Ricci, *et al.*,⁴⁹ carried out neutron diffraction and deep inelastic neutron scattering experiments on water embedded in silica substrates and found similarities between the confined water and super-cooled water in the bulk. Recent MD simulations⁵⁰ on water in reverse micelles reveal very slow and curvature-dependent rotational relaxation of water molecules in the reverse micelle, with stretched exponential relaxation and exponents well below 0.5, indicative of glassy dynamics and a wide range of hydrogen bonding environments. While slow rotational relaxation over 10s of picoseconds has been observed in MD simulations for other polar molecules in confined environments,⁵¹ the relaxation of nanoconfined water in reverse micelles occurs on the nanosecond time scale or longer. We observe a similar size-dependent trend for clusters of water molecules confined in proteins.

We focus in this chapter on hydrogen bond and rotational dynamics of the interfacial water molecules of HbI. Crystallographic structures of HbI (Figure 4-1) reveal that the unliganded, deoxy form contains 17 interfacial water molecules, while the liganded, oxy form has 11. In addition to the release of 6 water molecules in going from the deoxy to oxy structure, there are notable tertiary and modest quaternary structural changes, the latter important in the late stages of the allosteric transition.^{52,53} Following ligand binding and displacement of the Fe atom the two heme groups are found further apart.^{54,55} Phe97 changes its structural conformation during the transition from the deoxy to oxy state by shifting the phenyl ring to the interfacial hydrated region, so that the

volume of the interface is smaller for the oxy form than deoxy-HbI and thus holds less water.⁵⁶ Side chains of Asn100 and Lys96 and the propionate group of the hemes form specific hydrogen bonding arrangements with interfacial waters, arrangements that differ for the liganded and unliganded states. The pivotal role of interfacial water molecules in the allosteric mechanism has been examined by mutation studies.^{55,57} Mutations that affect protein-water interactions at the edge of the water cluster in the interface also change the Hill coefficient,^{55,58,59} proving that ordered water molecules act as key allosteric mediators in HbI.⁶⁰ Molecular dynamics (MD) simulations by Karplus and co-workers⁶¹ elaborated on the role of interfacial water molecules and found that there existed a well-defined water channel that connects the interface between the subunits to bulk water. Although details of the function of the confined waters in HbI cooperativity have not yet been completely worked out, it is clear that the interfacial water is required for efficient communication between the globules.⁵⁵

The MD simulations reported here indicate that rotational relaxation of water molecules at the interface occurs on the 10 to 1000 ps time scale. Since the interfacial water molecules are embedded inside the protein they can contribute to fast communication between the globules via vibrational energy transport, which occurs on the 1 – 10 ps time scale,^{23,60,62-67} even if the water is translationally and rotationally static on that short time scale.

The Chapter is organized as follows: In Section 2, we detail the computational methods. In Section 3A, we present and discuss the results of hydrogen bond time correlation functions and in Section 3B we provide results and discuss rotational reorientation dynamics. Concluding remarks are provided in Section 4.

4.2 Computational Details

The initial homodimeric hemoglobin structures (oxy 1HbI and deoxy 4SDH) were obtained from the Protein Data Bank (PDB). The first missing residue (proline) was added by using the software package Swiss PDB viewer.⁶⁸ We kept the 218 and 219 water molecules included in the crystallographic PDB data for the oxy and deoxy structures, respectively. A cubic box 70 Å on each side was filled with TIP3P water. Following the procedure of Zhou, *et al.*,⁶¹ we kept the 17 and 11 interfacial water molecules of 4SDH and 1HbI, respectively, which are within 10.4 Å of the center of geometry of the deoxy and oxy interface, respectively. The system, now with 12,310 water molecules, was energy-minimized by using the steepest descent algorithm to remove bad contacts. The added solvent water molecules in the simulation box were equilibrated at 300 K for 100 ps by fixing the positions of the protein atoms. The net charge of the system (+4) was neutralized in some of our simulations by addition of counter ions, but we found that the addition of counter ions to the simulation box did not affect the dynamics of either hydration water or interfacial water on the time scale of our simulations, and the results we report here were obtained without the counter ions. The particle-mesh Ewald (PME) method^{69,70} was employed to calculate all long-range non-bonded electrostatic interactions by using a 0.12 nm grid and a fourth-order interpolation. The SHAKE algorithm⁷¹ was employed to constrain all bonds to hydrogen and periodic boundary conditions were applied.

All the classical MD simulations and analyses were performed with the GROMACS software package. An NPT ensemble was applied by maintaining the temperature at 300 K with the velocity-rescaling thermostat⁷² and by maintaining the

pressure at 1 atm with the Parrinello-Rahman barostat.⁷³ The simulations were performed for 1 ns by integrating the equations of motion with a leapfrog algorithm with integration time steps of 1 fs and the system coordinates were stored every 10 fs. The simulation system consisted of 41,616 (41,612) atoms for oxy (deoxy) hemoglobin. We carried out 5 simulations of 1 ns each for analysis of the hydrogen bond time correlation function and rotational anisotropy. At some point during each of the five 1 ns simulations that we carried out for each system, one water molecule diffused out of the interface and into the hydration layer or the bulk. Once this water molecule escaped it was no longer considered for analysis of the interfacial water dynamics.

To examine how dynamics of water molecules is affected by the presence of hydration sites at the HbI interface we have studied the hydrogen bond lifetime correlation function $C_{\text{HB}}(t)$, defined as⁴

$$C_{\text{HB}}(t) = \frac{\langle h(t)h(0) \rangle}{\langle h \rangle}, \quad (4-1)$$

where h is unity if a hydrogen bond that exists at time 0 still exists at time t and is otherwise 0. We employed standard geometrical criteria for hydrogen bonds, i.e., the distance between donor and acceptor atoms ≤ 0.35 nm and the bond angle between the hydrogen and the acceptor atom is taken as $\leq 30^\circ$. Reorientation of the surface and interfacial water molecules was studied by considering the rotational anisotropy decay, $C_n(t)$, defined as

$$C_n(t) = \langle P_n(\hat{\mu}(t) \cdot \hat{\mu}(0)) \rangle, \quad (4-2)$$

where $\hat{\mu}$ is the unit vector pointing along the dipole moment of the water molecule, and P_n is the n^{th} degree Legendre polynomial. We specifically considered $n = 1$ and 2, which

have been studied in other simulations of hydration water dynamics,⁷⁴ as discussed below. For analysis of hydration water molecules, we took all water molecules within the 3 Å layer from the surface of the protein, excluding the interfacial waters.

4.3 Results and Discussion

A. Hydrogen bond time correlation functions

The hydrogen bond correlation functions computed for the hydration and interfacial water molecules are plotted in Fig. 4-2. Relaxation of hydrogen bonds is non-exponential, and we fit the results instead from 0.2 ps to 200 ps instead to a stretched exponential, $\exp[-(t/\tau)^\beta]$, which, as we see below, provides an excellent description of $C_{\text{HB}}(t)$ for the results for hydrogen bonds between interfacial waters and the protein and for hydrogen bonds between hydration waters and the protein. The values of τ and β used to fit the results are listed in Table 4-1. Values of β for the hydration water are essentially the same for the oxy and deoxy structure, 0.51 and 0.52, respectively. The time constants are also similar, both near 40 ps. Unsurprisingly, hydrogen bond lifetimes for water-protein hydrogen bonds at the surface of the proteins do not change in going from the unliganded to the liganded structures, since this transition mainly involves tertiary changes at the interface.

The hydrogen bonding dynamics of the interfacial waters, however, depends noticeably on ligation. We observe that the hydrogen bonds between interfacial waters and protein, which generally rearrange far slower than the protein-water hydrogen bonds at the surface of the protein, rearrange more slowly for the oxy structure, where we find β to be 0.37, compared with 0.63 for the deoxy structure. While τ is nearly the same for both, about 14 ps, it is the much smaller value of β for the interfacial waters of the oxy structure that corresponds to the slow hydrogen bond rearrangement times observed in Fig. 4-2, reflecting a large distribution of hydrogen bonding environments. The faster

rearrangement of the hydrogen bonds between interfacial waters and protein found for the deoxy structure, particularly at longer times, is consistent with the more facile removal of water molecules at the interface, which occurs when ligands bind to the hemes and 6 water molecules are expelled.

Hydrogen bonds between water molecules in the bulk rearrange on a picosecond time scale, as seen in Fig. 4-2, much faster than hydrogen bond rearrangements between water and protein at the interface or in the hydration layer. Hydrogen bond rearrangements in bulk water are non-exponential, but we could not fit the data well to a stretched exponential. The non-exponential decay for bulk water has been observed in a number of previous simulations and has been related to vibrational energy transfer⁷⁸⁻⁸³. Luzar and Chandler^{84,85} identified the concerted formation and breaking of hydrogen bonds between water molecules as a possible contributor to non-exponential decay observed in hydrogen bond dynamics. In general, the strength of hydrogen bonds vary in different environments accounting for differences in kinetic behavior.⁸⁶

A number of earlier computational studies have been carried out to describe water dynamics near the surface of a protein. Concerning hydrogen bond rearrangement and protein-water hydrogen bond lifetimes, Tarek and Tobias^{5,87,88} reported MD simulations on hydration water dynamics around ribonuclease A to provide insights into neutron-scattering experiments conducted on hydrated proteins. As part of their analysis they calculated $C_{\text{HB}}(t)$ for hydrogen bonds between the protein and water and found it to be non-exponential, relaxing more slowly than hydrogen bond rearrangements between water molecules.^{5,88} At about 300 K the former hydrogen bond lifetimes are on the order of 10 ps, but slow dramatically to the nanosecond time scale at temperatures approaching

the protein dynamical transition. In related work, water mobility in layers around the protein were examined by Cannistraro and coworkers,^{7,89} who carried out MD simulations on solvated plastocyanin and analyzed survival-time-correlation functions for water layers around the protein, finding β from 0.60 to 0.70. Dynamics of hydrogen bonds between water molecules as a function of distance from the surface of a protein in its native conformer and denatured states have been studied by MD simulations.^{37,90} Slower hydrogen bond rearrangements near the hydrophobic groups have been observed and attributed⁹⁰ to sizable steric hindrance to water mobility near such groups in the systems studied, consistent with the influence of surface curvature on local water dynamics.⁹ For hydrogen bonds between water molecules moving between two protein domains similar trends in the lifetime with distance from the surface of the protein are found, but the lifetimes between the hydrogen bonds are longer than for similar distances around only one domain.⁹

Other relevant computational studies include recent MD simulations of Johnson, *et al.*,⁸ who adopted both fixed-charge and polarizable potentials to calculate the translational diffusion coefficient of hydrated water around the amphiphilic N-acetyl-leucine-methyl-amide and hydrophilic N-acetyl-glycine-methyl-amide with different concentrations, and to reproduce the correct temperature trends in solution structure and the quasi-elastic neutron scattering dynamical data. To describe the two different translational motions that they observed in their simulations they fit their data using a combination of a stretched exponential and an exponential function; β was found to range from 0.64 - 0.87 at temperatures of 248, 288 and 298 K. Also, Pal, *et al.*,⁹¹ reported a theoretical study on the structure and dynamics of water layers present on the surface of

cationic micelle decyltrimethylammonium bromide (DeTAB) by using atomistic MD simulations. They fitted the hydrogen bond lifetime correlation function by a multi-exponential function. For the interfacial water molecules existing in the DeTAB micellar solution they found the mean relaxation time to be 11.7 ps and for bulk water 6.5 ps.

Many of these computational studies have been compared with results of experimental neutron scattering data. Neutron scattering experiments by Yoshida, *et al.*,⁹² and Dellerue, *et al.*,⁹³ examining the surface water around β -lactoglobulin and around C-phycocyanin revealed stretched exponential relaxation with $\beta \approx 0.5$. Jansson, *et al.*,⁹⁴ investigated by quasi-elastic neutron scattering the protein solvent dynamics and found for water that β ranges from 0.38-0.45. This turns out to be comparable to results for water confined in sizable non-biological matrices. For example, another quasi-elastic neutron scattering experiment, carried out by Faraone, *et al.*,⁹⁵ focused on dynamics of the confined water in MCM-41-S, a nanoporous silica matrix. They found the intermediate scattering function is well described by a stretched exponential with $\beta \approx 0.5$.

B. Orientation dynamics of surface and interfacial water molecules

Rotational anisotropy decay of order $n = 1$ and 2 , defined by Eq. (4-2), was computed for the interfacial waters, for all other water molecules within 3 \AA of the protein surface, and for bulk water molecules, and the results for C_1 and C_2 are plotted in Figure 4-3 and 4-4, respectively. The rotational anisotropy decay for the interfacial water molecules fit well to a stretched exponential, $\exp[-(t/\tau)^\beta]$, and the corresponding values of τ and β for fits from 0.2 to 200 ps are listed in Table 4-2. The anisotropy decay for the hydration water molecules, while non-exponential, fits poorly to a stretched exponential and we note only that the data fit well to the sum of an exponential and stretched exponential, as has been reported in other MD simulation studies of hydration water.⁹⁶ A time constant for the hydration water rotational anisotropy decay can be found from the plot if desired. The rotations of the hydration water molecules relax more rapidly than do the rotations of the interfacial waters, and rotation of all waters near the protein relax more slowly than do the rotations of water molecules in the bulk.

For both C_1 and C_2 the rotational decay is slower for the interfacial water molecule of oxy-HbI than for deoxy-HbI, the same trend we observed for the hydrogen bond lifetimes discussed above. Considering first C_1 , the time constant for the interfacial waters of oxy-HbI is about 2 ns , compared with 141 ps for the interfacial waters of deoxy-HbI. For C_2 , these values are 71 and 16 , respectively. Values of β are 0.2 and 0.6 for the interfacial waters of oxy- and deoxy-HbI, respectively for C_1 , and they are 0.1 and 0.4 , respectively, for C_2 . While both C_1 and C_2 have been calculated in previous MD simulations for water in a number of environments, the latter is of additional interest in that it can, with some approximation, be compared with results of polarization-resolved

pump-probe experiments.⁹⁷ The value for β is especially small for the interfacial waters of oxy-HbI and the time constants particularly large, reflecting the heterogeneity of the hydrogen bonding environments of the tightly bound water molecules at the interface, particularly for the liganded protein, which has a smaller number of interfacial waters due to its smaller interface volume.

Slow rotational relaxation of water in contact with proteins has been found in many earlier MD simulations. This includes work by Marchi, *et al.*,⁷⁴ on water around lysozyme. They reported the rotational relaxation of water in the vicinity of lysozyme is 3-7 times slower than that in the bulk, depending on how the hydration shell is defined in the calculation. They presented comparisons of the second rank dipolar correlation function, $C_2(t)$, for simulations of lysozyme in TIP3P water and SPC/E water. The relaxation time of bulk water for TIP3P is 0.76 ps and for SPC/E is 1.85 ps, while for hydrated water it varies from 2.19 ps to 4.90 ps using the TIP3P water model and from 5.68 ps to 13.0 ps for the SPC/E water model, depending on the cutoff parameter used with the different water levels. MD simulations by Oleinikova, *et al.*,⁹⁶ examined the dynamics of hydration water at the surface of lysozyme at various hydration levels. They reported fits to stretched exponentials for the rotational correlation function at low and high hydration levels. At low hydration level, corresponding to only 200 water molecules at the surface of lysozyme, they found that a stretched exponential equation of the form $\exp[-(t/\tau)^\beta]$ with $\beta \approx 0.325$ and relaxation time ≈ 48 ps fit the data for C_l well, whereas for higher hydration level C_l was described by the sum of a stretched exponential and simple exponential. We note also that Cannistraro and coworkers⁸⁹ employed simple stretched exponentials to fit the decay of the first and second rank rotational reorientation

autocorrelation function to 50 ps for water molecules at various distances from the surface of plastocyanin, specifically water molecules in the range 0 – 4 Å, 0 – 6 Å and 0 – 14 Å. The time constants become smaller as a larger hydration layer is chosen. For the smallest hydration layer, they found time constants of about 10 and 5 ps, respectively, for C_1 and C_2 . Tobias, *et al.*,⁵ examined the second rank reorientation autocorrelation function for water around H α LA, both in the native structure and molten globule states. Water around either structure relaxed more slowly than water molecules in the bulk, but the rate depends on structure, relaxing slightly faster for the molten globule than native structure.

Particularly relevant to the calculation of rotational reorientation of water confined in proteins are the MD simulations of Ladanyi, Skinner and coworkers⁵⁰ on water confined in reverse micelles. They investigated the vibrational spectroscopy and dynamics of water confined inside reverse micelles by calculating correlation functions of spectral densities, hydrogen bonding and anisotropy decay. They fit the latter to stretched exponentials over times from 0.2 ps to 1 ns, and observed very slow relaxation times and β from 0.17 to 0.37, comparable to values we find for the interfacial waters of HbI, revealing, as in the protein, a highly heterogeneous hydrogen bonding environment and cooperative, glass-like rearrangements of the confined waters molecules. Rotational reorientation was particularly sluggish for small micelles, commensurate with the slower relaxation we find when comparing rotational relaxation for the water molecules at the interface of oxy-HbI with those at the interface of deoxy-HbI. The volume of the interface is smaller for the oxy form because of the position of the Phe97 side chains

(Fig. 4-1), leaving space for only 11 water molecules, as opposed to the 17 interfacial water molecules of deoxy-HbI.

4.4 Concluding remarks

MD simulations were performed for solvated oxy and deoxy HbI systems and the trajectories analyzed to study hydrogen bond and rotational dynamics of water in a cluster that is confined to the interface between the two globules and the water in the hydration layer surrounding the protein. The dynamics of hydration water is retarded by the surface of the protein, as expected, and neither the hydrogen bonds between water and protein nor the rotation of the waters near the protein were affected by protein-ligand binding, which gives rise mainly to rearrangements of side chain positions at the interface. However, ligand binding strongly affects the much slower dynamics of the 11 – 17 water molecules clustered at the interface between the two globules of the protein. The interfacial water molecules of the liganded structure of HbI exhibit slower hydrogen bond rearrangement and rotational reorientation than do those of the deoxygenated protein.

The slower relaxation of the interfacial waters found for oxy-HbI compared to the interfacial waters of deoxy-HbI correlates with the smaller volume available at the interface of oxy-HbI, where only 11 water molecules are present, compared to 17 at the interface of deoxy-HbI. That the water molecules clustered at the interface of the oxygenated structure are more tightly bound than those of the unliganded structure is consistent with the allosteric transition between the two forms, since in the course of the cooperative transition from the deoxy to oxy structures 6 water molecules are expelled from the interface, and these molecules are presumably less tightly bound than the remaining 11. The observation of more tightly bound water in the smaller interface volume of the liganded protein is consistent with results of a recent computational study

of dynamics of water confined in reverse micelles,⁵⁰ where rotational relaxation and hydrogen bond rearrangements were found to be correlated with a smaller confining volume.

4.5 Acknowledgements

Support from NSF grant CHE-0910669 is gratefully acknowledged.

Table 4-1 Listed are values of β and τ for the stretched exponential, $\exp[-(t/\tau)^\beta]$, fit to $C_{HB}(t)$ for hydrogen bonds between hydration water molecules around oxy and deoxy HbI and the protein, and for hydrogen bonds between interfacial water molecules and the protein.

	Hydration (β)	Hydration (τ /ps)	Interfacial (β)	Interfacial (τ /ps)
Oxy-HbI	0.52	14.2	0.37	39.0
Deoxy-HbI	0.52	14.4	0.63	42.0

Table 4-2 Listed are values of β and τ for the stretched exponential, $\exp[-(t/\tau)^\beta]$, fit to the rotational reorientation correlation function of order $n = 1$ and 2 for interfacial water molecules of oxy and deoxy HbI.

		β	τ/ps
Oxy-HbI	C_1	0.21	1970.6
	C_2	0.10	70.8
Deoxy-HbI	C_1	0.56	140.8
	C_2	0.44	15.6

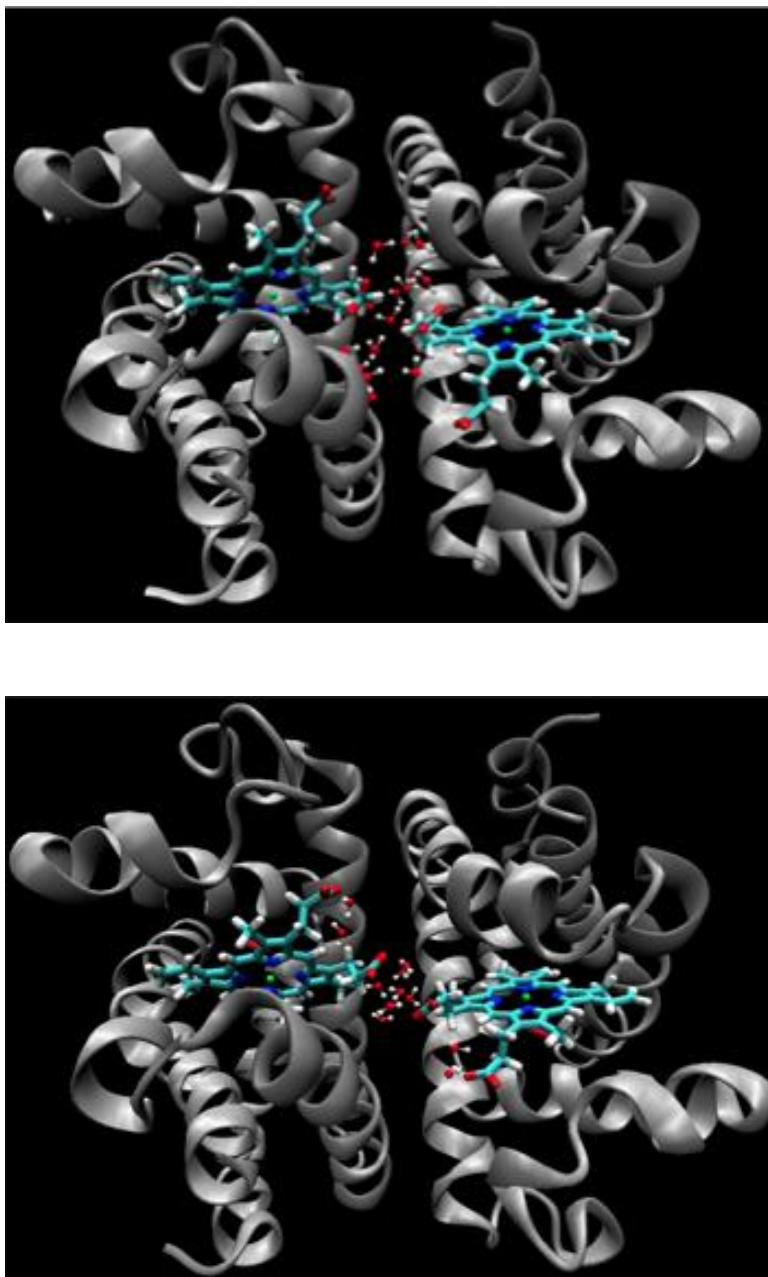


Figure 4-1 Structures of deoxy- (top) and oxy-homodimeric (bottom) hemoglobin from *Scapharca inaequalvis* (HbI). The unliganded protein and liganded protein have 17 and 11 water molecules at the interface, respectively.

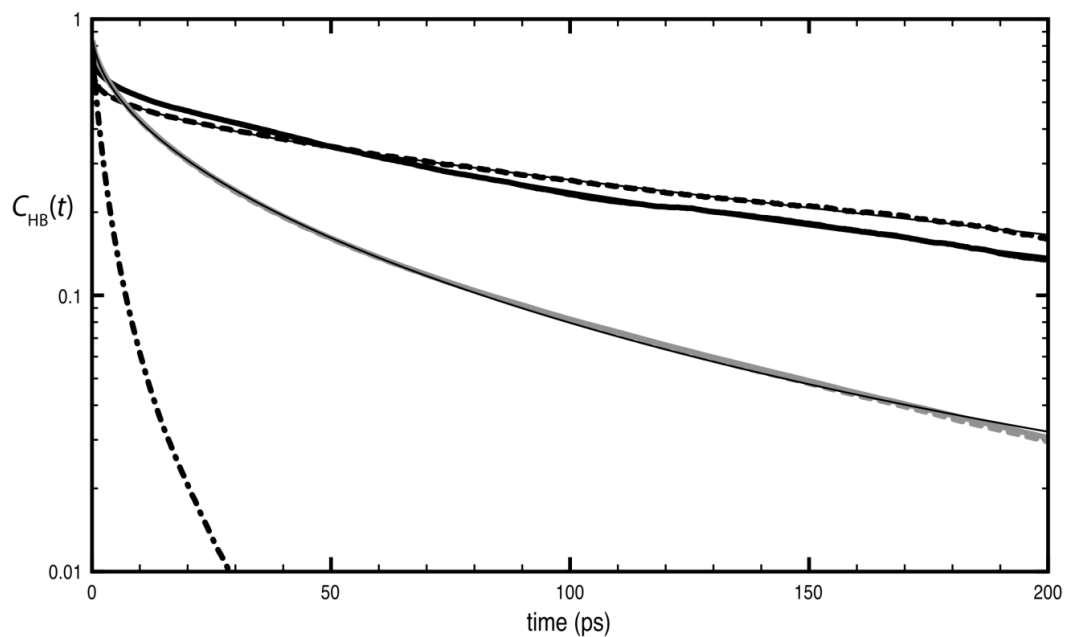


Figure 4-2 The top two black curves are hydrogen bond time correlation functions, $C_{HB}(t)$, for hydrogen bonds between water and protein for the 11 interfacial water molecules of oxy-HbI (dashed curve) and the 17 water molecules of deoxy-HbI (solid curve). Thin solid lines through $C_{HB}(t)$ are stretched exponentials fit to the computed $C_{HB}(t)$ from 0.2 to 200 ps; the parameters for the fits are listed in Table 4-1. Gray solid and dashed curves, which are nearly identical, correspond to $C_{HB}(t)$ for hydration waters with the protein for the deoxy- and oxy- system, respectively. Thin lines through the computed $C_{HB}(t)$ are fits to stretched exponentials. The dot-dashed curve corresponds to $C_{HB}(t)$ computed for hydrogen bonds of pure water.

