Title

Decoding the molecular-scale determinants of functional properties at aqueous interfaces

Author

Robinson Brown, Dennis Charles

Publication Date

2023

UNIVERSITY of CALIFORNIA

Santa Barbara

Decoding the molecular-scale determinants of functional properties at aqueous interfaces

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy
in Chemical Engineering

by

Dennis C. Robinson Brown

Committee in charge:
Professor M. Scott Shell, Chair
Professor Songi Han
Professor Angela A. Pitenis
Professor Todd M. Squires

December 2023

The dissertation of Dennis C. Robinson Brown is approved.

Songi Han

Angela A. Pitenis

Todd M. Squires

M. Scott Shell, Committee Chair

September 2023

ACKNOWLEDGEMENTS

This work is dedicated to my mom, Donna, who helped shape the person I am today by fostering my curiosity from a very young age—even when I was precocious to the point of irritation. Were it not for her example, I wouldn't have had the confidence or resolve to pursue a PhD in the first place.

My brothers Jake and Michael who are my oldest friends. As much fun as if was growing up with the two of you, our companionship in adulthood is even better.

My grandma, Lynda, who is impossibly kind and brave in face of hardship. I try (and fail) to match her level of empathy every day.

My late grandpa and namesake, Dennis, who had an excellent dry sense of humor and boundless persistence.

I'm owe a lot of gratitude to my first research mentor and now father-in-law, Mitko Petsev. Without his guidance, I wouldn't have even known what grad school is much less get accepted into some of the top programs in Chemical Engineering.

Of course, I owe a lot of my success to the excellent guidance to my advisor, Prof. Scott Shell. Scott excels at nurturing the strengths and developing underdeveloped research skills in early graduate students. As my research progressed past candidacy, Scott gave a lot of latitude to explore new and exciting research directions.

I'm also grateful for the guidance of Prof. Songi Han my unofficial co-advisor. Without Songi's boundless curiosity, a lot of the projects that I'm currently involved in wouldn't have been pursued.

My experience at UCSB wouldn't have been what it was without the brilliant collaborators and friends within and outside of the Shell Group and UCSB including: Prof.

Michael Doherty, Dr. Audra DeStefano, Sally Jiao, Dr. Vikram Khanna, Kimberlee Keithley, Dr. Pat Leggieri, Dr. Debasish Das Mahanta, Prof. Jacob Monroe, Prof. Josh Moon, Prof. Oscar Nordness, Alex Padilla, Andrew Pascoe, Marshall Reviere, Dr. Ashley Roach, Dr. Nick Sherck, Dr. Rahul Sujanani, Austen Tigges, Thomas Webber, Everett Zofchak, the MWET Team, and many others.

This work is also dedicated to my wife and soulmate, Nia, who has made my life fuller than I would have otherwise thought possible. I wouldn't have been able to see my PhD though without her by my side. I can't wait to continue exploring the world with you. P.S. Sorry for always falling into niche scientific discussions with your dad when we visit home.

CURRICULUM VITA OF DENNIS C. ROBINSON BROWN

EDUCATION

2023: Ph. D., Chemical Engineering, University of California, Santa Barbara

2018: B.S., Chemical Engineering, University of New Mexico (*Summa Cum Laude*)

2018: B.S., Physics, University of New Mexico (*Summa Cum Laude*)

PROFESSIONAL EXPERIENCE

2018-2023: Graduate Research Assistant, University of California, Santa Barbara

2015-2018: Undergraduate Research Assistant, Sandia National Laboratories

2015-2018: Undergraduate Research Assistant, University of New Mexico

PUBLICATIONS

(* equal contribution)

1. **D.C. Robinson Brown**, T. Webber, S. Jiao, and coworkers (2023). "Relationships between molecular structural order parameters and equilibrium water dynamics in aqueous mixtures". In: *JPC B.* doi: https://doi.org/10.1021/acs.jpcb.3c00826?ref=PDF
2. D.D. Mahanta*, **D.C. Robinson Brown***, S. Pezzotti, and coworkers (2023). "Local Water Structures Govern the Mixing Thermodynamics of Glycerol-Water Solutions". In: *Chem. Sci.* doi: https://doi.org/10.1039/d3sc00517h
3. J.I. Monroe, S. Jiao, **D.C. Robinson Brown**, and coworkers (2021). "Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution". In: *PNAS*. doi: https://doi.org/10.1073/pnas.2020205118
4. N. Sherck, T. Webber, **D.C. Robinson Brown**, and coworkers (2020). "End-to-End Distance Probability Distributions of Dilute Poly (ethylene oxide) in Aqueous Solution". In: JACS. doi: https://doi.org/10.1021/jacs.0c08709
5. J.I. Monroe, M. Barry, A. Destefano, P.A. Gokturk, S.Jiao, **D.C. Robinson Brown**, and coworkers (2020). "Water Structure and Properties at Hydrophilic and Hydrophobic Surfaces". In: Annu Rev Chem Biomol Eng. doi: https://doi.org/10.1146/annurev-chembioeng-120919-114657.
6. J. Moon* and T. Webber* and **D. C. Robinson Brown*** and coworkers (2023). "(Accepted at *Chem. Sci.*) Macroscopic water transport in poly (ethylene oxide) solution is tuned by nanoscale interactions".
7. **D. C. Robinson Brown**, T. Webber, T. Casey, and coworkers (2023). "(In Review at *PCCP*) Computation of Overhauser Dynamic Nuclear Polarization parameters and their correlation with water dynamics, structure and thermodynamics".
8. S. Jiao, **D. C. Robinson Brown** and coworkers (2023). "(Accepted at *Langmuir*) Relationships between water behavior and solute affinity at polypeptoid brush surfaces".

9. D. D. Mahanta*, **D. C. Robinson Brown***, T. Webber, and coworkers (2023). "(In Preparation) Structure Dynamics-Thermodynamics Connections in Cryoprotectant Aqueous Solutions".
10. O. Nordness*, **D. C. Robinson Brown*** and coworkers (2023). "(In Preparation) Directly probing ion-pairing and equilibrium dynamics in variable dielectric environments".

ORAL PRESENTATIONS

1. **D.C. Robinson Brown** (2023), "Decoding the molecular-scale determinants of functional properties at aqueous interfaces," Sandia National Laboratories, Albuquerque, NM.
2. **D.C. Robinson Brown** (2023), "Decoding the molecular-scale determinants of functional properties at aqueous interfaces," Sandia National Laboratories, Albuquerque, NM.
3. **D.C. Robinson Brown** (2023), "Decoding the molecular-scale determinants of functional properties at aqueous interfaces," Cedars-Sinai, Los Angeles, CA.
4. **D. C. Robinson Brown** (2022). "Decoding the molecular-scale determinants of antifouling at polymer membrane selective layers". Clorox-Amgen Graduate Student Symposium. Santa Barbara, California
5. **D. C. Robinson Brown** (2022). "Decoding the molecular-scale determinants of antifouling at polymer membrane selective layers". North American Membrane Society Conference. Tempe, Arizona
6. **D. C. Robinson Brown** (2016). "Improving Kelvin probe work function measurement reliability". The Rio Grande Symposium on Advanced Materials. Albuquerque, NM

ABSTRACT

Decoding the molecular-scale determinants of functional properties at aqueous interfaces

by

Dennis C. Robinson Brown

Current water purification membrane technologies cannot readily treat the high concentration and multi-component produced water (PW) from oil and gas operations. This stems, in part, from membrane fouling induced a diverse array of organic and inorganic solutes. At present, the design of antifouling membrane materials relies on macroscale heuristics such as ensuring the smoothness, charge neutrality, and hydrophilicity of the membrane surface. For instance, hydrophilic coatings such as polyethylene oxide (PEO)-based hydrogels dramatically increase resistance to natural organic matter. This anti-fouling property is hypothesized to originate from the formation of a bound water layer at the membrane surface that resists adsorption of hydrophobic molecules; however, the molecular basis for this antifouling property is not well understood.

This work leverages detailed atomistic molecular simulations to elucidate the molecular scale determinants of water properties at aqueous interfaces. First, we implement a synergistic computational-experimental approach to unveil the persistent connection between water's collective molecular structure and equilibrium water dynamics in aqueous solutions. Then, high-throughput molecular dynamics simulations and a statistical learning workflow reveal persistent connections between water structure and functional thermophysical properties in aqueous environments. Further, unsupervised learning (principal component analysis) reveals hidden signatures of water structuring on small (<1 nm) and large length scales. Finally, free energetic calculations detail foulant interactions with a model antifouling PEO brush surface. Further analysis demonstrates that foulant-surface interactions are driven by a combination of direct interaction (energetic) and solution restructuring (entropic) contributions.

Table of Contents

Chapter 1: Introduction

Water is a profoundly important constituent of all manner of naturally occurring and engineered systems. The near ubiquitous role of water in nature owes partially due to its unique molecular properties. Most well-known amongst these properties is expansion of water upon freezing due to the existence of a density maximum at 4°C. Water also exhibits a several other anomalous trends in thermophysical properties such as the self-diffusivity, compressibility, etc., particularly upon supercooling. Much of this strange behavior can be attributed to water's unique collective network structure. Specifically, unlike any other room temperature liquid, water adopts a tetrahedral structure driven by the tendency for each water molecule to participate in four hydrogen bonds (two accepting and two donating). This structural motif has profound implications on both naturally abundant (e.g., proteins) and man-made systems.

Dating back to the 1970's(1–4), a large body of research has been dedicated to determining the precise modes and mechanisms through which water mediates structure and function in the world around us. The earliest approaches to studying aqueous systems—and fluids more broadly—leveraged relatively simple thermodynamic equations of state. These works employed theoretical approaches such as Scaled Particle Theory (SPT)(1) to describe fluid behavior in the vicinity of hard particles. Using these methods, the molecular scale details were, by necessity, neglected in favor of assessing qualitative trends in thermophysical properties. Modern, simulation-based approaches are required to resolve the effect of atomic-level details on water behavior.

Monte Carlo (MC) and Molecular Dynamics (MD) are common simulation methods for studying liquid state systems with molecular detail. In brief, MC methods are probabilistic in nature and directly sample properties in fluids to estimate the ensemble average. However, purely MC-based approaches to molecular simulation have been largely supplanted by MD simulation methods, in part, due to widely available simulation environments and force field parameters. MD simulations directly integrate Newton's Second Law to deterministically (or quasi-deterministically) produce a time-resolved trajectory of all particles in a system. Unlike MCMC, MD yield system dynamics for the calculation of equilibrium (e.g., water self-diffusivity) or nonequilibrium (e.g., pressure-driven flux) dynamic quantities. However, MC procedures are vital to determining free energetic processes such as hydration, binding, and folding processes in proteins. Hence, modern approaches of molecular simulation—including the studies detailed in this thesis—embed MC procedures within MD workflows.

The application of these advanced simulation methods yielded accurate predictions and molecular-scale insights on the atomic-level determinants of water properties(5–9). Beyond these fundamental areas of inquiry, water plays a prominent role in biophysics(10–18), polymer physics(19–26), and—unsurprisingly—water purification membrane sciences(27–29, 29–33). For instance, shifts in water structure and dynamics near biomacromolecules is suggestive of the underlying mechanisms undergirding protein-protein interactions(34), protein-ligand interactions(35–38), protein aggregation(39, 40), protein folding(14, 14, 41, 42), and many other phenomena. Though fundamental investigations of water have long pervaded biophysical research, many fields have yet to fully adopt this approach to understanding their systems. Beyond biological systems, water's molecular behavior has deep implications on the function of engineered aqueous systems. Recent efforts leveraging water's molecular scale properties in conjunction with free energy calculation(35, 43, 44) and statistical learning(44–46) methods have shed new light on the modes and mechanisms through which water mediates properties at engineered heterogeneous interfaces.

Water purification membranes used in ultrafiltration, nanofiltration (NF), and reverse osmosis (RO) operations are prime examples of highly heterogeneous engineered materials. The molecular behavior of water is perhaps most germane to RO and NF applications which remove small molecule and ionic foulants. Further, current generation water purification membranes are ill-equipped to treat high concentration and multicomponent waste streams like Produced Water (PW) from oil and gas operations(28, 47). Current production scale commercial design strategies for membrane materials rely on three heuristic, macroscale design rules(28): (1) mitigate surface roughness, (2) reduce membrane surface charge, and (3) maximize surface hydrophilicity. Surface roughness induces accumulation of foulants on the micron scale(48). On the other hand, surface charge and hydrophilicity control surface fouling at the nanoscale. The presence of a net surface charge tends to drive unfavorable condensation of ionic species commonly found in PW(28), but the inclusion of net zero charge functionalities (zwitterions) has been demonstrated to improve membrane antifouling capacity(49–51). Enhancing membrane hydrophilicity is hypothesized to lead to the development of a layer bound water at the membrane surface and hence improve rejection of hydrophobic contaminants (e.g., small organic molecules and emulsified oils)(28). However, the precise molecular mechanisms underlying this antifouling property is not yet fully understood. By directly characterizing water's molecular properties in relevant systems, the membranes community can discover molecular-scale design rules for next-generation purification membranes.

We aim to advance the fundamental understanding of the *molecular-scale determinants of functional properties at aqueous interfaces* by leveraging detailed molecular simulation in synergy with our experimental collaborators. In pursuit of this goal, **we hypothesize a profound connection between structure, dynamics, and solvation thermodynamics at molecular time and length scales**. In particular, we center the recurrent—and often predictive—connections between water structure-dynamics and structure-thermodynamics.

We provide a brief outline of the following five chapters. In **Chapter 2**, we leverage atomistic MD simulations to interpret the water dynamics underlying spectroscopic quantities from Overhauser Dynamic Nuclear Polarization (ODNP) measurements[1] of glycerol-water mixtures. In doing so, we provide a demonstration of the *structure-dynamics-thermodynamics* hypothesis. **Chapter 3** details a finer investigation of the water structure-dynamics connection in the membrane-relevant system: PEO-water solutions. We highlight the limitations of a continuum mechanics approach, the Stokes-Einstein relationship, in predicting water dynamics trends with increasing PEO concentration in a synergistic study with Pulsed-Field Gradient NMR (PFG-NMR)[2] and ODNP. In **Chapter 4**, we delve deeper into the molecular structure of the hydration shell by interrogating two components of hydration waters: waters hydrogen bonded to a cosolvent (“bound” waters) and non-solute bound waters (“wrap” waters). We use molecularly detailed structural characterizations from

---

[1] All ODNP experiments were performed by Thomas Webber.

[2] All PFG-NMR experiments were performed by Pete Richardson and Josh Moon.

MD simulations to further understand spectroscopic signatures from THz Calorimetry experiments[3]. In **Chapter 5**, we synthesize simulation-derived structural information from the previous three chapters (**Chapters 2-4**) to quantitatively predict water dynamics and solvation thermodynamics in a wide array of cosolvent chemistries and conditions (varying concentration and temperature). Using a simple principal component analysis (PCA), we uncover hidden structural signatures of water structure at small (<1 nm) and large (>1 nm) length scales. **Chapter 6** summarizes our efforts to develop molecular-scale intuition for interactions between model anti-fouling surfaces and a wide array of PW-relevant uncharged small molecules. We close the chapter by utilizing the PCA components from **Chapter 5** to better understand the molecular structural trends underlying these solute-surface interactions. In **Chapter 7,** we provide commentary on the preceding chapters and offer extensions to the current work for predicting hydration water properties at chemically and topologically heterogeneous surfaces—namely, protein-water interfaces.

References

1. Stillinger FH. 1972. Structure in Aqueous Solutions of Nonpolar Solutes from the Standpoint of Scaled-Particle Theory. *J. Solution Chem.* 2(2/3):141–58
2. Stillinger FH, Weber TA. 1982. Hidden structure in liquids. *Phys. Rev. A*. 25(2):978–89
3. Pratt LR, Chandler D. 1977. Theory of the hydrophobic effect. *The Journal of Chemical Physics*. 67(8):3683–3704
4. Pratt LR, Chandler D. 1980. Hydrophobic solvation of nonspherical solutes. *The Journal of Chemical Physics*. 73(7):3430–33
5. Debenedetti PG, Stillinger FH. 2001. Supercooled liquids and the glass transition. *Nature*. 410(6825):259–67
6. Errington JR, Debenedetti PG. 2001. Relationship between structural order and the anomalies of liquid water. *Nature*. 409(6818):318–21

---

[3] All THz spectroscopy experiments were performed by Debasish Das Mahanta.

7. Lynden-Bell RM, Debenedetti PG. 2005. Computational Investigation of Order, Structure, and Dynamics in Modified Water Models. *J. Phys. Chem. B*. 109(14):6527–34
8. Willard AP, Chandler D. 2010. Instantaneous Liquid Interfaces. *J. Phys. Chem. B*. 114(5):1954–58
9. Gartner TE, Piaggi PM, Car R, Panagiotopoulos AZ, Debenedetti PG. 2022. Liquid-Liquid Transition in Water from First Principles. *Phys. Rev. Lett.* 129(25):255702
10. Barnes R, Sun S, Fichou Y, Dahlquist FW, Heyden M, Han S. 2017. Spatially Heterogeneous Surface Water Diffusivity around Structured Protein Surfaces at Equilibrium. *J. Am. Chem. Soc.* 139(49):17890–901
11. Dahanayake JN, Mitchell-Koch KR. 2018. Entropy connects water structure and dynamics in protein hydration layer. *Physical Chemistry Chemical Physics*. 20(21):14765–77
12. Gekko K, Timasheff SN. 1981. Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry*. 20(16):4667–76
13. Heyden M. 2019. Heterogeneity of water structure and dynamics at the protein-water interface. *J. Chem. Phys.* 150(9):094701
14. Kim SJ, Born B, Havenith M, Gruebele M. 2008. Real-Time Detection of Protein-Water Dynamics upon Protein Folding by Terahertz Absorption Spectroscopy. *Angew. Chem. Int. Ed.* 47(34):6486–89
15. Qiao B, Jiménez-Ángeles F, Nguyen TD, Olvera de la Cruz M. 2019. Water follows polar and nonpolar protein surface domains. *Proc Natl Acad Sci USA*. 116(39):19274–81
16. Schirò G, Fichou Y, Gallat F-X, Wood K, Gabel F, et al. 2015. Translational diffusion of hydration water correlates with functional motions in folded and intrinsically disordered proteins. *Nature Communications*. 6(1):
17. Swenson J, Jansson H, Bergman R. 2006. Relaxation Processes in Supercooled Confined Water and Implications for Protein Dynamics. *Physical Review Letters*. 96(24):
18. Pizzitutti F, Marchi M, Sterpone F, Rossky PJ. 2007. How Protein Surfaces Induce Anomalous Dynamics of Hydration Water. *The Journal of Physical Chemistry B*. 111(26):7584–90
19. Sherck N, Webber T, Brown DR, Keller T, Barry M, et al. 2020. End-to-End Distance Probability Distributions of Dilute Poly(ethylene oxide) in Aqueous Solution. *J. Am. Chem. Soc.* 142(46):19631–41
20. Abrahmsén-Alami S, Stilbs P, Alami E. 1996. Water Self-Diffusion in Aqueous Associative Polymer Solutions. *J. Phys. Chem.* 100(16):6691–97
21. Hager SL, Macrury TB. 1980. Investigation of phase behavior and water binding in poly(alkylene oxide) solutions. *Journal of Applied Polymer Science*. 25(8):1559–71
22. Huang L, Nishinari K. 2001. Interaction between poly(ethylene glycol) and water as studied by differential scanning calorimetry. *Journal of Polymer Science Part B: Polymer Physics*. 39(5):496–506
23. Trotzig C, Abrahmsén-Alami S, Maurer FHJ. 2007. Structure and mobility in water plasticized poly(ethylene oxide). *Polymer*. 48(11):3294–3305
24. Masaro L, Ousalem M, Baille WE, Lessard D, Zhu XX. 1999. Self-Diffusion Studies of Water and Poly(ethylene glycol) in Solutions and Gels of Selected Hydrophilic Polymers. *Macromolecules*. 32(13):4375–82
25. Shinyashiki N, Asaka N, Mashimo S, Yagihara S. 1990. Dielectric study on dynamics of water in polymer matrix using a frequency range 106–1010 Hz. *J. Chem. Phys.* 93(1):760–64

26. Hoshina H, Iwasaki Y, Katahira E, Okamoto M, Otani C. 2018. Structure and dynamics of bound water in poly(ethylene-vinylalcohol) copolymers studied by terahertz spectroscopy. *Polymer*. 148:49–60

27. Aryal D, Howard MP, Samanta R, Antoine S, Segalman R, et al. 2020. Influence of pore morphology on the diffusion of water in triblock copolymer membranes. *J. Chem. Phys.* 152(1):014904

28. Miller DJ, Dreyer DR, Bielawski CW, Paul DR, Freeman BD. 2017. Surface Modification of Water Purification Membranes. *Angewandte Chemie International Edition*. 56(17):4662–4711

29. Moon JD, Galizia M, Borjigin H, Liu R, Riffle JS, et al. 2020. Modeling water diffusion in polybenzimidazole membranes using partial immobilization and free volume theory. *Polymer*. 189:122170

30. Perrin J-C, Lyonnard S, Volino F. 2007. Quasielastic Neutron Scattering Study of Water Dynamics in Hydrated Nafion Membranes. *The Journal of Physical Chemistry C*. 111(8):3393–3404

31. Jang E-S, Kamcev J, Kobayashi K, Yan N, Sujanani R, et al. 2020. Influence of water content on alkali metal chloride transport in cross-linked Poly(ethylene glycol) diacrylate.2. Ion diffusion. *Polymer*. 192:122316

32. Yasuda H, Lamaze CE, Peterlin A. 1971. Diffusive and hydraulic permeabilities of water in water-swollen polymer membranes. *Journal of Polymer Science Part A-2: Polymer Physics*. 9(6):1117–31

33. Moilanen DE, Piletic IR, Fayer MD. 2007. Water Dynamics in Nafion Fuel Cell Membranes: The Effects of Confinement and Structural Changes on the Hydrogen Bond Network †. *J. Phys. Chem. C*. 111(25):8884–91

34. Julius K, Weine J, Berghaus M, König N, Gao M, et al. 2018. Water-Mediated Protein-Protein Interactions at High Pressures are Controlled by a Deep-Sea Osmolyte. *Phys. Rev. Lett.* 121(3):038101

35. Rego NB, Xi E, Patel AJ. 2021. Identifying hydrophobic protein patches to inform protein interaction interfaces. *Proc Natl Acad Sci USA*. 118(6):e2018234118

36. Schiebel J, Gaspari R, Wulsdorf T, Ngo K, Sohn C, et al. 2018. Intriguing role of water in protein-ligand binding studied by neutron crystallography on trypsin complexes. *Nat Commun*. 9(1):3559

37. Poornima CS, Dean PM. 1995. Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions. *J Computer-Aided Mol Des*. 9(6):500–512

38. Kim KH. 2021. Outliers in SAR and QSAR: 3. Importance of considering the role of water molecules in protein–ligand interactions and quantitative structure–activity relationship studies. *J Comput Aided Mol Des*. 35(3):371–96

39. Vagenende V, Yap MGS, Trout BL. 2009. Mechanisms of Protein Stabilization and Prevention of Protein Aggregation by Glycerol. *Biochemistry*. 48(46):11084–96

40. Okumura H, Itoh SG. 2020. Molecular dynamics simulations of amyloid- β (16–22) peptide aggregation at air–water interfaces. *The Journal of Chemical Physics*. 152(9):095101

41. Baldwin RL, Rose GD. 2016. How the hydrophobic factor drives protein folding. *Proc Natl Acad Sci USA*. 113(44):12462–66

42. Norman AI, Fei Y, Ho DL, Greer SC. 2007. Folding and Unfolding of Polymer Helices in Solution. *Macromolecules*. 40(7):2559–67

43. Rego NB, Xi E, Patel AJ. 2019. Protein Hydration Waters Are Susceptible to Unfavorable Perturbations. *J. Am. Chem. Soc.* 141(5):2080–86

44. Dallin BC, Kelkar AS, Van Lehn RC. 2023. Structural features of interfacial water predict the hydrophobicity of chemically heterogeneous surfaces. *Chem. Sci.* 14(5):1308–19

45. Monroe JI, Jiao S, Davis RJ, Robinson Brown D, Katz LE, Shell MS. 2021. Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution. *Proc Natl Acad Sci USA*. 118(1):e2020205118

46. Jiao S, Katz LE, Shell MS. 2022. Inverse Design of Pore Wall Chemistry To Control Solute Transport and Selectivity. *ACS Cent. Sci.* 8(12):1609–17

47. Landsman MR, Sujanani R, Brodfuehrer SH, Cooper CM, Darr AG, et al. 2020. Water Treatment: Are Membranes the Panacea? *Annual Review of Chemical and Biomolecular Engineering*. 11(1):559–85

48. Zhong Z, Li D, Zhang B, Xing W. 2012. Membrane surface roughness characterization and its influence on ultrafine particle adhesion. *Separation and Purification Technology*. 90:140–46

49. Cheung D, Lau KHA. 2018. Atomistic Study of Zwitterionic Peptoid Antifouling Brushes. *Langmuir*. 35(5):1483–94

50. He Y, Hower J, Chen S, Bernards MT, Chang Y, Jiang S. 2008. Molecular Simulation Studies of Protein Interactions with Zwitterionic Phosphorylcholine Self-Assembled Monolayers in the Presence of Water. *Langmuir*. 24(18):10358–64

51. Jiang S, Cao Z. 2010. Ultralow-Fouling, Functionalizable, and Hydrolyzable Zwitterionic Materials and Their Derivatives for Biological Applications. *Adv. Mater.* 22(9):920–32

Chapter 2: Computation of Overhauser Dynamic Nuclear Polarization parameters and their correlation with water dynamics, structure, and thermodynamics

Dennis C. Robinson Brown, Thomas R. Webber, Thomas M. Casey, John M. Franck, M. Scott Shell, Songi Han. *Computation of Overhauser Dynamic Nuclear Polarization processes reveals fundamental correlation between water dynamics, structure, and solvent restructuring entropy.* In preparation

2.1 Introduction

Local hydropathies near molecular interfaces to water (e.g., proteins and polymers) modulate the surface activity for solute binding(1–5), among many other properties(2, 6–19). Locally hydrophobic regions of proteins are key to facilitating folding and inter-protein interactions, and are often characteristic of active sites(1, 16, 17, 20). Experimentally characterizing the structural, dynamical, and thermodynamic properties of water in fully hydrated environments at the molecular scale is challenging. Several experimental methods can probe the behavior of hydration layer waters,(6, 16, 19, 21, 22) but experimental measurements are often limited to detecting average water properties of the entire ensemble(23, 24). There is extensive literature on molecular dynamics (MD) simulation studies that probe hydration dynamics of different solution systems with heterogeneous water-protein interfaces(7–9, 11, 16, 17, 21, 25–33), but relatively few directly compare and validate with experimental techniques,(16, 17, 21, 22) in part, due to the difficulty of resolving *local* water dynamics experimentally. Undoubtedly, it is critical to leverage atomistic MD simulations synergistically with experiments to comprehensively and robustly characterize the heterogeneous hydration environments on soft material and biomolecular surfaces(1, 3, 4, 9, 17, 20)~~.~~

Advanced techniques such as quasi-elastic neutron scattering (QENS)(17, 34, 35), Overhauser Dynamic Nuclear Polarization (ODNP)(16, 21, 36), Terahertz (THz) spectroscopy and pump-probe infrared (IR) spectroscopy(41) can directly yield surface-specific water properties. Among those, only ODNP is sensitive to translational hydration water properties around localized sites or surfaces that are fully surrounded by bulk water in biological solution-state environments. Fundamentally, ODNP is a magnetic resonance technique that quantifies electron-^{1}H cross-relaxation by measuring (i) the enhancement of ^{1}H NMR signals induced by dynamic nuclear polarization (DNP) via the transfer of polarization from an unpaired electron of a nitroxide spin probe to the ^{1}H nucleus of water, and (ii) the longitudinal spin lattice relaxation time, T_1, reporting on all ^{1}H relaxation mechanisms induced by the dipolar coupling between the ^{1}H nuclear spin and the electron spin of the nearby spin probe. The so derived electron-^{1}H cross-relaxivity at a magnetic field of 0.35 Tesla and electron Larmor frequency of 9.8 GHz is sensitive to translational movement of hydration waters near the electron spin probe on instantaneous timescales ranging from about 10 picoseconds to nanoseconds and within 8 to 15 angstroms of nitroxide-based spin probes that can be tethered to specific surface sites.

ODNP is sensitive to chemically or topologically distinct local environments in highly heterogeneous surfaces immersed in bulk water, demonstrating that the diversity in these environments is reflected in a diversity of local water dynamics and structure. A recent synergistic ODNP-MD simulation study of sites with varying hydropathy on a CheY protein surface displayed a positive correlation between site-specific hydrophobicity and translation water dynamics[43]. This study demonstrated a connection between the ODNP spectroscopic quantities and computed thermodynamic properties of water. However, no study has directedly computed ODNP parameters for the purpose of deriving hydration water dynamical information and coupling to molecular dynamical, structural, and thermodynamic properties of hydration water near surfaces or solutes.

In the present work, we leverage molecular dynamics simulations to characterize the dynamical, structural, and thermodynamic properties of water in water-glycerol mixtures with increasing glycerol content [**Figure 2.1**]. Water-glycerol mixtures allow us to probe the complex interrelationship between water properties at the molecular scale while varying solution viscosity by a known quantity. Moreover, glycerol is of fundamental interest due to its ubiquity in biological studies for its role in cryopreservation of proteins and has been recently shown to alter not only the dynamical, but also structural and thermodynamic properties of water(43). Regardless of the system, the study of molecular determinants of surface hydration requires dual experimental and computational insight. Experimental techniques like ODNP offer valuable, but incomplete, insight on the molecular scale properties of hydration waters. Synergistic fully-atomistic simulations can aide in elucidating the molecular details that are often inaccessible to experimental methods.

In this study, we perform atomistic MD simulations of glycerol-water to test the ability of MD simulations to reproduce ODNP spectroscopic quantities and to connect ODNP spectroscopic and MD-derived translational, rotational, and hydrogen bonding dynamical quantities. We furthermore exploit the atomistic information content of MD simulations to quantify the relationship between hydration water dynamics, structure, and thermodynamics in glycerol-water mixtures across a wide range of compositions.

2.2. Molecular Dynamics Methods[4]

We perform molecular dynamics simulations using the OPC 4-site water model(44), the Blieck-Chelli (BC) model for glycerol(45–47) and a 4-Hydroxy-TEMPO nitroxide spin probe. Both the OPC water model and the BC model for glycerol accurately reproduce the diffusivities of pure water(44) and pure glycerol(45, 47) under ambient conditions (298.15 K and 1-bar). For the spin probe, its partial charges are obtained using the AMBER18 Antechamber package(48) informed by quantum chemical calculations using the Gaussian 16 software(49). All other inter- and intramolecular parameters derive from the second-generation generalized Amber forcefield (GAFF2)(50, 51). The results of this parametrization scheme yield similar parameters to those obtained in previous publications(52). All Coulombic interactions are modeled with the particle-mesh Ewald summation scheme (PME)(53).

We simulate glycerol-water-spin probe systems with glycerol mole fractions (x_{glyc}) of 0, 0.01, 0.033, 0.05, 0.075, 0.1, 0.15, 0.2 and 0.3 using the GPU-optimized OpenMM molecular simulation software(54). We first energy minimize each system, then equilibrate in the NPT ensemble using a Langevin thermostat(54) paired with a Monte Carlo barostat(54) at 290 K and 1 atm. Following equilibration, the NPT run continues for 250 ns with system configurations saved every 10 ns. Each saved configuration serves as the starting point for an independent 1 nanosecond NVE simulation for dynamic properties, with system coordinates saved every 0.1 ps. To calculate hydration water dynamics timescales and ODNP spectroscopic

---

[4] ODNP experimental methods will be detailed in the upcoming publication from which this chapter was derived.

quantities, we compute 95% confidence intervals by bootstrapping the results obtained from 20 independent MD simulations.

We characterize the effect of glycerol on solvation thermodynamics via the solvation free energy ΔG_{solv} of methane—an ideal small hydrophobic solute—in glycerol-water. To estimate ΔG_{solv} for this series of mixtures, we implement an expanded ensemble simulation procedure in which we gradually scale intermolecular interaction parameters between the methane molecule and the glycerol-water mixture. We smoothly scale Lennard Jones (LJ) and Columbic interaction parameters via a scalar parameter λ from $\lambda = 0$ (non-interacting, or ideal gas molecule) [panels (1) and (3) in **Figure 2.11(a)**] to 1 (fully interacting methane) [panels (2) and (4) in **Figure 2.11(a)**]. To estimate the Gibbs free energy of solvation ΔG_{solv} [**Figure 2.11(b)**], we then apply the multistate Bennett acceptance ratio (MBAR) method distributed via the pymbar(55) Python library.

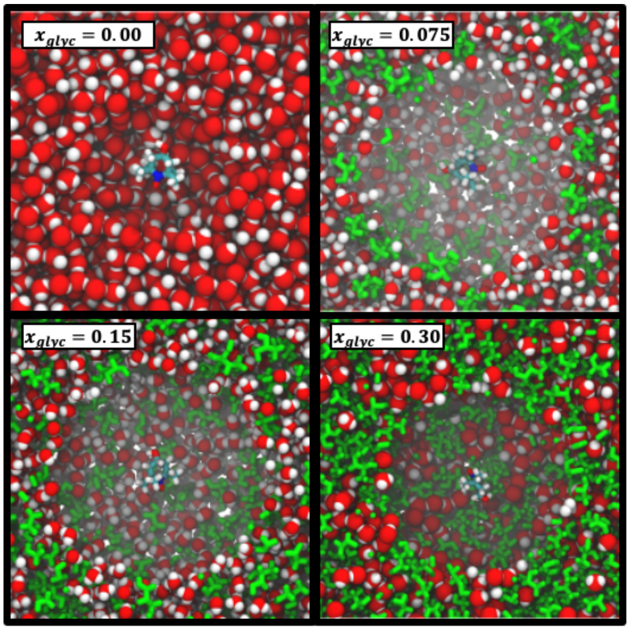


Figure 2.1 Snapshots of simulation boxes for a range of glycerol-water compositions. Here, OPC water molecules are represented as VDW spheres and glycerol is shown in green (licorice representation). 4-OH-TEMPO is not visible amidst the densely packed glycerol mixture.

2.3 Theory

2.3.1 Computing Spectroscopic Quantities

Dipolar autocorrelation functions. ODNP is an NMR technique that quantifies pairwise magnetic dipolar cross-relaxation between an electron spin on a free radical spin probe molecule (here of 4-Hydroxy-TEMPO) and the nuclear spin of the water proton. The spin-spin dipolar coupling energy between the radical electron and a water proton that are dynamically diffusing in solution state is governed by the following semi-classical Hamiltonian(56)

$$\hat{H}(r(t)) = \left[\left(\frac{2}{\sqrt{6}}\hat{I}_z\hat{S}_z - \frac{1}{2\sqrt{6}}\hat{I}_-\hat{S}_+ - \frac{1}{2\sqrt{6}}\hat{I}_+\hat{S}_-\right)F_2^{(0)}(\boldsymbol{r}(t))\right]c_\delta + \left[\left(-\frac{1}{2}\hat{I}_z\hat{S}_+ - \frac{1}{2}\hat{I}_+\hat{S}_z\right)F_2^{(1)}(\boldsymbol{r}(t)) + \left(\frac{1}{2}\hat{I}_z\hat{S}_- + \frac{1}{2}\hat{I}_-\hat{S}_z\right)F_2^{(1)}(\boldsymbol{r}(t))^*\right]c_\delta + \left[\frac{1}{2}\hat{I}_+\hat{S}_+F_2^{(2)}(\boldsymbol{r}(t)) + \frac{1}{2}\hat{I}_+\hat{S}_-F_2^{(2)}(\boldsymbol{r}(t))^*\right]c_\delta, \tag{2.1}$$

where $F_2^{(m)}(\boldsymbol{r}(t))$ (and $F_2^{(m)}(\boldsymbol{r}(t))^*$) are spherical harmonic functions (and their complex conjugates) [**Section 2.6.1**] that are dependent on the displacement vector between the electron spin and the proton nuclear spin, $\boldsymbol{r}(t)$. $\hat{I}_i$ and $\hat{S}_i$ are the quantum mechanical spin operators for the proton nuclear spin and electron spin, respectively. The constant $c_\delta = \mu_0\hbar\gamma_S\gamma_I/4\pi$ appears in **Equation 2.1** where $\mu_0 = 4\pi \times 10^{-7}\frac{N}{A^2}$, $\gamma_I = 4.26 \times 10^7\frac{Hz}{T}$, and $\gamma_S = \gamma_I/1.52 \times 10^{-3}$ are the vacuum permeability, gyromagnetic ratio of the proton spin, and the gyromagnetic ratio of the electron spin, respectively.

Using this semi-classical framework, we compute the equilibrium translational diffusion of water molecules by constructing the time-autocorrelation functions (ACFs), $C_{ODNP}^{(m)}(t)$, that depend purely on the relative positions of water hydrogen and the unpaired electron of a spin probe via the classical spherical harmonic functions included in **Equation 2.1**

$$C^{(m)}_{ODNP}(t) = \sum_i F_2^{(m)*}\left(\vec{r}_{O_r-H_w,i}(t)\right) F_2^{(m)}\left(\vec{r}_{O_r-H_w,i}(0)\right)$$

(2.2)

where $\vec{r}_{O_r-H_w,i}(t)$ is the displacement vector between the water hydrogen i and the oxygen radical of the spin probe and $m = 0, 1, 2$ the order, indicating the number of spin quanta undergoing flip-flop transition of the coupled I-S spin system. The $C^{(m)}_{ODNP}(t)$ functions decay with time as water molecules diffuse from their initial position relative to the spin probe, just as the dipolar coupling between the proton and electron spin weakens with increased spatial separation. The $C^{(m)}_{ODNP}(t)$ functions are complex-valued, but the complex part is negligible for isotropic systems(57) and thus ignored in this study [more detailed description in **Section 2.6.1**]. Further, the system isotropy ensures that $C^{(m)}_{ODNP}(t) = C^{(n)}_{ODNP}(t)$ for all n and m. Hence, we refer to only a single ACF for computing ODNP properties, $C_{ODNP}(t)$, for the remainder of the discussion. For the present work, we find that all measured $C_{ODNP}(t)$ are well-described by a tri-exponential fit:

$$C_{ODNP,fit}(t) = a_1 e^{-t/\tau_1} + a_2 e^{-t/\tau_2} + a_3 e^{-t/\tau_3} \qquad (2.3)$$

where $\tau_1 > \tau_2 > \tau_3$ and the coefficients, a_i $(i = 1,2,3)$, sum to unity. Note, that the fitting parameters are independently determined for each glycerol-water mixture.

Spectral density function. With analytical models for $C_{ODNP}(t)$ at hand [**Equation 2.3**], we next derive the spectral density function, $J(\omega)$, required to compute ODNP spectroscopic quantities for a given glycerol-water mixture. Specifically, we define the three spectral densities via the real part of the Fourier transform of $C_{ODNP,fit}(t)$ that contains three decay terms, and hence three □ values.

$$J(\omega;\{\tau_i\}_{i=1,2,3}) = Re\mathcal{F}\{C_{ODNP,fit}(t)\} = Re\mathcal{F}\{\sum_{i=1}^{3} a_i e^{-t/\tau_i}\} = \sum_{i=1}^{3} \frac{a_i\tau_i}{1+(\omega\tau_i)^2}$$

(2.4)

Here, the three terms in the sum account for the contribution of long (τ_1), intermediate (τ_2) and short (τ_3) timescales to local water diffusion. In the low-frequency regime ($\omega \ll \omega_S$), the spectral densities are more affected by the long timescale contribution (collective motion). On the other hand, the short timescale contribution (instantaneous motion) dominates the spectral densities at high frequency ($\omega \gg \omega_I$). To computationally derive the spectroscopic quantities measured by ODNP, we determine the transition rate via the amplitude of the spectral density, $J(\omega)$, for an electron spin-proton spin (e-n) transition that occurs at a given electron spin resonance (ESR) frequency, ω. Alternatively, $J(\omega)$ may be predicted using a theoretical continuum model(58)—the force-free hard sphere (FFHS) model—for water diffusion relative to the spin probe's radical electron. While previous work has demonstrated that the FFHS approach yields experimental water dynamics in qualitative agreement with simulation results(21, 59), the approach presented in **Equation 2.4** more directly determines the spectral density from e-n dipolar correlations with molecular detail and in dynamic systems that cannot be modeled by the FFHS model.

Figure 2.2(a) illustrates the dynamic spin-spin dipolar interaction between the unpaired electron of 4-OH-TEMPO and water protons underpinning the calculation of $C_{ODNP,fit}(t)$. In **Figure 2.2(b)**, we depict $C_{ODNP}(t)$ for a representative model system, a water-glycerol mixture with $x_{glyc} = 0.1$. We find that $C_{ODNP}(t)$ and its model fit, $C_{ODNP,fit}(t)$, are nearly indistinguishable with $R^2 = 0.99$. In **Table 2.1**, we summarize the tri-exponential fitting parameters for each glycerol-water mixture. **In Figure 2.2(c)**, we apply **Equation 2.4** to compute the spectral density function $J(\omega)$ from $C_{ODNP}(t)$ for the same mixture with $x_{glyc} = 0.1$. Here, the amplitude of the spectral density values at frequencies ω_I and ω_S [$J(\omega_I = 14.8\ MHz)$ and $J(\omega_S = 9.8\ GHz)$] are sensitive to translational dynamics on nanosecond and picosecond timescales, respectively.

ODNP spectroscopic probes. From the amplitudes of the spectral density function, we directly compute several spectroscopic quantities derived from ODNP experiments. One such quantity is the cross-relaxivity k_σ—the rate of transitions for the mutual flip of electron and proton spins in opposite directions—that strongly depends on the translational mobility of water on instantaneous timescales and within 8-15 angstroms of the spin probe(16, 36, 58). As defined in prior work(58, 60, 61), k_σ originates from the zero and double quantum transition of the dipolar coupled electron and proton spin pairs (expressed in the Hamiltonian of **Equation 2.1**), as given by

$$k_\sigma = \frac{c_\delta^2}{12C_{SL}}[\,6J(\omega_s + \omega_I) - J(\omega_S - \omega_I)] \qquad (2.5)$$

where C_{SL}, $\omega_S = \gamma_S B_0$, and $\omega_I = \gamma_I B_0$ are the molar spin label concentration [**Table S2.2**], Larmor precession frequency of the electron spin, and Larmor precession frequency of the proton spin, respectively. The ODNP experiments are conducted at a static magnetic field of $B_0 = 0.35\ T$, which sets $\omega_S = 9.8\ GHz$ and $\omega_I = 14.8\ MHz$. For the purposes of the present study, we determine the k_σ values from the simulation-derived spectral densities [**Equation 2.4**].

To provide an illustration, we depict the theoretical FFHS-derived cross-relaxivity, k_σ^{FFHS}, at $B_0 = 0.35\ T$ as a function of the correlation time in **Figure 2.2(d)**. The k_σ^{FFHS} values exhibit non-monotonic behavior with increasing translational correlation time, τ_c, of hydration waters. Specifically, k_σ^{FFHS} increases with correlation times up to $\tau_c \approx 100\mathrm{ps}$. Upon further retardation of translational dynamics, k_σ^{FFHS} decreases monotonically with increasing τ_c. Notably, this maximum in k_σ^{FFHS} is found at τ_c that is 3 times greater than the expected correlation time, $\tau_c \approx 33\mathrm{ps}$, for pure water at standard temperature (298.15 K) and pressure (1 bar) conditions (STP)(22). The precise location of the peak in k_σ^{FFHS} strongly depends on the Larmor precession frequencies by way of the applied magnetic field strength, B_0.

In systems that exhibit sufficiently slow water diffusion $[D_{H_2O,local} < \frac{1}{3} D_{H_2O}(@STP)]$, such as water hydrating catalyst support surfaces, hydrophilic materials interfaces(42), or protein surfaces and interiors(16), k_σ approximately linearly correlates with the diffusivity of hydration waters. However, for systems spanning a wide range of translational dynamics that includes systems that diffuse faster than pure water $[D_{H_2O,local} \approx D_{H_2O}(@STP)]$ or slower by up to a factor of three compared to pure water $[D_{H_2O,local} < \frac{1}{3} D_{H_2O}(@STP)]$ as found in bulk solutions of water and viologens, as in water-glycerol mixtures studied in the present work.

Another ODNP spectroscopic quantity of interest is the proton self-relaxivity k_ρ—the rate at which water proton polarization returns to thermal equilibrium modulated by the electron spin dipolar coupled to the proton nuclear spin, as follows:

$$k_\rho = \frac{c_\delta^2}{12C_{SL}}[J(\omega_S - \omega_I) + 3J(\omega_I) + 6J(\omega_S + \omega_I)] \qquad (2.6)$$

Unlike k_σ, k_ρ is the rate of transition for all proton nuclear spin flip events induced by dipolar coupling to the electron spins, not only the mutual proton-nuclear spin flips, and hence depends on the spectral density at the sum and difference of the electron and proton Larmor frequency and the proton Larmor frequency alone. The amplitude of the spectral density at the sum and difference frequencies, $\omega_S \pm \omega_I$, is close to simply the electron spin Larmor frequency, ω_S , and is sensitive to instantaneous translational dynamics correlation times of sub-nanoseconds, while the amplitude of the spectral density at the nuclear Larmor frequency, ω_I , is sensitive to longer 1-10-ns timescales for the slower, collective, motion of water. Because both k_σ and k_ρ contain the pre-factor $\frac{c_\delta^2}{12C_{SL}}$ the coupling factor that is defined by the ratio of k_σ and k_ρ eliminates the dependence on C_{SL}.

$$\xi = \frac{k_\sigma}{k_\rho} = \frac{6J(\omega_S+\omega_I)-J(\omega_S-\omega_I)}{J(\omega_S-\omega_I)+3J(\omega_I)+6J(\omega_S+\omega_I)} \qquad (2.7)$$

In experimental systems with difficult-to-quantify spin label concentration, the elimination of the pre-factor in **Equation 2.7** facilitates direct comparison between the experimental and MD-derived ξ. Conveniently, ξ is readily obtained from ODNP measurements of ^{1}H NMR signal enhancements and ^{1}H T_1 spin lattice relaxation times(36, 62, 63). As with k_ρ, ξ depends on both instantaneous and collective hydration water dynamics(22). When assuming simple diffusion as reflected in the FFHS model to determine the form of the spectral density function(22, 58), the value of ξ^{FFHS} monotonically depends on the characteristic correlation time for the translational diffusion dynamics of hydration water, simplifying its analysis, and as depicted in **Figure 2.2(d)**.