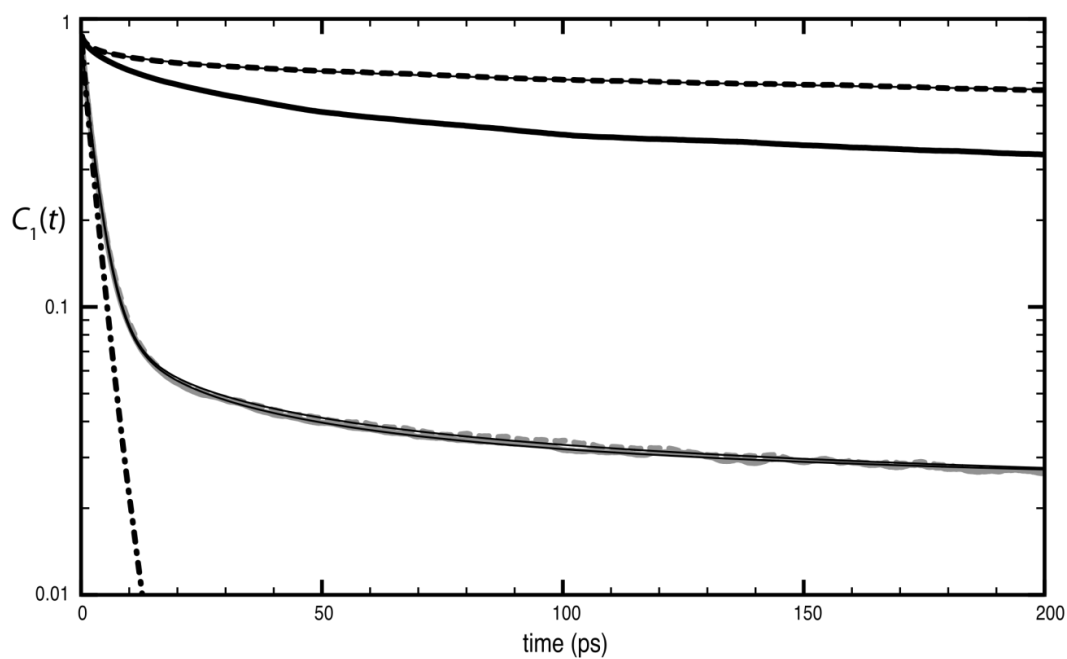


Figure 4-3 The top two black curves are rotational anisotropy decay functions ($n = 1$), $C_1(t)$, for the 11 interfacial water molecules of oxy-HbI (dashed curve) and the 17 water molecules of deoxy-HbI (solid curve). Thin solid lines through $C_1(t)$ for the interfacial water molecules are stretched exponentials fit to the computed $C_1(t)$ from 0.2 to 200 ps; the parameters for the fits are listed in Table 4-2. Gray solid and dashed curves, which are nearly identical, correspond to $C_1(t)$ for hydration waters within 3 Å of the protein surface for the deoxy- and oxy- system, respectively. The dot-dashed curve corresponds to $C_1(t)$ computed for hydrogen bonds of pure water.

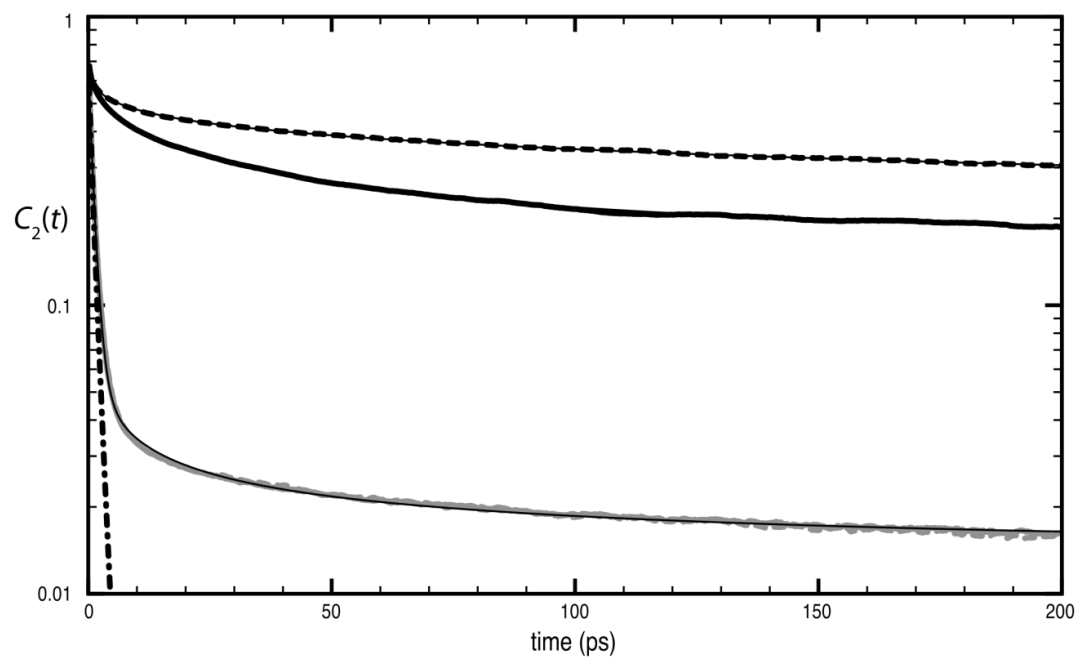


Figure 4-4 The top two black curves are rotational anisotropy decay functions ($n = 2$), $C_2(t)$, for the 11 interfacial water molecules of oxy-HbI (dashed curve) and the 17 water molecules of deoxy-HbI (solid curve). Thin solid lines through $C_2(t)$ for the interfacial water molecules are stretched exponentials fit to the computed $C_2(t)$ from 0.2 to 200 ps; the parameters for the fits are listed in Table 4-2. Gray solid and dashed curves, which are nearly identical, correspond to $C_2(t)$ for hydration waters within 3 Å of the protein surface for the deoxy- and oxy- system, respectively. The dot-dashed curve corresponds to $C_2(t)$ computed for hydrogen bonds of pure water.

4.6 References

- (1) Persson, E.; Halle, B. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 6266.
- (2) Ball, P. *Chem. Rev.* **2008**, *108*, 74.
- (3) Pal, S. K.; Peon, J.; Zewail, A. H. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 1763.
- (4) Bagchi, B. *Chem. Rev.* **2005**, *105*, 3197.
- (5) Tobias, D. J.; Sengupta, N.; Tarek, M. *Proteins: Energy, Heat and Signal Flow* **2009**, D. M. Leitner and J. E. Straub, eds. (Taylor and Francis, New York, 2009).
- (6) Stanley, H. E.; Kumar, P.; Han, S.; Mazza, M. G.; Stokely, K.; Buldyrev, S. V.; Franzese, G.; Mallamace, F.; Xu, L. *J. Phys.: Condens. Matter* **2009**, *21*, 504105.
- (7) Bizzarri, A. R.; Cannistraro, S. *J. Phys. Chem B* **2002**, *106*, 6617.
- (8) Johnson, M. E.; Malardier-Jugroot, C.; Murarka, R. K.; Head-Gordon, T. *J. Phys. Chem. B* **2009**, *113*, 4082.
- (9) Hua, L.; Huang, X.; Zhou, R.; Berne, B. J. *J. Phys. Chem. B* **2006**, *110*, 3704.
- (10) Wiesner, S. K.; Prendergast, E.; Halle, B. *J. Mol. Biol.* **1999**, *286*, 233.
- (11) Leitner, D. M.; Havenith, M.; Gruebele, M. *Int. Rev. Phys. Chem.* **2006**, *25*, 553.
- (12) Grebenkov, D. S.; Goddard, Y. A.; Diakova, G.; Korb, J.-P.; Bryant, R. G. *J. Phys. Chem. B* **2009**, *113*, 13347.
- (13) Pizzitutti, F.; Marchi, M.; Sterpone, F.; Rossky, P. J. *J. Phys. Chem. B* **2007**, *111*, 7584.
- (14) Furse, K. E.; Corcelli, S. A. *J. Phys. Chem. Lett.* **2010**, *1*, 1813.
- (15) Furse, K. E.; Corcelli, S. A. *J. Phys. Chem. B* **2010**, *114*, 9934.
- (16) Yu, I.; Tasaki, T.; Nakada, K.; Nagaoka, M. *J. Phys. Chem. B* **2010**, *10.1021/jp1043267*.
- (17) Yu, I.; Nagaoka, M. *Chem. Phys. Lett.* **2004**, *388*, 316.
- (18) Mitra, L.; Smolin, N.; Ravindra, R.; Royer, C.; Winter, R. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1249.
- (19) LeBard, D. N.; Matyushov, D. V. *J. Phys. Chem. B* **2010**, *114*, 9246.
- (20) Dadarlat, V. M.; Post, C. B. *Biophys. J.* **2006**, *91*, 4544.

- (21) Yu, X.; Park, J.; Leitner, D. M. *J. Phys. Chem. B* **2003**, *107*, 12820.
- (22) Chalikian, T. V. *Annu Rev Biophys Biomol Struct* **2003**, *32*, 207.
- (23) Lervik, A.; Bresme, F.; Kjelstrup, S.; Bedeaux, D.; Rubi, J. M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 1610.
- (24) Carey, C.; Cheng, Y.-K.; Rosicky, P. J. *Chem. Phys.* **2000**, *258*, 415.
- (25) Yu, H.; Rick, S. W. *J. Am. Chem. Soc.* **2009**, *131*, 6608.
- (26) Frölich, A.; Gabel, F.; Jasnin, M.; Lehnert, U.; Oesterhelt, D.; Stadler, A.; Tehei, M.; Weik, M.; Wood, K.; Zaccai, G. *Faraday Discussions* **2009**, *141*, 117.
- (27) Knab, J. R.; Chen, J.-Y.; Markelz, A. G. *Biophys. J.* **2006**, *90*, 2576.
- (28) Born, B.; Kim, S.-J.; Ebbinghaus, S.; Gruebele, M.; Havenith, M. *Faraday Discussions* **2009**, *141*, 161.
- (29) Makarov, V. A.; Feig, M.; Andrews, B. K.; Pettitt, B. M. *Biophys. J.* **1998**, *75*, 150.
- (30) Shenogina, N.; Koblinski, P.; Garde, S. *J. Chem. Phys.* **2008**, *129*, 155105.
- (31) Frauenfelder, H.; Fenimore, P. W.; Chen, G.; McMahon, B. H. *Proc. Natl. Acad. Sci. (USA)* **2006**, *103*, 15469.
- (32) Heugen, U.; Schwaab, G.; Bründermann, E.; Heyden, M.; Yu, X.; Leitner, D. M.; Havenith, M. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 12301.
- (33) Heyden, M.; Bründermann, E.; Heugen, U.; Niehues, G.; Leitner, D. M.; Havenith, M. *J. Am. Chem. Soc.* **2008**, *130*, 5773.
- (34) Schmidt, D. A.; Birer, Ö.; Funkner, S.; Born, B.; Gnanasekaran, R.; Schwaab, G.; Leitner, D. M.; Havenith, M. *J. Am. Chem. Soc.* **2009**, *131*, 18512.
- (35) Ebbinghaus, S.; Kim, S.-J.; Heyden, M.; Yu, X.; Heugen, U.; Gruebele, M.; Leitner, D. M.; Havenith, M. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 20749.
- (36) Kim, S. J.; Born, B.; Havenith, M.; Gruebele, M. *Angewandte Chemie* **2008**, *120*, 6486.
- (37) Ebbinghaus, S.; Kim, S. J.; Heyden, M.; Yu, X.; Gruebele, M.; Leitner, D. M.; Havenith, M. *J. Am. Chem. Soc.* **2008**, *130*, 2374.
- (38) Ebbinghaus, S.; Meister, K.; Born, B.; DeVries, A. L.; Gruebele, M.; Havenith, M. *J. Am. Chem. Soc.* **2010**, *132*, 12210.

- (39) Lombardo, T. G.; Giovambattista, N.; Debenedetti, P. G. *Faraday Discussions* **2009**, *141*, 359.
- (40) Dzubiella, J.; Swanson, J. M. J.; McCammon, J. A. *Phys. Rev. Lett.* **2006**, *96*, 087802.
- (41) Bresme, F.; Wynveen, A. *J. Chem. Phys.* **2007**, *126*, 044501.
- (42) Giovambattista, N.; Debenedetti, P. G.; Rossky, P. J. *J. Phys. Chem. C* **2007**, *111*, 1323.
- (43) Hua, L.; Huang, X.; Liu, P.; Zhou, R.; Berne, B. J. *J. Phys. Chem. B* **2007**, *111*, 9069.
- (44) Cheng, Y.-K.; Rossky, P. J. *Nature* **1998**, *392*, 696.
- (45) Luise, A.; Falconi, M.; Desideri, A. *Proteins: Struct., Funct., Genet.* **2000**, *39*, 56.
- (46) Yu, H.; Rick, S. W. *J. Phys. Chem. B* **2010**, *114*, 11552.
- (47) Chiancone, E.; Vecchini, P.; Verzilli, D.; Ascoli, F.; Antonini, E. *J. Mol. Biol.* **1981**, *152*, 577.
- (48) Royer, W. E.; Love, W. E.; Fenderson, F. F. *Nature* **1985**, *316*, 277.
- (49) Ricci, M. A.; Bruni, F.; Giuliani, A. *Faraday Discuss.* **2009**, *141*, 347.
- (50) Pieniazek, P. A.; Lin, Y.; Chowdhary, J.; Ladanyi, B. M.; Skinner, J. L. *J. Phys. Chem. B* **2009**, *113*, 15017.
- (51) Morales, C. M.; Thompson, W. H. *J. Phys. Chem. A* **2009**, *113*, 1922.
- (52) Knapp, J. E.; Pahl, R.; Srajer, V.; Royer, W. E. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 7649.
- (53) Elber, R. *Biophys. J.* **2007**, *92*, L85.
- (54) Knapp, J. E.; Gibson, Q. H.; Cushing, L.; Royer, W. E. *Biochem.* **2001**, *40*, 14795.
- (55) Royer, W. E.; Pardanani, A.; Gibson, Q. H.; Peterson, E. S.; Friedman, J. M. *Proc. Natl. Acad. Sci.* **1996**, *93*, 14526.
- (56) Royer, W. E.; Hendrickson, W. A.; Chiancone, E. *Science* **1990**, *249*, 518.
- (57) Ceci, P.; Giangiacomo, L.; Boffi, A.; Chiancone, E. *J. Mol. Biol.* **2002**, *277*, 6929.

- (58) Pardanani, A.; Gambacurta, A.; Ascoil, F.; Royer, W. E. *J. Mol. Biol.* **1998**, *284*, 729.
- (59) Pardanani, A.; Gibson, Q. H.; Colotti, G.; Royer, W. E. *J. Biol. Chem.* **1997**, *272*, 13171.
- (60) Leitner, D. M.; Straub, J. E. *Proteins: Energy, Heat and Signal Flow (Taylor and Francis Press, New York)* **2009**.
- (61) Zhou, Y.; Zhou, H.; Karplus, M. *J. Mol. Biol.* **2003**, *326*, 593.
- (62) Fujisaki, H.; Straub, J. E. *Proc. Natl. Acad. Sci. (USA)* **2005**, *102*, 6726.
- (63) Leitner, D. M. *Adv. Chem. Phys.* **2005**, *130B*, 205.
- (64) Leitner, D. M. *Ann. Rev. Phys. Chem.* **2008**, *59*, 233.
- (65) Yu, X.; Leitner, D. M. *J. Phys. Chem. B* **2003**, *107*, 1698.
- (66) Nguyen, P. H.; Park, S. M.; Stock, G. *J. Chem. Phys.* **2010**, *132*, 025102.
- (67) Takayanagi, M.; Okumura, H.; Nagaoka, M. *J. Phys. Chem. B* **2007**, *111*, 864.
- (68) Guex, N.; M. C. Peitsch *Electrophoresis* **1997**, *18*, 2714.
- (69) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- (70) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.
- (71) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comp. Phys.* **1997**, *23*, 327.
- (72) Bussi, G.; Donadio, D.; Parrinello, M. *J. Chem. Phys.*, **2007**, *126*, 014101.
- (73) Parrinello, M.; Rahman, A. *J. Appl. Phys.* **1981**, *52*, 7182.
- (74) Marchi, M.; Sterpone, F.; Ceccarelli, M. *J. Am. Chem. Soc.* **2002**, *124*, 6787.
- (75) Ota, N.; Agard, D. A. *J. Mol. Biol.* **2005**, *351*, 345.
- (76) Yu, X.; Leitner, D. M. *Journal of Chemical Physics* **2005**, *122*, 054902.
- (77) Yu, X.; Leitner, D. M. *J. Chem. Phys.* **2005**, *123*, 104503.
- (78) Yu, X.; Leitner, D. M. *Phys. Rev. B* **2006**, *74*, 184305.
- (79) Enright, M. B.; Leitner, D. M. *Phys. Rev. E* **2005**, *71*, 011912.

- (80) Enright, M. B.; Yu, X.; Leitner, D. M. *Phys. Rev. E* **2006**, *73*, 051905.
- (81) Yu, X.; Leitner, D. M. *J. Chem. Phys.* **2003**, *119*, 12673.
- (82) Leitner, D. M. *Phys. Rev. B* **2001**, *64*, 094201.
- (83) Leitner, D. M. *Phys. Rev. Lett.* **2001**, *87*, 188102.
- (84) Luzar, A. *J. Chem. Phys.* **2000**, *113*, 10663.
- (85) Luzar, A.; Chandler, D. *Nature* **1996**, *379*, 55.
- (86) Xu, H.; Berne, B. J. *J. Phys. Chem. B* **2001**, *105*, 11929.
- (87) Tarek, M.; Tobias, D. J. *Biophys. J.* **2000**, *79*, 3244.
- (88) Tarek, M.; Tobias, D. J. *Phys. Rev. Lett.* **2002**, *88*, 138101.
- (89) Rocchi, C.; Bizzarri, A. R.; Cannistraro, S. *Phys. Rev. E* **1998**, *57*, 3315.
- (90) Heyden, M.; Havenith, M. *Methods* **2010**, *52*, 74.
- (91) Pal, S.; Bagchi, B.; Balasubramanian, S. *J. Phys. Chem. B* **2005**, *109*, 12879.
- (92) Yoshida, K.; Yamaguchi, T.; Bellissent-Funel, M.-C.; Longeville, S. *Eur. Phys. J. Special Topics* **2007**, *141*, 223.
- (93) Dellerue, S.; Bellissent-Funel, M.-C. *Chem. Phys.* **2000**, *258*, 315.
- (94) Jansson, H.; Kargl, F.; Fernandez-Alonso, F.; Swenson, J. *J. Chem. Phys.* **2009**, *130*, 205101.
- (95) Faraone, A.; Liu, K.; Mou, C.; Zhang, Y.; Chen, S. *J. Chem. Phys.* **2009**, *130*, 134512.
- (96) Oleinikova, A.; Smolin, N.; Brovchenko, I. *Biophys. J.* **2007**, *93*, 2986.
- (97) Piletic, I. R.; Moilanen, D. E.; Spry, D. B.; Levinger, N. E.; Fayer, M. D. *J. Phys. Chem. A* **2006**, *110*, 4985.
- (98) Nienhaus, K.; Knapp, J. E.; Palladino, P.; Royer, W. E.; Nienhaus, G. U. *Biochem.* **2007**, *46*, 14018.
- (99) Botan, V.; Backus, E. H. G.; Pfister, R.; Moretto, A.; Crisma, M.; Toniolo, C.; Nguyen, P. H.; Stock, G.; Hamm, P. *Proc. Natl. Acad. Sci. (USA)* **2007**, *104*, 12749.

5 The Dielectric Response to Photoexcitation of GFP: A Molecular Dynamics Study

Abstract

The dielectric response to photoexcitation of the Green Fluorescent Protein (GFP) chromophore and contributions to it from water and GFP are computed by molecular simulations using a force field for the chromophore parametrized by *ab initio* calculations of ground and excited states. The chromophore is embedded in the β -barrel where it is surrounded by about ten waters, which are found to play a significant role in slow dielectric relaxation. Dynamics of hydrogen bonds between water and GFP is examined and found to be slower and more heterogeneous for hydrogen bonds inside the β -barrel than in the hydration layer around GFP.

Keywords: Dielectric Relaxation; Time-dependent Stokes Shift; Biological Water

5.1 Introduction

The response of water to photoexcitation of a solute immersed in it, and the corresponding time-dependent Stokes shift (TDSS) of the solute, have been understood at a molecular level with the help of molecular dynamics (MD) simulations for some time (1). The dielectric response to the photoexcitation of a chromophore of a biological macromolecule such as a protein or DNA typically proceeds at a much slower pace than the response of water near a much smaller solute (2-10), and the origin of the sluggish TDSS has been the subject of some debate (10-12). The slow response has been attributed by some to the response of hydration water to photoexcitation and charge redistribution of the chromophore (5,6). Recent MD simulations of solvated proteins and nucleic acids, however, indicate that the TDSS is coupled to the relaxation of the biomolecule itself rather than the water (12). Of course, the dynamics of water and a biomolecule are coupled (13-23) but the water is often thought to play only a secondary role in the response, contributing to the plasticity of the biomolecule, which in turn mainly controls the slow response to photoexcitation. However, some clusters of water molecules may be particularly tightly bound to a biomolecule (14,24), and if they are in close proximity to a chromophore they may mediate the dielectric response. We address here the contribution of water molecules to the overall dielectric response to photoexcitation of the green fluorescent protein (GFP) chromophore.

Numerous studies have probed the contribution of water to the response of photoexcitation of large biomolecules. In early work, Zewail and coworkers suggested a significant, even dominant role of water in solvation response, concluded from femtosecond spectroscopic experiments with various probes in proteins and DNA, as well

as MD simulations (25,26). Later, however, Halle and Nilsson concluded that water response was relatively rapid, and that any long-time response was due to the protein (27,28). Molecular simulations by Singer and coworkers reached a similar conclusion, though the contribution of water to a protein's plasticity was also argued to be a factor in the overall response (10). Golosov and Karplus revealed that solvation response to photoexcitation of the chromophore can be very sensitive to the location of the chromophore in the protein, i.e., both protein and water dynamics contribute to solvation response in a site-dependent manner (9). Detailed MD simulation studies of Stokes shift dynamics of chromophores in solvated DNA have also been carried out and multiple time scales identified (12,29). As with proteins, the slowest time scale response to photoexcitation of the chromophore appears largely due to the biomolecule (12).

The GFP chromophore is embedded inside the β -barrel surrounded by roughly ten water molecules that move in and out of the hydration layer (Fig. 5-1). These relatively tightly bound waters are critical to the mechanism for GFP fluorescence. One of them lies along the pathway for excited state proton transfer (ESPT) (30), which gives rise to the chromophore anion several ps following photoexcitation (30,31). In this study, we neglect ESPT and focus on relaxation over many time scales and probe the extent to which confined waters surrounding a chromophore in a protein contribute to slow dielectric relaxation following photoexcitation. We adopt a force field model for the chromophore parametrized by *ab initio* calculations of the ground and excited states.

Computational methods, including information about the force field model used for the chromophore, are described in the following section. Section 3 presents results and discussion of the computational studies of the dielectric response of the protein and

water to photoexcitation of the GFP chromophore, and dynamics of hydrogen bonds between water molecules and protein in the β -barrel and hydration layer around the protein. Concluding remarks are given in Section 4.

5.2 Computational Methods

The initial structure of GFP was obtained from the Protein Data Bank (1EMA) with 10 confined crystallographic water molecules before energy minimization for 1000 steps with the steepest descent algorithm. Then it was solvated in a cubic water box of 73 Å on each side filled with TIP3P water (32,33). The system containing 11,816 water molecules was energy-minimized again using the same algorithm followed by equilibration to 300 ps using the AMBER force field embedded in the GROMACS package (34,35) with periodic boundary conditions. In the first 100 ps the protein was restrained before it was released in the next 200 ps. The SHAKE algorithm was used to constrain all bonds that include hydrogen atoms. Non-bonded interactions were gradually brought to zero using a shift function for electrostatic interactions and a switch function for van der Waals interactions at 10 and 12 Å cutoffs, respectively. A standard leapfrog algorithm with an integration time step of 0.5 fs was employed to integrate Newton's equations of motion. The charges of the chromophore in neutral form at both ground (S_0) and vertical excited states (S_1), were obtained from an *ab initio* CASSCF (12/11)/6-311G** calculation with Gaussian03 (36). The charges were then introduced into the AMBER force field, giving rise to a change of the transition dipole moment of 1.36 D from the ground to excited state. Two 10 ns classical MD simulations were run using the micro-canonical (NVE) ensemble at 300 K for the ground and excited state, respectively, with the atomic coordinates stored every 1 fs by using GROMACS software package (37). We note that we carry out classical MD simulations on GFP, rather than, e.g., Quantum Mechanics/Molecular Mechanics (QM/MM) methods, as have also been adopted for GFP (38,39). MD simulations are more convenient for this study because of

the long times, several ≈ 10 ns MD trajectories, required to enhance sampling and converge the dielectric response of GFP, an issue discussed by Furse and Corcelli concerning calculation of the TDSS in DNA (12).

Hydrogen bond time correlation functions, $C_{HB}(t)$, were computed for bonds between water molecules and the protein at 300 K. For this study, we define $C_{HB}(t)$ as the probability that, if a hydrogen bond between donor, D, and acceptor, A, exists at $t = 0$, then is still exists at time, t , even if the bond broke at some intermediate time (40). Formally, the definition is given by

$$C_{HB}(t) = \frac{\langle h(t)h(0) \rangle}{\langle h(0) \rangle}, \quad (5-1)$$

where $h(t) = 1$ if a hydrogen bond between DA exists and at time, t , and is otherwise 0. We adopt a standard criterion for hydrogen bonds, specifically a D-A distance of 3.5 Å and a D-H-A angle greater than 150° (5).

For calculation of the solvent response to photoexcitation, the atomic charge distribution of the GFP chromophore between its ground S_0 and excited S_1 state was shifted instantaneously, leading to sudden perturbation on the internal electronic environment of the protein. We assumed that the intermolecular potential is pairwise-additive and hence the energy difference (ΔE) between S_0 and S_1 can be represented as

$$\Delta E = \sum_j w_{0j} \quad (5-2)$$

where Δw_{0j} is the change in the potential between the chromophore (molecule 0) and the j^{th} solvent molecule or protein region. It is also assumed that the charge distributions of the chromophore and polar solvent molecules are represented as sets of partial charges and Δw_{0j} can be further expressed as a sum of Coulombic interactions

$$\Delta W_{0j} = \sum_{\alpha \in 0} \sum_{\beta \in j} \frac{\Delta q_{\alpha} q_{\beta}}{4\pi\epsilon_0 r_{0\alpha,j\beta}} \quad (5-3)$$

where Δq_{α} is the change in the partial charge of chromophore site α from the electronic transition, q_{β} is the partial charge on the solvent site β and $r_{0\alpha,j\beta}$ is the distance between the chromophore sites α and β on the j^{th} solvent molecule (41).