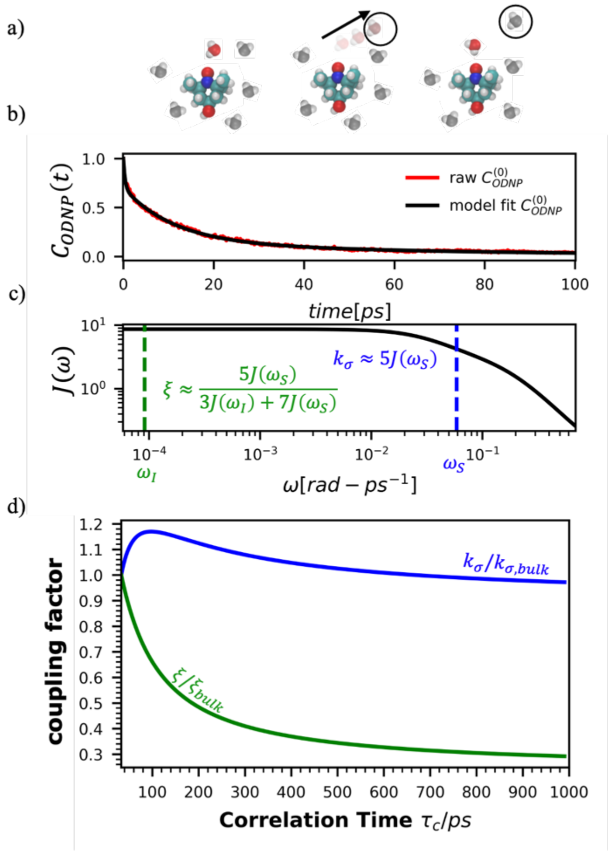


Figure 2.2: Schematic of ODNP spectroscopic quantities calculation from classical MD trajectories. (a) The snapshot shows *4-OH-TEMPO* in a 0.1 mole fraction glycerol mixture at 290K with bulk water represented in VMD as a medium and the hydration waters within 3.5 angstroms of the spin probe in VDW sphere representation. Glycerol molecules are omitted for clarity. The diagram to the right of the snapshot illustrates the ODNP mechanism for the nearest water molecule to the spin probe at some time t. (b) The ODNP time autocorrelation function $C_{ODNP}(t)$ at a glycerol mole fraction of 0.1 (red line) is fit to a tri-exponential model [**Equation 2.3**] (black line) as described in the text. (c) The real part of the Fourier transforms of $C_{ODNP,fit}(t)$ gives spectral density function $J(\omega)$. Subsequently, $J(\omega)$ values at radical electron $J(\omega_S)$ and proton Larmor frequencies $J(\omega_I)$ are identified on the plot by the blue and green vertical lines, respectively. Approximate functional forms of ODNP spectroscopic quantities k_σ and ξ are in blue and green text, respectively. (d) Based on the Force-Free Hard Sphere model*(58)*, we depict the expected trends in $k_\sigma/k_{\sigma,bulk}$ (blue line) and ξ/ξ_{bulk} (green line). Here, $k_{\sigma,bulk}$ and ξ_{bulk} are the expected bulk values of k_σ and ξ in pure STP water as defined by Franck et al.*(22)*

Proton spin-lattice relaxation times. We computationally extract the water proton spin-lattice relaxation times $T_{10}[0]$, which depend on both the system-average translational and rotational dynamics of water. We directly compute $T_{10}[0]$ according to the relationship presented by Bloembergen, Purcell and Pound(64) which assumes that effects of J-coupling, spin-rotation, and chemical shift anisotropy of the ^{1}H NMR signals are negligible, and that T_1 only depends on the system-average dynamics of water that is strongly modulated by the viscosity in a glycerol-water mixture(64). Because the primary relaxation mechanism involved in $T_{10}[0]$ is dipolar coupling between water protons modulated by water dynamics, we again compute the time autocorrelation functions, this time between the water protons $C_{T_{10}}^{(m)}(t) = \sum_i F_2^{(m)*}\left(\vec{r}_{H_w-H_w,i}(t)\right) F_2^{(m)}\left(\vec{r}_{H_w-H_w}, i(0)\right)$. The right-hand-side of the equation is identical in every way to **Equation 2.2**, but with the proton-proton displacement vector $\vec{r}_{H_w-H_w}$ replacing $\vec{r}_{O_r-H_w}$. To efficiently compute $C_{T_{10}}^{(m)}(t)$, we only consider displacement vectors $\vec{r}_{H_w-H_w,i}$ between a randomly chosen "probe" water proton and all other nearby water protons that the probe water proton encounters. Just as with $C_{ODNP}(t)$, the $C_{T_{10}}^{(m)}(t)$ decays monotonically as water protons diffuse away from the "probe" proton. As in **Equation 2.4**, we fit **Equation 2.6** to a multiexponential model and analytically Fourier transform the multiexponential $C_{T_{10}}^{(m)}(t)$ functions to obtain the spectral density functions, as follows:

$$K^{(m)}(\omega) = Re\mathcal{F}\left\{C_{T_{10}}^{(m)}(t)\right\} = Re\mathcal{F}\left\{\sum_{i=1}^{2} b_i^{(m)} e^{-t/\eta_i^{(m)}}\right\} = \sum_{i=1}^{2} \frac{b_i^{(m)}\eta_i^{(m)}}{1+\left(\omega\eta_i^{(m)}\right)^2} \quad (2.8)$$

where $b_1^{(m)}$, τ_1 and τ_2 (with $\tau_1 > \tau_2$) are fit parameters and $b_2^{(m)} = 1 - b_1^{(m)}$. We apply a bi-exponential model here because $C_{T_{10}}^{(m)}(t)$ decays much faster than $C_{ODNP}(t)$, allowing us to fit $C_{T_{10}}^{(m)}(t)$ with $R^2 > 0.95$ for the entire range of glycerol concentrations with only two exponentials, as shown in **Table S2.2**. Given these analytical spectral densities, we directly compute the longitudinal relaxation time using a formula from Bloembergen and coworkers(64)

$$\frac{1}{T_{10}[0]} = \frac{9}{8}\left(\frac{\mu_0 \hbar \gamma_I^2}{4\pi}\right)^2 \left[K^{(1)}(\gamma_I B_0) + \frac{1}{2}K^{(2)}(2\gamma_I B_0)\right] \tag{2.9}$$

Further, we physically interpret the translational and rotational contributions to $T_{10}[0]$ via the decomposition $T_{10}^{-1}[0] = (T_{10}^{-1})_{inter} + (T_{10}^{-1})_{intra}$. Here, $(T_{10}^{-1}[0])_{inter}$ and $(T_{10}^{-1}[0])_{intra}$ refer to the intermolecular and intramolecular contributions to $T_{10}^{-1}[0]$, respectively. The intramolecular contribution to the relaxation rate, $(T_{10}^{-1}[0])_{intra}$, can be ignored when using a rigid water model (such as the OPC 4-site model used in this study) that depends only on the system-average rotational dynamics of water given that the distance between the two hydrogen atoms is fixed. For the intermolecular contribution to $(T_{10}^{-1}[0])_{inter}$ between water, the spherical harmonic functions $F_2^{(m)*}\left(\vec{r}_{H_w - H_{w,i}}(t)\right)$ strictly depend on the displacement between water protons on separate water molecules, and not the re-orientation of the displacement vector between protons on the same water molecules. As such, $C_{T_{10}}^{(m)}(t)$ reduces to the same functional form as $C_{ODNP}(t)$ with water protons acting as probes (such as the electron in a nitroxide spin probe) for the surrounding water protons. Because water protons are indistinguishable from each other, the resulting correlation functions are manifestly system-average properties of water. The corresponding intermolecular relaxation rate, $(T_{10}^{-1}[0])_{inter}$, calculated from **Equations 2.8 and 2.9** is hence sensitive to the system-average translational dynamics of water.

2.3.2 Additional probes of water dynamics

Translational dynamics. We characterize the translational dynamics of water by computing the so-called survival probability $C_{survival}(t)$(32, 59) ACF that quantifies the timescale for waters to remain near the spin probe:

$$C_{survival}(t) = \frac{\sum_{i=1}^{N_w} S_i(0)S_i(t)}{\sum_{i=1}^{N_w} S_i(0)^2}$$

(2.10)

where N_w gives the number of water molecules in the simulation box, and $S_i(t)$ an indicator function that is 1 if the molecule i is inside a cutoff radius of 8-Å from the unpaired electron of the spin probe—approximately the width of the first two hydration shells around the radical oxygen. We apply absorbing boundary conditions such that only the water molecules that remain continuously within the cutoff radius from the initial time $t = 0$ to time t contribute to $C_{survival}(t)$. We find that neither choosing smaller cutoff radii (for instance, the first hydration shell near 5-Å) nor removing the absorbing boundary conditions qualitatively affect the trends in any ACF described here [**Figure 2.9**].

Rotational dynamics. To quantify the rotational dynamics of hydration waters, we compute the orientational ACF (OACFs)(26–28, 30, 65) that measures a characteristic time for water reorientation:

$$C_{OACF}^{(l)}(t) = \frac{1}{N_w}\sum_{i=1}^{N_w} P_l(\vec{u}_i(0) \cdot \vec{u_i}(t))$$

(2.11)

where N_w gives number of waters within the 8-Å cutoff radius at initial time $t = 0$, $P_l(\cdot)$ is the l-th Legendre polynomial function, and $\vec{u}_i(t)$ is the unit vector of water dipole i. For this present work, we only consider the second order OACF $C_{OACF}^{(2)}$ due to its relevance to longitudinal spin-relaxation rates.

Hydrogen bond dynamics. To probe the dynamics of water-water hydrogen bonding, we compute the hydrogen bond survival probability(29)

$$C_{HB}(t) = \frac{1}{N_{HB}} \sum_{i=1}^{N_{HB}} h_i(0) h_i(t) \tag{2.12}$$

where N_{HB} is the number of water-water hydrogen bonds containing waters within the cutoff radius at initial time $t = 0$, and $h_i(t)$ is a function that assumes a value of 1 if hydrogen bond i is intact at time t. We define hydrogen bonds via the widely-used geometric criteria of Luzar and Chandler(66), namely, distance and angular cutoff values of 3.5 angstroms and 120 degrees, respectively [see the inset schematic in **Figure 2.4(c)**].

Estimating relaxation time constants. To quantify the shifts in water's equilibrium dynamics with varying x_{glyc}, we compute several relaxation time constants: the ODNP derived translational diffusion correlation time (τ_{ODNP}), survival correlation time due to translational diffusion ($\tau_{survival}$), rotational diffusion correlation time (τ_{OACF}), and hydrogen bond correlation time (τ_{HB}). More specifically, we estimate these time constants by integrating the multiexponential fits to all ACFs detailed above

$$\tau_i = \int_{t=0}^{\infty} C_i(t) dt \tag{2.13}$$

where $i = ODNP, survival, OACF,$ or HB.

2.4 Results

2.4.1 Direct comparison of ODNP and MD-derived spectroscopic quantities

To directly probe the effect of hydration water retardation on the ODNP parameters (including $T_{10}[0]$), we systematically increase the solution viscosity by adding glycerol to water at mole percentages ranging from $x_{glyc} = 0$ to 0.3. In **Figure 2.3(a)**, we highlight the composition-dependent amplitude of the spectral density, $J(\omega)$, at the Larmor precession frequencies of protons ($\omega_I = \gamma_I B_0 \approx 14.8\, MHz$) and unpaired electrons ($\omega_S = \gamma_S B_0 \approx 9.8\, GHz$) at $B_0 = 0.35\, T$. The increase in $J(\omega_I)$ from pure water to $x_{glyc} = 0.3$ reflects on the increase in the relaxation rate of the proton spins, which according to **Equation 2.7**, leads to a monotonic decay of the coupling factor, as depicted in **Figure 2.3(b)**. The inset of **Figure 2.3(a)** shows the approximate spectral density contribution to the cross-relaxation rate of $5J(\omega_S)$. The approximation stems from the limit of $\omega_S \gg \omega_I$ and hence—by **Equation 2.5**—k_σ goes as $\frac{5c_\delta^2}{12C_{SL}} J(\omega_S)$. $5J(\omega_S)$ shows the same trend as the direct computation of k_σ [**Figure 2.3(c)**], increasing up to a glycerol concentration of $x_{glyc} = 0.1$ and decreasing at $x_{glyc} > 0.1$.

The MD simulation-derived relative coupling factor $\xi_r = \xi/\xi_{pure}$ exhibits a similar monotonic decrease as the experimental coupling factor with increasing glycerol concentration throughout the whole range of the glycerol-water mixture [**Figure 2.3(b)**]. Here, we normalize ξ by the coupling factor of pure water to better compare directly to the experimental results. Though simulation systematically underestimates the experimental ξ_r values by 5 to 60%, we observe qualitative agreement between both measures of ξ_r for $x_{glyc} < 0.30$. Notably, the 95% confidence interval (CI) broadens with increasing glycerol concentration. At $x_{glyc} = 0.3$, MD simulations underestimate ξ_r by 60%, with the experimental value lying outside the 95%

confidence interval (CI) of the MD-computed value. While ODNP measurements measure e-n dipolar correlations averaged over 10^{14} spin probes, our simulated systems contain a single spin probe molecule. Hence, we observe inherent noise in $C_{ODNP}(t)$ due to limited sampling of the dipolar interactions between the spin probe and the water nuclei. Increasing the concentration of glycerol exacerbates this sampling limitation due a systematic decrease in the average number of hydration waters. Increasing the NVE simulation length computes $C_{ODNP}(t)$ with improved resolution, which we assess by performing 20 additional independent 3-ns long (versus the original 1-ns) simulations at $x_{glyc} = 0.30$. While increasing NVE simulation time slightly decreases the width of the 95% CI on ξ_r (approximately by 6%), the simulations continue to systematically underestimate ξ_r at $x_{glyc} = 0.3$. This suggests that the disagreement between experiment and simulation ξ_r does not solely stem from under sampling of long timescale collective water motions.

Instead, the quantitative disagreement between the experimental and simulation results may stem in large part from the use of classical, fixed-charge molecular models. For instance, such MD models can only approximate interactions between water and 4-OH-TEMPO, while in reality the electron spin is delocalized between the N and O of the nitroxide radical, and the water hydrogen may polarize differently compared to bulk water, resulting in different O-H bond length and H-O-H angle. Further, OPC is a rigid water model and hence does not contain bonded (O-H) or angular (∠H-O-H) interaction terms. In particular, the inability to capture O-H vibrations with rigid water molecules likely affects the measurement of proton-nitroxide pair distances and thus the resulting spectroscopic quantities. Applying flexible(67, 68), or even polarizable(69, 70), water models may rectify some of the discrepancy between ODNP experiment and MD simulations.

Additionally, ξ_r is sensitive to the motion of bound waters—such as those buried within soft materials—on longer timescales (100-ps to 10-ns). While we do not expect these longer timescale dynamics to dramatically impact hydration water dynamics in glycerol-water mixtures, the magnitude of the slow time constant τ_1 exceeds 100-ps for $x_{glyc} = 0.3$. As the timescales for diffusive and viscous relaxation approach the nanosecond regime, accurately quantifying spectroscopic quantities becomes more difficult for simulations. For glycerol-water, we find that precise quantification of τ_1 becomes challenging for $x_{glyc} > 0.15$ [**Figure 2.8**].

In the inset of **Figure 2.3(a)**, we highlight the spectral density $J(\omega_S)$ at the Larmor precession frequency of the electron spin ($\omega_S = \gamma_S B_0 \approx 9.8\ GHz$) with increasing glycerol concentration. Notably, $J(\omega_S)$ is non-monotonic, increasing up to a glycerol concentration of $x_{glyc} = 0.1$, then decreasing at higher concentrations [**Figure 2.3(a)** inset]. Accordingly, the computed k_σ values in **Figure 2.3(c)** exhibit the same non-monotonic trend as $J(\omega_S)$ with increasing glycerol concentration. Notably, both the experimental and computed values of the relative cross-relaxivity $k_{\sigma,r} = k_\sigma / k_{\sigma,pure}$ exhibit the non-monotonic trends with increasing viscosity, with $k_{\sigma,r}$ values initially increasing up to a critical glycerol concentration ($x_{glyc} \approx 0.10$) and then decreasing thereafter [**Figure 2.3(c)**]. In the glycerol-water system with compositions from $x_{glyc} = 0$ to 0.1, the experimental and MD-derived $k_{\sigma,r}$ are in near quantitative agreement. Here, simulation underestimates $k_{\sigma,r}$, but the 95% CI of the MD-derived $k_{\sigma,r}$ values bracket the experimental results. For glycerol concentrations beyond $x_{glyc} = 0.1$, simulations again underestimate the $k_{\sigma,r}$, values, with the experimental $k_{\sigma,r}$

consistently lying above the 95%-CI of the simulation-derived values. We further note a dramatic increase in the relative error in computing $k_{\sigma,r}$ at high concentrations.

We attribute the underestimation of $k_{\sigma,r}$ to the inability of classical MD simulations to reproduce spectral density amplitudes at high frequency ($\omega \gg \omega_I$). We also note that the agreement between ODNP and MD results is much improved for ξ_r. We believe that this improved agreement results from the dominant influence of spectral density amplitudes at the lower frequencies [e.g., $J(\omega_I) \gg J(\omega_S)$] in the denominator of ξ [**Equation 2.5**]. The agreement of MD-derived values of both ξ_r and $k_{\sigma,r}$ with experiments for concentrations between $x_{glyc} = 0$ and 0.1 demonstrates that atomistic MD simulations accurately model trends in the translational dynamics of hydration waters on timescales of tens to hundreds of picoseconds.

We also verify the ability of MD simulations to reproduce spectroscopic measures of glycerol-water bulk dynamics via computation of $T_{10}[0]$. The relative proton longitudinal relaxation time, $(T_{10}[0])_r = T_{10}[0]/T_{10}[0]_{pure}$, exhibits a similar monotonic decrease as ξ_r with increasing glycerol content [**Figure 2.3(d)**]. Further, we observe a striking, near-quantitative agreement between experimental and MD-derived $(T_{10}[0])_r$ for $x_{glyc} < 0.10$. The MD-derived $(T_{10}[0])_r$ data systematically underestimate the experimental value for $x_{glyc} > 0.075$ with the experimental data falling outside of the 95% CI of the MD-calculated values. For $x_{glyc} \geq 0.1$, the computed values underestimate the experimental results by between 20 and 50%. In combination with the ξ_r and $k_{\sigma,r}$ results, the $(T_{10}[0])_r$ results suggest that classical atomistic MD simulations can reliably capture the low-frequency contributions to spectroscopic quantities, but less well the high-frequency contribution from instantaneous motion reflected in quantities such as k□.

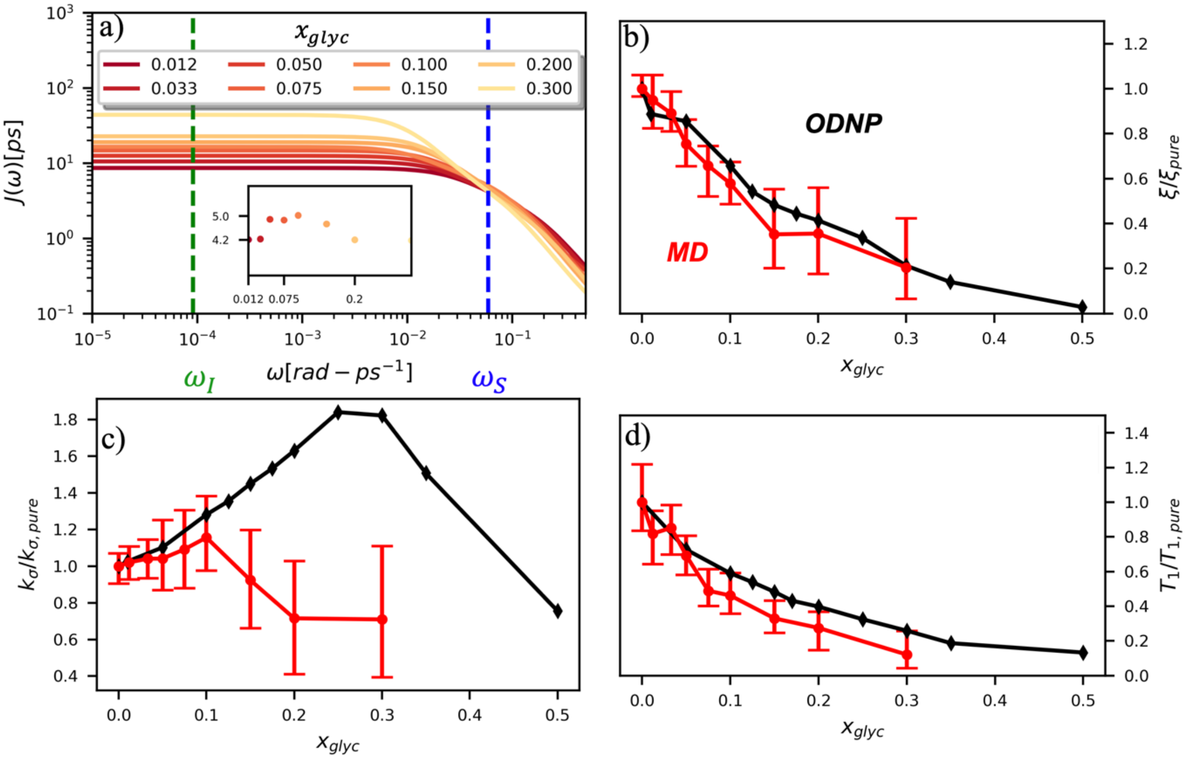


Figure 2.3: ODNP spectroscopic quantities measured experimentally and computed from classical MD simulations. (a) ODNP spectral density functions as a function of glycerol concentration where lighter-colored lines correspond to higher glycerol content with the Larmor precession frequency of the proton and radical electron indicated by the green and blue vertical lines, respectively. Comparing experimentally and computationally determined (b) coupling factor ξ, (c) cross-relaxivity k_σ, and (d) $T_{10}[0]$ as a function of glycerol content. ODNP experiments and MD simulations yield spectroscopic quantities with similar trends with increasing glycerol concentration.

2.4.2 Connecting other hydration water dynamics probes to ODNP measurements

By effectively modeling spectroscopic quantities, we can directly connect ODNP measurements to the microscopic dynamics and structural properties of water. MD simulations also enable computation of other modes of hydration water dynamics that cannot be measured experimentally. For instance, the coupling of these various modes of water dynamics have been studied extensively in liquid water(31, 71–73). One such example is the extended jump model of Laage and Hynes(26, 27), describing the process of water reorientation as dependent simultaneously on the breaking/forming of hydrogen bonds precipitated by large rotational jumps while being rate-limited by translational motion. In this work, we measure characteristic time scales for hydration water translation, rotation, and hydrogen bonding in the glycerol-water mixtures, as summarized in **Table S2.3**, using computed autocorrelation functions (ACFs): $C^{(0)}_{ODNP}(t)$, $C_{survival}(t)$, $C^{(2)}_{OACF}(t)$ and $C_{HB}(t)$. Further, we explicitly characterize hydration water dynamics via the different relaxation time constants for ODNP-derived diffusion (τ_{ODNP}), translational diffusion underlying survival probability ($\tau_{survival}$), rotational diffusion (τ_{OACF}), and hydrogen bonding (τ_{HB}).