The solvation response function is defined as (42)

$$S(t) = \frac{\nu(t) - \nu(\infty)}{\nu(0) - \nu(\infty)} \quad (5-4)$$

where $\nu(t)$ is the time evolution of the frequency of maximum fluorescence. Within the linear response approximation, the normalized solvation response function, $S(t)$, is equal to the equilibrium solvation time correlation function, $C(t)$, which is (1,2,42,43)

$$C(t) = \frac{\langle (\Delta E(t) - \langle \Delta E \rangle) (\Delta E(0) - \langle \Delta E \rangle) \rangle}{\langle (\Delta E(0) - \langle \Delta E \rangle)^2 \rangle} \quad (5-5)$$

where $\Delta E(t) = E_e - E_g$ is the difference in the chromophore-bath interaction energy when the chromophore is in the excited state, e , and the ground state, g , at time, t ; the brackets denote equilibrium ensemble average, which in practice we obtain by carrying out sufficiently long MD simulations in the ground state. In our calculations of $C(t)$ the bath is the rest of the system, specifically the rest of the protein and all waters.

To extract the contribution to the total response from each component, α , where a component can be the protein or water, we have, following Ref. (12) decomposed $C(t)$ into

$$C(t) = \sum_{\alpha} \frac{\langle \delta \Delta E_{\alpha}(t) \delta \Delta E(0) \rangle}{C(0)} = \sum_{\alpha} C_{\alpha}(t) \quad , \quad (5-6)$$

where $\delta\Delta E_\alpha(t) = \Delta E_\alpha(t) - \langle \Delta E_\alpha \rangle$. We have thereby broken the total solvent-response interaction energy into a sum, i.e., $\Delta E(t) = \sum_\alpha \Delta E_\alpha(t)$. Auto- and cross-correlations between components are calculated as (12)

$$C(t) = \sum_\alpha \sum_\beta \frac{\langle \delta\Delta E_\alpha(0) \delta\Delta E_\beta(t) \rangle}{C(0)} = \sum_\alpha \sum_\beta C_{\alpha\beta}(t) \quad . \quad (5-7)$$

5.3 Results and Discussion

To quantify the hindered motion of the water molecules inside the GFP β -barrel and to compare with the motion of bulk and hydration water, we have computed hydrogen bond lifetime correlation function, $C_{HB}(t)$, defined by Eq. (5-1), widely used as a measure of the dynamics of water molecules in a hydrogen bond network (44). In Fig. 5-2 we plot $C_{HB}(t)$ for hydrogen bonds between water molecules and the chromophore in the β -barrel of GFP, the confined water molecules and GFP, and water molecules in the hydration layer outside the protein, and for water molecules in the bulk. As we have observed for clusters of confined waters in other proteins (14), $C_{HB}(t)$ can be fit well to a stretched exponential, $\exp[-(t/\tau)^\beta]$, over times from 0.2 to 100 ps. Values of time constant τ and exponent β are listed in Table 5-1. For the most tightly confined water in Fig. 5-2, those hydrogen bonding to the chromophore, we find the values of τ and β to be 17.85 ps and 0.307, respectively. These results, as well as the lifetimes of hydrogen bonds between confined waters and other parts of the β -barrel, are comparable to the values we found recently for the 10 to 20 water molecules confined to the interface between two globules of a homodimeric hemoglobin (14) and are similar to values obtained for water molecules confined in reverse micelles (45). The small value of β is typical of glass-like dynamics in a highly heterogeneous environment. We observe hydrogen bonds between water molecules and GFP in the hydration layer to exhibit significantly shorter lifetimes, though we could still fit the data over the first 100 ps to a stretched exponential, and the hydrogen bonds are much longer lived than hydrogen bonds between water molecules in the bulk.

Though $C_{HB}(t)$ appears noisy beyond 100 ps for some of the bonds, we plot it in the inset to 1 ns to assess the stretched exponential fits made to 100 ps. We observe that $C_{HB}(t)$ for water-protein hydrogen bonds in the β -barrel fits very well to the same stretched exponential used to fit the data to 100 ps. This appears to be the case for the water-chromophore hydrogen bonds, too, but the data becomes noisy beyond about 400 ps. $C_{HB}(t)$ for water-protein hydrogen bonds in the hydration layer does not fit to a stretched exponential over long times, at least not to the same stretched exponential we fit to 100 ps. It also does not appear to exhibit a power law variation in the tail, as there is no linear variation in the tail in this log-log plot, as there is for bulk water, the latter expected from diffusion theory for a reversible geminate pair (46).

Results for the dielectric response to photoexcitation of the chromophore, $C(t)$, computed for solvated GFP at 300 K are plotted in Fig. 5-3 for times out to 100 ps. We compare this with a second system, specifically the dielectric response by water following photoexcitation of the GFP chromophore separated from the protein (Fig. 5-1b). In this way, we observe differences between $C(t)$ for a solvated chromophore, the same chromophore as in GFP, and the GFP chromophore in its protein and solvent environment. It is apparent from the very different responses in these two systems that $C(t)$ for GFP must be controlled by either the protein environment, the confined waters, or both. At very early times, $C(t)$ is comparable for both systems, indicating that the fast inertial response to changes in partial charges on the chromophore that occurs in water also occurs in the protein and in the confined water molecules.

To provide a more quantitative analysis, we have matched the time behavior of $C(t)$ obtained from the simulations to a four-exponential fitting function,

$$C_{fit}(t) = a_1 \exp(-t/\tau_1) + a_2 \exp(-t/\tau_2) + a_3 \exp(-t/\tau_3) + a_4 \exp(-t/\tau_4) \quad (5-8)$$

In Table 5-2 we list the values of the constants in Eq. (5-8) that we obtain when we fit this equation to the data plotted in Fig. 5-2. For $C(t)$ obtained for the chromophore inside the protein (i) and chromophore-bulk water interaction energies (ii), we find distinct values for these constants, as we would expect from the dissimilarity of the curves in Fig. 5-3. We find for both (i) and (ii) that approximately 43 % of the decay occurs with a time constant of tens of femtoseconds. The fast time scale is often considered to correspond to the inertial fast dynamics (1,2), and while the actual time constants in this regime are quite different, this could be due to uncertainties in the fitting during the rapid decay.

The other components of $C(t)$ correspond to significantly longer time scales for the chromophore in GFP compared to the chromophore in aqueous solution. In the former case the slowest time scale, with a contribution of 18% to the overall decay, is 0.3 ns, whereas the slowest time constant in the latter case is about 5 ps, consistent with expectations of dye molecules in water (1). Thus, for the chromophore embedded in solvated GFP the dielectric response to chromophore photoexcitation occurs over time scales of tens of femtoseconds to hundreds of picoseconds, consistent with time scales for TDSS obtained in the studies on biomolecules discussed in the introduction and in other recent experimental studies (47). The main distinction of GFP, as we address below, is the significant role of water in the slow response. We note that the third and fourth components for (i), which occur with time constants of tens of ps and beyond, are much longer than the time for ESPT in GFP. These components would thus contribute little to the TDSS in GFP, and indeed only negligible TDSS is observed in GFP (3). We note also that, at long times, it is also possible that $C(t)$ varies as a power law.

We turn now to the respective contributions of protein and water to the dielectric response of the solvated protein to photoexcitation of the chromophore. Specifically, we calculate the decomposition of $C(t)$ in terms of the contributions of protein and water given by Eq. (5-6) and plot the results in Fig. 5-4(a). We observe in this figure that while the contribution to $C(t)$ due to chromophore-protein interactions is generally larger than the contribution due to chromophore-water interactions, the latter is quite similar to the former for times beyond 40 ps. Therefore the confined waters surrounding the chromophore contribute significantly, about as much as the protein, to the slow dielectric response to photoexcitation of the chromophore. These tightly bound waters essentially solvate the chromophore, and we have seen that hydrogen bonds between these waters and the interior of GFP persist for 10s of picoseconds, so it is quite reasonable that these waters impact the long-time response to photoexcitation so significantly.

In Fig. 5-4(b) we plot the auto- and cross-correlation functions computed with Eq. (5-7). The negative cross-correlation function for GFP and water plotted in Fig. 5-4(b) indicates an anti-correlation between the confined water and protein underlying the energy gap fluctuations, and is comparable in magnitude to the anti-correlation of water and DNA observed by Furse and Corcelli (12).

5.4 Concluding Remarks

We have computed the dielectric response of the GFP-solvent environment to photoexcitation of the GFP chromophore, as well as the response of water to photoexcitation of the bare chromophore in aqueous solution for comparison. Decomposition of the dielectric response indicates a comparable role of water and protein in slow relaxation. This finding is in contrast to a number of studies of the dielectric response of a solvated biomolecule to photoexcitation of a chromophore, where the slow response has been found to be largely due to the biomolecule (12). Indeed, a recent MD simulation of the TDSS in GFP found a slow response attributable to the protein (23), though a decomposition, such as Eq. (6), to extract the contribution of the waters was not carried out. The difference between GFP and many of the other biomolecules studied thus far is that there are about 10 tightly bound water molecules in the β -barrel of GFP that surround the chromophore. Our analysis of the hydrogen bond dynamics between these confined waters around the chromophore and the protein indicate highly heterogeneous hydrogen bonding with lifetimes of tens of picoseconds, consistent with the slow dielectric relaxation times computed for the water.

Table 5-1 Parameters for the stretched exponential, $\exp[-(t/\tau)^\beta]$, used to fit the hydrogen bond correlation functions computed for hydrogen bonds between water molecules and the chromophore, the interior of the β -barrel, and the hydration layer around GFP (excluding the interior of the β -barrel). The fits are also plotted in Fig. 5-2.

<i>Hydrogen bonding region</i>	τ /ps	β
Chromophore	17.85	0.307
β -barrel	10.86	0.355
Hydration layer	6.08	0.401

Table 5-2 Parameters used to fit Eq. (5-8) to the computed dielectric response for the chromophore in solvated GFP, and the chromophore in aqueous solution.

Interactions	a_1	τ_1/ps	a_2	τ_2/ps	a_3	τ_3/ps	a_4	τ_4/ps
Chromophore in solvated GFP	0.43	0.08	0.17	1.93	0.15	20.0	0.18	333
Chromophore in bulk water	0.43	0.01	0.27	0.07	0.28	0.36	0.11	5.28

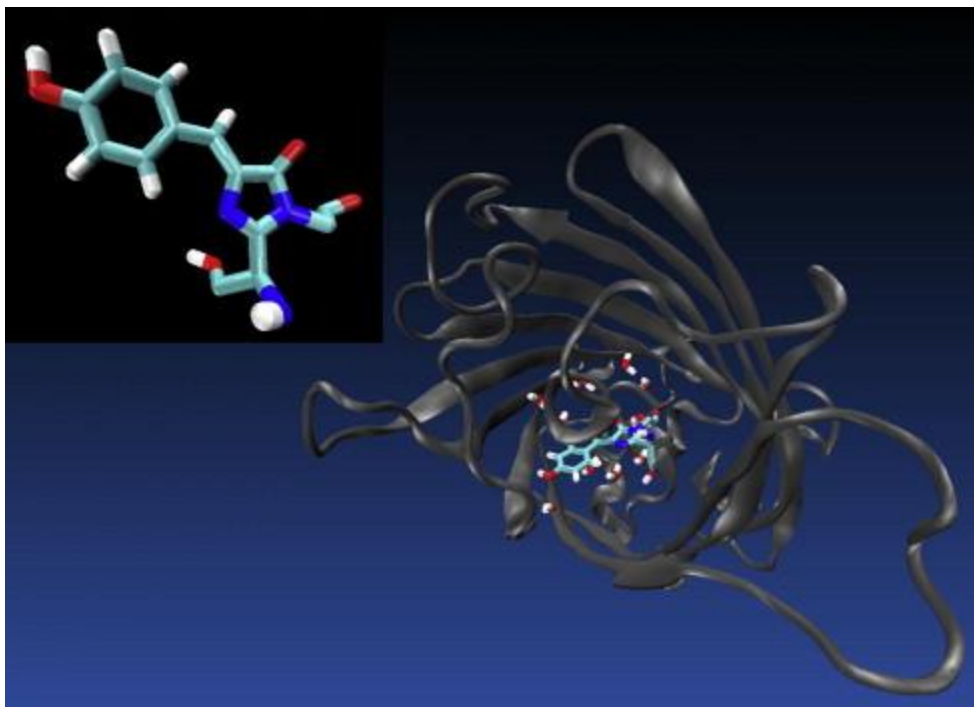


Figure 5-1 Illustration of green fluorescence protein from *Aequorea victoria* in its PDB structure, produced with VMD (48). The chromophore, surrounded by 10 confined crystallographic water molecules, is shown in atomic detail in cyan. Several additional water molecules are at times present in the β -barrel during the MD simulations. Inset (black background) shows the isolated chromophore used in MD simulations to determine the dielectric response of the water due to photoexcitation of this molecule in aqueous solution.

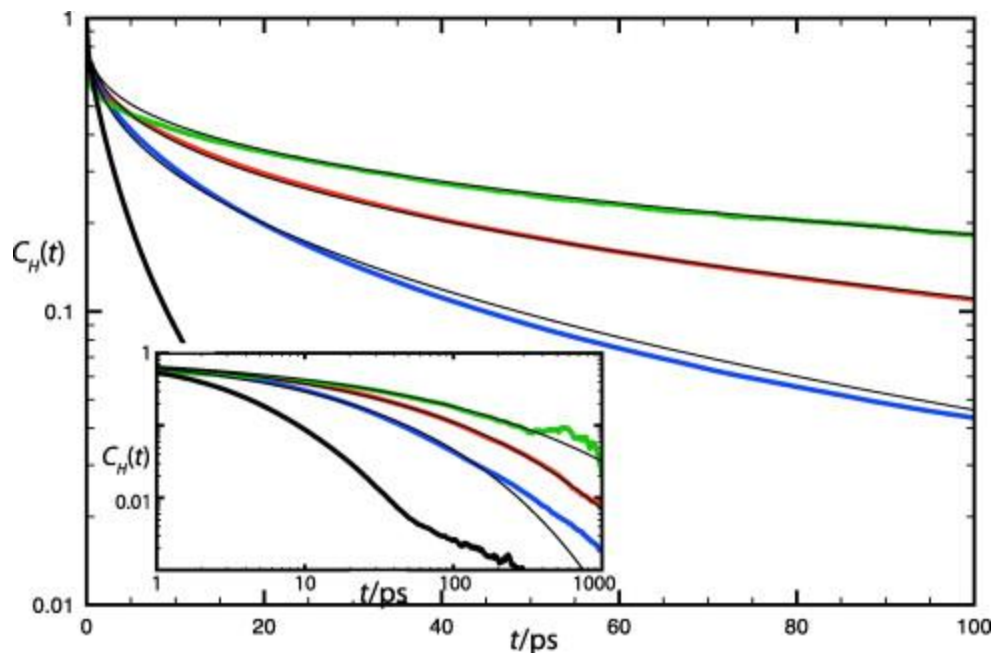


Figure 5-2 Hydrogen bond correlation function, $C_{HB}(t)$, for hydrogen bonds between water molecules and protein atoms in the hydration layer (blue), in the β -barrel (red) and hydrogen bonds between the chromophore and water (green) at 300 K, plotted to 100 ps. $C_{HB}(t)$ for hydrogen bonds between water molecules in the bulk is shown for comparison (black). Thin black curves are fits to a stretched exponential, $\exp[-(t/\tau)^\beta]$, for which the parameters are listed in Table 5-1. Inset shows a log-log plot of the time-dependence of C_{HB} to 1 ns (see text).

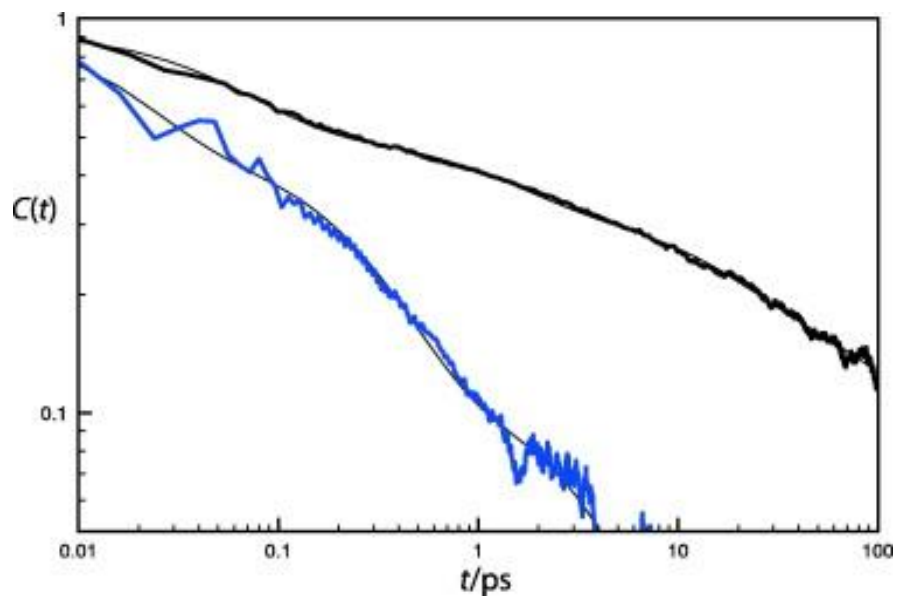


Figure 5-3 Normalized equilibrium time autocorrelation functions of fluctuations in the solvation energy differences at 300 K (solid curves) for the interactions between the chromophore and the rest of the system (black), the chromophore and bulk water (blue). Thin black curves are fits to the data using Eq. (5-8). The parameters for the fits are listed in Table 5-2.

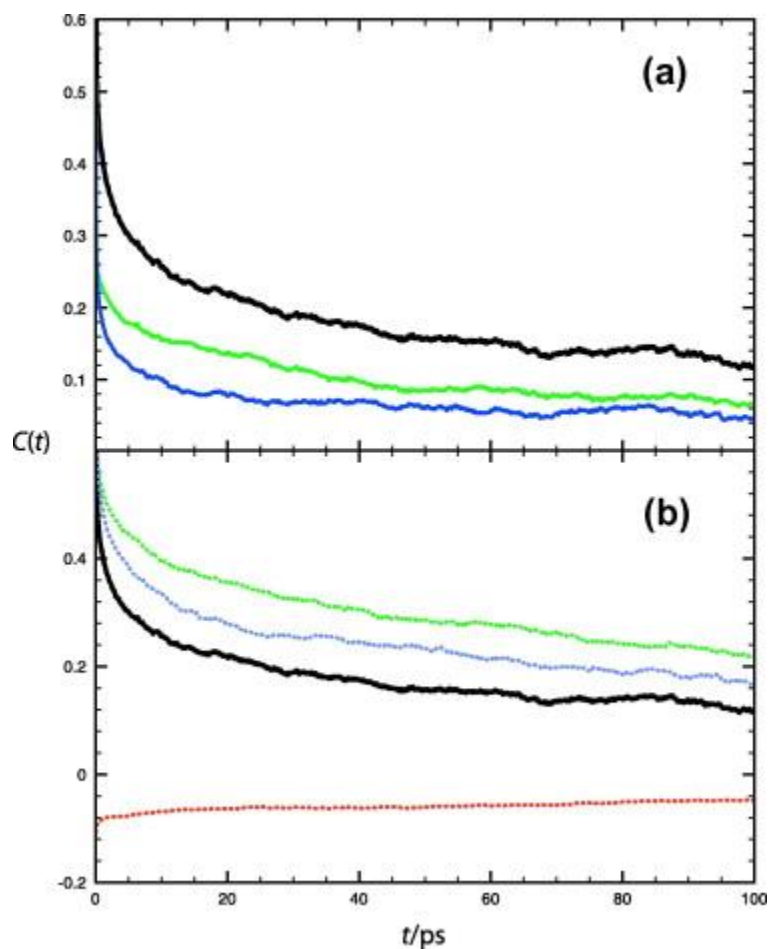


Figure 5-4 (a) Dielectric response for chromophore bound to GFP (black) and its linear response decomposition, Eq. (5-6), in terms of the components present in the simulation, confined water (solid blue), and protein matrix (solid green); (b) auto-correlation functions for protein (dotted green) and water (dotted blue) and the cross-correlation function (dotted red) between chromophore-water and chromophore-protein interactions, both given by Eq. (5-7). They are shown with the dielectric response for chromophore bound to GFP (black).

5.5 References

- (1) R Jimenez, G R Fleming, P V Kumar, M Maroncelli, *Nature* 369 (1994) 471 - 72.
- (2) D Xu, C Martin, K Schulten, *Biophys. J.* 70 (1996) 453 - 60.
- (3) P Abbyad, W Childs, X Shi, S G Boxer, *Proc. Natl. Acad. Sci. (USA)* 104 (2007) 20189 - 94.
- (4) P Abbyad, X Shi, W Childs, T B McAnaney, B E Cohen, S G Boxer, *J. Phys. Chem. B* 111 (2007) 8269 - 76.
- (5) B Bagchi, *Chem. Rev.* 105 (2005) 3197 - 219.
- (6) L Z Samir Kumar Pal, Ahmed H. Zewail, *PNAS* 100 (2003) 8113-18.
- (7) K E Furse, S A Corcelli, *J. Am. Chem. Soc.* 130 (2008) 13103-09.
- (8) K E Furse, S A Corcelli, *J. Phys. Chem. Lett.* 1 (2010) 1813-20.
- (9) A A Golosov, M Karplus, *J. Phys. Chem. B* 111 (2007) 1482-90.
- (10) A A H Tanping Li, Sherwin J. Singer *J. Phys. Chem. B* 112 (2008) 16121-34.
- (11) B Halle, L Nilsson, *J. Phys. Chem. Lett.* 113 (2009) 8210-13.
- (12) K E Furse, S A Corcelli, *J. Phys. Chem. Lett.* 1 (2010) 1813 - 20.
- (13) S Ebbinghaus, S-J Kim, M Heyden, X Yu, U Heugen, M Gruebele, D M Leitner, M Havenith, *Proc. Natl. Acad. Sci. USA* 104 (2007) 20749 - 52.
- (14) R Gnanasekaran, Y Xu, D M Leitner, *J. Phys. Chem. B* 114 (2010) 16989 - 96.
- (15) U Heugen, G Schwaab, E Bründermann, M Heyden, X Yu, D M Leitner, M Havenith, *Proc. Natl. Acad. Sci. USA* 103 (2006) 12301 - 06.
- (16) D M Leitner, M Havenith, M Gruebele, *Int. Rev. Phys. Chem.* 25 (2006) 553 - 82.
- (17) X Yu, J Park, D M Leitner, *J. Phys. Chem. B* 107 (2003) 12820 - 9
- (18) S J Kim, B Born, M Havenith, M Gruebele, *Angew. Chemie Int. Ed.* 120 (2008) 6486 - 89.
- (19) D J Tobias, N Sengupta, M Tarek, *Proteins: Energy, Heat and Signal Flow* D. M. Leitner and J. E. Straub, eds. (2009) (Taylor and Francis, New York) 361 - 86.
- (20) D N LeBard, D V Matyushov, *J. Phys. Chem. B* 114 (2010) 9246 - 58.
- (21) H Frauenfelder, P W Fenimore, G Chan, B H McMahon, *Proc. Natl. Acad. Sci. (USA)* 103 (2006) 15469 - 72.
- (22) K Meister, S Ebbinghaus, Y Xu, J G Duman, A DeVries, M Gruebele, D M Leitner, M Havenith, *Proc. Natl. Acad. Sci. USA* 110 (2013) 1617 - 22.
- (23) D R Martin, D V Matyushov, *J. Phys. Chem. B* 116 (2012) 10294 - 300.
- (24) L Hua, X Huang, R Zhou, B J Berne, *J. Phys. Chem. B* 110 (2006) 3704 - 11.
- (25) S K Pal, J Peon, A H Zewail, *Proc. Natl. Acad. Sci. USA* 99 (2002) 1763 - 8.
- (26) J Peon, S K Pal, A H Zewail, *Proc. Natl. Acad. Sci. USA* 99 (2002) 10964 - 69.
- (27) L N Bertil Halle, *J. Phys. Chem. B* 113 (2009) 8210-13.
- (28) B H Lennart Nilsson, *PNAS* 102 (2005) 13867-72.
- (29) S Pal, P K Maiti, B Bagchi, J T Hynes, *J. Phys. Chem. B* 110 (2006) 26396 - 402.
- (30) M Chattoraj, B A King, G U Bublitz, S G Boxer, *Proc. Natl. Acad. Sci. USA* 93 (1996) 8362 - 7.
- (31) C Fang, R R Frontiera, R Tran, R A Mathies, *Nature* 462 (2009) 200 - 05.
- (32) J A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir,

- K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, 1998 102 (1998) 3586-616.
- (33) W L Jorgensen, J Chandrasekhar, J D Madura, R W Impey, M L Klein, *J. Chem. Phys.* 79 (1983) 926-35.
- (34) E J Sorin, V J Pande, *Biophys. J.* 88 (2005) 2472 - 93.
- (35) A J DePaul, E J Thompson, S S Patel, K Haldeman, E J Sorin, *Nucleic Acids Res.* 38 (2010) 4856 - 67.
- (36) M J Frisch, G W Trucks, H B Schlegel, G E Scuseria, M A Robb, J R Cheeseman, J A Montgomery, T Vreven, K N Kudin, J C Burant, S S I J. M. Millam, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT. (2004).
- (37) B Hess, D v d Spoel, E Lindahl, *J. Chem. Theory Compu.* 4 (2008) 435 - 47.
- (38) A Sinicropi, T Andruniow, N Ferré, R Basosi, M Olivucci, *J. Am. Chem. Soc.* 127 (2005) 11534 - 5.
- (39) A H Steindal, J M H Olsen, K Ruud, L Frediani, J Kongsted, *Phys. Chem. Chem. Phys.* 14 (2012) 5440 - 51.
- (40) A Luzar, D Chandler, *Nature* 379 (1996) 55-57.
- (41) B M Ladanyi, *Theoretical Methods in Condensed Phase Chemistry*, S. D. Schwartz, ed.(Kluwer Academic Publishers, Amsterdam) (2000) 207-33.
- (42) W H Thompson, *Ann. Rev. Phys. Chem.* 62 (2011) 599 - 619.
- (43) B B Laird, W H Thompson, *J. Chem. Phys.* 135 (2011) art. no. 084511.
- (44) B Bagchi, *Chem. Rev.* 105 (2005) 3197-219.
- (45) P A Pieniazek, Y Lin, J Chowdhary, B M Ladanyi, J L Skinner, *J. Phys. Chem. B* 113 (2009) 15017-28.
- (46) O Markovitch, N Agmon, *J. Chem. Phys.* 129 (2008) art. no. 084505.
- (47) A Jha, K Ishii, J B Udgaonkar, T Tahara, G Krishnamoorthy, *Biochem.* 50 (2011) 397 - 408.
- (48) A D William Humphrey, Klaus Schulten, *J. Molec. Graphics* 14 (1996) 33-38.