To complement ODNP-derived measurements of water dynamics, we probe the translational mobility of hydration waters local to the spin probe via the survival probability ACF, $C_{survival}(t)$ [see **Section 2.3.2**]. **Figure 2.4(a)** depicts slower decay of $C_{survival}(t)$ as glycerol concentration increases, indicating the retardation of translational dynamics near the spin probe. We quantify this translational retardation via a characteristic time constant of translational diffusion $\tau_{survival}$ by fitting $C_{survival}(t)$ to a bi-exponential model $c_1 e^{-t/\alpha_1} + (1 - c_1)e^{-t/\alpha_2}$ and integrating over time. Here, c_1, α_1, and α_2 are the fitting parameters with $\alpha_1 > \alpha_1$. In agreement with the visible shift in the decay rate of $C_{survival}(t)$, $\tau_{survival}$ monotonically increases with glycerol concentration [**Figure 2.4(c)**]. There, we illustrate the simultaneous increase in $\tau_{survival}$ and decrease in ξ_r with glycerol concentration. This strong correlation [$R^2 = 0.96$] suggests a correspondence between the ODNP coupling factor and measures of the instantaneous translation dynamics of hydration water.

In addition to the translational mobility, we also examine the hydration water orientational dynamics via the orientational autocorrelation functions (OACFs) [see **Section 2.3.2**]. While we anticipate strong coupling between the translational and rotational mobility of water molecules in dilute solutions, the degree of translation-rotation coupling in crowded glycerol-water mixtures is less intuitive given that the breakdown of orientational-translational coupling has been previously observed for water under confinement(72). As seen in **Figure 2.4(b)**, $C_{OACF}^{(2)}$ decays more slowly with increasing glycerol content, signifying the retardation of rotational motion of hydration waters. This systematic slowdown of rotational dynamics is further illustrated by the monotonic increase in the characteristic time constants for rotational diffusion τ_{OACF}. As with $\tau_{survivial}$ we derive τ_{OACF} by fitting $C_{OACF}^{(2)}$ to a bi-exponential model. The simultaneous increase in characteristic timescales $\tau_{survivial}$ and τ_{OACF} [**Figure 2.4(e-f)**] across the whole range of glycerol concentrations establishes that, even in highly crowded glycerol-water mixtures, translational and rotational water diffusion remain strongly coupled. The persistent connection between rotational and translational dynamics suggests that—in contrast to water under geometric confinement (i.e., micelles)—the effect of glycerol on water stems from a distinct physical mechanism such as the development of a collective glycerol-water hydrogen bond network.

Glycerol is often assumed to simply decrease water translational dynamics due to increased mixture viscosity, but the effect of glycerol on water structural dynamics is typically not considered nor well understood. To investigate the relevance of such considerations, we probe the dynamics of hydrogen bonding of spin probe hydration waters via the hydrogen bond survival probability $C_{HB}(t)$. Like $C_{survival}(t)$ and $C_{OACF}^{(2)}(t)$, we report slower decay of $C_{HB}(t)$ with rising glycerol content [**Figure 2.4(c)**]. As with $\tau_{survival}$ and τ_{OACF}, we compute water-water hydrogen bond lifetimes τ_{HB} by integrating a bi-exponential fit to $C_{HB}(t)$. The resultant τ_{HB} values monotonically increase with increasing glycerol concentration, again in a manner strongly correlated with the ODNP results. The simultaneous increase in τ_{OACF} and τ_{HB}

suggests that glycerol enhances the lifetime of orientationally-coordinated and hydrogen bonded microstructures. Such an enhancement water structure with increasing glycerol content was demonstrated in our recently published study (published during preparation of this manuscript). In this work, we found that increased glycerol concentration not only slows the water diffusivity, but also enhances water orientational structure by increasing the tetrahedrality of water(74). We expand this analysis in **Section 2.4.3**, discussing connections between water tetrahedrality and various characteristic time constants in glycerol-water.

With these analyses, we elucidate the as-of-yet unknown persistent coupling between metrics for water translational, rotational, and hydrogen bonding dynamics in glycerol-water, including ODNP-derived and computational translational water diffusivity. We support this finding in **Figure 2.10** by depicting the connections between the hydration water dynamics relaxation times $\tau_{survival}$, τ_{HB}, τ_{OACF}, and τ_{ODNP} with spectroscopic quantities k_{σ}, ξ, and $T_{10}[0]$. Apart from k_{σ}, we observe strong correlation between each of these quantities and the others ($R^2 > 0.9$). The persistent connection between the characteristic time constants of hydration water dynamics suggests that glycerol, even in significant quantities, does not decouple the translational and rotational modes of water motion, in contrast to dynamic decoupling of water molecules reported under nanoscale confinements(72) and in supercooled water(71, 73, 75). In the context of nanoconfinement between silica planar surfaces, Romero-Vargas Castrillón et. al find that water rotational diffusion is bulk-like near the center of the nanochannel while translational diffusion is suppressed relative to bulk(72). The authors attribute this rotation-translation decoupling to the strong inverse dependences of translational diffusion on density (above bulk near the center channel) and the water-water hydrogen bond network (same as the bulk near the center channel). In contrast, increasing glycerol concentration yields simultaneous increases in mixture density and water-water network hydrogen bond lifetimes (increasing τ_{HB}), as well as water structuring as measured by tetrahedrality, relative to pure water.

Due to the non-monotonic behavior of k_σ [**Figure 2.3(c)**] with composition, it correlates less well ($R^2 < 0.9$) with the monotonically varying relaxation times and spectroscopic quantities. Traditionally, glycerol is thought to act as a classical viscogen without specifically considering its effect on water's molecular structure. With our simulation studies, we observe an increase in the rotational time constant τ_{OACF} and hydrogen bond lifetimes τ_{HB} that suggest tetrahedral enhancement with increasing glycerol concentration. To directly quantify tetrahedrality, we require a more direct structural metric, which we pursue next.

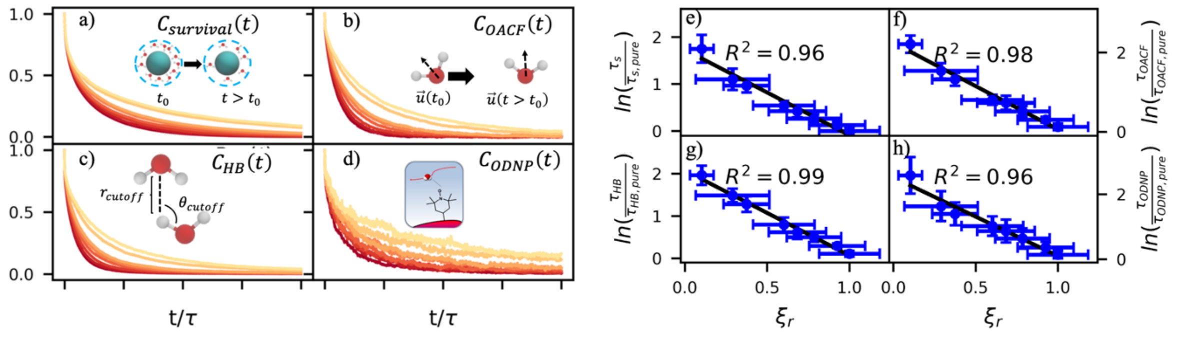


Figure 2.4: Molecular dynamics probes of hydration water dynamics correlate strongly with ODNP coupling factors. (a) The survival probability $C_{survival}(t)$ is the fraction of hydration shell waters that remain continuously within the second hydration shell of the spin probe radical oxygen. (b) The orientational autocorrelation function (OACF) $C_{OACF}(t)$ measures the rotation of hydration water dipole vectors away from their initial position. (c) The hydrogen bonding survival probability $C_{HB}(t)$ gives a time scale for water-water hydrogen bond breaking with a hydrogen bond being defined by cutoff radius r_{cutoff} and cutoff angle θ_{cutoff}. (d) The ODNP correlation function $C_{ODNP}(t)$ is used to estimate ODNP spectroscopic quantities. We derive characteristic time constants for (e) translational diffusion $\tau_{survival}$, (f) rotational diffusion τ_{OACF}, (g) hydrogen bond lifetimes, and (h) ODNP diffusion τ_{ODNP} by integrating bi-exponential model fits to the ACFs [(a), (b), (c), and (d), respectively]. Further, these time constants all correlate strongly with relative coupling factor ξ_r.

2.4.3 Glycerol Enhances the Population of Tetrahedral Waters

A number of order parameters have been used to assess the tetrahedrality of water in simulation studies(76–78). Many such parameters are influenced by the fact that water is undercoordinated near surfaces and in confined environments due to geometrical constraints(23, 76, 77, 79). For example, the tetrahedral order parameter, q, from Errington and Debenedetti(78) effectively characterizes water's tetrahedrality in many contexts, but in particular in dilute aqueous mixtures. However, the interpretation of q is more challenging when the average spatial separation between water molecules increases (e.g., concentrated glycerol-water mixtures) because it relies on four nearest water neighbors, which in such cases can populate non-first shell distances such that they adopt less correlated orientational order (decreased tetrahedrality).

As an alternative to q, three-body angle distributions avoid this geometric bias by consider only nearest neighboring water molecules within a distance cutoff of each central water molecule and computing the distribution of neighbor-central-neighbor triplet angles[1–3](76, 77) averaged over all simulation time steps. We can characterize the population of tetrahedrally coordinated waters, p_{tet}, by integrating $P(\theta)$ over the tetrahedral region of the distribution $p_{tet} = \int_{100^\circ}^{120^\circ} P(\theta) d\theta$. In addition to readily accommodating cases when water is under-coordinated, three-body angle distributions have been shown to aptly capture changes in water structure in response to shifts in thermodynamic properties(76, 77) and solute chemistry(79). Furthermore, shifts in three-body angle distribution have also proved to be predictive of solvation thermodynamics for a wide range of colloidal particle sizes(76).

In **Figure 2.5(a)**, we illustrate shifts in the populations of waters in tetrahedral and icosahedral (simple fluid) environments at 109.5 degrees and 64 degrees, respectively, that report on water's structural orientational environment. In **Figure 2.5(a)**, we show that the increase in glycerol concentration from $x_{glyc} = 0.01$ to 0.3 is accompanied by an increase of 2% in the overall p_{tet} and an equivalent decrease in the population of icosahedrally-coordinated waters. Curiously, prior work by Monroe and Shell demonstrated that increasing pure water density tended to decrease tetrahedrality as measured by the three body angle distribution,(76) while glycerol-water behaves in the opposite manner: as mixture density increases (increasing x_{glyc}), p_{tet} increases. This finding further supports that glycerol, rather than acting as a simple viscogen, enhances water's structural environment through enhanced water-glycerol hydrogen bonding.

Remarkably, the enhancement of water tetrahedrality strongly correlates with the retardation of hydration water dynamics probes. In **Figure 2.5(b-e)**, we discover nearly linear relationships ($R^2 > 0.96$) between p_{tet} and several probes of water dynamics, including logarithms of characteristic time constants for translational diffusion $\tau_{survival}$, rotational diffusion τ_{OACF}, hydrogen bonding lifetimes τ_{HB}, and the relative ODNP coupling factor ξ_r. This suggests that structural enhancements, driven by the addition of glycerol, impose systematic retardation of water dynamics. Moreover, the remarkably strong correlation between the dynamic ODNP coupling factor ξ_r and the structural metric p_{tet} suggests that ODNP measurements indirectly reports on the effect of mixture properties on tetrahedrality of water in bulk solution. The strong correspondence equilibrium water dynamics and water's molecular structure in glycerol-water is in keeping with a previous work by Shell and coworkers(74). In that work, we discovered that the system-average water self-diffusivity is accurately predicted by only two structural metrics—including p_{tet}. In the present work, we expand upon this observation finding strong relationships between not only the system-average translational dynamics of water, but local dynamics for water rotation and water-water hydrogen bonding.

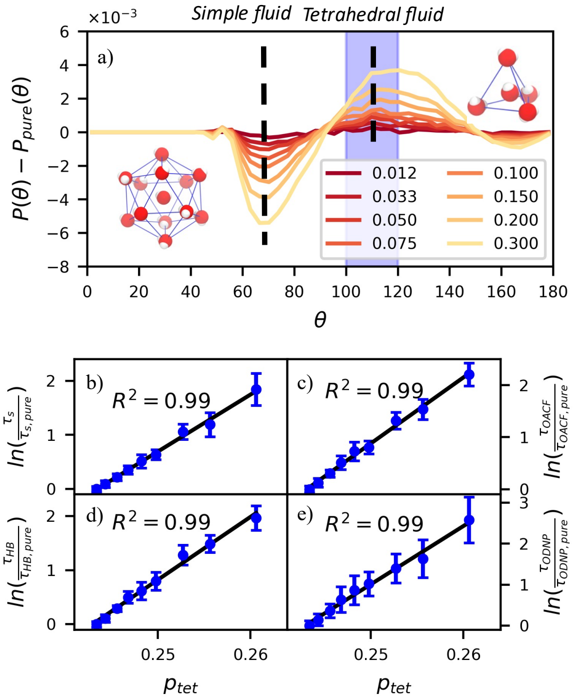


Figure 2.5: The three-body-angle distribution shows enhanced water tetrahedrality with increasing glycerol concentration. (a) Increasing glycerol concentration in the mixture increases the incidence of tetrahedrally-coordinated waters relative to pure water $[P(109.5°) - P_{pure}(109.5°)]$ while decreasing the incidence of icosahedrally-coordinated (simple-fluid like) waters $[P(64°) - P_{pure}(64°)]$. The increasing population of tetrahedral waters with glycerol concentration correlates strongly to the relative diffusivity of pure water $D_{H_2O}/D_{H_2O,pure}$ at a given mixture composition. Characteristic time constants for (b) translational diffusion $\tau_{survival}$, (c) rotational diffusion τ_2, (d) hydrogen bond lifetimes τ_{HB}, and (e) the relative coupling factor $\xi_r = \xi/\xi_{pure}$ correlate strongly with $R^2 > 0.99$ to the population of tetrahedral waters $p_{tet} = \int_{100°}^{120°} P(\theta)d\theta$.

2.4.4 Impact of Glycerol on Solvation Thermodynamics

We apply this understanding of water dynamics and structure to contextualize the glycerol-water thermodynamics by characterizing the tendency for a model small hydrophobic molecule (methane) to transfer from an ideal gas to a glycerol-water solution phase in the infinite dilution limit. We directly quantify this by computing the excess free energy of solvation ΔG_{solv}^{ex} as discussed in the **Section 2.2**. $\Delta G_{solv}^{ex} < 0$ suggests favorable solvation of a solute relative to ideal gas phase while the opposite is true for $\Delta G_{solv}^{ex} > 0$. In **Figure 2.11(b)**, we observe a monotonically increasing and positive ΔG_{solv}^{ex} with increasing x_{glyc} for $x_{glyc} < 0.1$. This increase in ΔG_{solv}^{ex} is in part explained by the enhancement of water structure [**Figure 2.5**] that presumably generates an increase in the entropic penalty for restructuring water-glycerol. However, for $x_{glyc} > 0.1$, ΔG_{solv}^{ex} begins to plateau to a constant $\Delta G_{solv}^{ex} \approx 4.25\ k_B T$. Notably, none of the structural or dynamic metrics above exhibit such a plateau.

To better understand the trend depicted in **Figure 2.11(b)**, we apply a free energy decomposition $\Delta G_{solv}^{ex} = \langle U \rangle_{sw} + S_{res}$ previously described by Monroe and Shell(76). Here, $\langle U \rangle_{sw}$ is the mean interaction energy between methane and glycerol-water in the solvated state (2) and S_{res} is strictly positive and largely gives the entropic penalty to create a void large enough to accommodate the solvation of methane from the gas phase (1) into the solution phase (2). We quantify $\langle U \rangle_{sw}$ by directly calculating the interaction energy between methane and glycerol-water $\langle U \rangle_{sw}$ while S_{res} follows from S_{res}=$\Delta G_{solv}^{ex} - \langle U \rangle_{sw}$. We observe a monotonic decrease in $\langle U \rangle_{sw}$ with increased glycerol concentration [**Figure 2.6(a)**], indicating a more favorable enthalpy of solvation for methane in glycerol-water. On the other hand, we demonstrate that S_{res} increases monotonically with glycerol concentration [**Figure 2.6(b)**], which indicates an increase in restructuring penalty to methane solvation. The entropic penalty ($S_{res} > 0$) and enthalpic gain ($\langle U \rangle_{sw} < 0$) for methane solvation in glycerol-water are consistent with the long-known hydration behavior of small hydrophobic molecules in aqueous environments(80, 81). As ΔG_{solv}^{ex} becomes less favorable (more positive) for $x_{glyc} < 0.1$, $\langle U \rangle_{sw}$ decreases such that further increases in S_{res} yield constant ΔG_{solv}^{ex} for $x_{glyc} > 0.1$.

While the excess solvation free energy, ΔG_{solv}^{ex}, does not show a simple correlation with dynamic or structural metrics, the entropic and energetic contributions display striking connections to dynamic metrics. Specifically, we note that the increasing penalty for restructuring glycerol-water S_{res} displays a strong negative correlation with the MD-computed relative coupling factor ξ_r [$R^2 = 0.99$; **Figure 2.6(c)**]. Given that increases in S_{res} are driven by an enhancement in the underlying glycerol-water solution structure, the connection between S_{res} and ξ_r again suggests a persistent structure-dynamics connection in glycerol-water. However, the $S_{res} - \xi_r$ relationship does not simply indirectly demonstrate what **Figure 2.5** does directly. In fact, this direct relationship between water dynamics and restructuring entropy is reminiscent of entropic theories for microscopic dynamics like the viscosity-entropy relationship proposed by Adam and Gibbs(82).

Adam-Gibbs theory proposes that retardation of liquid dynamics (such as liquid viscosity, η, or self-diffusivity, D) stems from a decrease in the number of available configurational states of a liquid via the configurational entropy S_{conf}. Notably, a recent study by Handle and Sciorno(83) demonstrated that a persistent $\ln D - \left(TS_{conf}\right)^{-1}$ connection—per Adam-Gibbs theory—in simulations of pure TIP4P/2005 water. The observed correlation between the probe of water dynamics (ξ_r) and restructuring entropy hints at a theoretical analog to Adam-Gibbs relating equilibrium water dynamics to solvation thermodynamics. If such a S_{res}-dynamics relationship persists for contexts beyond glycerol-water, it may be possible to forecast the solvation thermodynamics of small molecules in aqueous mixtures without computationally laborious free energy calculations. Probing dynamics-structure-thermodynamics correspondence at heterogeneous surfaces would be fascinating extensions of the present analysis. For instance, one could specifically interrogate how water-glycerol's molecular scale structure and dynamics mediate hydration free energies at protein-water interfaces to elucidate the cosolvent-mediated mechanisms of protein cryoprotection[84–87].

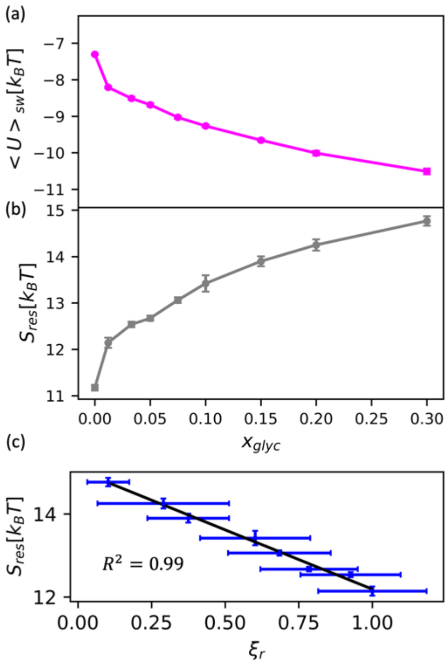


Figure 2.6: Decomposition of the solvation free energy of methane into glycerol-water mixtures. We calculate the solvation free energy for a methane molecule via expanded ensemble calculations, decomposing the resultant solvation free energy [**Figure 2.11(b)**] into (a) enthalpic contribution via the direct energy term $\langle U \rangle_{sw}$ and (b) entropy of solution restructuring S_{res}. (a) $\langle U \rangle_{sw}$ decreases as more glycerol is added to the mixture. (b) S_{res} increases as more glycerol is added to the mixture. (c) We observe a strong correlation ($R^2 = 0.99$) between the MD-computed ODNP coupling factor ξ_r and S_{res}.

2.5 Conclusions

In the present study, we reproduce critical ODNP spectroscopic measures of translational hydration water dynamics using classical atomistic molecular simulations to model equilibrium dynamics in glycerol-water mixtures with semi-quantitative agreement. Further, using the molecular scale detail revealed through MD simulations, we discover strong correlations between ODNP-measured coupling factors and computational probes of translational, rotational, and hydrogen bonding dynamics. The strong relationships found between MD-derived measures of water structure and dynamics are exciting for the potential of ODNP to serve as a surrogate probe of underlying solution structure. Finally, the clear connection between water tetrahedrality, water dynamics, and solvent restructuring entropy suggests a novel framework to describe and quantify the molecular scale mechanisms underlying hydrophobic hydration.

Though the existing literature on local water properties and water-mediated behavior is extensive, our analyses are unique because we directly link simulation to experiment and develop a dynamics-structure-thermodynamics connection. The generality of the experimental and computational methods discussed here will enable further investigations of other systems of broad interest such as water-alcohol and multicomponent mixtures used for protein stabilization (e.g., water-glycerol-DMSO). Further, the dynamics-structure-thermodynamics framework will aide further understanding of the molecular scale mechanisms (e.g., water tetrahedrality and diffusivity) underlying hydration properties in a wide range of chemically and topologically heterogeneous interfaces such as at protein-water interfaces.

2.6 Appendix

2.6.1 Further Details on spectral density calculations

Mathematical Form of the Spherical Harmonic Functions. As stated in the main text, we calculate the spectroscopic quantities ξ, k_σ and $T_{1,0}[0]$ via a time autocorrelation function of the form

$$C^{(m)}(t) =< F_2^{(m)*}\big(\vec{r}(t)\big)F_2^{(m)}\big(\vec{r}(0)\big) >= \frac{\sum_i^N F_l^{(m)*}\big(\vec{r}_i(t)\big)F_l^{(m)}\big(\vec{r}_i(0)\big)}{\sum_i^N F_l^{(m)*}\big(\vec{r}_i(0)\big)F_l^{(m)}\big(\vec{r}_i(0)\big)} \qquad (2.14)$$

where spherical harmonic functions $F_l^{(m)}\big(\vec{r}(t)\big)$ are

$$F_2^{(0)}\big(r(t)\big) = \sqrt{\frac{3}{2}}\frac{\left(r^2-3r_z^2\right)}{r^5}, \qquad (2.15)$$

$$F_2^{(1)}\big(r(t)\big) = 3\,\frac{r_z\left(r_x+ir_y\right)}{r^5}, \qquad (2.16)$$

and

$$F_2^{(2)}\big(r(t)\big) = -\frac{3}{2}\frac{\left(r_x - i r_y\right)^2}{r^5}. \qquad (2.17)$$

For ξ and k_σ, N and $\vec{r}_i(t)$ are the number of water hydrogens in the simulation box and the displacement vector between the *4-Hydroxy-TEMPO* radical oxygen and the i-th water hydrogen at time t, respectively. In the case of $T_{1,0}[0]$, N and $\vec{r}_i(t)$ are the number of water hydrogens in the simulation box (excluding the randomly chosen probe hydrogen) and the displacement vector between the probe water hydrogen and the i-th water hydrogen at time t, respectively.