6 Dielectric Response in Photoactive Yellow Protein: Quantum Mechanics/Molecular Mechanics and Molecular Dynamics

Simulation Studies

Abstract

Quantum mechanics/molecular mechanics and molecular dynamics simulations of the dielectric response due to vibrational energy relaxation from the chromophore to the protein and solvent environment following photoexcitation of photoactive yellow protein (PYP) are presented. The computed protein response to photoexcitation appears over a wide range of timescales. And the slow response in time dependent Stokes shift is found to be largely due to the contribution from the protein matrix rather than from the hydration water around the protein.

Keywords: Time Dependent Stokes Shift; Vibrational Energy Relaxation; Photoactive Yellow Protein.

6.1 Introduction

A variety of time-resolved optical experiments probing protein photochemistry have revealed coherent oscillations produced by photoexcitation of protein chromophores that survive for hundreds of femtoseconds to several picoseconds (1-5). The oscillations correspond to modes typically below 200 cm^{-1} and last up to a few vibrational periods. For example, ultrafast fluorescence decay measurements of photoactive yellow protein (PYP) reveal oscillations that correspond to modes of *ca.* 50 cm^{-1} (670 fs period) and *ca.* 120 cm^{-1} (280 fs period) (5). The former, which has a lifetime of about 400 fs, has been attributed to coupled chromophore-protein modes while the latter, with a lifetime around 700 fs, mainly to a vibration of the chromophore (5,6). In this Chapter, we examine by quantum mechanics/molecular mechanics (QM/MM) and molecular dynamics (MD) simulations the excited state dynamics of PYP that yield insights into the dynamic coupling between the chromophore and protein environment following photoexcitation. We use the results of the QM/MM and MD simulations on PYP in the ground and excited electronic states to compute the dielectric response, i.e., the response of the protein and solvent environment to photoexcitation and corresponding charge redistribution of the chromophore.

Photoactive yellow protein, discovered in the bacterium *Halorhodospira halophila*, is a water-soluble photoreceptor protein containing 125 amino acids that absorbs blue light, initiating a signaling state (7). The PYP chromophore, a *p*-coumaric acid anion, is covalently bonded on one end to a cystine residue and hydrogen bonds at the other end, the O⁻ phenyl- group, to two residues (Fig. 6-1). The chromophore undergoes *trans* \rightarrow *cis* photoisomerization within a few picoseconds upon

photoexcitation, modifying hydrogen bonding between the chromophore and nearby residues (8) and triggering the PYP photocycle. The time-resolved absorption spectra reveal three time scales, the fastest about 0.7 ps (relative population about 50 %), the second about 3 ps (relative population about 40 %) and the longest about 15 ps (relative population about 10 %) (9). These times have been suggested to correspond to different conformational states of the protein, and thus different chromophore environments yielding somewhat different reaction paths (9,10). The fastest and most efficient reaction path corresponds to the sub-picosecond component (9). The oscillations observed during ultrafast fluorescence decay measurements provide clues of the dynamic interactions between the chromophore and protein environment during the course of reaction. They are found to be absent in the fluorescence decay of denatured PYP, in which photochemical isomerization proceeds more slowly and less efficiently (2).

The response of water to photoexcitation of a solute immersed in it and the corresponding time-dependent Stokes shift (TDSS) of the solute have been understood at a molecular level with the help of MD simulations for some time (11,12). The solvation response function is usually expressed in normalized form as

$$S(t) = \frac{\nu(t) - \nu(\infty)}{\nu(0) - \nu(\infty)} \quad (6-1)$$

where $\nu(t)$ is the time evolution of the frequency of maximum fluorescence. The Stokes shift for PYP is 2180 cm^{-1} (9). Within the linear response approximation, the normalized solvation response function, $S(t)$, is equal to the equilibrium solvation time correlation function, $C(t)$, which is (12-14)

$$C(t) = \frac{\langle \delta\Delta E(t)\delta\Delta E(0) \rangle}{\langle (\delta\Delta E)^2 \rangle} \quad (6-2)$$

where $\Delta E(t) = E_e - E_g$ is the difference in the chromophore-bath interaction energy when the chromophore is in the excited state, e , and the ground state, g , at time, t ; $\delta\Delta E(t) = \Delta E(t) - \langle \Delta E \rangle$; and the brackets denote equilibrium ensemble average. The linear response approximation to the dynamic Stokes shift becomes more questionable on times near and beyond the reaction time of ≈ 1 ps, and we focus here mainly on times below ≈ 1 ps, though even at the short times we address the dielectric response to changes in charge on atoms of the chromophore calculated with Eq. (6-2) may only be a crude approximation to the TDSS. As with many recent MD simulations of the TDSS of proteins, the results of which typically exhibit multiexponential decay with slow components due to protein relaxation (14-16), we, too, observe decay components beyond the picosecond time scale. Therefore, we calculate the TDSS to times beyond the reaction time but we do not address them here since they correspond to times after reaction. Our computation of $C(t)$ with Eq. (6-2), the dielectric response of the protein and solvent environment to charge redistribution on the chromophore, parallels earlier work by Schulten and coworkers (17), who examined the protein response to photoexcitation of bacteriorhodopsin.

An outline of this chapter is as follows: Computational methods are reviewed in Section 2. The results and discussion of our QM/MM and MD simulations are presented in Section 3. Concluding remarks are given in Section 4.

6.2 Computational Methods

The initial structure of wild type photoactive yellow protein 1OTB was obtained from the Protein Data Bank (PDB). All MD simulations were carried out using the GROMACS software package (19). PYP was minimized for 3000 steps with the steepest descent algorithm using the AMBER03 force field (20). Then it was solvated in a cubic water box of 60Å on each side with TIP3P water (21). The system, which comprised 6499 water molecules, was neutralized by adding 7 Na⁺ ions and 1 Cl⁻ ion, and minimized again using the steepest decent algorithm. The system was then equilibrated for 300 ps using the AMBER03 force field with periodic boundary conditions. In the first 100 ps the protein was restrained before it was released in the next 200 ps. The SHAKE algorithm was employed to constrain bonds with hydrogen atoms throughout the simulation. Non-bonded interactions were gradually brought to zero by using a PME-switch for electrostatic interactions and a switch function for van der Waals at 10 and 12 Å cutoffs, respectively. A standard leapfrog algorithm with an integration time step of 0.5 fs was used to integrate Newton's equations of motion. To show the dielectric response of the solvent environment directly to charge redistribution on the chromophore, we built up the second system, the separated chromophore in bulk water. All the MD simulations were done in the micro-canonical (NVE) ensemble for 8 ns at energies corresponding to a temperature near 300 K with atomic coordinates stored every 1 fs. The charges of the chromophore (deprotonated cinnamic acid) in anionic form at both ground (S_0) and vertical excited state (S_1) were obtained from an *ab initio* CASSCF (10,12)/6-31G** calculation with Gaussian03 (22). This force field model for the PYP chromophore is similar to one used by us in previous MD simulations of PYP (23,24). In QM/MM

calculation, the chromophore was carefully built up using xLEAP module in AMBER9 simulation package, and we followed the similar simulation protocol as in its MD counterpart. In both systems (chromophore in solvated PYP and in bulk water), we modeled the chromophore using the semi-empirical PM3 method with the remaining part of the system treated classically.

6.3 Results and discussion

Results for the dielectric response to photoexcitation of the chromophore, $C(t)$, computed for solvated PYP at 300 K are plotted in Fig. 6-2 for times out to 100 ps. We compare this with the second system, specifically the dielectric response by water following photoexcitation of the PYP chromophore separated from the protein (Fig. 6-1 right). In this way, we observe differences between $C(t)$ for a solvated chromophore, the same chromophore as in PYP (Fig. 6-3), and the PYP chromophore in its protein and solvent environment. To provide a more quantitative analysis, we have matched the time behavior of $C(t)$ obtained from the simulations to a tri-exponential fitting function,

$$C_{fit}(t) = a_1 \exp(-t/\tau_1) + a_2 \exp(-t/\tau_2) + a_3 \exp(-t/\tau_3) \quad (6-3)$$

In Table 6-1 we list the values of the constants in Eq. (6-3) that we obtain when we fit this equation to the data plotted in Fig. 6-2 and 6-3. Overall, the results are consistent with fits to the TDSS computed for other proteins to times of a few picoseconds, where, over a few picoseconds, there is a fast, inertial component on the order of tens of femtoseconds, and a slow component on the order of 1 ps (14). A slower component of 100s of picoseconds or even nanoseconds (16) is typically also found fitting the TDSS to much longer times, but we address here only times of a few picoseconds, since for longer times most photoexcited PYP has reacted.

For $C(t)$ obtained for the chromophore inside the protein (i) and chromophore-bulk water interaction energies (ii), we find distinct values for these constants, as we would expect from the dissimilarity of the curves in Fig. 6-3, indicating $C(t)$ for PYP must be controlled by the protein environment, as suggested also by a careful analysis based on high-level electronic structure calculations (18). We find for both (i) and (ii)

that approximately 70 % in MD of the decay occurs with a time constant of tens of femtoseconds. The fast time scale is often considered to correspond to the inertial fast dynamics (1,2) with the actual time constants in this regime are quite similar in MD and QM/MM calculations, respectively. The other components of $C(t)$ correspond to significantly longer time scales for the chromophore in PYP compared to the chromophore in aqueous solution. For the QM/MM calculation, in the former case the slowest time scale, with a contribution of 2% to the overall decay, is 109.0 ps, whereas the slowest time constant in the latter case is about 9760 ps, consistent with expectations of dye molecules in water (1). Thus, for the chromophore embedded in solvated PYP the dielectric response to chromophore photoexcitation occurs over time scales of tens of femtoseconds to hundreds of picoseconds and beyond, consistent with time scales for TDSS obtained in the studies on biomolecules discussed in the introduction and in other recent experimental studies (25).

6.4 Concluding Remarks

We have carried out QM/MM and MD simulations of the dielectric response from the chromophore to the protein and solvent environment following photoexcitation of PYP, as well as the response of water to photoexcitation of the bare chromophore in aqueous solution for comparison. Decomposition of the dielectric response indicates an important role of the protein microenvironment in slow relaxation. This finding is in agreement with a number of studies of the dielectric response of a solvated biomolecule to photoexcitation of a chromophore, where the slow response has been found largely due to the biomolecule (12). A somewhat comprehensive picture to understand TDSS in photoreceptor proteins can be built upon when combining findings of the TDSS in GFP, where there are about ten tightly bound water molecules in the beta barrel of GFP that surround the chromophore. It seems to us that it is the hydrogen bonding in the microenvironment of the chromophore, regardless whether this constitutes confined water molecules and/or residues, which makes the main contribution accounting for the slow dielectric response to the photoexcitation of the chromophore.

6.5 Acknowledgements

Support from NSF CHE-0910669 is gratefully acknowledged.

Table 6-1 Parameters used to fit Eq. (6-3) to the computed dielectric response for the chromophore in solvated PYP (Fig. 6-2), and the chromophore in bulk water (Fig. 6-3).

Interactions	a_1	τ_1/ps	a_2	τ_2/ps	a_3	τ_3/ps
Chromophore in solvated PYP (MD)	0.74	0.04	0.21	2.61	0.05	168.56
Chromophore in solvated PYP (QM/MM)	0.51	0.06	0.47	1.85	0.02	108.91
Chromophore in bulk water (MD)	0.69	0.04	0.14	0.10	0.21	0.28
Chromophore in bulk water (QM/MM)	0.52	0.06	0.49	1.69	0.01	9759.43

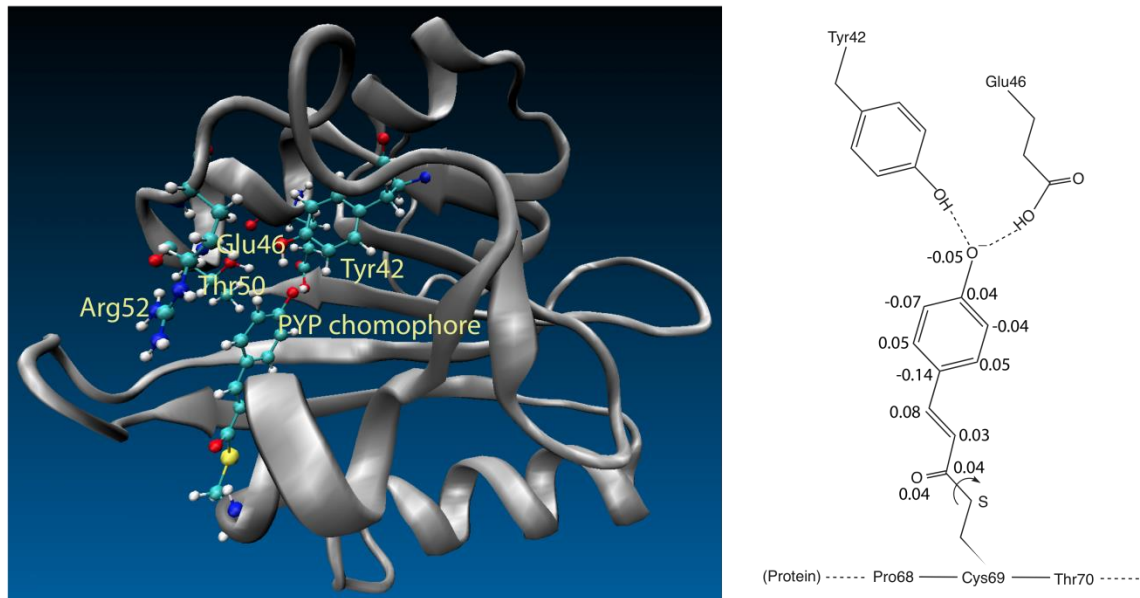


Figure 6-1 Ribbon diagram of PYP (left) showing all atoms of the chromophore, which is attached to Cys69, and nearby residues Glu46, Tyr42, to which it hydrogen bonds, and Thr50, which also appears close the chromophore anion during MD and QM/MM simulation. Arg52, which also lies near the chromophore, is also shown. The hydrogen bonding between the chromophore and protein residues is indicated on the right, as is the ground state charge minus the excited state charge on chromophore atoms, in units of electron charge.

PYP TDSS 300 K

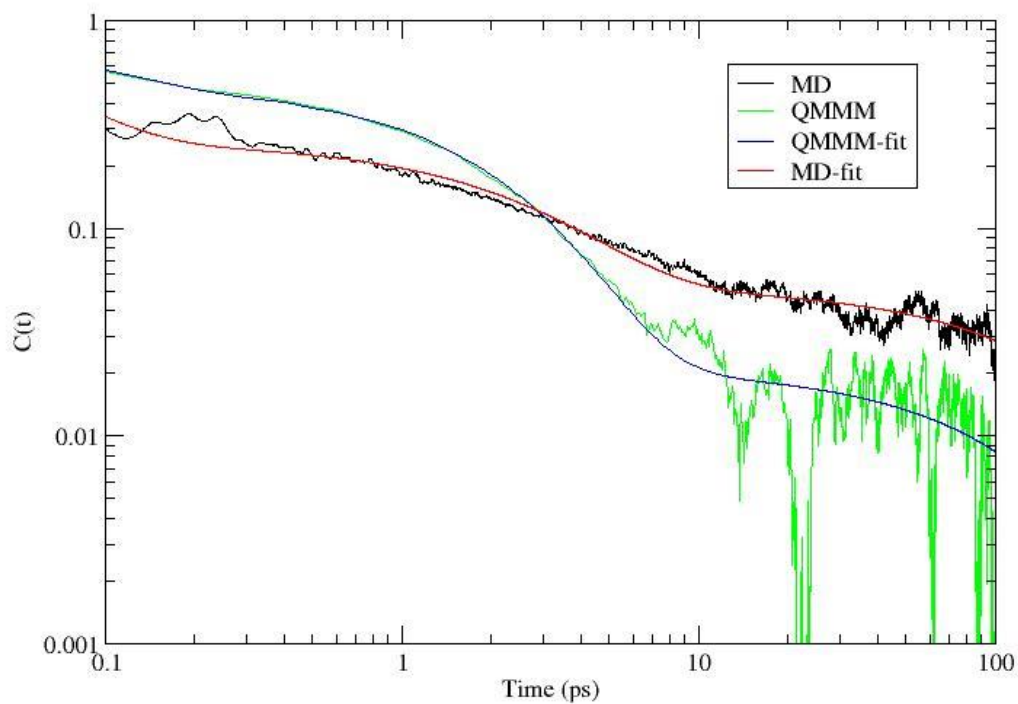


Figure 6-2 The dielectric response of the protein and solvent to photoexcitation of the PYP chromophore, calculated using the interaction energy, E , between the chromophore and the rest of the system (thick solid curves, black for MD and red for QM/MM). The two results both fit well, other than some oscillations, to a tri-exponential, Eq. (6-3) (thin curves, blue for MD and green for QM/MM).

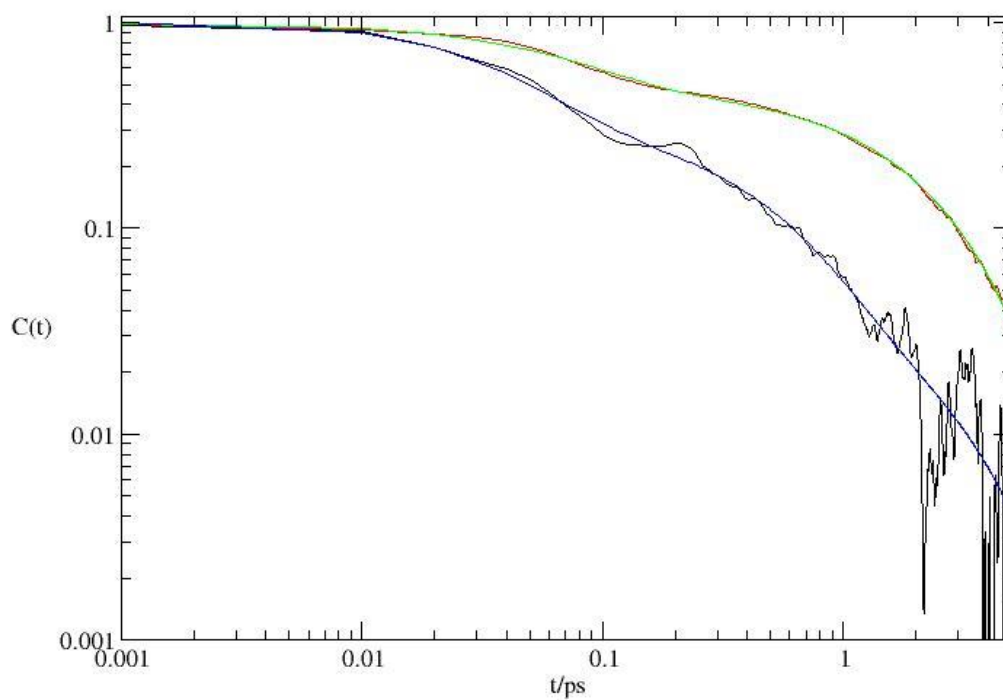


Figure 6-3 Normalized equilibrium time autocorrelation functions of fluctuations in the solvation energy differences at 300 K (solid curves) for the interactions between the chromophore and bulk water (black for MD and red for QM/MM). Blue and green curves are fits to the MD and QM/MM data, respectively, using Eq. (6-3). The parameters for the fits are listed in Table 6-1.

6.6 References

- (1) Q Wang, R W Schoenlein, L A Peteanu, R A Mathies, C V Shank, *Science* 266 (1994) 422 - 24.
- (2) N Mataga, H Chosrowjan, S Taniguchi, N Hamada, F Tokunaga, Y Imamoto, M Kataoka, *Chem. Phys. Phys. Chem.* 5 (2003) 2454 - 60.
- (3) N Mataga, H Chosrowjan, Y Shibata, Y Imamoto, M Kataoka, F Tokunaga, *Chem. Phys. Lett.* 352 (2002) 220 - 25.
- (4) C Fang, R R Frontiera, R Tran, R A Mathies, *Nature* 462 (2009) 200 - 05.
- (5) R Nakamura, N Hamada, H Ichida, F Tokunaga, Y Kanematsu, *J. Chem. Phys.* 127 (2007) art. no. 215102, pp. 1 - 7.
- (6) H Chosrowjan, S Taniguchi, N Mataga, M Unno, S Yamauchi, N Hamada, M Kumauchi, F Tounaga, *J. Phys. Chem. B* 108 (2004) 2686 - 98.
- (7) K J Hellingwerf, J Hendriks, T Gensch, *J. Phys. Chem. A* 107 (2003) 1082 - 94.
- (8) M Mizuno, H Kamikubo, M Kataoka, Y Mizutani, *J. Phys. Chem. B* 115 (2011) 9306 - 10.
- (9) P Changuenet-Barret, P Plaza, M M Martin, H Chosrowjan, S Taniguchi, N Mataga, Y Imamoto, M Kataoka, *J. Phys. Chem. C* 113 (2009) 11605 - 13.
- (10) M Vengris, M A v d Horst, G Zgrablic, I H M v Stokkum, S Haacke, M Chergui, K J Hellingwerf, R v Grondelle, D S Larsen, *Biophys. J.* 87 (2004) 1848-57.
- (11) R Jimenez, G R Fleming, P V Kumar, M Maroncelli, *Nature* 369 (1994) 471 - 72.
- (12) B Bagchi, *Chem. Rev.* 105 (2005) 3197-219.
- (13) K E Furse, S A Corcelli, *J. Phys. Chem. Lett.* 1 (2010) 1813-20.
- (14) A A Golosov, M Karplus, *J. Phys. Chem. B* 111 (2007) 1482-90.
- (15) T Li, A A Hassanali, S J Singer, *J. Phys. Chem. B* 112 (2008) 16121-34.
- (16) D V Matyushov, DOI: 10.1021/jp200409z (2011) 1 - 10.
- (17) D Xu, C Martin, K Schulten, *Biophys. J.* 70 (1996) 453 - 60.
- (18) C. M. Isbon, A. W. Gotz, M. A. Clark, R. C. Walker, T. J. Martinez, *J. Chem. Theory Comput.* 8 (2012) 5092 - 5106.
- (19) D v d Spoel, E Lindahl, B Hess, G Groenhof, A. E. Mark, H J C Berendsen, *J. Comp. Chem.* 26 (2005) 1701 - 18.
- (20) Y Duan, C Wu, S Chowdhury, M C Lee, G Xiong, W Zhang, R Yang, P Cieplak, R Luo, T Lee, J Caldwell, J Wang, P Kollman, *J. Comp. Chem.* 24 (2003) 1999 - 2012.
- (21) W L Jorgensen, J Chandrasekhar, J D Madura, R W Impey, M L Klein, *The Journal of Chemical Physics* 79 (1983) 926-35.
- (22) M J Frisch, G W Trucks, H B Schlegel, G E Scuseria, M A Robb, J R Cheeseman, J A Montgomery, T Vreven, K N Kudin, J C Burant, S S I J. M. Millam, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg,

- V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez, and J. A. Pople, Gaussian, Inc., Wallingford CT. (2004).
- (23) X Yu, D M Leitner, Chem. Phys. Lett. 391 (2004) 181 - 86.
- (24) G G Maisuradze, X Yu, D M Leitner, J. Biolog. Phys. Chem 7 (2007) 25 - 29.
- (25) A. Jha, K. Ishii, J.B. Udgaonkar, T. Tahara, G. Krishnamoorthy, Biochemistry 50 (2011) 397-408.

7 Communication Maps and Non-equilibrium Simulation Computed for Green Fluorescent Protein and Photoactive Yellow Protein: Studies of Energy Transport in Proteins

Abstract

Frequency-resolved communication maps provide a coarse-grained picture of energy transport in nanoscale systems. We calculate communication maps for Green Fluorescent Protein (GFP) from the jellyfish *Aequorea victoria* and Photoactive Yellow Protein (PYP) from the purple phototropic eubacterium *Halorhodospira halophile*, and sample them to elucidate energy transfer pathways between the chromophore and other parts of the protein. In GFP we focus on the role of the cluster of water molecules confined within the beta barrel that encapsulates the chromophore. We complement analysis of communication maps with molecular simulations of energy flow. Both approaches reveal that in PYP, excess energy from the chromophore flows mainly to regions of the surrounding residues that hydrogen bond to the chromophore. In GFP, energy is carried disproportionately by some of the confined water molecules, consistent with the larger thermal conductivity of water compared to proteins.

Keywords: Communication Map, Energy Flow, Green Fluorescent Protein, Photoactive Yellow Protein.

7.1 Introduction

Properties of water surrounding proteins and its role in protein dynamics, thermodynamics and function have been closely examined in recent years.¹⁻¹⁶ Water molecules confined inside proteins, far less mobile than bulk or hydration water,¹⁷⁻¹⁹ also contribute to function in a variety of ways, including charge transfer, binding and allostery.^{20-28, 30-36, 38-45} In this Chapter, we examine communication between the chromophores of GFP and PYP and their microenvironment by vibrational energy flow with focus on energy transport channels between the chromophore and the contribution of the confined water in the beta-barrel in GFP. We adopt and further develop a coarse-graining procedure, communication maps,²⁹ to locate the energy transport channels in GFP and PYP. We complement that study with molecular simulations of vibrational energy transport from the chromophores to other parts of the protein. Analysis of the communication maps and the molecular simulations reveals that in GFP, vibrational energy transport occurs through both the confined waters and side chains that hydrogen bond with the chromophore; in PYP, the energy is mainly transported by neighboring residues that hydrogen bond to the chromophore.

Here we explore the specific pathways that enable vibrational signaling in this protein by adopting a coarse-graining procedure, frequency-resolved communication maps,²⁹ which we recently introduced. This is one of a number of methods developed in the past few years to locate energy transport channels in biomolecules.^{37, 46-55} Frequency-resolved communication maps, like some of these methods,^{50, 51} provides information about how specific vibrational modes contribute to pathways for energy transport. It is a coarse graining procedure that in the limit of the size of the biomolecule itself goes over

to the frequency resolved energy diffusion coefficients from which thermal transport coefficients can be determined, as described in Sec. II. In this study we thermally average the frequency resolved maps to compute temperature-dependent communication maps, and we sample the communication maps to find pathways for energy transport beginning in a certain part of the protein. With this analysis we find that in GFP, vibrational energy transport from the chromophore to its neighboring residues through the confined water molecules in the beta barrel. Communication between residues and water, and between residues close to the chromophore, occurs mainly via vibrations with frequency below 200 cm^{-1} . In PYP, the chromophore interacts with some surrounding residues at higher frequency, specifically Glu46 and Tyr42. In addition to the communication maps we carry out molecular simulations to examine the flow of energy from hot chromophores of GFP and PYP to their surroundings.

In the following section we summarize calculation of frequency-dependent communication maps and an approach to sample the maps to locate energy transport channels. We also discuss our method for molecular simulations to study vibrational energy transport from a chromophore with excess energy to other parts of the protein. In Sec. III we present and discuss our results. Concluding remarks are given in Section IV.

7.2 Methods

A. Molecular dynamics simulations and normal mode analysis

The initial GFP (1EMA) and PYP (1OTB) structures (in Fig. 7-1) were obtained from the Protein Data Bank (PDB). This structure was introduced into a 50 \AA^3 cubic box of equilibrated TIP3P water model⁵⁶ and simulated using AMBER force field in GROMACS program⁵⁷. The entire GFP system was composed of 3603 atoms, of which 3573 belong to GFP and the rest to the 10 confined water molecules. And the PYP protein has 1929 atoms. The particle-mesh Ewald (PME) method⁵⁸ was employed to account for long-range electrostatic effects and a constant pressure of 1.0 atm was maintained. The SHAKE algorithm was employed to constrain all bonds. Periodic boundary conditions were used with standard cutoffs. The excess potential energy due to bad contacts and strain was then reduced using the Adopted Basis Newton-Raphson (ABNR) minimization method for about 1000 steps so that the norm of the gradient was 5.0 kcal/mol. The molecular dynamics simulation was carried out using the Verlet algorithm. The temperature was increased slowly to 300 K in intervals of 10 K for about 6 ps. Equilibration dynamics followed for about 100 ps by keeping the temperature constant at 300 K with a time step of 1 fs. Both systems were minimized using the ABNR method for about 50,000 steps, which satisfy the norm of the gradient 1×10^{-6} kcal/mol. Normal mode analysis was then carried out on the systems.

B. Frequency-resolved local energy diffusivities and communication maps

We recently introduced frequency resolved communication maps²⁹ and in this subsection we summarize the method, which is based on a coarse-graining of the vibrational mode diffusivity. The coefficient of thermal conductivity, which relates the net energy flux to the thermal gradient, is expressed in terms of the mode diffusivity, D_α , of mode α with frequency ω_α as

$$\kappa = \sum_{\alpha} C_{\alpha} D_{\alpha}, \quad (7-1)$$

where C_α is the heat capacity of mode α per unit volume. The mode diffusivity can be computed by propagation of vibrational wave packets comprised of modes filtered around frequency ω_α .^{35, 59-61} D_α can also be expressed in terms of the heat current operator, \mathbf{S} , in the harmonic approximation as follows:^{62, 63} The local energy density, $h(x)$, for instance the energy density at an amino acid or cluster of water molecules, is obtained by summing over all atoms, l , in this region, A , so that

$$h(x) = \sum_{l \in A} h_l. \quad (7-2)$$

The condition of local energy conservation is

$$\frac{\partial h(x)}{\partial t} + \nabla \cdot \mathbf{S}(x) = 0. \quad (7-3)$$

The total heat current operator for a particular region, AA' , of the object of interest is

$$\mathbf{S}^{\{AA'\}} = \frac{1}{V} \int_{x \in A, A'} d^3x \mathbf{S}(x). \quad (7-4)$$

The region AA' could encompass the whole molecule but for most of our applications here will denote pairs of residues or a residue and water cluster. In the harmonic approximation, the heat current operator can be written as^{62, 64}

$$\mathbf{S} = \sum_{\alpha, \beta} \mathbf{S}_{\alpha\beta} a_{\alpha}^{\dagger} a_{\beta}. \quad (7-5)$$

where α and β are two modes of the protein. The coefficient, $\mathbf{S}_{\alpha\beta}$, corresponding to the protein as a whole can be expressed in terms of the Hessian matrix, \mathbf{H} , and eigenmodes, \mathbf{e} , of the object.⁶² We break the coefficient up into contributions from various regions. For the region, AA' , it is

$$\mathbf{S}_{\alpha\beta}^{\{AA'\}} = \frac{i\hbar(\omega_{\alpha} + \omega_{\beta})}{4V\sqrt{\omega_{\alpha}\omega_{\beta}}} \sum_{r, r' \in (x, y, z)} \sum_{l, l' \in AA'} e_l^{\alpha} H_{rr'}^{ll'}(\mathbf{R}_l - \mathbf{R}_{l'}) e_{l'}^{\beta}, \quad (7-6)$$

where \mathbf{R}_l is the position of atom l and r is a coordinate (x , y or z). When the region AA' spans the protein, Eq. (7-6) expresses the matrix elements of the heat current operator for the whole system in harmonic approximation given in Ref. [62]. Similarly, we can sum over all regions to obtain $\mathbf{S}_{\alpha\beta}$ for the whole molecule. Considering only the local region AA' we write the local energy diffusivity in mode α in harmonic approximation as

$$D_{\alpha}^{\{AA'\}} = \frac{\pi V^2}{3\hbar^2 \omega_{\alpha}^2} \sum_{\beta \neq \alpha} |\mathbf{S}_{\alpha\beta}^{\{AA'\}}|^2 \delta(\omega_{\alpha} - \omega_{\beta}). \quad (7)$$

When the local region AA' spans the molecule Eq. (7-7) gives the mode diffusivity that appears in Eq. (7-1) for the coefficient of thermal conductivity. For a practical calculation on a finite-sized system we substitute a Lorentzian of width η for the delta function. The width of the Lorentzian should be large enough to envelop several vibrational modes, enough so that the mode diffusivity converges. For the results presented here we

calculated the mode diffusivity using $\eta = 15 \text{ cm}^{-1}$, which is large enough to include many modes in the averaging; results for D_α did not change significantly with larger η . Communication maps can then be constructed by plotting $D_\alpha^{\{AA'\}}$ for all regions AA' at frequencies, ω_α .

Below we use frequency-resolved communication maps to locate energy transport channels that include a particular site on the protein, in this case the chromophore. The procedure can be carried out for a particular frequency or temperature. We carried out a thermal average over the frequency-resolved communication maps for the structures at 300 K, using for the thermal averaging the communication maps constructed at 50 cm^{-1} , 100 cm^{-1} , and continuing in 50 cm^{-1} intervals until 400 cm^{-1} , assigning a Boltzmann weight to the frequency resolved communication maps. Each of these frequency-resolved maps is actually an average over maps for 4 modes closest in frequency to the designated frequency and an average over four protein structures. Our emphasis was on the early steps of energy transfer from the chromophore. For this purpose, we randomly selected a residue of the protein or water molecule in the interface. If x is the distance between the center of mass of the chromophore and the center of mass of a residue, we accepted a transition if a random number, r , between 0 and 1 was smaller than $(D_r^{\{AA'\}}/x^2)/(D_r^{\{AA'\}}/x^2)_{\max}$, where $D_r^{\{AA'\}}$ is the local energy diffusivity between the chromophore (A) and the residue (A') obtained after the thermal average at temperature T , and $(D_r^{\{AA'\}}/x^2)_{\max}$ is the maximum of this ratio between the chromophore and all residues.

C. Energy transport

To complement the analysis of energy transport using communication maps we also studied vibrational energy flow by molecular simulation, with focus on energy transport from a chromophore. We in effect heated the chromophore to 300 K and followed the transfer of energy over the next few picoseconds to the rest of the molecule and water, all of which was initially at 0 K. Similar starting conditions have been applied in earlier molecular simulations to elucidate energy transport pathways in proteins.⁴⁶ We carried out this analysis in harmonic approximation to provide a realistic representation of the thermal population of the vibrational modes of the protein.

We propagated the vibrational wave packets as follows: Consider the kinetic energy, $E_n(t)$, of atom n at time t . We start with a quenched structure and introduce an excitation in the form of a wave packet. The center of kinetic energy, $\mathbf{R}_0(t)$, and variance, $\langle \mathbf{R}^2(t) \rangle$, are

$$\mathbf{R}_0(t) = \frac{\sum_n \mathbf{R}_n E_n(t)}{\sum_n E_n(t)}, \quad (7-8a)$$

$$\langle \mathbf{R}^2(t) \rangle = \frac{\sum_n (\mathbf{R}_n - \mathbf{R}_0(t))^2 E_n(t)}{\sum_n E_n(t)}. \quad (7-8b)$$

We propagate a wave packet expressed as a superposition of the normal modes, starting with an initial wave packet taken to be a traveling wave, where the displacement of atom n initially has the form

$$\mathbf{U}_n(t) = \mathbf{B}_n \exp\left(-\frac{(\mathbf{R}_n - \mathbf{R}' - \mathbf{v}_0 t)^2}{2g^2}\right) e^{i(\mathbf{Q}_0 \cdot \mathbf{R}_n - \omega_0 t)}, \quad (7-9)$$

where \mathbf{R}' is the position of the atom at the center of the wave packet, \mathbf{v}_0 is the initial velocity, g is the width, \mathbf{Q}_0 is the wave vector of the initial excitation, ω_0 is the central frequency of the initial excitation, and \mathbf{B}_n is the amplitude. With this initial wave packet displacements and velocities at $t = 0$ are determined. Specific values for the variables used in the calculations are given below Eq. (7-11). Displacements and velocities for $t > 0$ are then computed in terms of normal modes as^{59-61, 65, 66}

$$\mathbf{U}_n(t) = \sum_{\alpha} \mathbf{e}_n^{\alpha} \cos(\omega_{\alpha} t) \sum_{n'} \mathbf{e}_{n'}^{\alpha} \cdot \mathbf{U}_{n'}(0) f_{\omega_{\alpha}} + \sum_{\alpha} \mathbf{e}_n^{\alpha} \frac{\sin(\omega_{\alpha} t)}{\omega_{\alpha}} \sum_{n'} \mathbf{e}_{n'}^{\alpha} \cdot \mathbf{V}_{n'}(0) f_{\omega_{\alpha}}, \quad (7-10a)$$

$$\mathbf{V}_n(t) = \sum_{\alpha} \mathbf{e}_n^{\alpha} \sin(\omega_{\alpha} t) \sum_{n'} \mathbf{e}_{n'}^{\alpha} \cdot \mathbf{V}_{n'}(0) f_{\omega_{\alpha}} - \sum_{\alpha} \mathbf{e}_n^{\alpha} \omega_{\alpha} \cos(\omega_{\alpha} t) \sum_{n'} \mathbf{e}_{n'}^{\alpha} \cdot \mathbf{U}_{n'}(0) f_{\omega_{\alpha}}. \quad (7-10b)$$

In Eq. (7-10) we include a coefficient, f_{ω} , to filter the wave packet in frequency around $\bar{\omega} \pm \delta\omega$, where we use

$$f_{\omega} = \exp\left(-(\omega - \bar{\omega})^2 / 2\delta\omega^2\right). \quad (7-11)$$

The thermal energy flow from the chromophore to the rest of the system was carried out as follows: The initial wavepacket, expressed as a superposition of the normal modes, was centered on the atom that is closest to the center of the chromophore. We use a frequency filter whose width, $\delta\omega$, is 50 cm^{-1} , and took the central frequencies to be 10 cm^{-1} , 50 cm^{-1} , 100 cm^{-1} , and continuing in 50 cm^{-1} interval until 400 cm^{-1} , which was high enough in frequency to obtain converged results in thermal averaging for temperatures to 300 K. The width of the initial wave packet is $g = 3 \text{ \AA}$. The magnitude of the wave vector of the initial excitation, Q_0 , is 0.63 \AA^{-1} and it points $+45^\circ$ from the x-, y- and z-axis; we take $\omega_0 = 9.4 \text{ ps}^{-1}$, and $\mathbf{v}_0 = 20 \text{ \AA ps}^{-1}$, which is reasonably close to the

speed of sound in proteins. We checked that our results did not vary significantly with modest changes in these initial conditions. All components of \mathbf{B}_n for all atoms are taken to be the same, and the magnitude is unimportant as it cancels out when we compute the center of energy and its variance.

7.3 Results and Discussion

A. Communication maps

We have computed frequency resolved communication maps for GFP and PYP. The thermal average results for modes of GFP and PYP from 50 cm^{-1} to 400 cm^{-1} are plotted in Fig. 7-2. The communication maps plotted are averages over the 4 modes closest in frequency to the indicated frequency, as described in the previous section. For ground state GFP, we also plot in Fig. 7-2 a distance map, showing residues for which the centers of mass lie within 6, 8 and 10 \AA of each other. There are several observations we make from these plots and the others we have computed. In going from lower to higher frequencies the communication maps become sparser, consistent with the typically smaller participation of residues in the vibrations of the molecule as frequency increases.^{38, 43, 67, 68} Because vibrational modes of proteins become more localized to specific regions of the protein at higher frequency, even above 150 cm^{-1} ,^{38, 43} vibrational signaling via higher frequency modes is more local. We also observe that the chromophore communicates less with other parts of the molecule for frequencies above about 400 cm^{-1} , a trend we also observed in an earlier analysis of myoglobin.²⁹

There is qualitative similarity between the communication maps and the distance map, where we observe that pairs of residues between which energy flow is relatively facile tend to be relatively close to one another. Communication maps can be used to determine energy transport channels between specific sites on the protein. We illustrate this for communication maps of GFP and PYP at 300 K, which are obtained as a thermal average over the frequency-dependent maps in 50 cm^{-1} interval up to 400 cm^{-1} , above which the communication maps contribute very little to the thermal average due to the

small values of $D_{\alpha}^{\{AA'\}}$ at higher frequency. The communication maps for GFP and PYP at 300 K is plotted in Fig. 7-3 and 7-4, respectively. We sample the communication maps for GFP and PYP at 300 K to examine the early steps of energy transfer from the chromophore, both at ground and excited state. In GFP, most steps from the chromophore were taken to the confined water molecules, then to neighboring Thr61 and Val60, as seen in Fig. 7-5. Considering Thr61 and Val60, and the chromophore, which are involved in the detailed hydrogen bonding arrangement with the confined waters shown in Fig. 7-5, most of the response to “excess energy” in the chromophore occurs in these residues and in the waters. Sampling the communication map reveals that most of excess energy in the chromophore is funneled to the surrounding residues that hydrogen bond to the chromophore. In PYP, our results show that quite a few neighboring residues that hydrogen bond to the chromophore actively transport the excess energy in the chromophore at early times, namely, Tyr42, Glu46, Thr50, as seen in Fig. 7-6. Moreover, it is also quite noticeable that there is another energy transfer pathway that dissipates the excess energy effectively through Arg52, Tyr98 and Gln99, all of which do not hydrogen bond to the chromophore. All of the pathways highly agree with those discovered by Ishikura and Yamato⁴⁴, who similarly focused on the interresidue energy conductivity in terms of time-correlation function of interatomic energy flux, suggesting the significant roles of these residues in the long-range intramolecular signaling of PYP.

B. Vibrational energy transport

We turn now to molecular simulations of vibrational energy transport in GFP and PYP. Here, as in some earlier simulations aiming to elucidate pathways for vibrational energy flow in allosteric proteins,⁴⁶ we introduce “heat” locally in the otherwise “cold” molecule, and follow the flow of energy at later times. We carry this out in harmonic approximation, which yields realistic thermal populations of the vibrational modes, generating wave packets that initially displace atoms of the chromophore of one globule. As described in Section II.3, we filter the wave packets to a range of frequencies of width 50 cm^{-1} around the designated central frequency, a procedure that we carry out for frequencies up to 400 cm^{-1} . Information about energy transport at 300 K is obtained by subsequently taking a thermal average over the frequency-dependent wave packets. The temperature of the initially excited region of the protein, in this case the chromophore, is then 300 K while the rest of the protein is initially 0 K. A series of snapshots for one structure of GFP and PYP is shown in Figs. 7-7, 7-8, and 7-9 from 1 to 3 ps. In GFP, at 1 ps, we find that a large portion of the energy in the system is contained in Val60, Thr61 and the chromophore and the confined water molecules, and at 3 ps we still find much of the energy of the system in these residues. In PYP, it appears to us that the pathway mainly composed of Arg52, Tyr98 and Gln99 is dominant over others, at both ground and excited states. This indicates that our method is more sensitive to the long-range electrostatic interaction since Arg52 can form a counter ion to the phenolic oxygen anion of the chromophore.

As we observed energy flow directly and disproportionately to the interfacial water molecules from the “hot” chromophore within the first few picoseconds,

significantly more than to any single residue of the homodimeric hemoglobin in our previous study⁵⁹ by considering rescaling the energy to account for the greater mass of the waters, we did observe similar trend of the confined water in our GFP study here after considering the same energy rescaling, but merely in short time scales, suggesting a somewhat different microenvironment in the beta barrel where the confined waters lie and perhaps the dominant role of hydrogen bonding in the early events of energy transfer after excitation of the chromophore. We expect our conclusion concerning the role played by the confined water molecules in vibrational energy transport through the protein to be generally valid, and should be observable using a wide range of force fields.

7.4 Concluding remarks

Frequency-resolved communication maps provide a coarse-grained view of regions of a protein and water through which vibrational energy flow is facile as a function of the frequency of the vibrations that carry the energy. In this study we have calculated and examined communication maps for the Green Fluorescent Protein, GFP, from the jellyfish *Aequorea victoria*, a protein that exhibits bright green fluorescence when exposed to light in the blue to ultraviolet range and Photoactive Yellow Protein (PYP) from the purple phototropic eubacterium *Halorhodospira halophile*, another protein that probably functions as a photoreceptor for a negative phototaxis response. Interresidue and residue-water communication occurs mainly via vibrational modes below about 150 cm^{-1} . This spectral region, where the normal mode density of GFP and PYP, as other proteins is relatively large, is accessible experimentally by terahertz radiation,⁶ which could be adopted to facilitate communication between the globules. We have also sampled the communication maps and determine where excess energy flows from a particular part of the protein, in this case one of the chromophore groups of GFP and PYP. Recent time-resolved ultrafast spectroscopic studies indicate that fluorescence is triggered by the interaction between the cluster of confined water molecules inside the beta barrel, the chromophore side changes, and side chains of neighboring residues. Sampling the 300 K communication map for GFP and PYP we find that over most of the excess energy in the chromophore flows directly to these parts of the protein.

We have complemented the study of energy flow in GFP and PYP by communication maps with molecular simulations. We find that energy flows disproportionately to some of the confined water molecules during the first few

picoseconds. After accounting for the difference in mass between the water cluster and average amino acid, about twice as much energy flows to the waters as to residues a similar distance from the hot chromophore. This is consistent with the local energy diffusivity between the neighboring residues and confined waters that appears in the communication maps, where the values are about a factor of two larger than the local energy diffusivities between residues in the surrounding. These results reflect the larger thermal transport coefficients of water compared to protein molecules, in particular the thermal conductivity, which is more than twice as large as the thermal conductivity of proteins. A cluster of water molecules embedded in a biomolecule can therefore generally serve as a conduit for efficient energy transport.

7.5 Acknowledgments

Support from the National Science Foundation (NSF CHE-0910669) is gratefully acknowledged.

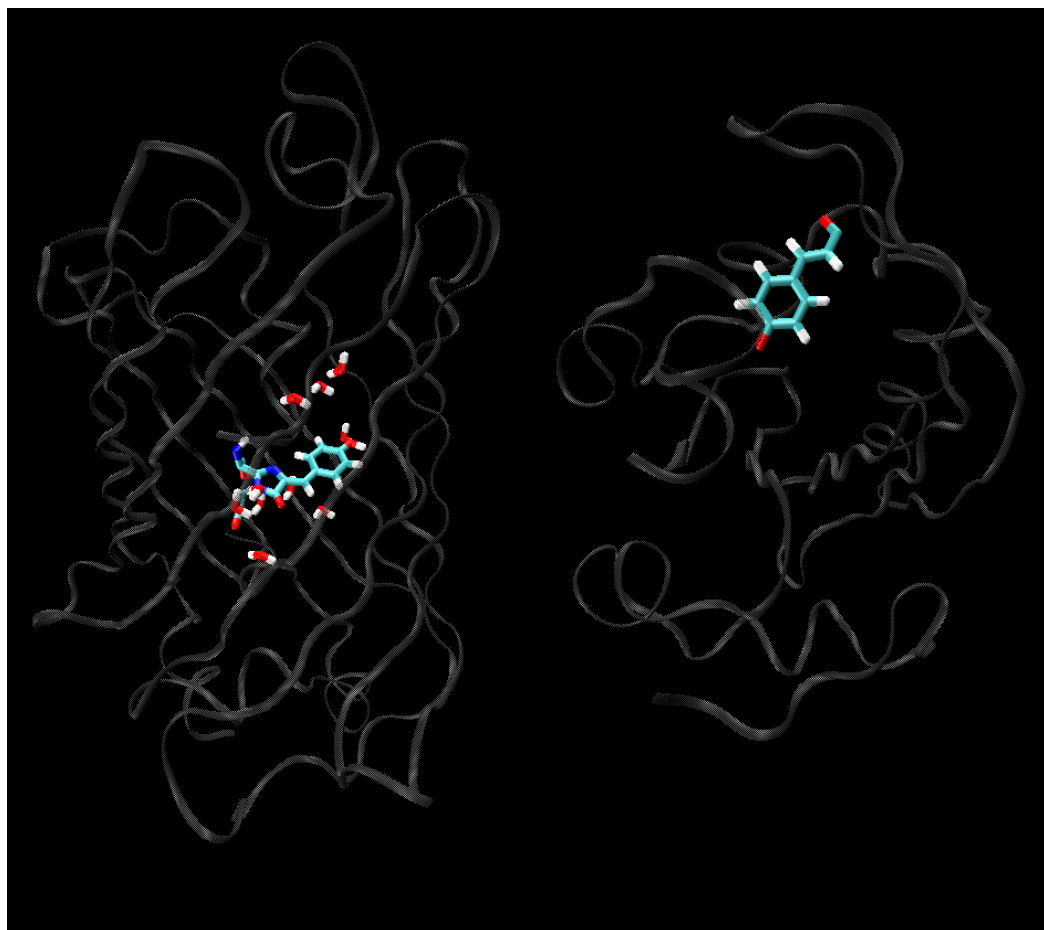


Figure 7-1 Ribbon diagrams of crystallographic structures of GFP (left) and PYP (right). The chromophores in both proteins are highlighted in licorice representation. The GFP chromophore is surrounded by 10 confined crystallographic water molecules.