Concentration dependence of the spectral density amplitudes. As shown indirectly in Figure 3(a) in the main text, we find that the spectral density functions at the Larmor precession frequency of the spin probe radical election $J^{(m)}(\omega_S)$ exhibits non-monotonic behavior as glycerol concentration increases. On the contrary, the spectral density functions at the precession frequency of a water proton $J^{(m)}(\omega_I)$ increases monotonically with glycerol concentration.

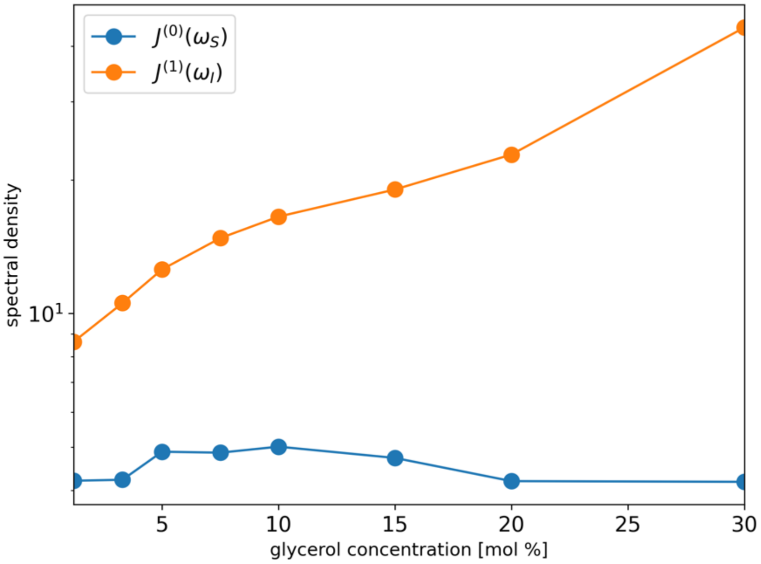


Figure 2.7. Comparing the amplitude of spectral density functions at ω_S (blue) and ω_I (orange) as a function of glycerol concentration. Amplitudes of the form $J^{(m)}(\omega_I)$ dominate across the entire range of concentrations, reducing the sensitivity of ξ to the non-monotonicity of amplitudes of the form $J^{(m)}(\omega_S)$.

Increase in the relative error of estimating long time behavior of $C^{(m)}_{ODNP}$. As discussed in the main text, we fit the ODNP autocorrelation functions to a tri-exponential model function

$$C^{(m)}_{ODNP,fit} = \mathrm{a}_1^{(m)}\mathrm{e}^{-\mathrm{t}/\tau_1^{(m)}} + \mathrm{a}_2^{(m)}\mathrm{e}^{-\mathrm{t}/\tau_2^{(m)}} + (1 - a_1^{(m)} - a_2^{(m)})\mathrm{e}^{-\mathrm{t}/\tau_3^{(m)}} \qquad (2.18)$$

where $\tau_1^{(m)} > \tau_2^{(m)} > \tau_3^{(m)}$.

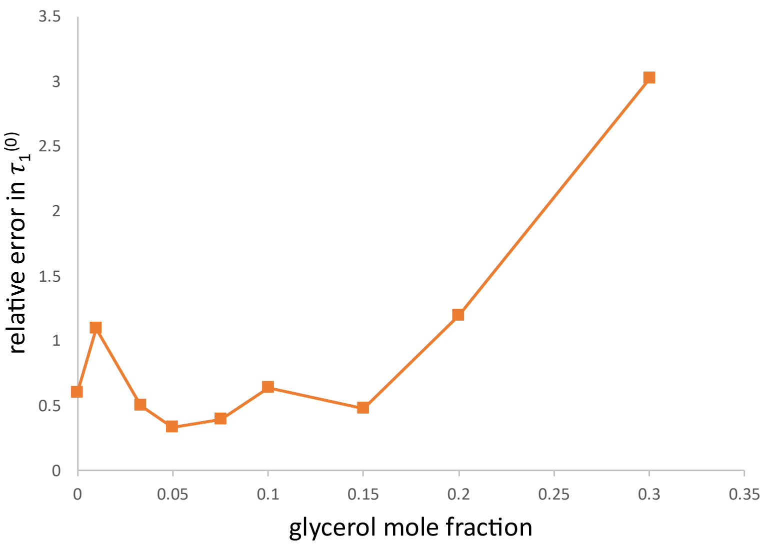


Figure 2.8. The relative error in the long timescale fitting parameter $\tau_1^{(0)}$ for the 0-th order ODNP correlation function $C_{ODNP,fit}^{(0)}$ increases dramatically for glycerol mole fractions higher than 0.15.

Table 2.1. Tri-exponential fitting function parameters for $C_{ODNP}^{(0)}$ at all simulated mixture compositions. Here, the uncertainties are 95% confidence intervals on the fitting parameters resulting from a bootstrapping procedure.

	$C_{ODNP}^{(0)}$ fitting parameters: $C_{ODNP,fit}^{(0)} = a_1 e^{-t/\tau_1} + a_2 e^{-t/\tau_2} + a_3 e^{-t/\tau_3}$					
x_{glyc}	a_1	a_2	a_3	τ_1/ps	τ_2/ps	τ_3/ps
0.00	$0.21_{-0.10}^{+0.12}$	$0.51_{-0.09}^{+0.09}$	$0.28_{-0.13}^{+0.16}$	$25.18_{-6.10}^{+10.61}$	$5.65_{-2.23}^{+1.93}$	$0.24_{-0.11}^{+0.16}$
0.01	$0.27_{-0.10}^{+0.07}$	$0.50_{-0.07}^{+0.07}$	$0.23_{-0.12}^{+0.10}$	$27.58_{-4.23}^{+7.54}$	$4.90_{-1.31}^{+3.61}$	$0.17_{-0.09}^{+0.67}$
0.033	$0.24_{-0.14}^{+0..12}$	$0.49_{-0.12}^{+0.09}$	$0.27_{-0.19}^{+0.15}$	$36.66_{-9.91}^{+29.88}$	$7.06_{-3.01}^{+4.63}$	$0.30_{-0.22}^{+1.02}$
0.05	$0.29_{-0.15}^{+0.08}$	$0.49_{-0.07}^{+0.07}$	$0.22_{-0.17}^{+0.11}$	$32.66_{-6.06}^{+19.94}$	$5.74_{-1.52}^{+4.62}$	$0.15_{-0.06}^{+0.56}$
0.075	$0.28_{-0.18}^{+0.18}$	$0.46_{-0.13}^{+0.20}$	$0.26_{-0.22}^{+0.27}$	$50.76_{-16.05}^{+88.97}$	$9.46_{-4.89}^{+7.05}$	$0.15_{-0.08}^{+0.29}$
0.10	$0.29_{-0.16}^{+0.14}$	$0.45_{-0.17}^{+0.18}$	$0.25_{-0.23}^{+0.23}$	$58.51_{-18.78}^{+19.56}$	$10.39_{-3.58}^{+6.80}$	$0.29_{-0.24}^{+0.63}$
0.15	$0.32_{-0.13}^{+0.13}$	$0.47_{-0.14}^{+0.15}$	$0.21_{-0.19}^{+0.20}$	$98.46_{-39.60}^{+75.27}$	$10.04_{-4.96}^{+10.89}$	$0.09_{-0.09}^{+0.77}$
0.20	$0.21_{-0.12}^{+0.21}$	$0.55_{-0.20}^{+0.15}$	$0.24_{-0.23}^{+0.26}$	$114.27_{-45.35}^{+149.44}$	$20.54_{-7.06}^{+7.59}$	$0.13_{-0.12}^{+0.46}$
0.30	$0.47_{-0.29}^{+0.16}$	$0.31_{-0.17}^{+0.17}$	$0.21_{-0.34}^{+0.23}$	$139.69_{-35.17}^{+133.72}$	$17.42_{-11.86}^{+49.01}$	$0.11_{-0.11}^{+1.88}$

Table 2.2. Comparing spectroscopic quantities obtained from ODNP experiments and MD simulations. Here, the MD uncertainties are 95% confidence intervals on the spectroscopic quantities resulting from a bootstrapping procedure. The uncertainties in the experimental values are the result of repeated measurements.

	MD Simulations				ODNP Experiments*		
x_{glyc}	$C_{SL}[mM]$	$k_\sigma/k_{\sigma,bulk}$	ξ/ξ_{bulk}	$T_{1,0}(0)$ $/T_{1,0}(0)_{bulk}$	$k_\sigma/k_{\sigma,bulk}$	ξ/ξ_{bulk}	$T_{1,0}(0)$ $/T_{1,0}(0)_{bulk}$
0.01	26.3	$1.03^{+0.12}_{-0.12}$	$0.97^{+0.12}_{-0.13}$	$0.817^{+0.22}_{-0.22}$	0.37 ± 0.01	0.51	0.87
0.05	22.8	$1.05^{+0.23}_{-0.20}$	$0.76^{+0.13}_{-0.13}$	$0.69^{+0.19}_{-0.16}$	0.40 ± 0.01	0.50	0.73
0.10	19.1	$1.17^{+0.25}_{-0.21}$	$0.59^{+0.14}_{-0.17}$	$0.46^{+0.17}_{-0.13}$	0.47 ± 0.01	0.38	0.60
0.15	16.2	$0.93^{+0.29}_{-0.28}$	$0.36^{+0.13}_{-0.12}$	$0.33^{+0.13}_{-0.10}$	0.53 ± 0.01	0.28	0.48
0.20	13.7	$0.72^{+0.32}_{-0.31}$	$0.28^{+0.22}_{-0.17}$	$0.28^{+0.11}_{-0.13}$	0.59 ± 0.01	0.24	0.40
0.30	10.1	$0.71^{+0.41}_{-0.32}$	$0.10^{+0.07}_{-0.05}$	$0.12^{+0.14}_{-0.08}$	0.66 ± 0.01	0.12	0.26

**The experimental spin-label concentration is held constant 200μM for each glycerol-water mixture.*

2.6.2 Additional information on MD-derived relaxation time constants

Table 2.3. Relaxation time constants for $C^{(0)}_{ODNP}(t)$, $C_{survival}(t)$, $C^{(2)}_{OACF}(t)$ *and* $C_{HB}(t)$. Here, the uncertainties are 95% confidence intervals on the fitting parameters resulting from a bootstrapping procedure.

	Relaxation time scales τ_i/ps			
x_{glyc}	$\tau^{(0)}_{ODNP}$	$\tau_{survival}$	τ_{OACF}	τ_{HB}
0.00	$8.22^{+0.67}_{-0.59}$	$14.33^{+0.48}_{-0.37}$	$2.72^{+0.12}_{-0.12}$	$6.57^{+0.24}_{-0.22}$
0.01	$9.90^{+1.68}_{-1.03}$	$15.71^{+0.63}_{-0.80}$	$3.07^{+0.21}_{-0.22}$	$7.30^{+0.43}_{-0.44}$
0.033	$12.02^{+2.19}_{-1.68}$	$17.79^{+0.88}_{-0.85}$	$3.68^{+0.21}_{-0.21}$	$8.82^{+0.34}_{-0.38}$
0.05	$14.51^{+3.59}_{-2.68}$	$20.45^{+1.39}_{-1.23}$	$4.54^{+0.57}_{-0.46}$	$10.82^{+1.12}_{-1.04}$

0.075	$18.85^{+6.36}_{-2.69}$	$23.89^{+2.80}_{-2.35}$	$5.64^{+0.98}_{-0.62}$	$12.19^{+1.94}_{-0.89}$
0.10	$21.65^{+5.90}_{-5.48}$	$26.98^{+2.26}_{-2.18}$	$6.08^{+0.73}_{-0.63}$	$14.58^{+1.70}_{-2.37}$
0.15	$35.11^{+19.22}_{-10.45}$	$41.17^{+5.70}_{-4.71}$	$10.17^{+1.60}_{-1.34}$	$23.64^{+4.29}_{-2.82}$
0.20	$49.52^{+22.61}_{-17.75}$	$47.10^{+10.45}_{-7.68}$	$12.62^{+2.39}_{-1.74}$	$29.05^{+4.53}_{-4.02}$
0.30	$72.64^{+19.51}_{-16.18}$	$90.17^{+26.43}_{-16.49}$	$24.75^{+4.50}_{-5.31}$	$46.79^{+10.53}_{-7.86}$

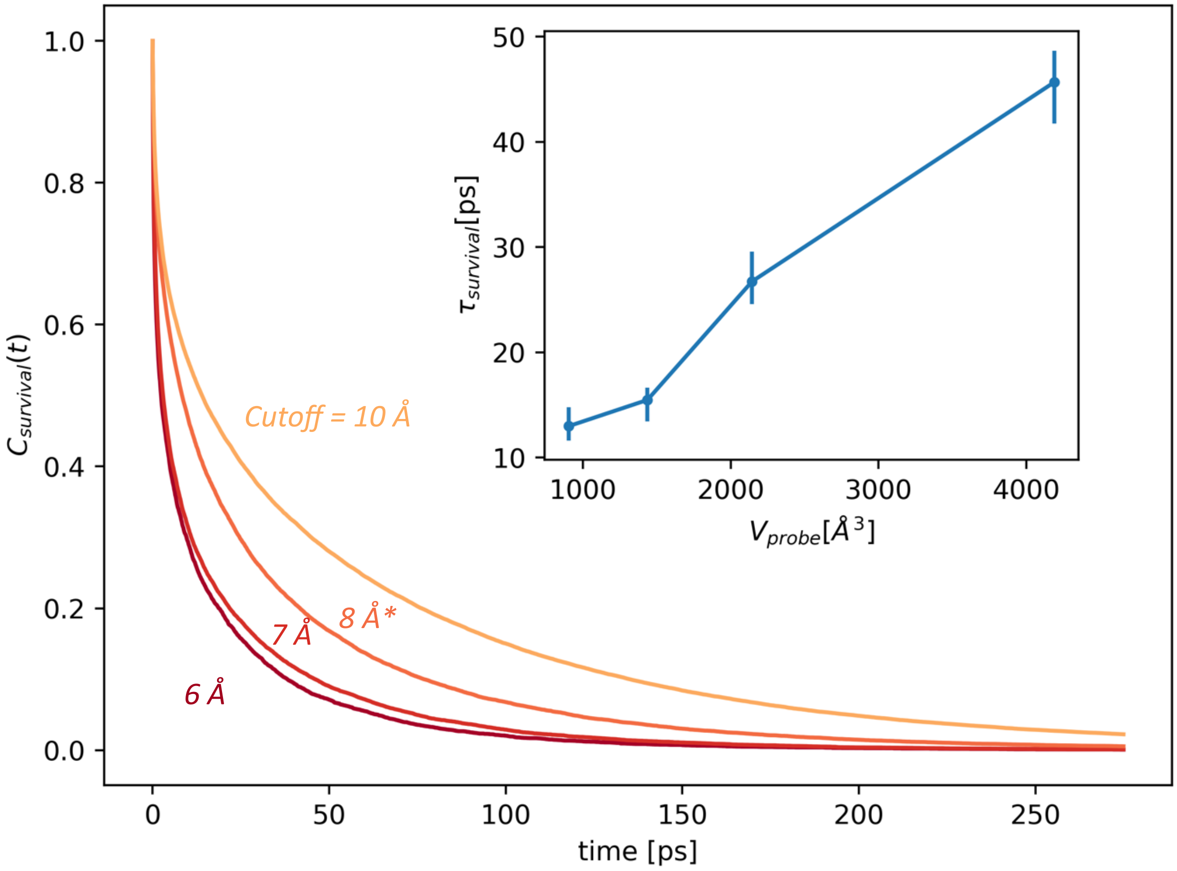


Figure 2.9. Investigating the effect of cutoff volume on the survival probability calculation for $x_{glyc} = 0.10$. We observe a systematic increase in the decay time for $C_{survival}(t)$ as the cutoff radius from the *4-Hydroxy TEMPO* radical oxygen is increased through values of: 6, 7, 8, and 10 Å. These cutoff radii correspond to spherical cutoff volumes of 904.77, 1436.76, 2144.66, and 4188.79 Å^3, respectively. In the inset plot, we show the monotonic increase in the relaxation timescale $\tau_{survival} = \int_0^\infty C_{survival}(t)dt$ with increasing cutoff volume.

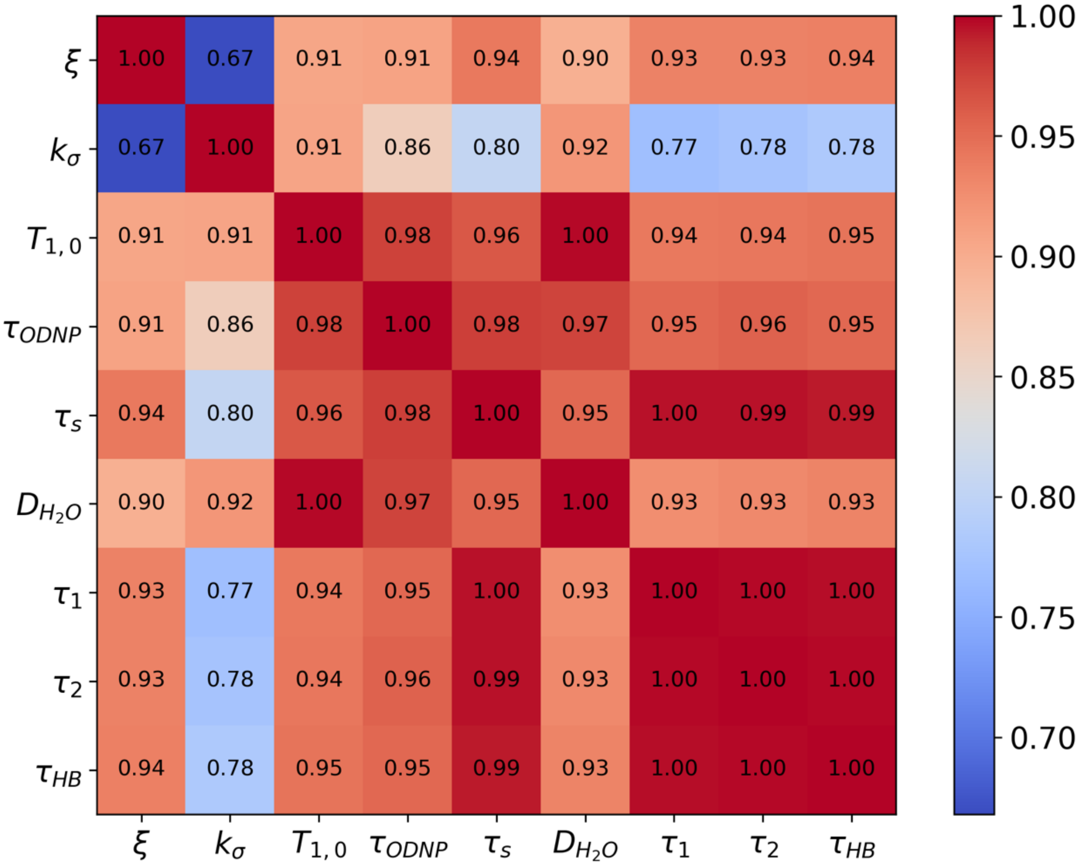


Figure 2.10. A cross-correlation heat map of all relaxation time constants and measurable dynamics probes. There is excellent correlation between all quantities ($R^2 > 0.9$) except for k_σ.

2.6.3 Decomposition of excess solvation free energy into solution restructuring and direct energetic terms

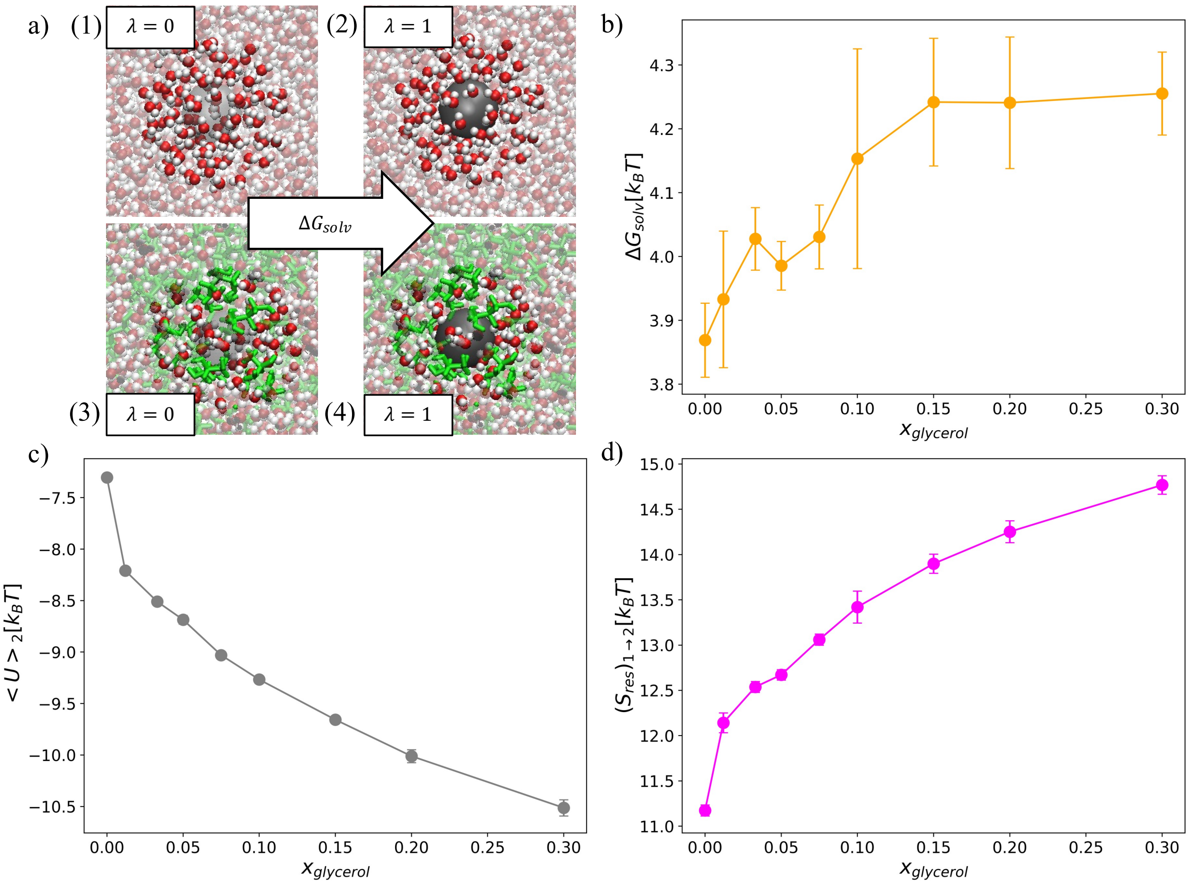


Figure 2.11. Decomposing the solvation free energy of methane into glycerol-water mixtures. (a) We depict expanded ensemble calculations schematically for glycerol mole fractions of 0 [panels (1) and (2)] and 0.1 [panels (3) and (4)]. The methane molecule is smoothly scaled from a non-interacting $\lambda = 0$ [panels (1) and (3)] to a fully interacting methane molecule $\lambda = 1$ [panels (2) and (4)]. (b) The solvation free energy, ΔG_{solv}, for bringing a methane from vacuum into solution with a glycerol-water mixture shows a non-monotonic trend with increasing glycerol content. The solvation free energy is decomposed into (c) enthalpic contribution via the direct energy term $< U >_2$ and (d) entropy of solvent restructuring $(S_{res})_{1\rightarrow 2}$. (c) $< U >_2$ decreases as more glycerol is added to the mixture. (d) $(S_{res})_{1\rightarrow 2}$ increases as more glycerol is added to the mixture.

References

1. Rego NB, Xi E, Patel AJ. 2021. Identifying hydrophobic protein patches to inform protein interaction interfaces. *Proc Natl Acad Sci USA*. 118(6):e2018234118
2. Jamadagni SN, Godawat R, Garde S. 2009. How Surface Wettability Affects the Binding, Folding, and Dynamics of Hydrophobic Polymers at Interfaces. *Langmuir*. 25(22):13092–99
3. Godawat R, Jamadagni SN, Garde S. 2009. Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proceedings of the National Academy of Sciences*. 106(36):15119–24

4. Monroe JI, Jiao S, Davis RJ, Robinson Brown D, Katz LE, Shell MS. 2021. Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution. *Proc Natl Acad Sci USA*. 118(1):e2020205118
5. Young T, Abel R, Kim B, Berne BJ, Friesner RA. 2007. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *PNAS*. 104(3):808–13
6. Wirtz H, Schäfer S, Hoberg C, Havenith M. 2018. Differences in Hydration Structure Around Hydrophobic and Hydrophilic Model Peptides Probed by THz Spectroscopy. *J Infrared Milli Terahz Waves*. 39(9):816–27
7. Moskowitz I, Snyder MA, Mittal J. 2014. Water transport through functionalized nanotubes with tunable hydrophobicity. *The Journal of Chemical Physics*. 141(18):18C532
8. Laage D, Stirnemann G, Hynes JT. 2009. Why Water Reorientation Slows without Iceberg Formation around Hydrophobic Solutes. *The Journal of Physical Chemistry B*. 113(8):2428–35
9. Monroe JI, Shell MS. 2018. Computational discovery of chemically patterned surfaces that effect unique hydration water dynamics. *Proceedings of the National Academy of Sciences*. 115(32):8093–98
10. Lum K, Chandler D, Weeks JD. 1999. Hydrophobicity at Small and Large Length Scales. *The Journal of Physical Chemistry B*. 103(22):4570–77
11. Mittal J, Hummer G. 2008. Static and dynamic correlations in water at hydrophobic interfaces. *Proceedings of the National Academy of Sciences*. 105(51):20130–35
12. Garde S, Hummer G, García AE, Pratt LR, Paulaitis ME. 1996. Hydrophobic hydration: Inhomogeneous water structure near nonpolar molecular solutes. *Physical Review E*. 53(5):R4310–13
13. Acharya H, Vembanur S, Jamadagni SN, Garde S. 2010. Mapping hydrophobicity at the nanoscale: Applications to heterogeneous surfaces and proteins. *Faraday Discussions*. 146:353–65
14. Xi E, Venkateshwaran V, Li L, Rego N, Patel AJ, Garde S. 2017. Hydrophobicity of proteins and nanostructured solutes is governed by topographical and chemical context. *PNAS*. 114(51):13345–50
15. Giovambattista N, Debenedetti PG, Rossky PJ. 2007. Hydration Behavior under Confinement by Nanoscale Surfaces with Patterned Hydrophobicity and Hydrophilicity. *J. Phys. Chem. C*. 111(3):1323–32
16. Barnes R, Sun S, Fichou Y, Dahlquist FW, Heyden M, Han S. 2017. Spatially Heterogeneous Surface Water Diffusivity around Structured Protein Surfaces at Equilibrium. *J. Am. Chem. Soc.* 139(49):17890–901
17. Schirò G, Fichou Y, Gallat F-X, Wood K, Gabel F, et al. 2015. Translational diffusion of hydration water correlates with functional motions in folded and intrinsically disordered proteins. *Nature Communications*. 6(1):
18. Davis JG, Gierszal KP, Wang P, Ben-Amotz D. 2012. Water structural transformation at molecular hydrophobic interfaces. *Nature*. 491(7425):582–85
19. Böhm F, Schwaab G, Havenith M. 2017. Mapping Hydration Water around Alcohol Chains by THz Calorimetry. *Angew. Chem. Int. Ed.* 56(33):9981–85
20. Rego NB, Xi E, Patel AJ. 2019. Protein Hydration Waters Are Susceptible to Unfavorable Perturbations. *J. Am. Chem. Soc.* 141(5):2080–86

21. Schrader AM, Monroe JI, Sheil R, Dobbs HA, Keller TJ, et al. 2018. Surface chemical heterogeneity modulates silica surface hydration. *Proceedings of the National Academy of Sciences*. 115(12):2890–95

22. Franck JM, Ding Y, Stone K, Qin PZ, Han S. 2015. Anomalously Rapid Hydration Water Diffusion Dynamics Near DNA Surfaces. *J. Am. Chem. Soc.* 137(37):12013–23

23. Monroe J, Barry M, DeStefano A, Aydogan Gokturk P, Jiao S, et al. 2020. Water Structure and Properties at Hydrophilic and Hydrophobic Surfaces. *Annual Review of Chemical and Biomolecular Engineering*. 11(1):523–57

24. Franck JM, Han S. 2019. Chapter Five - Overhauser Dynamic Nuclear Polarization for the Study of Hydration Dynamics, Explained. In *Methods in Enzymology*, ed AJ Wand. 615:131–75. Academic Press

25. Heyden M. 2019. Heterogeneity of water structure and dynamics at the protein-water interface. *J. Chem. Phys.* 150(9):094701

26. Laage D. 2006. A Molecular Jump Mechanism of Water Reorientation. *Science*. 311(5762):832–35

27. Laage D, Hynes JT. 2008. On the Molecular Mechanism of Water Reorientation. *J. Phys. Chem. B*. 112(45):14230–42

28. Xiao S, Figge F, Stirnemann G, Laage D, McGuire JA. 2016. Orientational Dynamics of Water at an Extended Hydrophobic Interface. *J. Am. Chem. Soc.* 138(17):5551–60

29. Martiniano HFMC, Galamba N. 2013. Insights on Hydrogen-Bond Lifetimes in Liquid and Supercooled Water. *J. Phys. Chem. B*. 117(50):16188–95

30. Stirnemann G, Castrillón SR-V, Hynes JT, Rossky PJ, Debenedetti PG, Laage D. 2011. Non-monotonic dependence of water reorientation dynamics on surface hydrophilicity: competing effects of the hydration structure and hydrogen-bond strength. *Physical Chemistry Chemical Physics*. 13(44):19911

31. Tan P, Huang J, Mamontov E, García Sakai V, Merzel F, et al. 2020. Decoupling between the translation and rotation of water in the proximity of a protein molecule. *Phys. Chem. Chem. Phys.* 22(32):18132–40

32. Qiao B, Jiménez-Ángeles F, Nguyen TD, Olvera de la Cruz M. 2019. Water follows polar and nonpolar protein surface domains. *Proc Natl Acad Sci USA*. 116(39):19274–81

33. Dahanayake JN, Mitchell-Koch KR. 2018. Entropy connects water structure and dynamics in protein hydration layer. *Physical Chemistry Chemical Physics*. 20(21):14765–77

34. Chen S-H, Gallo P, Sciortino F, Tartaglia P. 1997. Molecular-dynamics study of incoherent quasielastic neutron-scattering spectra of supercooled water. *Physical Review E*. 56(4):4231–43

35. Perrin J-C, Lyonnard S, Volino F. 2007. Quasielastic Neutron Scattering Study of Water Dynamics in Hydrated Nafion Membranes. *The Journal of Physical Chemistry C*. 111(8):3393–3404

36. Armstrong BD, Han S. 2009. Overhauser Dynamic Nuclear Polarization To Study Local Water Dynamics. *Journal of the American Chemical Society*. 131(13):4641–47

37. Xu Y, Havenith M. 2015. Perspective: Watching low-frequency vibrations of water in biomolecular recognition by THz spectroscopy. *J. Chem. Phys.* 143(17):170901

38. Zhang J. 2019. Water Dynamics in the Hydration Shell of Amphiphilic Macromolecules. *J. Chem. Phys. B*. 123(13):2971–77

39. Hoshina H, Iwasaki Y, Katahira E, Okamoto M, Otani C. 2018. Structure and dynamics of bound water in poly(ethylene-vinylalcohol) copolymers studied by terahertz spectroscopy. *Polymer*. 148:49–60

40. Kim SJ, Born B, Havenith M, Gruebele M. 2008. Real-Time Detection of Protein-Water Dynamics upon Protein Folding by Terahertz Absorption Spectroscopy. *Angew. Chem. Int. Ed.* 47(34):6486–89

41. Lawrence CP, Skinner JL. 2003. Vibrational spectroscopy of HOD in liquid D2O. III. Spectral diffusion, and hydrogen-bonding and rotational dynamics. *The Journal of Chemical Physics*. 118(1):264–72

42. Moon H, Collanton RP, Monroe JI, Casey TM, Shell MS, et al. 2022. Evidence for Entropically Controlled Interfacial Hydration in Mesoporous Organosilicas. *J. Am. Chem. Soc.* 144(4):1766–77

43. Mahanta DD, Brown DR, Pezzotti S, Han S, Schwaab G, et al. 2023. Local solvation structures govern the mixing thermodynamics of glycerol–water solutions. *Chem. Sci.*

44. Izadi S, Anandakrishnan R, Onufriev AV. 2014. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* 5(21):3863–71

45. Blieck J, Affouard F, Bordat P, Lerbret A, Descamps M. 2005. Molecular dynamics simulations of glycerol glass-forming liquid. *Chemical Physics*. 317(2–3):253–57

46. Chelli R, Procacci P, Cardini G, Califano S. 1999. Glycerol condensed phases Part II.A molecular dynamics study of the conformational structure and hydrogen bonding. *Physical Chemistry Chemical Physics*. 1(5):879–85

47. Jahn DA, Akinkunmi FO, Giovambattista N. 2014. Effects of Temperature on the Properties of Glycerol: A Computer Simulation Study of Five Different Force Fields. *The Journal of Physical Chemistry B*. 118(38):11284–94

48. Case DA. 2018. AMBER 2018

49. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, et al. 2016. Gaussian 16 Rev. B.01

50. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. 2004. Development and testing of a general amber force field. *Journal of Computational Chemistry*. 25(9):1157–74

51. *Development of the Second Generation of the General AMBER Force Field*. ResearchGate. www.researchgate.net

52. Sezer D, Freed JH, Roux B. 2008. Parametrization, Molecular Dynamics Simulation, and Calculation of Electron Spin Resonance Spectra of a Nitroxide Spin Label on a Polyalanine α-Helix. *The Journal of Physical Chemistry B*. 112(18):5755–67

53. Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*. 98(12):10089–92

54. Eastman P, Pande V. 2010. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science & Engineering*. 12(4):34–39

55. Shirts MR, Chodera JD. 2008. Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of Chemical Physics*. 129(12):124105

56. Hausser KH, Stehlik D. 1968. Dynamic Nuclear Polarization in Liquids. In *Advances in Magnetic and Optical Resonance*. 3:79–139. Elsevier

57. Sezer D, Gafurov M, Prandolini MJ, Denysenkov VP, Prisner TF. 2009. Dynamic nuclear polarization of water by a nitroxide radical: rigorous treatment of the electron spin saturation and comparison with experiments at 9.2 Tesla. *Physical Chemistry Chemical Physics*. 11(31):6638

58. Hwang L-P, Freed JH. 1975. Dynamic effects of pair correlation functions on spin relaxation by translational diffusion in liquids. *The Journal of Chemical Physics*. 63(9):4017
59. Ortony JH, Qiao B, Newcomb CJ, Keller TJ, Palmer LC, et al. 2017. Water Dynamics from the Surface to the Interior of a Supramolecular Nanostructure. *J. Am. Chem. Soc.* 139(26):8915–21
60. Sezer D, Prandolini MJ, Prisner TF. 2009. Dynamic nuclear polarization coupling factors calculated from molecular dynamics simulations of a nitroxide radical in water. *Physical Chemistry Chemical Physics*. 11(31):6626
61. Franck JM, Pavlova A, Scott JA, Han S. 2013. Quantitative cw Overhauser effect dynamic nuclear polarization for the analysis of local water dynamics. *Progress in Nuclear Magnetic Resonance Spectroscopy*. 74:33–56
62. Armstrong BD, Soto P, Shea J-E, Han S. 2009. Overhauser dynamic nuclear polarization and molecular dynamics simulations using pyrroline and piperidine ring nitroxide radicals. *Journal of Magnetic Resonance*. 200(1):137–41
63. Armstrong BD, Han S. 2007. A new model for Overhauser enhanced nuclear magnetic resonance using nitroxide radicals. *The Journal of Chemical Physics*. 127(10):104508
64. Bloembergen N, Purcell EM, Pound RV. 1948. Relaxation Effects in Nuclear Magnetic Resonance Absorption. *Phys. Rev.* 73(7):679–712
65. Shin S, Willard AP. 2018. Characterizing Hydration Properties Based on the Orientational Structure of Interfacial Water Molecules. *Journal of Chemical Theory and Computation*. 14(2):461–65
66. Luzar A, Chandler D. 1993. Structure and hydrogen bond dynamics of water–dimethyl sulfoxide mixtures by computer simulations. *The Journal of Chemical Physics*. 98(10):8160–73
67. Wu Y, Tepper HL, Voth GA. 2006. Flexible simple point-charge water model with improved liquid-state properties. *The Journal of Chemical Physics*. 124(2):024503
68. González MA, Abascal JLF. 2011. A flexible model for water based on TIP4P/2005. *The Journal of Chemical Physics*. 135(22):224516
69. Gartner TE, Hunter KM, Lambros E, Caruso A, Riera M, et al. 2022. Anomalies and Local Structure of Liquid Water from Boiling to the Supercooled Regime as Predicted by the Many-Body MB-pol Model. *J. Phys. Chem. Lett.* 13(16):3652–58
70. Babin V, Leforestier C, Paesani F. 2013. Development of a “First Principles” Water Potential with Flexible Monomers: Dimer Potential Energy Surface, VRT Spectrum, and Second Virial Coefficient. *J. Chem. Theory Comput.* 9(12):5395–5403
71. Dueby S, Dubey V, Daschakraborty S. 2019. Decoupling of Translational Diffusion from Viscosity of Supercooled Water: Role of Translational Jump-Diffusion. *The Journal of Physical Chemistry B*
72. Romero-Vargas Castrillón S, Giovambattista N, Aksay IA, Debenedetti PG. 2009. Effect of Surface Polarity on the Structure and Dynamics of Water in Nanoscale Confinement. *The Journal of Physical Chemistry B*. 113(5):1438–46
73. Dubey V, Erimban S, Indra S, Daschakraborty S. 2019. Understanding the Origin of the Breakdown of the Stokes–Einstein Relation in Supercooled Water at Different Temperature–Pressure Conditions. *The Journal of Physical Chemistry B*. 123(47):10089–99
74. Robinson Brown DC, Webber TR, Jiao S, Rivera Mirabal DM, Han S, Shell MS. 2023. Relationships between Molecular Structural Order Parameters and Equilibrium Water Dynamics in Aqueous Mixtures. *J. Phys. Chem. B*. 127(20):4577–94

75. Debenedetti PG, Stillinger FH. 2001. Supercooled liquids and the glass transition. *Nature*. 410(6825):259–67

76. Monroe JI, Shell MS. 2019. Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *The Journal of Chemical Physics*. 151(9):094501

77. Chaimovich A, Shell MS. 2014. Tetrahedrality and structural order for hydrophobic interactions in a coarse-grained water model. *Physical Review E*. 89(2):

78. Errington JR, Debenedetti PG. 2001. Relationship between structural order and the anomalies of liquid water. *Nature*. 409(6818):318–21

79. Stock P, Monroe JI, Utzig T, Smith DJ, Shell MS, Valtiner M. 2017. Unraveling Hydrophobic Interactions at the Molecular Scale Using Force Spectroscopy and Molecular Dynamics Simulations. *ACS Nano*. 11(3):2586–97

80. Gallicchio E, Kubo MM, Levy RM. 2000. Enthalpy−Entropy and Cavity Decomposition of Alkane Hydration Free Energies: Numerical Results and Implications for Theories of Hydrophobic Solvation. *J. Phys. Chem. B*. 104(26):6271–85

81. Frank HS, Evans MW. 1945. Free Volume and Entropy in Condensed Systems III. Entropy in Binary Liquid Mixtures; Partial Molal Entropy in Dilute Solutions; Structure and Thermodynamics in Aqueous Electrolytes. *The Journal of Chemical Physics*. 13(11):507–32

82. Adam G, Gibbs JH. 1965. On the Temperature Dependence of Cooperative Relaxation Properties in Glass-Forming Liquids. *The Journal of Chemical Physics*. 43(1):139–46

83. Handle PH, Sciortino F. 2018. The Adam–Gibbs relation and the TIP4P/2005 model of water. *Molecular Physics*. 116(21–22):3366–71

84. Towey JJ, Soper AK, Dougan L. 2013. What happens to the structure of water in cryoprotectant solutions? *Faraday Discuss.* 167:159

85. Weng L, Chen C, Zuo J, Li W. 2011. Molecular Dynamics Study of Effects of Temperature and Concentration on Hydrogen-Bond Abilities of Ethylene Glycol and Glycerol: Implications for Cryopreservation. *J. Phys. Chem. A*. 115(18):4729–37

86. Weng L, Stott SL, Toner M. 2019. Exploring Dynamics and Structure of Biomolecules, Cryoprotectants, and Water Using Molecular Dynamics Simulations: Implications for Biostabilization and Biopreservation. *Annu. Rev. Biomed. Eng.* 21(1):1–31

87. Towey JJ, Soper AK, Dougan L. 2012. Molecular Insight Into the Hydrogen Bonding and Micro-Segregation of a Cryoprotectant Molecule. *The Journal of Physical Chemistry B*. 116(47):13898–904

Chapter 3: Multi-length scale water dynamics in aqueous PEO solutions

Joshua D. Moon[5], Thomas R. Webber, Dennis C. Robinson Brown, Pete M. Richardson, Thomas M. Casey, Rachel A. Segalman, M. Scott Shell, Songi Han. *Nanoscale water-polymer interactions tune macroscopic diffusivity of water in aqueous poly(ethylene oxide) solutions.* In Review at Chemical Science

3.1 Introduction

Polymer membrane materials play a critical role in developing energy-efficient water purification processes(1). Membranes selectively permeate water over solutes across various length scales from ions to microbes. The permeability of water and selectivity for water over other solutes in membranes are profoundly affected by the complex interplay of dynamic and thermodynamic properties of water in a given membrane (2–6). Macroscale water transport through porous filtration technologies like ultra- and nanofiltration is primarily controlled by membrane pore morphology (e.g., pore geometry and tortuosity)(7, 8). In non-porous media such as in reverse osmosis membranes, water transport has primarily been connected to water's molecular structure in the hydrated membrane. The dependence of non-porous membrane performance on molecular scale phenomena is suggested by the success of free volume and molecular obstruction models for transport (9–12).

Further, previous work has establish that the dynamics of water in polymer solutions is mediated by water-polymer surface interactions(13–18). While the connection between water-surface interactions and water dynamics is well-established in solution, the molecular scale properties of water at polymer surfaces in hydrated membranes are infrequently

---

[5] First-authorship shared between J.D.M., T.R.W., and D.C.R.B

considered and poorly understood. Through this chapter, we directly probe structural and dynamics water properties in the system-average and at the hydration later (a “molecular surface”) of linear poly (ethylene oxide) (OEO) in water. We study linear OEO for practical and application-driven reasons. Practically speaking, OEO is miscible with water and presents a relatively simple repeat unit structure. Heuristically, OEO-based coatings are widely leveraged to imbue materials in the context of water purification membranes, consumer products, and biomedicine(19–23) with an antifouling character. Though PEO’s chemical structure is nominally simple, even short EO oligomer chains explore a complex configurational landscape(24, 25) and eutectic formations over a wide range of water-PEO compositions(26). Several PEO-water studies have centered bulk thermodynamics phase behavior of PEO using differential scanning calorimetry(26–28) or system-averaged measures of water dynamics such as rotational dynamics as measured by dielectric spectroscopy(29–33). Fewer efforts have isolated and probed water properties within the hydration layer of PEO chains. However, some studies have characterized PEO hydration water structural and dynamic features relative to bulk water by leveraging terahertz (Thz) spectroscopy(18, 34) and quasi-elastic neutron scattering (QENS)(16, 18, 35, 36), respectively. However, considerable gaps in our understanding of hydration water properties and the determinants of water diffusion in PEO-water remain.

Probing water dynamics at the molecular scale in PEO-water is critical to deciphering the connection between membrane surface chemistry and functional membrane properties like water permeability. In this chapter, we outline experimental measures of water self-diffusivity in PEO-water using two techniques: Pulsed-Field Gradient (PFG)-NMR and Overhauser Dynamic Nuclear Polarization (ODNP). As described in **Chapter 2**, ODNP is a combined NMR-EPR technique that utilizes EPR enhancements of water's EPR signal to selectively probe water dynamics in the vicinity (within 1 nm) of an electron spin probe. In the case of PEO, the spin probe tethered to a subset of PEO chains rather than freely floating in solution. By probing water dynamics with both PFG-NMR and ODNP, we access water dynamics at length scales from nanometers to microns and picoseconds to milliseconds [**Figure 3.1**]. Simultaneously, we employ Molecular Dynamics (MD) simulations to gain molecular scale insight on the structure and dynamics of water in the hydration shell of PEO.

In previous works, transport models in hydrated membranes have largely centered the macroscale effects of (1) geometric obstructions to diffusion due to polymer configurations, (2) hydrodynamic drag resulting from diffusion in a viscous medium (such as PEO-water), or (3) free volume configurations(10, 12, 37, 38). On the other extreme, investigations of water dynamics at biomolecular surfaces such as proteins and DNA have primarily considered dynamics on the molecular scale(15, 39, 40). Few studies have sought to reconcile macroscale water transport with molecular scale water diffusivity. Free volume and obstruction-based models may suggest the effect of detailed molecular structure on dynamics, but the relationship is merely hypothetical without further investigation. We pose the question: "When and under which conditions do molecular surface effects impact the diffusion of water in PEO solutions?" Further, we leverage our multiscale approach to determine the suitably "molecular" transport model to properly describe water diffusion spanning pure water to nearly pure PEO.