GFP Distance Map 300 K

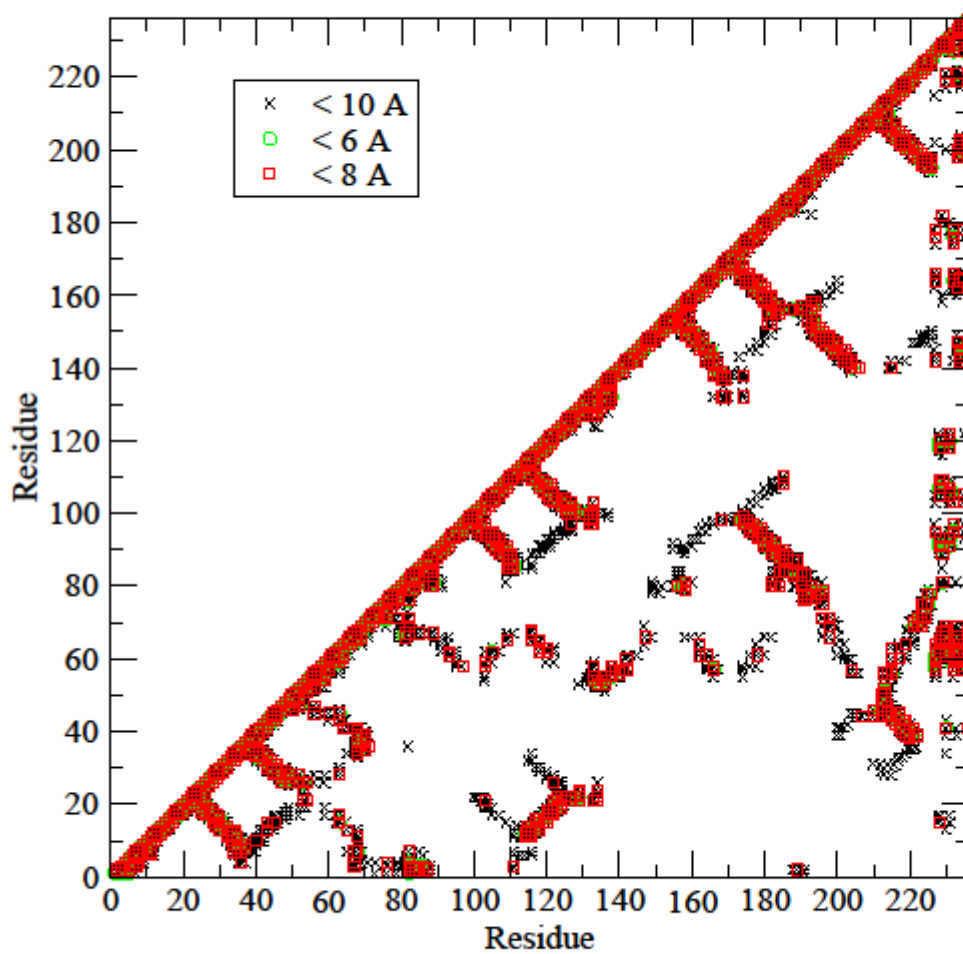
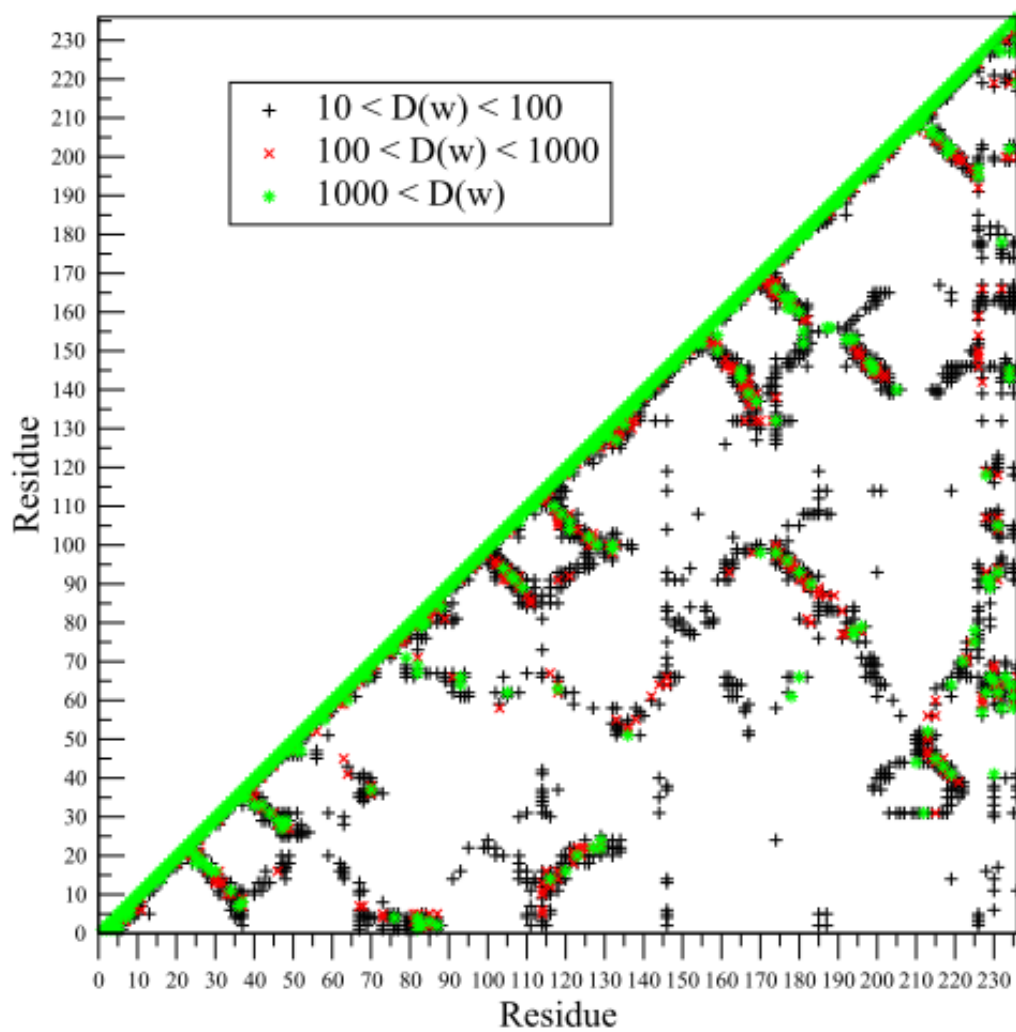


Figure 7-2 Thermal average distance maps are plotted for ground state GFP at 300 K. In GFP, residue 64 represents the chromophore, and the confined water molecules are represented as residues 227 – 236.

300 K Ground State GFP Communication Map



300 K Excited State GFP Communication Map

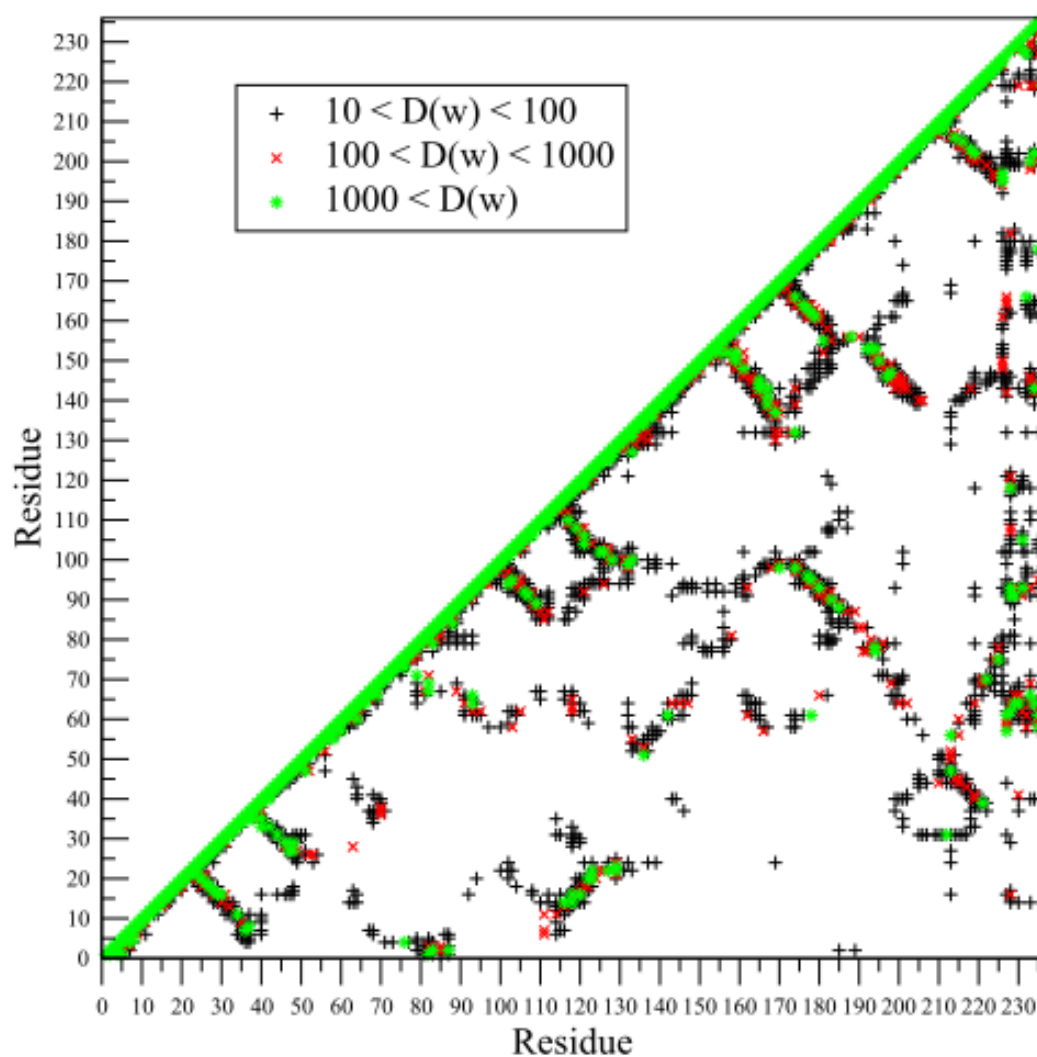
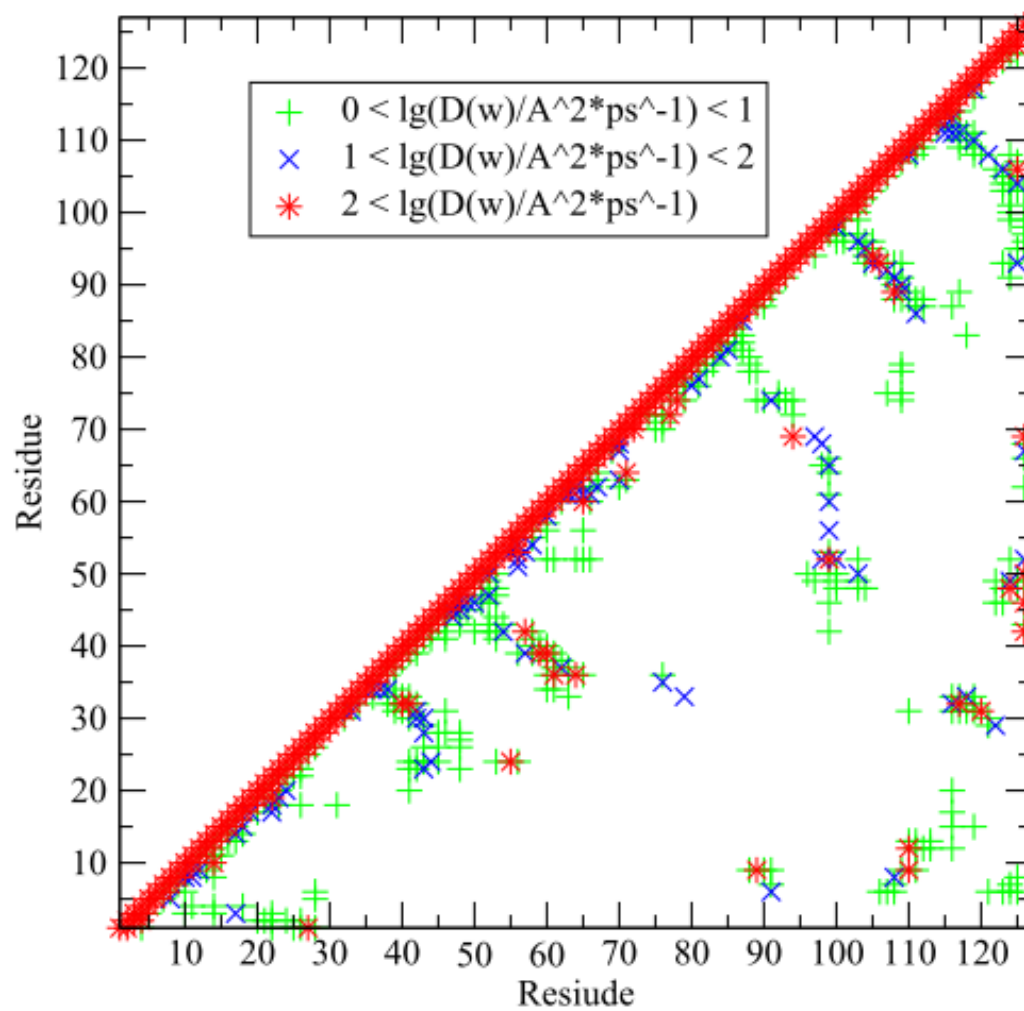


Figure 7-3 Thermal average communication maps are plotted for GFP at 300 K at ground and excited state, respectively. In GFP, residue 64 represents the chromophore, and the confined water molecules are represented as residues 227 – 236.

PYP Communication Map 300 K Ground State



PYP Communication Map 300 K Excited State

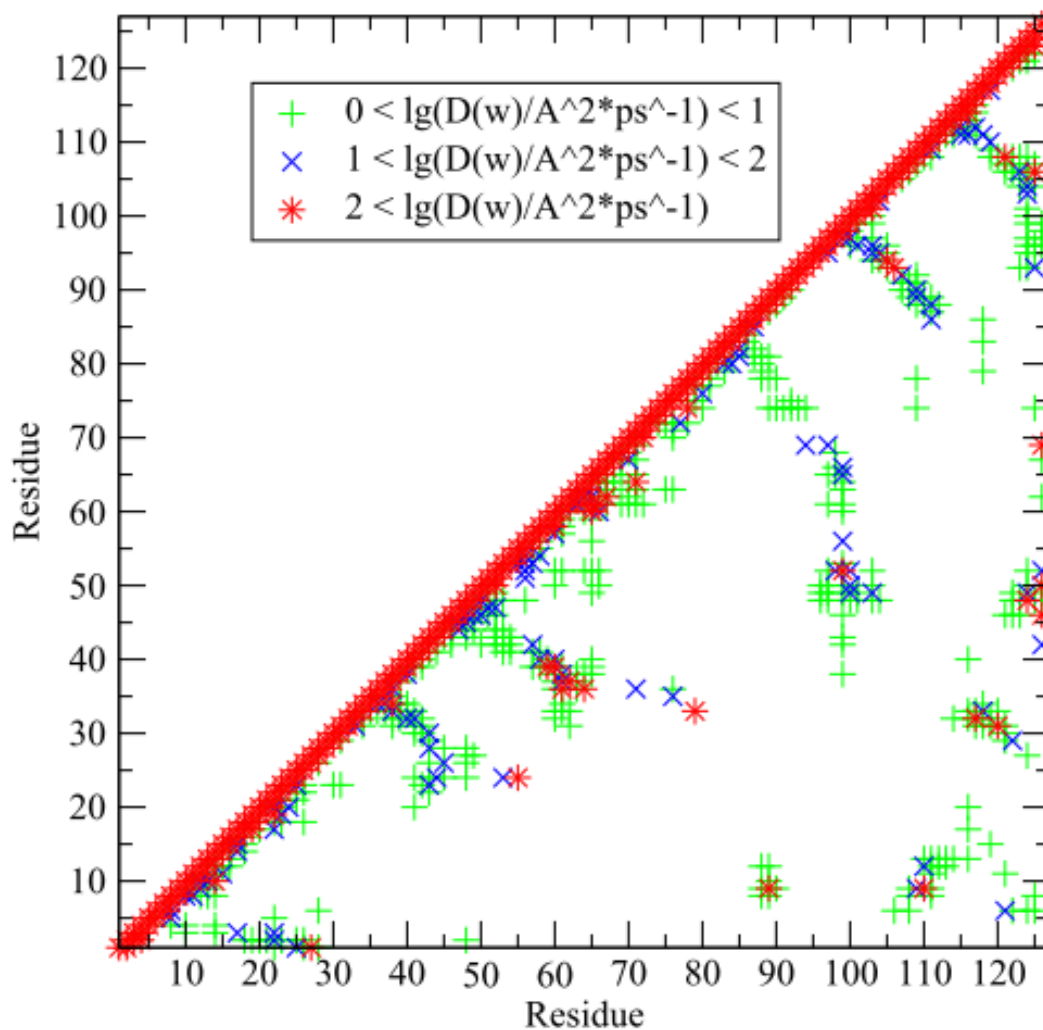
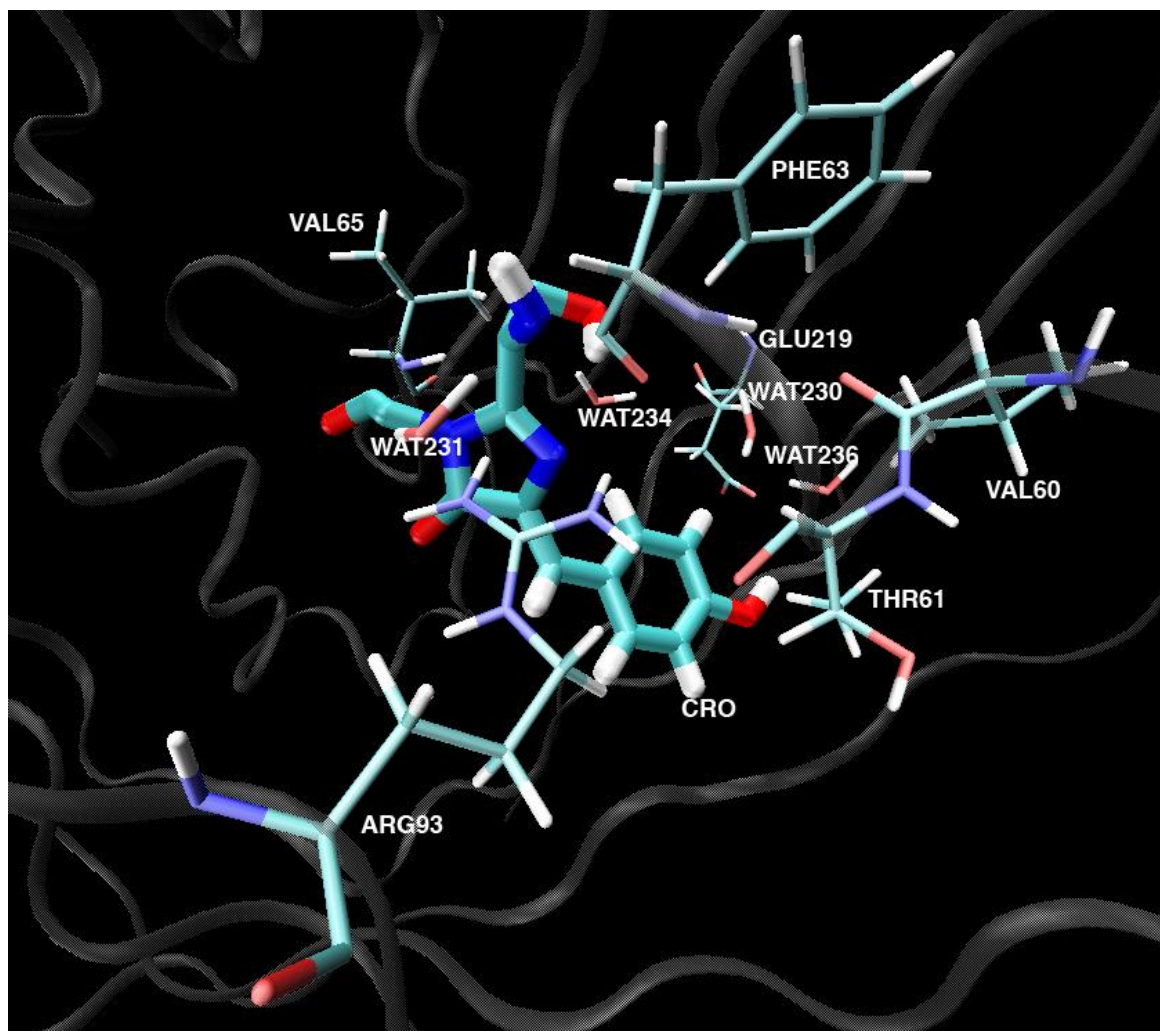


Figure 7-4 Thermal average communication maps are plotted for PYP at 300 K at ground and excited state, respectively. In PYP, the chromophore is represented as residue 126.



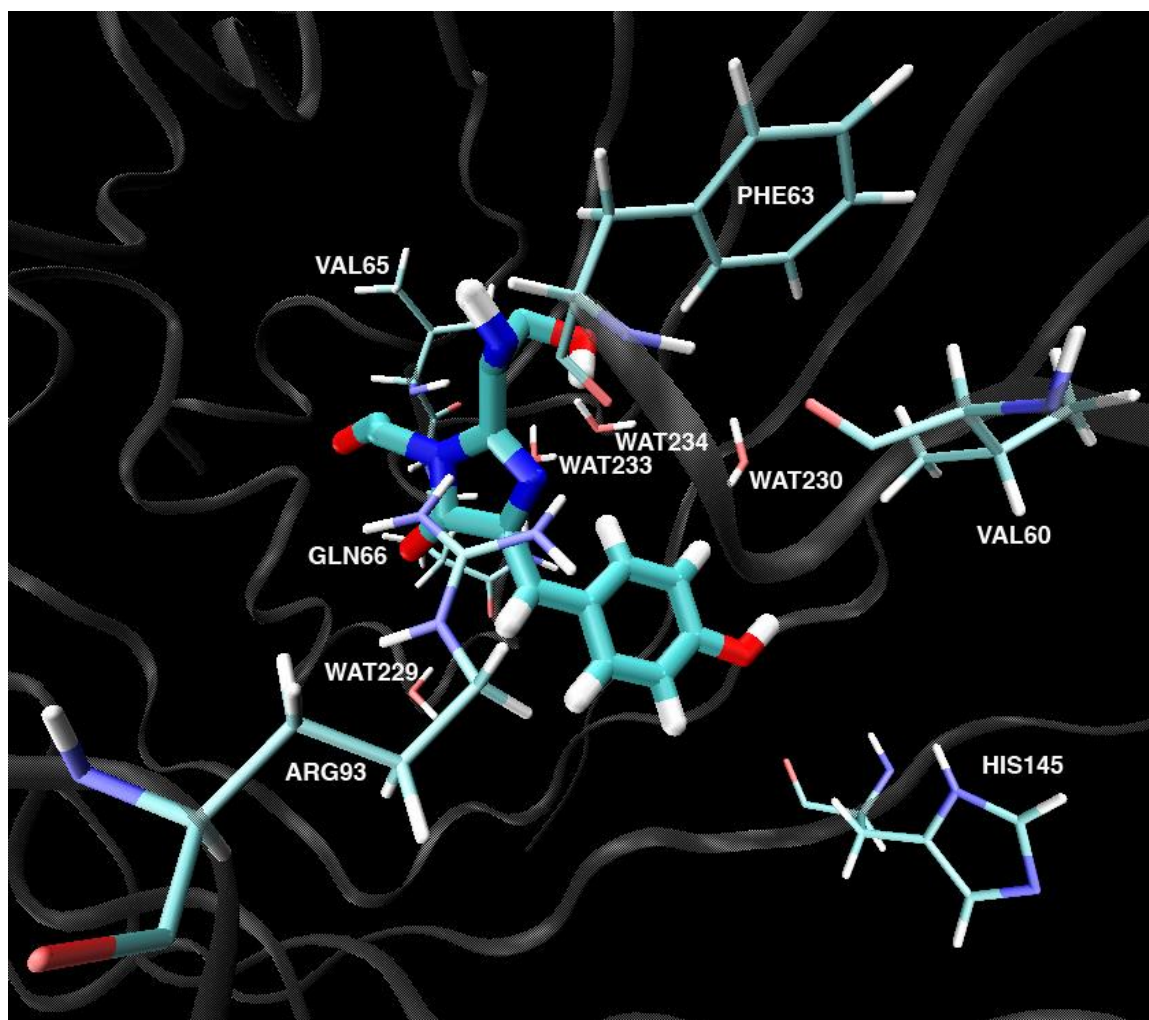
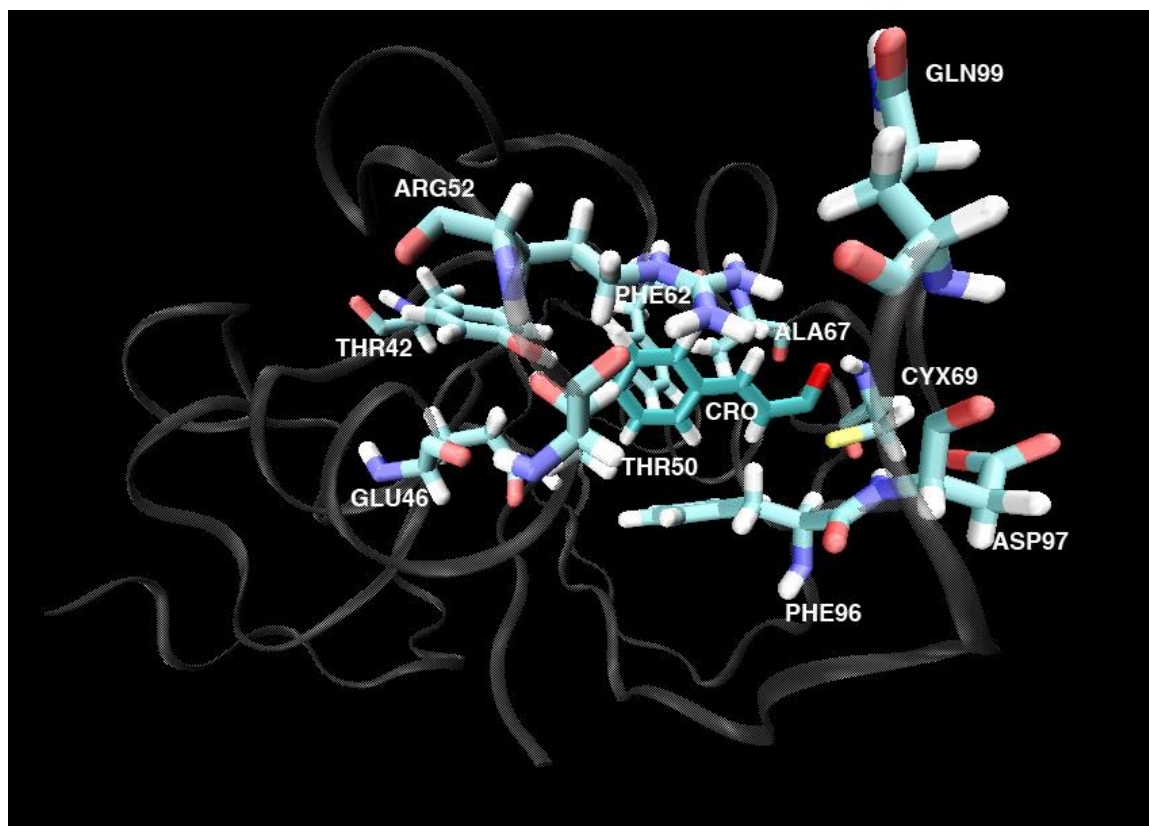


Figure 7-5 Top 10 hot residues interacting with the chromophore of GFP, obtained by frequency-dependent communication maps (top: ground state; bottom: excited state).