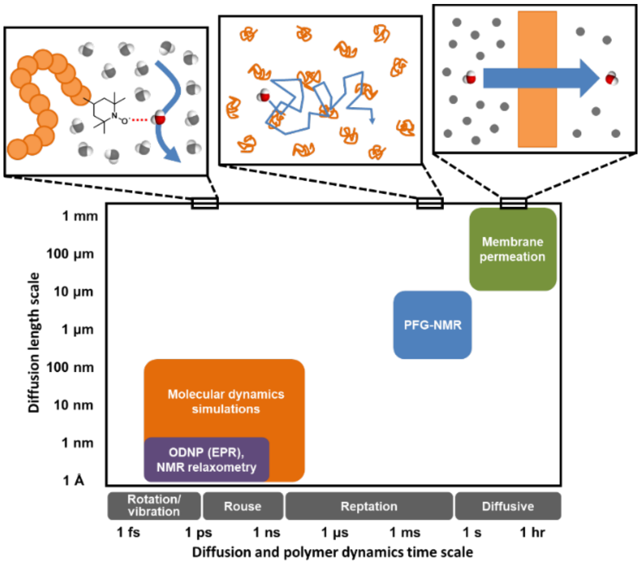


Figure 3.1 Distinct experimental and computational techniques probe small molecule diffusion processes occurring over vastly different length and time scales.(41) Techniques such as MD simulations and ODNP can uncover molecular-scale dynamic behavior of small molecules in polymers or polymer solutions while tools like PFG-NMR and membrane permeation experiments reveal macroscopic transport phenomena.

In the following sections of this chapter, we describe the equilibrium self-diffusivity of water in PEO-water by with a free volume model. Specifically, this “mixture fractional free volume” model implicitly accounts for the molecular structural details of PEO-water in contrast with hydrodynamic and obstruction-based models for water self-diffusivity. By leveraging atomistic MD simulations, we directly relate solution structural motifs to shifts in system-averaged (measured by PFG-NMR) and PEO hydration layer (measured by ODNP) water dynamics. Characterizing and understanding the drivers of local polymeric surface properties may be invaluable for tuning water dynamics at complex interfaces present in industrial polymer membranes.

3.2 Molecular Dynamics Simulations Methods[6]

We applied molecular dynamics models composed of the OPC 4-site water model(42), a generalized AMBER forcefield (GAFF2) [20,21] parametrized model for PEO(24), and a TEMPO spin probe functionalized PEO molecule. Both the OPC water model and the PEO model accurately reproduce the thermophysical properties of pure water and PEO under ambient conditions (298.15 K and 1 bar). Previous work demonstrated the ability of this functionalized PEO model to accurately describe the conformational landscape of PEO in close agreement with findings from experimental Double Electron-Electron Resonance (DEER) spectroscopy. (24, 45) The partial charges of the spin probe-functionalized PEO were obtained using the AMBER18 Antechamber package(46) informed by quantum chemical calculations using the Gaussian 16 software(47) as described in our previous work.(24) All other inter- and intramolecular parameters came from the second-generation generalized AMBER forcefield (GAFF2).(43, 44) The results of this parametrization scheme yielded similar parameters to those obtained in previous work.(48) All Coulombic interactions were described with particle-mesh Ewald summation scheme (PME).(49)

We considered PEO-water compositions of 0, 0.5, 1.5, 5, 10, 20, 33 and 50 wt% PEO using the GPU-optimized OpenMM molecular simulation software.(50) Each system was first energy minimized to remove overlapping atom positions in our PACKMOL(51) generated initial configurations. Systems were then equilibrated in the NPT ensemble using a Langevin Thermostat(52) paired with a Monte Carlo barostat (52) at 290 K and 1 bar. We determine the minimum necessary system equilibration period by estimating the time to convergence of for the density and temperature (under 1 ns for all systems). Following equilibration, the NPT run continued for 200 ns with system configurations saved every 10 ps. NPT-generated trajectories were used to calculate structural properties of the PEO-water mixtures such as radial

---

[6] Overhauser Dynamic Nuclear Polarization (ODNP) and Pulsed-Field Gradient-NMR (PFG-NMR) experiments are detailed in the upcoming paper from which this chapter is derived.

distribution functions, fractional free volumes (FFVs), and 3-body angle distributions. In addition to the system configurations, simulation states (atom positions and velocities) were saved every 10 ns. Each of these saved states served as the starting point for an independent 1-ns NVE simulation with system coordinates saved every 0.1 ps.

System-average water self-diffusivity, D_{H_2O}, values are estimated from the results of the 10 separate NVE simulations via the slope of the mean-square displacement (MSD) curve at long times

$$D_{H_2O} = \lim_{t\to\infty} \frac{1}{6t} MSD(t) = \lim_{t\to\infty} \frac{1}{6t} \langle |\vec{r}(t+\tau) - \vec{r}(\tau)|^2 \rangle$$

where $\vec{r}$, t, and τ are the position vector of a water oxygen, the current time step, and the initial time step. Here, $\langle\cdot\rangle$ denotes the ensemble average of a quantity across all water positions. We carry out the MSD slope determination for PEO heavy atoms to estimate PEO self-diffusivity, D_{PEO}. We compute the hydration layer water self-diffusivity, D_{local}, by considering the MSD of waters residing within the first two hydration shells of the PEO spin probe (8 Å) rather than all system waters. We estimate the local MSD slope in the diffusive region $t \in [10,40]$ ps to mitigate the effect of the hydration waters leaving the vicinity of the spin probe. To estimate the uncertainty in water and PEO diffusivities, 95% confidence intervals were computed by bootstrapping the mean-squared displacement (MSD) curves obtained from 10 independent MD simulations.

3.3 Computational/Experimental study of system-average water dynamics

While PFG-NMR measurements capture time-averaged diffusivities on the millisecond and micrometer scale, MD simulations probe water dynamics on the timescale of tens of picoseconds with characteristic distances on the order of nanometers. Solutions with PEO concentrations higher than 50 wt% are difficult to well-equilibrate due to the long relaxation time scales of polymer-rich systems. In addition to the system-average water self-diffusivities, we compute the water's self-diffusivity within an 8 Å radial distance from the spin labelled PEO chain, D_{local}. We discuss the trend in D_{local} in detail in **Section 3.** The details on the calculation of these two diffusivities is given in **Section 3.2**.

We first detail the system-average water dynamics using atomistic MD simulations and drawing comparison to the PFG-NMR results of our collaborators. **Figure 3.2(a)** depicts water self-diffusion coefficients, D_{H_2O}, from MD simulations at PEO concentrations ranging from 0 to 50 wt% near atmospheric temperature and pressure [further details given in **Section 3.2**]. We overlay PFG-NMR results at compositions up to 90 wt% PEO at 294 K and 1 bar (commensurate to the MD conditions). For both sets of measurements, the water self-diffusivities decrease exponentially from the pure water self-diffusivity—2.2 x 10^{-5} cm^2/s and 1.9 x 10^{-5} cm^2/s from PFG-NMR and MD, respectively—as indicated by prior measurements of water diffusivity in PEO solutions(53–57). Though MD and PFG-NMR yield different absolute D_{H_2O} values, the two sets of measurements yield similar scaling with concentration. Despite the substantial difference in length and timescales probed by the two techniques, computed values of D_{H_2O} exhibit nearly quantitative agreement with PFG-NMR self-diffusivities in dilute solutions below the overlap concentration $c^* = 46.4\ wt\%$. c^* is defined as the polymer concentration at which polymer coil volumes overlap; hence, c^* marks the dividing line between dilute and concentrated polymer mixtures [see **Appendix 3.7.1** for details on the determination of c^*](24, 58, 59). Both PFG-NMR and system-average MD capture the collective behavior of all water molecules in a solution and do not distinguish between "bulk" and hydration layer (i.e., near PEO) water dynamics.

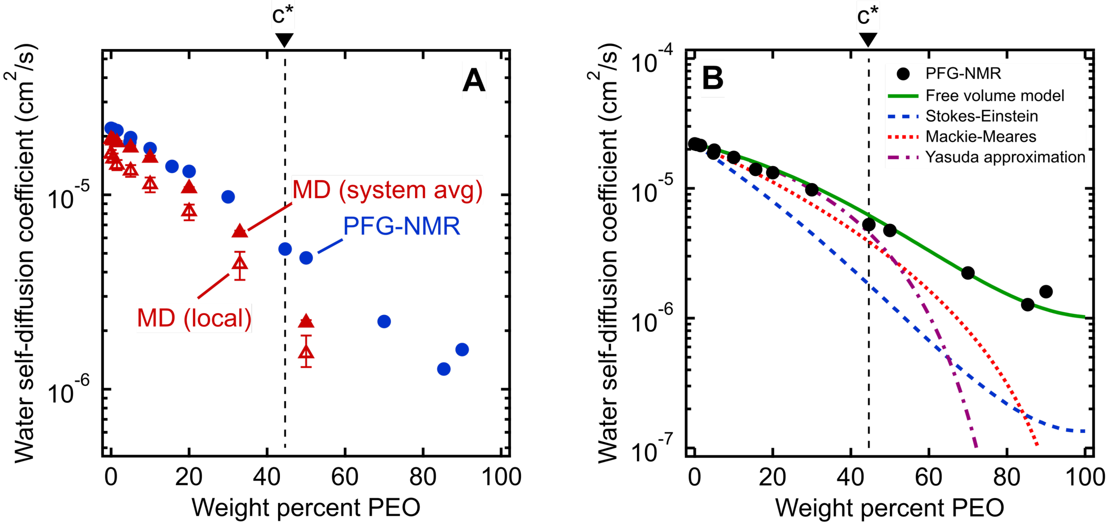


Figure 3.2 (A) Comparison of water self-diffusion coefficients in aqueous PEO solutions at room temperature (21 °C) from PFG-NMR (filled blue circles) with system-average water diffusivities (D_{H_2O}, filled red triangles) and local water diffusivities (D_{local}, unfilled red triangles) from MD simulations. (B) Comparison of PFG-NMR water diffusivities in aqueous PEO solutions (black circles) with model fits for Stokes-Einstein (blue dashed curve), Mackie-Meares (red dotted curve), free volume theory (green solid curve), and Yasuda's free volume approximation (purple dashed/dotted curve). PEO overlap concentration (c*) is marked in both figures by a dashed black line (see Supporting Information for c* derivation).

3.4 Analytical models of water diffusion in PEO solutions

3.4.1 Stokes-Einstein hydrodynamic model

Often, continuum models—independent of molecular scale detail—are utilized to capture diffusion trends in aqueous mixtures. Though water's molecular properties are highly determinative of its unique thermophysical properties, continuum models like the Stokes-Einstein relation can yield reasonable estimates of water self-diffusion coefficients near standard temperature and pressure (STP)(60). The Stokes-Einstein (S-E) relation connects water self-diffusivity to the mixture viscosity via the expression

$$D_{H_2O} = \frac{k_B T}{6\pi r \mu} \tag{3.1}$$

where D_{H_2O}, k_B, T, r, and μ are the water self-diffusion coefficient, the Boltzmann constant (1.38 x 10^{-23} J/K), the absolute temperature, the effective radius of a diffusing water molecule, and the dynamic viscosity of the solution. To evaluate the veracity of the S-E relation for PEO-water, we directly compare its output to the direct experimental measurement of water self-diffusivity from PFG-NMR [**Figure 3.2(b)**]. We fix the parameter $r = 0.97$ Å based on the measured D_{H_2O} (2.2 x 10^{-6} cm^2/s) and literature μ (1.01 x 10^{-3} Pa-s) of pure water at 1 bar and 293 K. We estimate the concentration-dependent PEO-water viscosity from literature data for at 293 K(61). The trend predicted from S-E deviates dramatically from the PFG-NMR estimate of D_{H_2O} even at low concentration of PEO. The predictive capacity S-E worsens with increasing PEO concentration and ultimately underestimates water self-diffusivity by an order of magnitude at 90 wt% PEO.

Fundamentally, the Stokes-Einstein relation assumes that PEO-water is a homogeneous mixture of non-interacting spheres of water and polymer. Particularly at high PEO concentrations, it appears that water dynamics become increasingly decoupled from the macroscopic properties of the system (e.g., viscosity). In an NMR study of water self-diffusivity in human serum albumin protein solutions, Lamanna and coworkers observe a similar violation of S-E with increasing protein concentration. They attribute this to increasing role that solute-solute interactions with protein concentration and a resultant decrease in the effective water hydrodynamic radius(62). We hypothesize that—particularly at high PEO concentrations—water dynamics are primarily dictated by water-water and water-PEO interactions rather than the macroscopic solution viscosity. Hence, we must consider PEO-water's collective molecular structure to properly describe water dynamics.

3.4.2 Free volume models

Unlike S-E, free volume models for diffusion treat solutions as a collection of molecular volume elements that, in aggregate, dictate dynamics. Hence, we anticipate better prediction of water dynamics from free volume models. The FFV is fraction of solution volume that is unoccupied by the effective van der Waals molecular (or atomic) volumes constituting the mixture. From this model standpoint, the FFV is the effective volume available for diffusion by water or polymer. We apply the following fractional free volume model(2)

$$D_{H_2O} = D_{pure}\, exp\left[-B\left(\frac{1}{FFV} - \frac{1}{FFV_0}\right)\right] \tag{3.2}$$

where D_{pure}, B, FFV, and FFV_0 are the self-diffusion coefficient of pure water, an empirical fitting parameter, the fractional free volume of a given PEO solution, and the fractional free volume of pure water. To apply **Equation 3.2** to the PFG-NMR experimental findings, FFV is estimated from literature solution densities and simple assumptions of the volume occupied by PEO and water [further detail in **Appendix 3.7.2**]. Performing a linear-least squares fitting procedure using the PFG-NMR water self-diffusivities, B assumes a value of 0.82(63). Per this FFV model, water diffusion is an activated process driven by thermal fluctuations in the arrangement of water, polymer, and free volume elements(37, 64, 65). Because the FFV depends on the spatial arrangement of the individual chemical elements (atoms and molecules) composing PEO-water, **Equation 3.2** is a fundamentally molecular theory for water diffusion.

We demonstrate that **Equation 3.2** yields accurate predictions of D_{H_2O} for the entire range of PEO concentrations from 0 to 90 wt% in **Figure 3.2(b)**. The reduction in D_{H_2O} with concentration is driven by the monotonic decrease in FFV from 0.251 at 0 wt% to 0.129 at 90 wt% PEO. Water self-diffusivity has previously been accurately described by **Equation 3.2** in polymer hydrogel networks(2, 66). This free volume theory explicitly accounts for the shifts in the PEO and water configurations via connection to the measured solution density across the entire concentration range. We note that the precise physical description of the (unoccupied) free volume of a mixture is nuanced with the FFV depending strongly on the underlying assumptions regarding the molecular volume of the mixture constituents and the lack of water-solute interactions. Though our application free volume theory is simple and purely entropic in nature, we predict water diffusion with remarkable fidelity across a wide range of PEO-water concentrations.

To further validate the validity of this FFV model for water diffusion, we directly compute FFV values in PEO solutions from MD simulations using an adapted procedure from Califano and coworkers.(67) Rather than approximating FFV from macroscopic thermodynamic quantities (i.e., solution densities) as performed above, these simulated FFV values are directly derived from atomistic configurational snapshots of PEO-water solutions. In brief, we insert a spherical test probe of radius 0.53 Å at the nodes of a grid and check for overlaps between the test particle and the atomic coordinates of a given configuration (further details are provided in the Supporting Information). Simulated FFV values agree qualitatively with experimentally-derived FFVs, while exhibiting ~20% smaller values compared to those derived from solution densities [**Figure 3.7**]. As discussed further in the Supporting Information, the magnitude of simulation-predicted FFVs depend on assumed values for the probe size; however, the relationship between the FFV and water self-diffusivity remains largely unchanged regardless of the probe size used. Furthermore, identical scaling is found between independently determined experimental and simulated sets of water diffusivities and FFV values as shown in **Figure 3.3** (B = 0.82 for PFG-NMR data and B = 0.83 for MD data). These findings reinforce the idea that the nanoscale structural properties of water in PEO solution are implicitly accounted for in the description of the free volume parameter, and that the varying FFV with increasing PEO concentration governs the water dynamics in PEO-water solutions at timescales of both picoseconds and milliseconds.

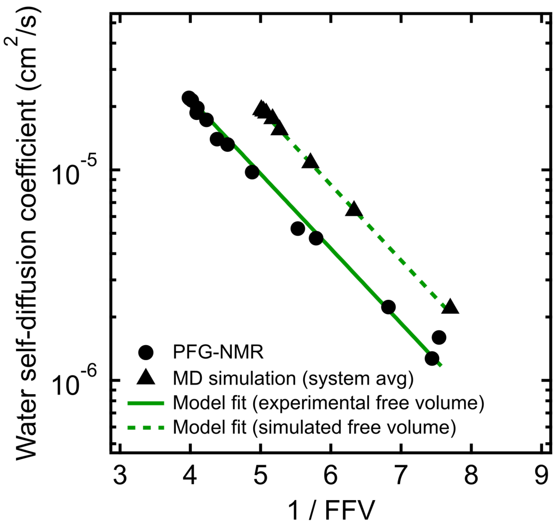


Figure 3.3 Free volume model for water self-diffusion coefficients in PEO determined by PFG-NMR using fractional free volume (FFV) values derived from experimental solution densities (black circles and solid green line) compared to free volume model fit for system-average water self-diffusion coefficients determined by MD simulations using FFV values derived from MD simulations (black triangles and dashed green line).

We present a further simplification of **Equation 3.**2—developed by Yasuda—in which the mixture FFV is assumed to be equivalent to the water volume fraction ϕ_w(9)

$$D_{H_2O} = D_{pure}\, exp\left[-B\left(\frac{1}{\phi_w} - 1\right)\right] \tag{3.3}$$

Often used in the context of highly swollen hydrogels, **Equation 3.3** assumes that water diffusion specifically occurs within the solution volume occupied by water regardless of the polymer concentration. This assumption leads to the unphysical conclusion that water is immobile as PEO concentration approaches 100 wt%. Instead, PEO is sufficiently mobile to accommodate co-solute diffusion even at very high polymer concentration. Hence, **Equation 3.3** performs well at $c < c^*$ but dramatically underestimates D_{H_2O} for $c > c^*$. As one final point of comparison, we consider another model used in swollen polymer networks known as the Mackie-Meares model [see **Appendix 3.7.3** for details]. Derived from a lattice polymer model, the Mackie-Meares model treats polymers as static physical obstacles to water diffusion rather than dynamic participants in diffusion(10). Ultimately, while the Mackie-Meares model yields marginally better quantitative agreement with experiment, the predicted D_{H_2O} still unphysically approaches zero as PEO concentration goes to 100 wt%. Simply accounting for non-zero FFV for pure PEO solutions—as is done in **Equation 3.2**—we predict D_{H_2O} with quantitative accuracy at all mixture compositions.

3.5 Computational/Experimental study of local water dynamics

While MD simulations are one route to characterizing water dynamics near the PEO, ODNP experiments are a valuable direct experimental probe of these local water dynamics. Specifically, ODNP experiments measure translational water diffusion on a 10 ps to 1 ns timescale [**Figure 3.1**] within an approximately ~1 nanometer radial distance from a spin labelled PEO chain end. The specific spin probe chemistry is detailed in our previous study on the conformational landscape of PEO in solution(68). Previous studies have demonstrated ODNP's ability to selectively characterize water dynamics at a wide variety of macromolecular surfaces in the context of proteins, lipid membranes, and silica nanoparticles(69–71). Further, local water self-diffusion coefficients, D_{ODNP}, are obtained by applying a continuum model diffusion model known as the force-free hard sphere (FFHS) model(72, 73). Finer details on ODNP theory and experimental protocols are described in several previous studies(74–78). For the present work, we provide ODNP water self-diffusivities in comparison and contrast to diffusion coefficients probed by PFG-NMR and MD simulations.

In **Figure 3.4(a)**, we display D_{ODNP} in contrast to system-average water self-diffusivities extracted from MD simulations and PFG-NMR [**Figure 3.2(a)**]. As an additional point of comparison to D_{ODNP}, we present the water self-diffusivity within the hydration layer of a spin labelled PEO, D_{local} (as described in **Section 3.2**). We first note that D_{ODNP} indicates slower water dynamics near PEO compared to either of the system-average diffusivities. Further, D_{ODNP} assumes nearly the same dramatically reduced value of 4.7 x 10^{-6} cm^2/s for concentrations in the dilute regime ($c < c^*$). Similarly, MD-derived D_{local} indicates 20-30% decreases in self-diffusivity compared to the system-average [**Figure 3.5**]. However, unlike the ODNP findings, MD suggests an exponential decrease in D_{local} for the entire concentration range (0 to 50 wt%). D_{ODNP} approaches a similar monotonic decrease to the PFG-NMR measurements for $c > c^*$.

The plateau in D_{ODNP} for $c < c^*$ suggests that ODNP is insensitive to changes in concentration within the dilute regime. This indicates that—from the ODNP reference frame—hydration waters are isolated from the effects of increasing PEO concentration for $c < c^*$. Upon entering the concentrated regime ($c > c^*$), PEO chain hydration environments begin to overlap with one another and hence induce the retardation of hydration water dynamics. The shift in D_{ODNP} scaling at c^* implies a molecular crowding effect that is invisible to system-average probes of water dynamics from PFG-NMR and MD simulations. However, the MD measure of hydration water dynamics, D_{local}, exhibits a similar mono-exponential decay [**Figure 3.5**] to the system-average trends.

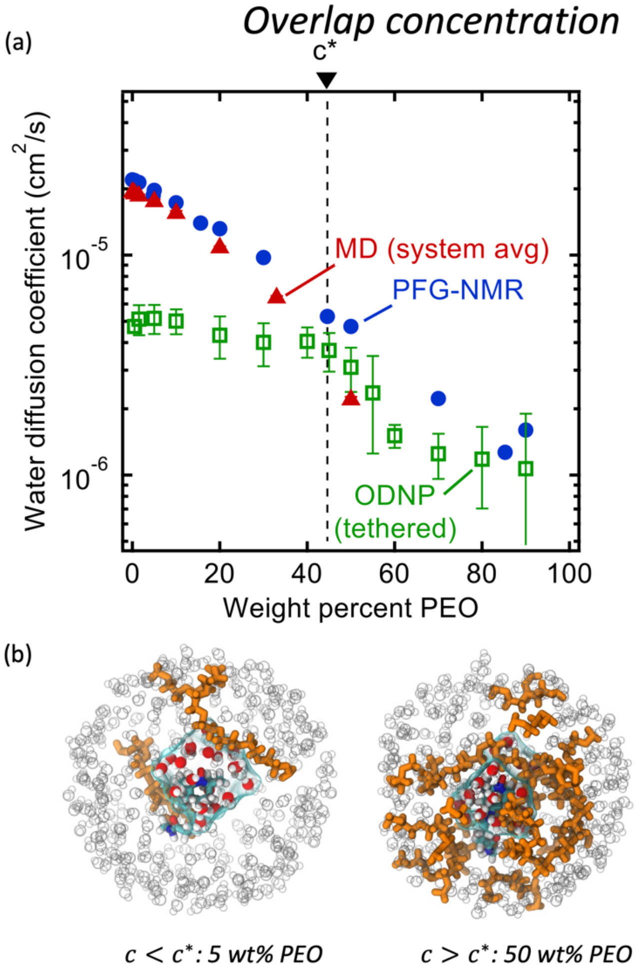


Figure 3.4. (a) Comparison of water diffusion coefficients in PEO solutions determined by PFG-NMR (filled blue circles), ODNP with PEO-tethered spin labels (unfilled gold squares), and ODNP with free TEMPO spin labels (filled gold squares). PEO overlap concentration (c*) is marked by a dashed black line. (b) MD snapshots illustrating first hydration waters around TEMPO spin label on PEO chain end for 5 and 20 wt% aqueous PEO solutions. Orange molecules represent PEO chains, grey circles represent water molecules beyond first hydration layer, and cyan surface represents volume enclosing first hydration shell waters.