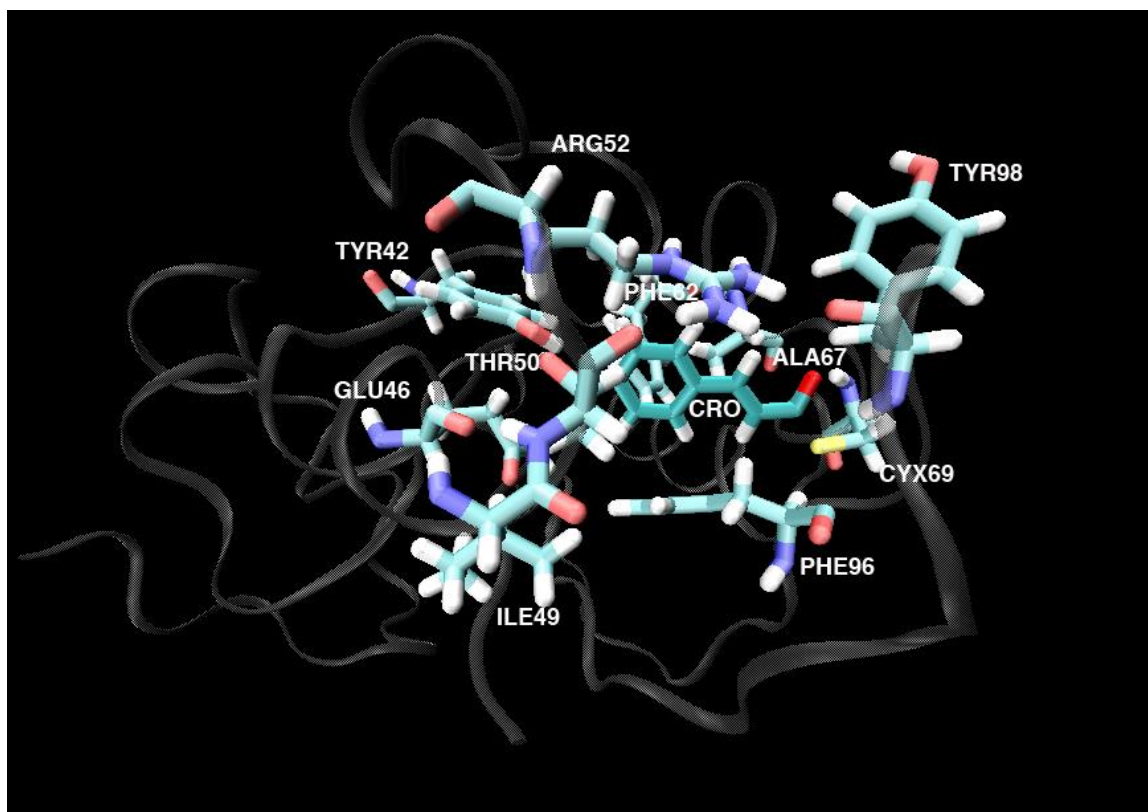
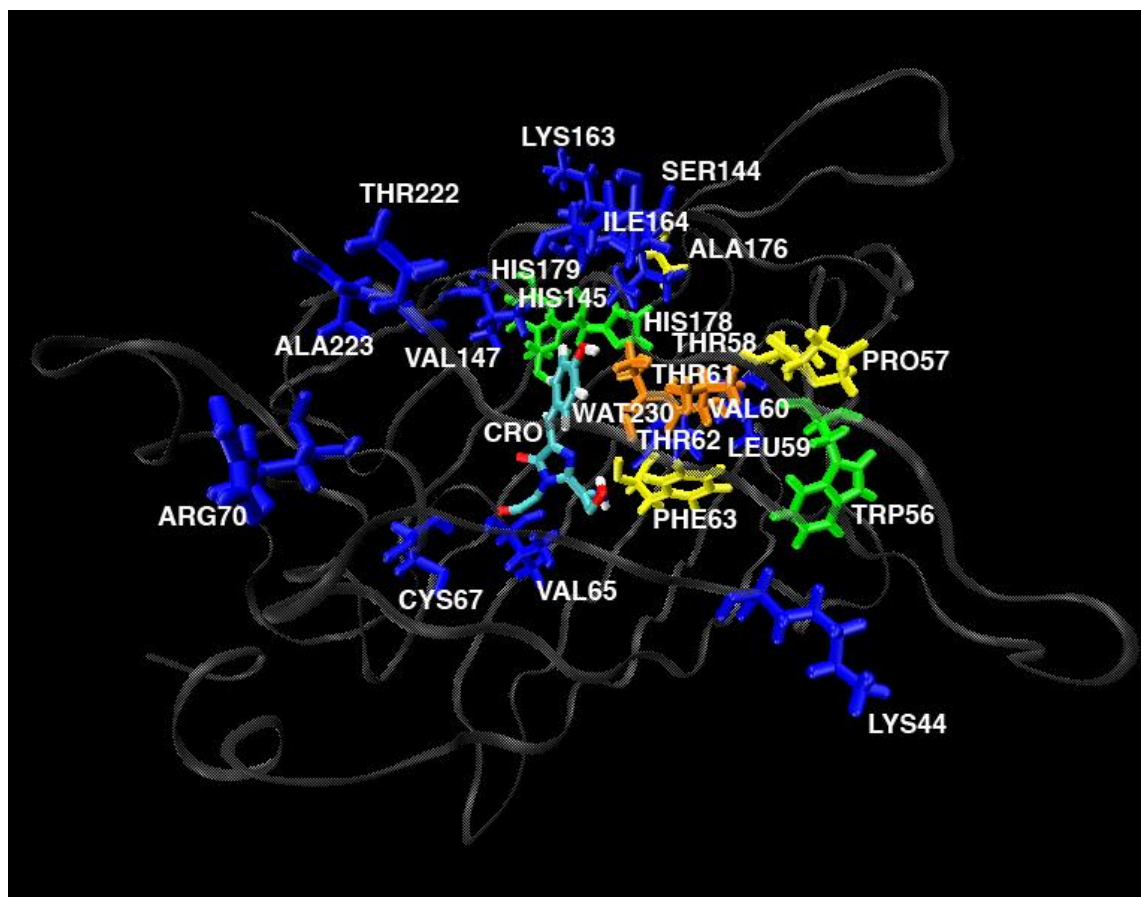
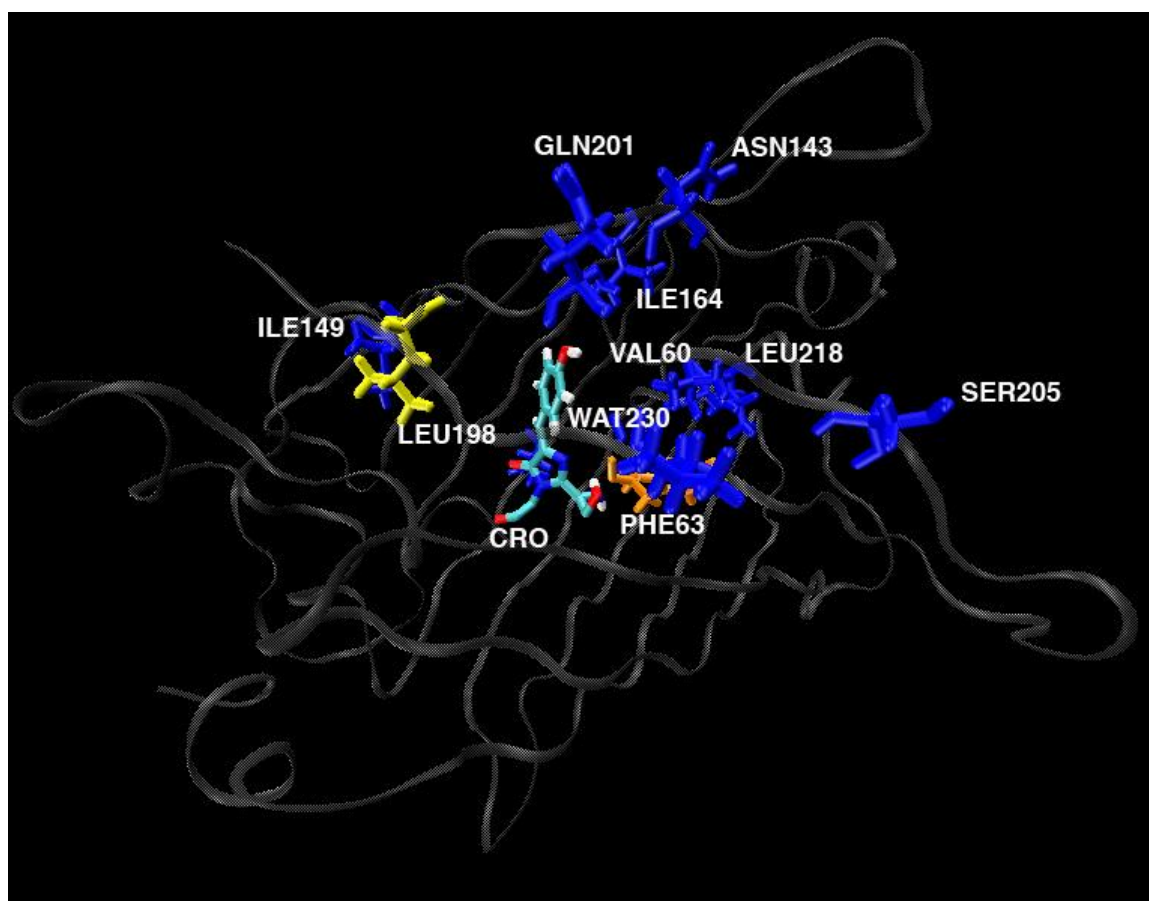


Figure 7-6 Top 10 hot residues interacting with the chromophore of PYP, obtained by frequency-dependent communication maps (top: ground state; bottom: excited state).





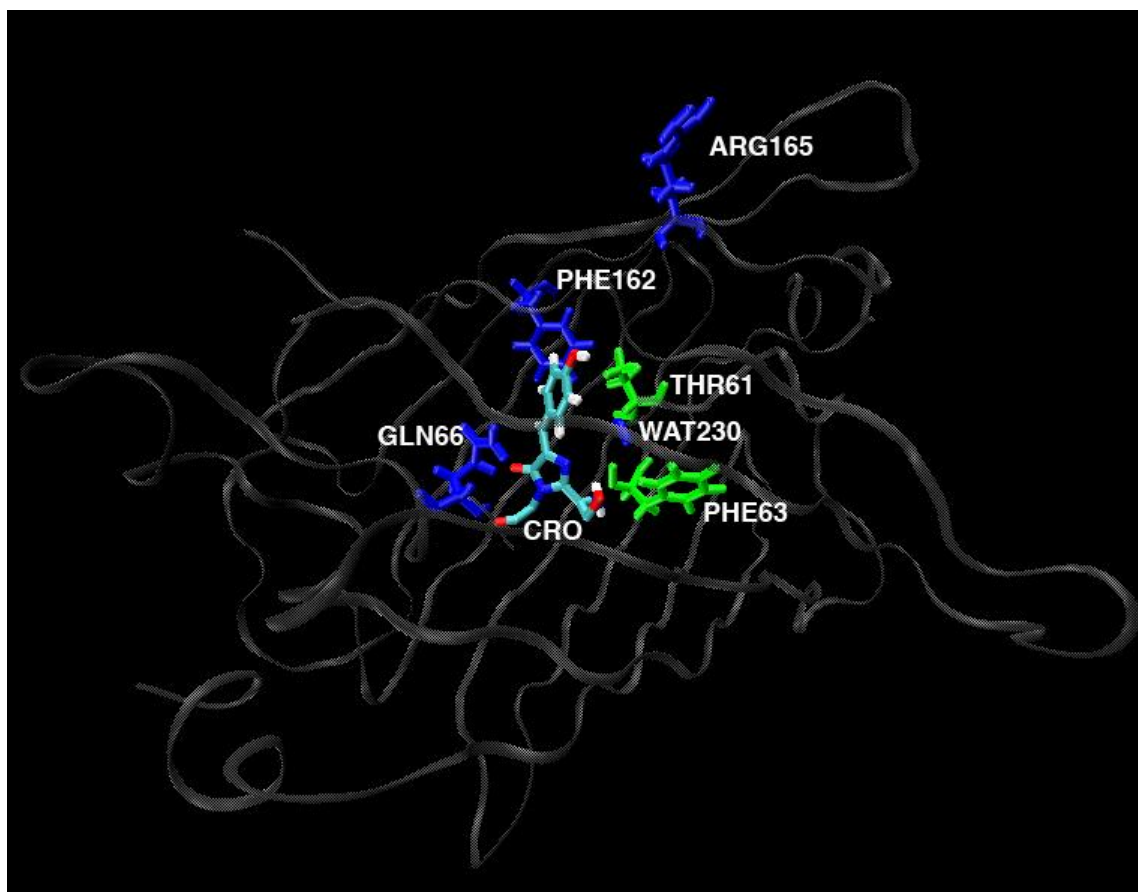
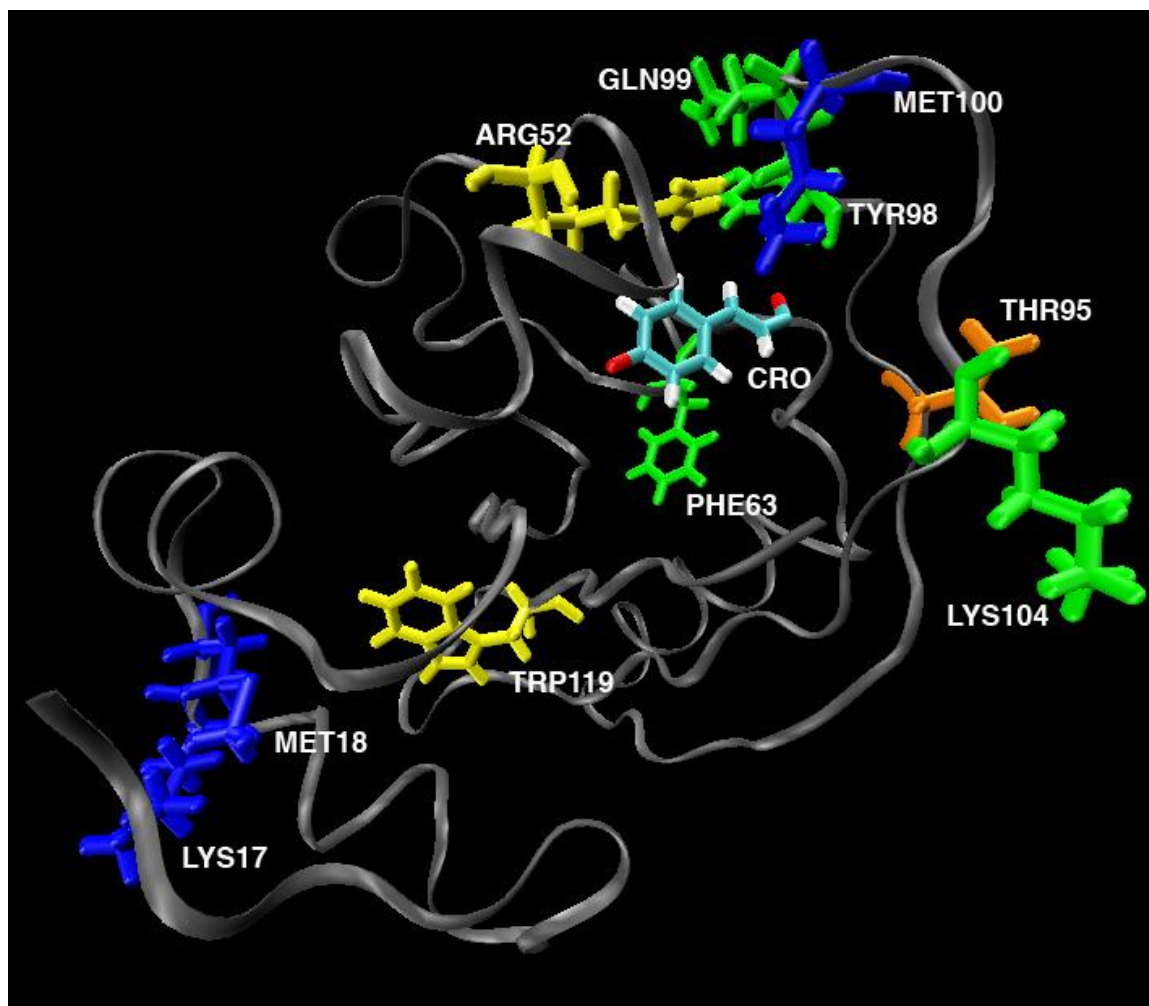
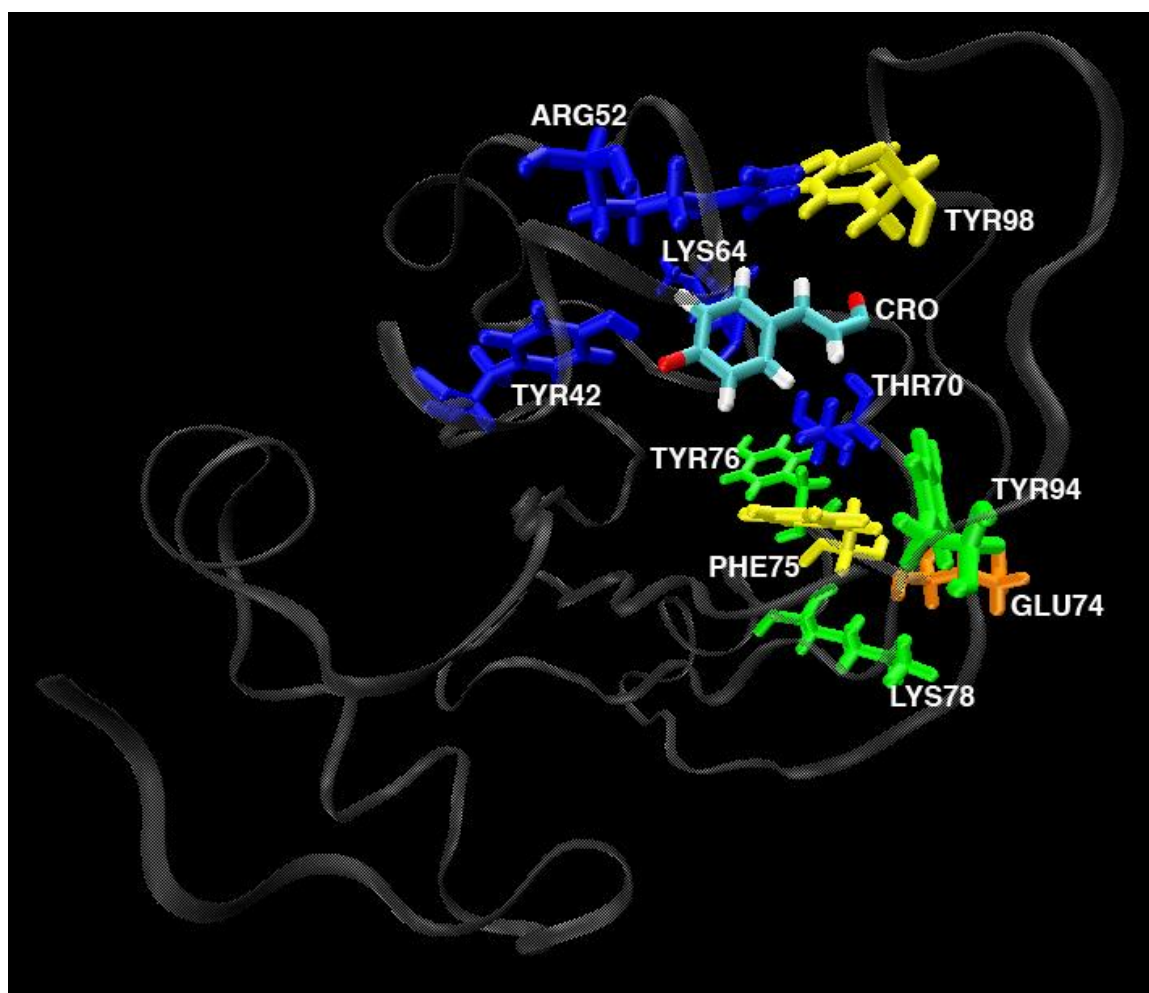


Figure 7-7 Hot regions of ground state GFP shown from 1 to 3 ps after the excitation of the chromophore (top: 1 ps; middle 2 ps; bottom: 3 ps). The colors indicated correspond to different percent kinetic energy of the whole system contained in a residue or confined water. Residues with at least 5% of the total energy are colored in red, 4% - 5% in orange, 3% - 4% in yellow, 2% -3% in green, 1% - 2% in blue.





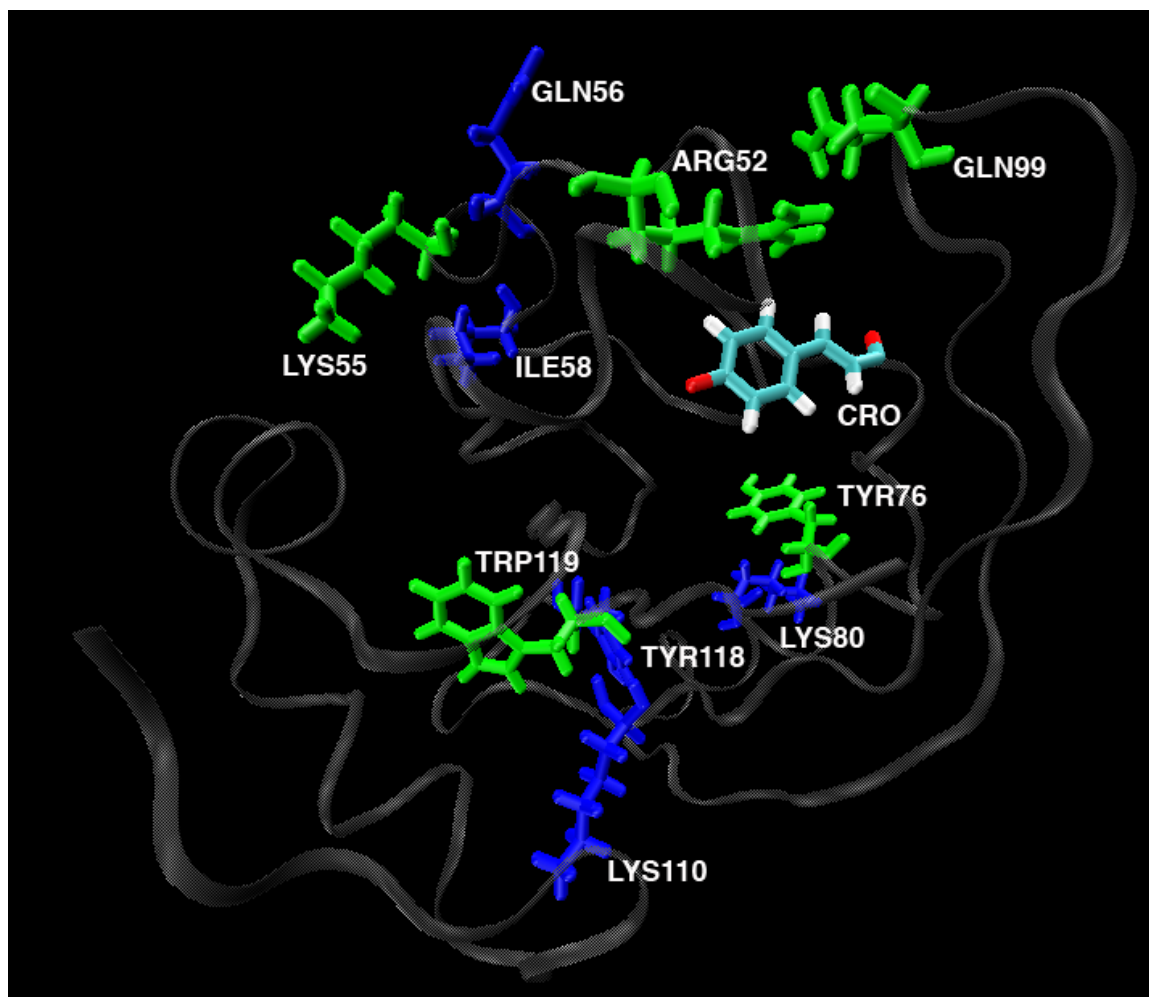
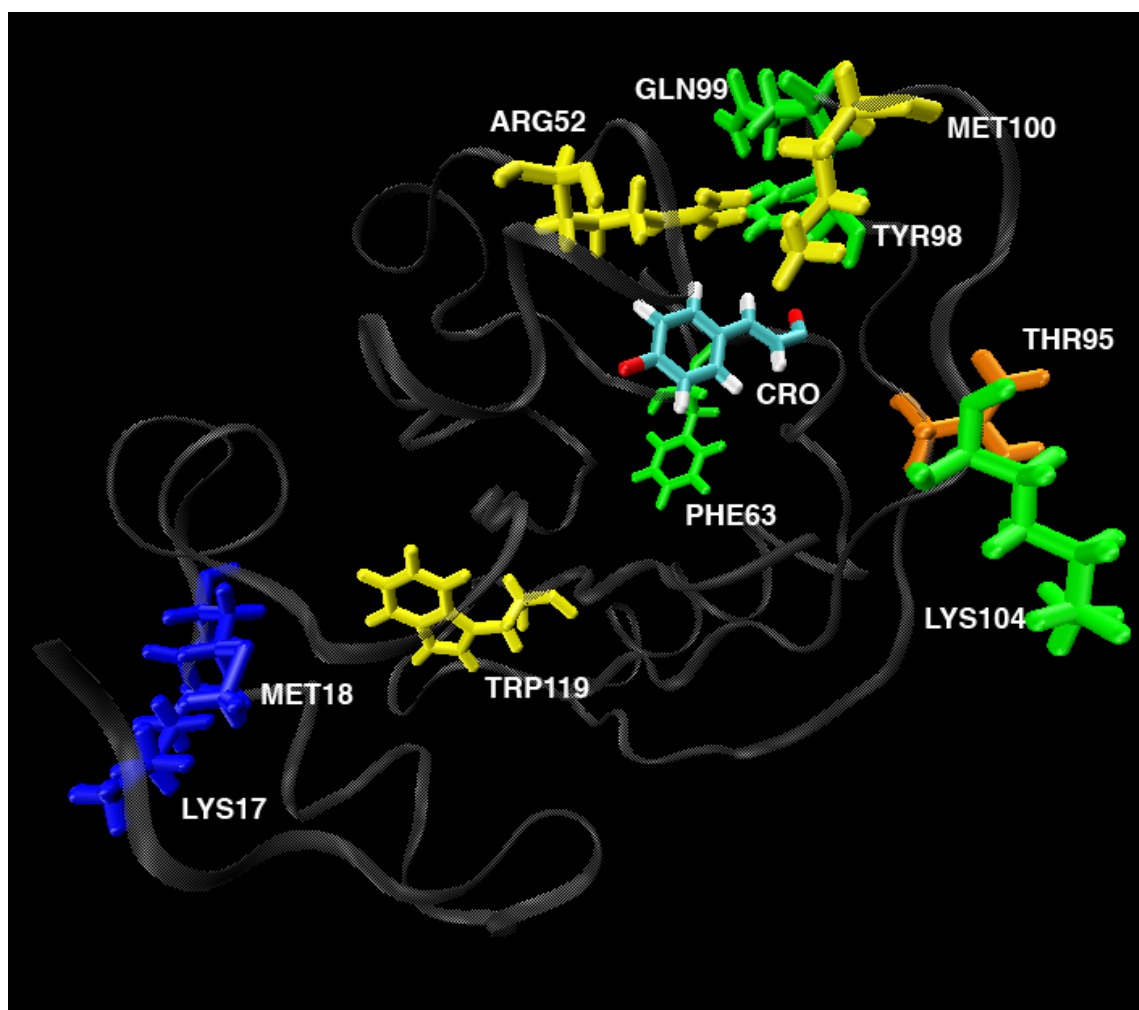
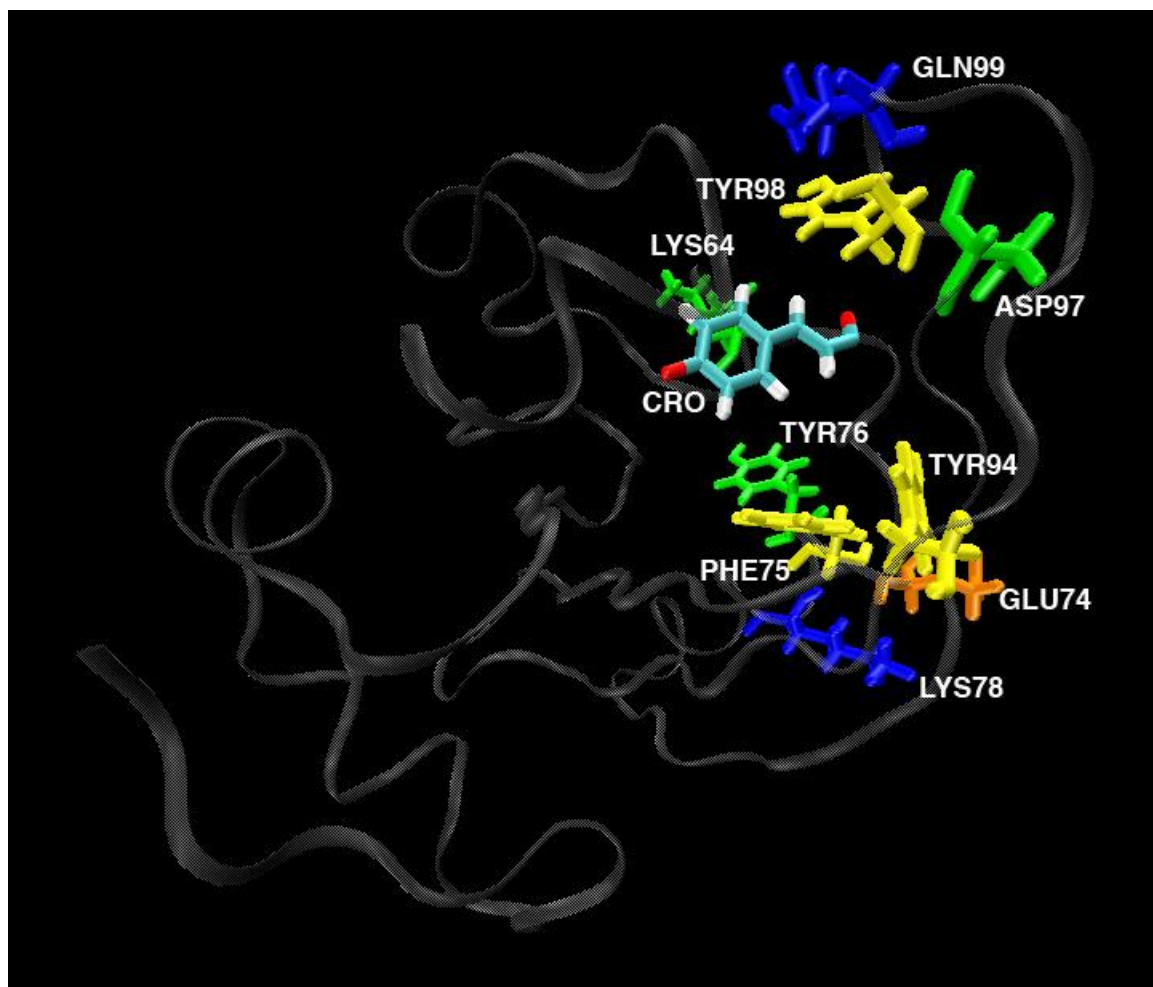


Figure 7-8 Hot regions of ground state PYP shown from 1 to 3 ps after the excitation of the chromophore (top: 1 ps; middle 2 ps; bottom: 3 ps). The colors indicated correspond to different percent kinetic energy of the whole system contained in a residue. Residues with at least 5% of the total energy are colored in red, 4% – 5% in orange, 3% - 4% in yellow, 2% -3% in green, 1% - 2% in blue.





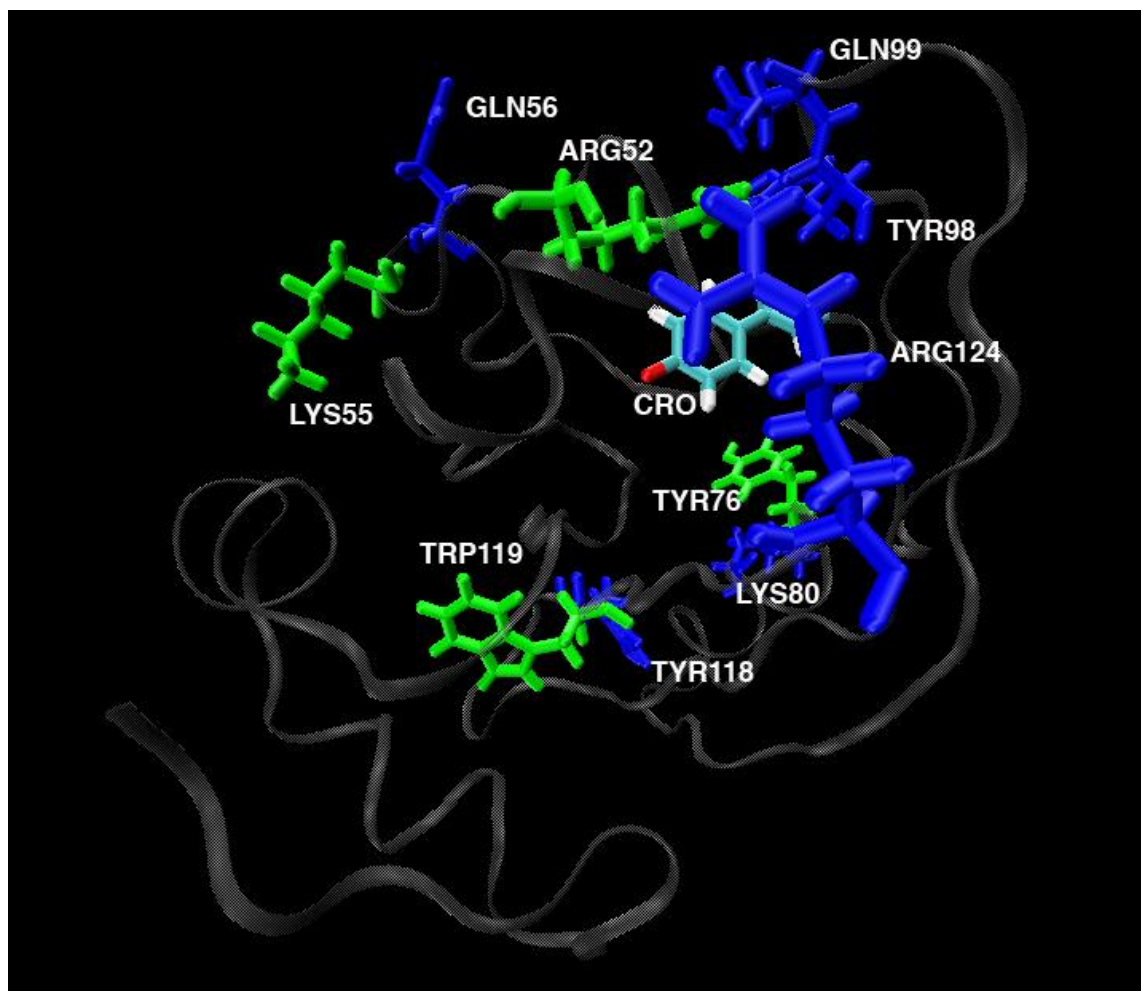


Figure 7-9 Hot regions of excited state PYP shown from 1 to 3 ps after the excitation of the chromophore (top: 1 ps; middle 2 ps; bottom: 3 ps). The colors indicated correspond to different percent kinetic energy of the whole system contained in a residue. Residues with at least 5% of the total energy are colored in red, 4% - 5% in orange, 3% - 4% in yellow, 2% - 3% in green, 1% - 2% in blue.

7.6 References

1. D. J. Tobias, N. Sengupta and M. Tarek, *Proteins: Energy, Heat and Signal Flow* **D. M. Leitner and J. E. Straub, eds.**, (Taylor and Francis, New York) 361 - 386 (2009).
2. S. J. Kim, B. Born, M. Havenith and M. Gruebele, *Angewandte Chemie* **120**, 6486 - 6489 (2008).
3. X. Yu, J. Park and D. M. Leitner, *J. Phys. Chem. B* **107**, 12820 - 12829 (2003).
4. K. Yoshida, T. Yamaguchi, M.-C. Bellissent-Funel and S. Longeville, *Eur. Phys. J. Special Topics* **141**, 223-226 (2007).
5. S. Ebbinghaus, S.-J. Kim, M. Heyden, X. Yu, U. Heugen, M. Gruebele, D. M. Leitner and M. Havenith, *Proc. Natl. Acad. Sci. USA* **104**, 20749 - 20752 (2007).
6. D. M. Leitner, M. Havenith and M. Gruebele, *Int. Rev. Phys. Chem.* **25**, 553 - 582 (2006).
7. V. M. Dadarlat and C. B. Post, *Biophys. J.* **91**, 4544 - 4554 (2006).
8. J. R. Knab, J.-Y. Chen and A. G. Markelz, *Biophys. J.* **90**, 2576 - 2581 (2006).
9. L. Mitra, N. Smolin, R. Ravindra, C. Royer and R. Winter, *Phys. Chem. Chem. Phys.* **8**, 1249 - 1265 (2006).
10. P.-H. Yang and J. A. Rupley, *Biochem.* **18**, 2654 - 2661 (1979).
11. A. R. Bizzarri and S. Cannistraro, *J. Phys. Chem B* **106**, 6617-6633 (2002).
12. M. E. Johnson, C. Malardier-Jugroot, R. K. Murarka and T. Head-Gordon, *J. Phys. Chem. B.* **113**, 4082-4092 (2009).

13. I. Yu, T. Tasaki, K. Nakada and M. Nagaoka, *J. Phys. Chem. B* **114**, 12392 - 12397 (2010).
14. D. N. LeBard and D. V. Matyushov, *J. Phys. Chem. B* **114**, 9246 - 9258 (2010).
15. H. Frauenfelder, P. W. Fenimore, G. Chen and B. H. McMahon, *Proc. Natl. Acad. Sci. (USA)* **103**, 15469 - 15472 (2006).
16. P. J. Steinbach and B. R. Brooks, *Proc. Natl. Acad. Sci. USA* **90**, 9135 - 9139 (1993).
17. M. S. P. Sansom, I. D. Kerr, J. Breed and R. Sankararamakrishnan, *Biophys. J.* **70**, 693 - 702 (1996).
18. L. Hua, X. Huang, R. Zhou and B. J. Berne, *J. Phys. Chem. B* **110**, 3704 - 3711 (2006).
19. R. Gnanasekaran, Y. Xu and D. M. Leitner, *J. Phys. Chem. B* **114**, 16989 - 16996 (2010).
20. J. C. Rasaiah, S. Garde and G. Hummer, *Annu Rev Phys Chem* **59**, 713 - 740 (2009).
21. P. M. Petrone and A. E. García, *J. Mol. Biol.* **338**, 419 - 435 (2004).
22. W. E. Royer, A. Pardanani, Q. H. Gibson, E. S. Peterson and J. M. Friedman, *Proc. Natl. Acad. Sci* **93**, 14526-14531 (1996).
23. M. Chattoraj, B. A. King, G. U. Bublitz and S. G. Boxer, *Proc. Natl. Acad. Sci. (USA)* **93**, 8362 - 8367 (1996).
24. A. Stuchebrukhov, *Proteins: Energy, Heat and Signal Flow* **D. M. Leitner and J. E. Straub, eds.** , (Taylor and Francis, New York) 71 - 104 (2009).
25. W. E. Royer, W. A. Hendrickson and E. Chiancone, *Science* **249**, 518-521 (1990).

26. E. Chiancone, P. Vecchini, D. Verzilli, F. Ascoli and E. Antonini, *J. Mol. Biol.* **152**, 577-592 (1981).
27. J. E. Knapp, R. Pahl, V. Srajer and W. E. Royer, *Proc. Natl. Acad. Sci. USA* **103**, 7649 - 7654 (2006).
28. Y. Zhou, H. Zhou and M. Karplus, *J. Mol. Biol.* **326**, 593-606 (2003).
29. D. M. Leitner, *J. Chem. Phys.* **130**, art. no. 195101, pp. 195101 - 195109. (2009).
30. W. E. Royer, H. Zhu, T. A. Gorr, J. F. Flores and J. E. Knapp, *J. Biological Chem.* **280**, 27477 - 27480 (2005).
31. R. Elber, *Biophys. J.* **92**, L85 - L87 (2007).
32. A. Pardanani, A. Gambacurta, F. Ascoil and W. E. Royer, *J. Mol. Biol.* **284**, 729-739 (1998).
33. A. Lervik, F. Bresme, S. Kjelstrup, D. Bedeaux and J. M. Rubi, *Phys. Chem. Chem. Phys.* **12**, 1610 - 1617 (2010).
34. H. Fujisaki and J. E. Straub, *Proc. Natl. Acad. Sci. (USA)* **102**, 6726 – 6731 (2005).
35. D. M. Leitner, *Adv. Chem. Phys.* **130B**, 205 - 256 (2005).
36. D. M. Leitner, *Ann. Rev. Phys. Chem.* **59**, 233 - 259 (2008).
37. D. M. Leitner and J. E. Straub, *Proteins: Energy, Heat and Signal Flow* (Taylor and Francis Press, New York) (2009).
38. X. Yu and D. M. Leitner, *J. Phys. Chem. B* **107**, 1698 - 1707 (2003).
39. P. H. Nguyen, S. M. Park and G. Stock, *J. Chem. Phys.* **132**, 025102 (2010).
40. M. Takayanagi, H. Okumura and M. Nagaoka, *J. Phys. Chem. B* **111**, 864 - 869 (2007).

41. A. Lervik, F. Bresme and S. Kjelstrup, *Soft Matter* **5**, 2407 - 2414 (2009).
42. M. Kobus, P. H. Nguyen and G. Stock, *J. Chem. Phys.* **134**, art. no. 124518 (2011).
43. K. Moritsugu, O. Miyashita and A. Kidera, *Phys. Rev. Lett.* **85**, 3970 - 3973 (2000).
44. T. Ishikura, T. Yamato, *Chem. Phys. Lett.* **432**, 533 - 537 (2006).
45. M. Nagaoka, I. Yu and M. Takayanagi, *Proteins: Energy, Heat and Signal Flow* **D. M. Leitner and J. E. Straub, eds.** , (Taylor and Francis, New York) 199 - 228 (2009).
46. N. Ota and D. A. Agard, *J. Mol. Biol.* **351**, 345 – 354 (2005).
47. F. Piazza and Y. H. Sanejouand, *Physical Biology* **6**, 046014 (2009).
48. Y. Kong and M. Karplus, *Proteins: Struct. Func. Bioinform.* **74**, 145 - 154 (2009).
49. P. H. Nguyen, P. Derreumaux and G. Stock, *J. Phys. Chem. B* **113**, 9340 - 9347 (2009).
50. K. Sharp and J. J. Skinner, *Proteins: Struct. Func. Bioinform.* **65**, 347 – 361 (2006).
51. M. Kaledin, A. L. Kaledin, A. Brown and J. M. Bowman, in *Normal Mode Analysis: Theory and Applications to Biological and Chemical Systems* **Q. Cui, I. Bahar, Eds.** ((CRC Press, Boca Raton)), 281 - 300 (2005).
52. T. Yamato, *Proteins: Energy, Heat and Signal Flow* **D. M. Leitner and J. E. Straub, eds.** , (Taylor and Francis, New York), 129 - 147 (2009).
53. T. Ishikura and T. Yamato, *Chem. Phys. Lett.* **432**, 533 – 537 (2006).
54. D. E. Sagnella and J. E. Straub, *J. Phys. Chem. B* **105**, 7057 – 7063 (2001).

55. L. Bu and J. E. Straub, *J. Phys. Chem. B* **107**, 12339 – 12345 (2003).
56. W. L. Jorgensen, J. Chandrasekar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).
57. A. J. DePaul, E. J. Thompson, S. S. Patel, K. Haldeman and E. J. Sorin, *Nucleic Acids Res.* **38**, 4856 - 4867 (2010).
58. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys.* **103**, 8577 - 8592 (1995).
59. R. Gnanasekaran, J. K. Agbo and D. M. Leitner, *J. Chem. Phys.* **135**, art. no. 065103 (2011).
60. X. Yu and D. M. Leitner, *J. Chem. Phys.* **123**, art. no. 104503 (2005).
61. X. Yu and D. M. Leitner, *Phys. Rev. B* **74**, art. no. 184305 (2006).
62. P. B. Allen and J. L. Feldman, *Phys. Rev. B* **48**, 12581 - 12588 (1993).
63. J. L. Feldman, M. D. Kluge, P. B. Allen and F. Wooten, *Phys. Rev. B* **48**, 12589-12602 (1993).
64. R. J. Hardy, *Phys. Rev.* **132**, 168 - 177 (1963).
65. X. Yu and D. M. Leitner, *J. Chem. Phys.* **119**, 12673 - 12679 (2003).
66. X. Yu and D. M. Leitner, *Chem. Phys. Lett.* **398**, 480 - 485 (2004).
67. C. L. Brooks, M. Karplus and B. M. Pettitt, *Adv. Chem. Phys.* **71**, 1 - 150 (1988).
68. T. Nishikawa and N. Go, *Proteins: Struct. Func. Genetics* **2**, 308 - 329 (1987).

8 Conclusion

The structure and dynamics of water in the vicinity of proteins in aqueous solutions are reported to be different from bulk water and play important roles in protein stability and function. Much effort has been made both theoretically and experimentally to understand the dynamics of water molecules near the surface of and confined in proteins in aqueous solutions.

We have carried out atomistic molecular dynamics simulation to investigate the water dynamics near and in the proteins and its functional roles in energy transfer in proteins. In Chapter 2, we examined the power spectrum of the velocity autocorrelation function for water molecules near the surface of the antifreeze protein (AFP) from the spruce budworm *Choristoneura fumiferana* and analyzed the hydrogen bond lifetimes for bonds between water molecules and the protein. We explored effects of the heterogeneity of the protein surface, in particular the distinctive properties of the water and protein-water interactions on the three planes of the protein, one of which is vital to the function of this AFP, and how the dynamics is affected by mutation. For the power spectra of the water in the hydration layer of the AFP and the subset of that water that hydrogen bonds to the protein we find a blue shift of the roughly 2 THz band compared to the same band in bulk water with a more pronounced shift for the water molecules that are hydrogen bonded to the protein residues. Although the power spectra for the water molecules within 5 Å of each of the planes of the protein appear quite similar, power spectra for the water molecules hydrogen bonded to different planes of the protein exhibit distinct

spectra in the range 1 – 4 THz. The differences among the power spectra for the water molecules hydrogen bonding to each of the three planes are influenced by mutation. The hydrogen bond time correlation function was computed for hydrogen bonds between water molecules and each of the planes of the protein. For wild type AFP we observe differences in the hydrogen bond lifetimes for bonds between water and the three planes. The longest lifetimes are found for hydrogen bonds between water molecules and Plane 1, the ice-binding plane of the protein. We observe that by introducing only four mutations to Plane 1, mutations that have been observed to substantially diminish the AFP activity of the protein, the hydrogen bond correlation function for bonds between water molecules and each of the three planes are similar to one another. Overall, mutation is seen to modify hydrogen bonding over a wide range of time scales observable both in the power spectra and analysis of hydrogen bond lifetimes. These measures of hydrogen bonding at the protein-water interface aid in quantifying the complexity and heterogeneity of the interactions between water and the antifreeze protein and reveal regions of the protein-water interface important for antifreeze activity.

In Chapter 3, we combined THz spectroscopy and molecular dynamics simulation results to provide evidence of key roles of water fast dynamics in molecular mechanism of antifreeze activity of DAFP-1, which has a well-defined surface composed of ice-binding and non-binding regions. In spite of traditional views on the short-range water interactions with regularly-spaced hydroxyl groups of threonine-rich ice-binding region, we found evidence of long-range slow water dynamics in the vicinity of the ice-binding region, similar to our recent observation in a hydrated enzyme system, suggesting that the retardation of the hydration water may dominate the antifreeze activity. In particular, the

estimated distance of the dynamical hydration layer from the protein surface is approximately 20 – 27 Angstrom, corresponding to typically 7th shell of water. We also added sodium citrate and mutated threonine residues on the ice-binding regions to explore how the addition and mutation affect the hydration water dynamics not far from the protein surface. What we found is a significant increase in the size of the dynamical hydration shell upon addition of sodium citrate, quite different from our observation in other type of AFP. There, negligible change in the size of dynamical hydration shell indicates significant roles of other factors such as rigidity of its helix structure and/or orientation of the threonines near the binding surface. Overall, the research survey of different types of AFP shows that quite variable mechanisms could exist in the same family of AFP. Seemingly, by integrating both long and short range interactions between water and proteins, Nature has achieved the same antifreeze activity function strategically.

In Chapter 4, we performed MD simulations to study hydrogen bond and rotational dynamics of interfacial water cluster confined between the two globules, as well as hydration water of both deoxygenated and oxygenated HbI. The hydration water dynamics is found to be retarded by the protein surface, but not by the ligand-binding through our hydrogen bond and rotational anisotropy studies. However, ligand binding strongly affects the significantly slower dynamics of the interfacial water molecules, compared to the water cluster in deoxygenated structure. This could be mainly due to the smaller volume of available at the interface of oxy-HbI. However, there are 6 water molecules in both structures forming strong hydrogen bonds with charged groups at the interface such as the heme propionate group. The difference between the number of hydrogen bonds formed by the interfacial water and side chains in deoxy and oxy

structure probably contribute to the slower dynamics of the interfacial water in oxy structure since more compact interfacial space definitely facilitate closer contact of interfacial water to its neighboring charged side chains, consistent with the expulsion of 6 water interfacial molecules in the allosteric transition between the two structures.

In Chapter 5 and 6, we computed dielectric response of the GFP-water environment and PYP protein matrix to photoexcitation of the GFP and PYP chromophore, respectively, as well as the response of water to photoexcitation of the bare GFP and PYP chromophore in aqueous solutions for comparison, individually. Decomposition of the response shows a wide range of characteristic time constants of water dynamics in both cases, indicative of a comparable role water and protein in slow relaxation, in contrast to recent discoveries on predominant roles of biomolecules in dielectric response studies of a solvated biomolecule to photoexcitation of a chromophore. This arises from the existence of 10 tightly bound water molecules in beta-barrel of GFP that surround the chromophore. The analysis of hydrogen bond dynamics between and the confined waters around chromophore and the protein indicate highly heterogeneous hydrogen bonding with lifetimes of tens of picoseconds. For PYP, we carried out TDSS studies on the response of protein and hydration water to photoexcitation of the chromophore by using a QM/MM calculation. The protein response to photoexcitation similarly appears in the TDSS on times of a wide range of scales. The integration of TDSS studies in both GFP and PYP leads us to believe that hydrogen-bonding in the microenvironment of the chromophore seems to have a notable influence on the dielectric response of the protein matrix and/or confined water to photoexcitation of the chromophore.

To validate this, in Chapter 7, we calculated static frequency-resolved thermal energy communication map on a residue-wise basis, as well as dynamic molecular simulations to investigate energy flow within proteins of GFP and PYP. The energy communication map for GFP reflects that those confined water molecules surrounding the chromophore have strong interaction thermally with the chromophore after its photoexcitation. Further analysis from energy flow studies show that the water plays as excess energy sink to stabilize the chromophore during the first few picoseconds after the photoexcitation, as demonstrated by much less long-range energy flow occurs from the GFP chromophore. In PYP, however, due to lack of confined water, excess energy obtained from photoexcitation has to be transported directly to those residues that hydrogen bond to the chromophore. Several energy flow channels were identified in our molecular simulation studies, consistent with the observed high energy possession of these residues in our communication map analysis.

The dynamics and energetics of macro-biomolecules are intimately coupled to the solvent. We studied by numerical simulation dynamical and energetic coupling between proteins and water molecules as solvent near the surface of and confined in the proteins. Overall, we found that hydration water around and confined water in proteins exhibit remarkably different dynamical properties from bulk water, consistent with the complex picture of protein-water interplay at their interface provided by others. Much effort on the complexity is exerted with the consideration of hydrogen bonds between proteins and solvents at the microscopic level, interpreting solvents as plasticizer of protein motions by percolating solvent dynamics into protein dynamics possibly through continuous forming and breaking hydrogen bonds. However, the view of aqueous and non-aqueous

solvents uniquely promote protein dynamics and function has been challenged recently by the observed bioactivity of some solvent-free protein-polymer surfactant hybrids from a combination of elastic incoherent neutron scattering and specific deuterium labeling to study protein and polymer separately.¹ The motions of the protein myoglobin within the hybrid are very close to those of its hydrated counterpart, and the polymer surfactant coating resembles hydration water, leading to the possibility of the protein plasticized by the polymer. Moreover, hydration water dynamics studies of intrinsically disordered proteins (IDP), a large family of proteins performing ligand recognition, enzyme activity and energy transduction without well-defined structure, surprisingly show more restricted hydration water motions near the IDP surface compared to those near globular, folded proteins, bringing new complexity to the traditional idea that protein and hydration-water mutually affect and shape each other.² Therefore, it is necessary to reconsider the protein-water relationship before we can achieve a much deeper understanding of the complexity between them. It is our hope that developments on synthetic chemical biology for artificial proteins and revolutionary techniques in photonics and electronics can advance our understanding of the relationship both *in vitro* and *in vivo* with much higher spatial and temporal resolutions.

8.1 References

1. Gallat, F.-X. *et al.* A Polymer Surfactant Corona Dynamically Replaces Water in Solvent-Free Protein Liquids and Ensures Macromolecular Flexibility and Activity. *J. Am. Chem. Soc.* **134**, 13168–13171 (2012).
2. Gallat, F.-X. *et al.* Dynamical Coupling of Intrinsically Disordered Proteins and Their Hydration Water: Comparison with Folded Soluble and Membrane Proteins. *Biophys. J.* **103**, 129–136 (2012).