There are several possible explanations for the apparent inconsistency between MD and ODNP characterization of PEO hydration waters. For instance, the accuracy of MD-derived thermophysical properties like self-diffusivities is model-dependent. While the model parameters appear to yield near quantitative agreement with the system-average PFG-NMR results, D_{local} may be more sensitive to the model assumptions. We use a simplified rigid and non-polarizable model for water. The dynamics probed by ODNP may be sensitive to intramolecular vibrational degrees freedom ignored by the OPC model; hence, future studies should consider flexible water model like TIP4P/2005f(79) or a polarizable model such as MB-pol(80). Further, ODNP-derived D_{ODNP} strongly depend on the choice of a spectroscopic quantity known as the saturation factor s_{max}. Typically, s_{max} is set to 1/3 or 1 for free and tethered spin probes. The assumption that $s_{max} = 1/3$ for free spin labels stems from direct experimental measurement of s_{max} for a free ODNP spin label in water(81). For tethered spin probes, s_{max} is assumed to be unity based lack of direct collisions between spin probes fixed to macromolecular surfaces like proteins and lipid bilayers(82, 83). However, it is not entirely clear whether s_{max} can accurately be set to unity for faster diffusing tethered PEO molecuels(77, 84). Future experimental efforts should seek to interrogate the chemistry and concentration-dependence of s_{max}.

3.6 Structural ordering of water revealed by MD simulations

To further investigate the connection between molecular structure and water dynamics, we employ MD simulations to probe the effect of PEO crowding on the structure of hydration waters. First, we compute 2-D radial distribution functions (RDFs) to determine how PEO concentration affects local water density near PEO chains relative to that of the bulk solution. Specifically, we consider the RDFs $g(r_{Or-Ow})$ between the radical oxygen on the spin probe (O_r) and all nearby water oxygens (O_w), where r_{Or-Ow} is the distance between a radical oxygen and water oxygen. Shifts in RDF peak position and amplitude with concentration indicate changes in the local environment such as changes in temperature, concentration, and water affinity (i.e., hydrophobicity). For the entire concentration range of 0 to 50 wt% PEO, RDFs do not exhibit changes in overall peak shape or location [**Figure 3.6(a)**]. Rather, there is a systematic increase in the amplitude of the first peak around 0.28 nm with increased PEO concentration, indicating an enhancement in the number density of water near the spin probe relative to bulk solution, where we define the bulk region as $r_{Or-Ow} > 1.5$ nm. We hypothesize that this increased local number density of water relative to bulk solution decreases the local free volume, effectively slowing water diffusion in the hydration layer below the bulk value.

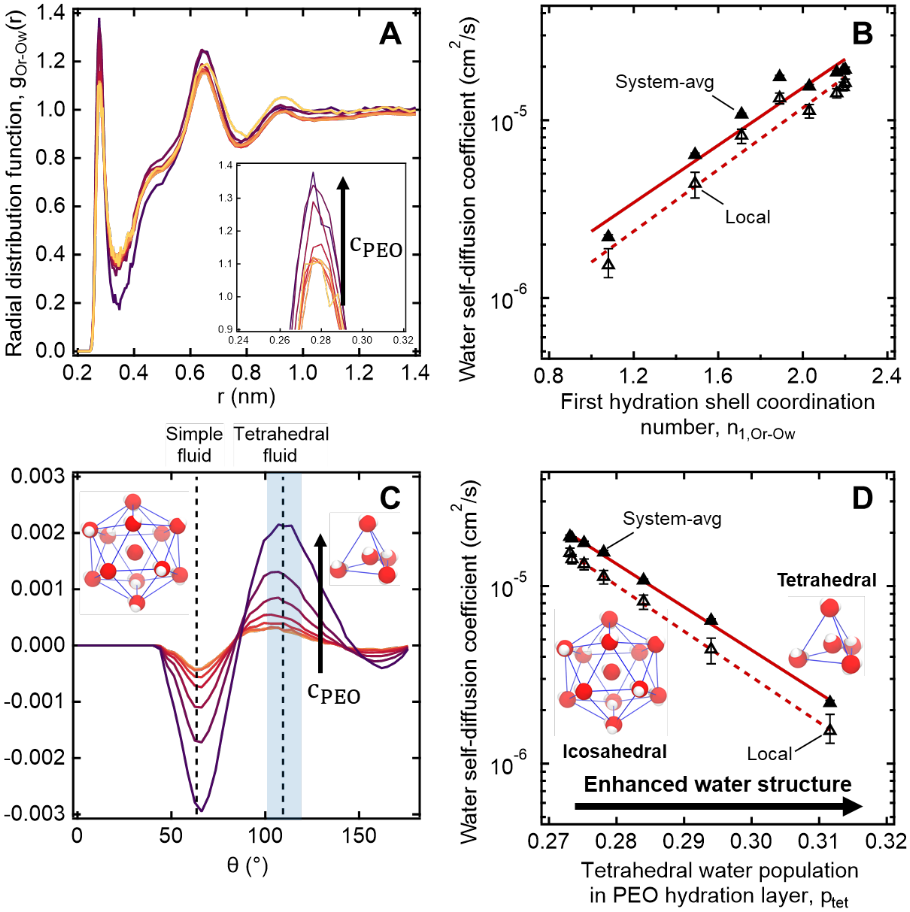


Figure 3.6 (A) Radial distribution functions of water molecules near PEO chains maintain similar shapes with increasing PEO concentration from 0.5 to 50 wt% with a systematic increase in the 1st and 2nd peak heights (darker colors correspond to higher concentrations). (B) MD-derived local and system-averaged water self-diffusivities in PEO solutions generally correlate with the coordination number of the first hydration shell. (C) Three-body angle distributions show enhancement in the tetrahedrality of water (109.5°) in the hydration layer near PEO chains with increasing PEO concentration from 0.5 to 50 wt% (darker colors correspond to higher concentrations). (D) The integral over the tetrahedral region of the three-body angle distributions $P(\theta)$, indicated by the shaded region in part A, correlates with both local and system-averaged water self-diffusivities in PEO solutions derived from MD simulations.

We further quantify local water density correlations by computing the coordination number of the first hydration shell, $n_{1,Or-Ow}$. Values for $n_{1,Or-Ow}$ are estimated by integrating $\rho g(r_{Or-Ow})$ over r_{Or-Ow} from 0 to 0.34 nm (i.e., the position of the first RDF minimum), where ρ is the number density of water in the bulk. $n_{1,Or-Ow}$ decreases from 2.2 to 1.08 as PEO concentration increases from 0 to 50 wt% as neighboring PEO chains begin to crowd out water molecules in the first hydration shell. As shown in **Figure 3.5(b)**, we observe a general correlation between both local and system-average water self-diffusivities and $n_{1,Or-Ow}$ (R^2=0.94), suggesting changes in water structure near PEO chains affect local water dynamics that in turn affect dynamics averaged over the bulk. While metrics like the FFV and $n_{1,Or-Ow}$ hint at systematic shifts in water structure, these metrics do not directly probe water's tetrahedral network structure.

To more directly describe changes in the tetrahedral structure of PEO-water, we compute three-body angle distributions $P(\theta)$, which present a detailed picture of water coordination by capturing the effects of small shifts in the solution environment on water structure.(85–87) To construct 3-body angle distributions, we compute and histogram the angles between hydration water oxygens (i.e., those within 4.2 Å of PEO heavy atoms) and their two nearest neighboring water oxygens. Peaks in **Figure 3.5(c)** represent differences between water populations of a particular coordination local to PEO chains in solution from those in pure water. We observe peaks at 64° and 109.5°, corresponding to a decrease in the population of icosahedrally-coordinated hydration waters near the PEO chains and an increase in the population of tetrahedrally-coordinated hydration waters, respectively. A shift from icosahedral to tetrahedral coordination indicates a shift away from simple fluid behavior toward orientations typically associated with enhancements in water ordering under supercooling(88) and in the hydration layers of small hydrophobic molecules.(89) As PEO concentration increases, the hydration layer exhibits a monotonic increase in the population of tetrahedrally-coordinated waters and a monotonic decrease in the population of icosahedrally-coordinated waters [**Figure 3.5(c)**].

To directly quantify the increase in hydration water ordering with PEO concentration, we approximate the population of tetrahedrally-coordinated waters as $p_{tet} = \int_{100^\circ}^{120^\circ} P(\theta) d\theta$.

For the entire range of PEO concentrations, there is a substantial increase in tetrahedral ordering as p_{tet} increases by 5% from 0 to 50 wt% PEO solutions, indicating an enhancement in the water network structure near PEO chains [**Figure 3.5(d)**]. Both local and system-averaged water self-diffusivities in PEO solutions from MD correlate inversely with the population of tetrahedral ordered hydration waters [see **Figure 3.5(d)**; R^2>0.99]. The relationship between equilibrium water dynamics and structure in PEO-water has been previously discussed in the computational literature.(16, 17) For instance, Borodin et. al. demonstrated that water self-diffusivity in PEO-water can be reconstructed from the portion of waters hydrogen-bonded to a PEO ether oxygen. While water-water hydrogen bonding is closely related to water's tetrahedral structure, metrics of tetrahedrality do not directly report on hydrogen bonding. We find that p_{tet} increases and the average number of water-water hydrogen bonds per water molecule decreases with increasing PEO concentration (see Supporting Information). This suggests that even as water molecules increasingly favor hydrogen bonding with PEO, the tetrahedral network structure of water persists.

The structure-dynamics connection apparent in **Figure 3.5(d)** paired with the literature, suggests that, in addition to reduced free volume and changes in local water density, enhanced tetrahedrality in more concentrated solutions leads to the concentration-dependent retardation of hydration water translational dynamics. Remarkably, water's network structure appears enhanced local to hydrophilic PEO, a phenomenon usually attributed to small hydrophobic molecules and moieties(89, 90, 90). This is in keeping with recent computational studies that demonstrated a more general enhancement water tetrahedrality in the hydration layer of various small length-scale (< 1 nm) chemical moieties irrespective of the chemical identity.(87, 91)

3.6 Conclusions

We apply a novel computational-experimental approach to reveal the effect of the hydration shell structure on water dynamics in PEO-water mixtures. The discrepancy between nanoscale water dynamics from Overhauser Dynamic Nuclear Polarization (ODNP) and microscale water dynamics from Pulsed-Field Gradient NMR (PFG-NMR) suggests the emergence of a molecular-crowding effect within the PEO hydration shell at concentrations above the PEO overlap concentration. MD simulations demonstrate strong correlation between hydration shell water dynamics and molecular structural metrics probing solution densification and enhancement of water's tetrahedral character with increasing PEO concentration. We faithfully capture trends in water self-diffusivity from both MD simulations and PFG-NMR experiments by leveraging a solution free volume model for water diffusion. On the other hand, continuum models—specifically, Stokes-Einstein—fail to capture essential trends in water dynamics. Ultimately, both structural and dynamic feature of water in aqueous polymer solutions must be studies to develop molecularly informed design rules for the complex polymeric materials (e.g., hydrogels) found in separation membranes. Leveraging these rational design rules may enable fine tuning of water polymer chemistry and topology to yield desired water transport properties and enable the development of next generation water purification technologies.

3.7 Appendices

3.7.1 Calculation of PEO's polymer overlap concentration

In our prior study of the configurational landscape of PEO, we demonstrated that the end-to-end distance of the polymer scales with the molecular weight, M, as

$$R_{ee} = aM^{\nu} \tag{3.4}$$

where the critical exponent $\nu = 0.59$ suggests that PEO behaves as a polymer good solvent and $a = 0.047\ nm$ is a pre-exponential factor. For a polymer in a good solvent, the radius of gyration is related to R_{ee} via the expression

$$R_g = \sqrt{R_{ee}^2/6.25} \tag{3.5}$$

We determine the overlap concentration from the known equation (58, 59)

$$c^* = \frac{M/N_A}{\frac{4\pi}{3} R_g^3} \tag{3.6}$$

where N_A is Avogadro's number. For the 13-mer PEO considered in this chapter (550 g/mol), the overlap concentration is 0.48 g/mL or 46.4 wt%.

3.7.2 Estimating fractional free volume for PEO and glycerol solutions

Fractional free volume (FFV) can be defined for an aqueous PEO solution as the fraction of solution volume that is unoccupied by the molecular volumes of the solution components:

$$FFV = \frac{V_{free}}{V_{solution}} = 1 - \frac{V_{O,w} + V_{O,PEO}}{V_{solution}} \tag{3.7}$$

where V_{free} and $V_{solution}$ are the unoccupied free volume of the solution and the total solution volume. $V_{O,w}$ and $V_{PEO,w}$ are the effective molecular volumes of water and PEO, respectively. Alternatively, **Equation 3.7** may be written in a mass specific volume form

$$FFV = 1 - \frac{w_w \hat{V}_{O,w}\ + w_{PEO} \hat{V}_{O,PEO}}{V_{solution}} \tag{3.8}$$

where w_w and w_{PEO} are water and PEO weight fractions, respectively. Here, $\hat{V}_{O,w}$, $\hat{V}_{O,PEO}$, and $\hat{V}_{solution}$ are the mass specific volumes of water, PEO, and the solution, respectively(92–

94). Specific occupied volume of a solution component, typically defined as the specific volume of the component at 0K, is commonly estimated for polymers as:

$$\hat{V}_{O,i} = 1.3\hat{V}_{vdw,i} \tag{3.9}$$

where $\hat{V}_{vDw,i}$ is the van der Waals volume of the polymer, and the 1.3 factor is an approximate "universal" packing factor relating the two volume terms.(95, 96) For polymers, $\hat{\mathrm{V}}_{\mathrm{vDw,i}}$ is often approximated using group contribution theory by summing the van der Waals volumes of individual polymer functional groups.(97–99) Using group contribution values given by Bondi, $\hat{\mathrm{V}}_{\mathrm{vDw,PEO}}$ for 550 g/mol PEO was estimated to be 0.594 cm^3/g, which yields a FFV for pure PEO of 0.129.(97) There is some degree of uncertainty in the literature about the most appropriate value to use for $\hat{\mathrm{V}}_{\mathrm{vDw,w}}$ for water.(92, 93, 100–104) The present study used a value of 0.577 cm^3/g based on the van der Waals volume of water in ice, which yields a FFV of pure water of 0.251.(102) Aqueous solution densities at 20 °C at varying concentrations of PEO with PEO molecular weights near 550 g/mol and at varying concentrations of glycerol were found in the literature.(105–107) Values of FFV for aqueous PEO solutions were calculated using **Equations. 3.8 and 3.9** and are shown in **Figure 3.7**. Non-linear trends of FFV with solution concentration are in part driven by a negative excess volume of mixing due to non-ideal mixing of water and PEO [**Figure 3.7**].

FFV values for aqueous glycerol solutions were likewise determined, where $\hat{\mathrm{V}}_{\mathrm{0,glycerol}}$ is substituted for $\hat{\mathrm{V}}_{\mathrm{0,PEO}}$ in Eqns. S5 and S6. $\hat{\mathrm{V}}_{\mathrm{0,glycerol}}$ was determined to have a value of 0.724 cm^3/g using group contribution theory, yielding a FFV value for pure glycerol of 0.088. FFV values for glycerol solutions are also shown in **Figure 3.7**.

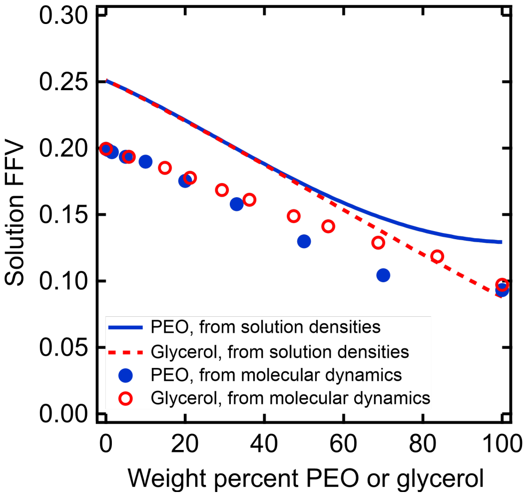


Figure 3.7. Comparison of fractional free volume (FFV) values derived from experimental solution densities (curves) and accessible FFV values simulated using Molecular Dynamics (circles) for PEO and glycerol solutions. Blue solid curve and filled circles are for PEO solutions, and red dotted curve and unfilled circles are for glycerol solutions.

3.7.3 Fitting water self-diffusion coefficients to fluid mechanics and diffusion models: Mackie-Meares model

The Mackie-Meares equation mentioned in the main text is given by (10)

$$D_{H_2O} = \frac{D_{pure}\phi_w^2}{(2-\phi_w)^2} \tag{3.10}$$

where D_{pure} and ϕ_w are the self-diffusion coefficient of pure water at 293K (2.20 x 10^{-5} $\mathrm{cm^2/s}$) and the water volume fraction. The volume fraction of water at a given water weight fraction, w_w, is given by

$$\phi_w(w_w) = \frac{w_w \hat{V}_w(w_w)}{\hat{V}(w_w)} \tag{3.11}$$

where $\hat{V}_w(w_w)$ and $\hat{V}(w_w)$ are the mass specific volumes of water and the solution at a given w_w. $\hat{V}(w_w)$ is simply determined from the inverse of the mass density of the solution density. $\hat{V}_w(w_w)$ is derived from a tangent line construction of $\hat{V}$ at a given w_w and the resulting values are presented in **Figure 3.8**.

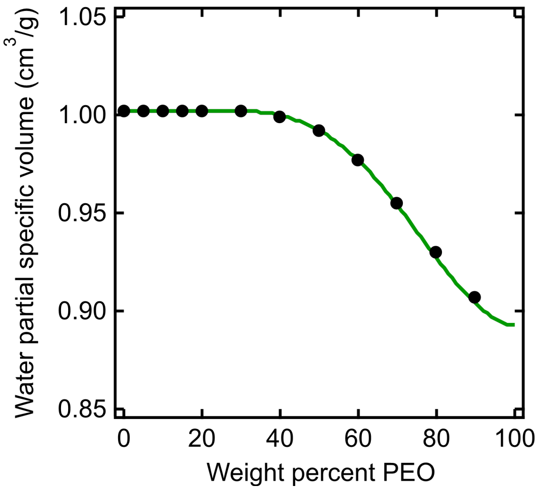


Figure 3.8. Water partial specific volumes for PEO-water solutions as a function of PEO concentration derived from solution densities.

3.7.4 MD free volume and diffusion simulations

The Califano method for estimating the accessible FFV relies on first inserting a 0.53 Å radius probe "molecule" at every node of a 3D grid (with 0.5 Å bin widths) cast over the simulation box and checking for overlaps with the VDW radii of atoms in the bin volume.(67) The probe size was chosen to match that used by experimental Positron Annihilation Lifetime Spectroscopy (PALS) in which polymer FFV is estimated using the lifetime of ortho-positroniums which have a 0.53 Å radius. If the node has no overlaps with atoms in the bin volume, it is considered an unoccupied node. A bin will be assigned a total free volume depending on the number of unoccupied nodes it contains. A bin with a single unoccupied node has a free volume of 1/8-th the volume of the probe molecule or 0.0779 Å^3. Two, three, and four unoccupied nodes correspond to free volume elements of 0.1020 Å^3, 0.1143 Å^3, and 0.1173 Å^3, respectively. These volumes are estimated by subtracting the average probe overlap volume from the sum of the probe volumes for the 2, 3, and four unoccupied node cases. If there are five or more unoccupied nodes, the entire bin is considered a free volume of size 0.125 Å^3. We note that that, outside of extreme cases (i.e., using a probe of size 0 Å), most combinations of bin size and probe size yield qualitatively similar values. For every simulated PEO-water composition (from 0 to 50 wt%), we check for probe overlaps with the van der Waals radii of any water or polymer molecule atoms (R_H = 1.2 Å, R_C = 1.7 Å, R_O = 1.52 Å).

We further note that accessible FFV values derived from MD are highly sensitive to the choice of probe size and grid spacing. We demonstrate the effect of probe size by inspecting the FFV-concentration [**Figure 3.9(a)**] and FFV-diffusivity [**Figure 3.9(b)**] scaling relationships. Ultimately, while the probe size impacts the magnitude of the FFV, probe size does not fundamentally alter the underlying FFV trends in PEO-water solutions.

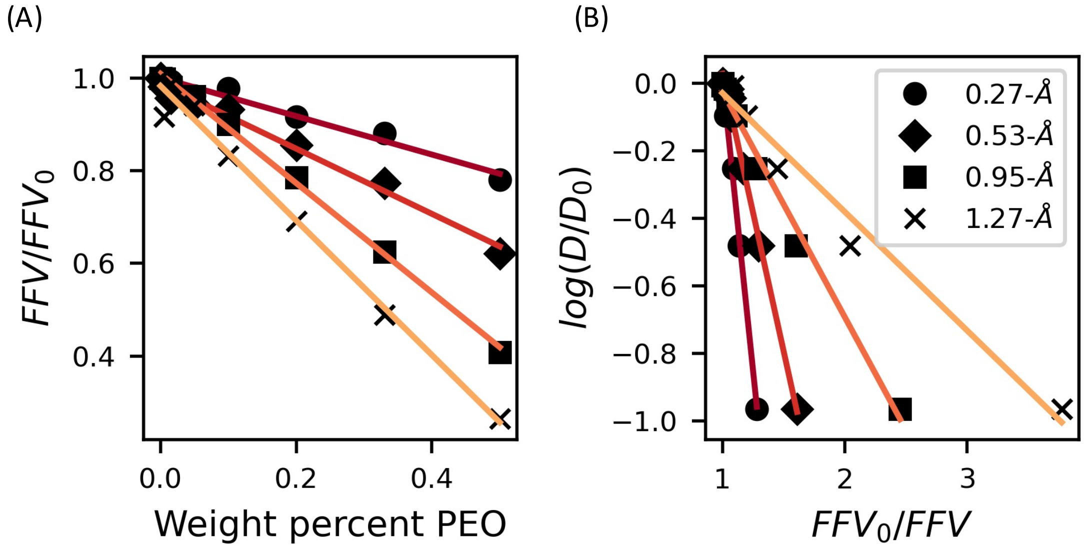


Figure 3.9. Systematically varying FFV probe size from 0.27 Å to 1.27 Å quantitatively alters the resulting FFV estimate while leaving the underlying relationships with (A) PEO concentration and (B) water self-diffusivity unchanged. Here, we normalize by the pure water fractional free volume $\boldsymbol{FFV_0}$ to emphasize the consistent exponential relationship between $\boldsymbol{D}$ and $\frac{1}{FFV}$. The 0.27 Å, 0.53 Å, 0.95 Å, and 1.27 Å probe size calculations yield $\boldsymbol{FFV_0}$ values of 0.332, 0.205, 0.036, and 0.004, respectively.

At dilute PEO concentrations (c<c*), MD-derived water self-diffusivities show near quantitative agreement with the PFG-NMR results, exhibiting less than 20% deviation. Much of this difference may be explained by the 4 °C lower simulated temperature. For instance, NMR studies have previously demonstrated as much as a 15% increase in water self-diffusivity going from 15 to 20°C (108). However, MD simulations give a water self-diffusivity just under half of the PFG-NMR measurement at 50 wt%. This more dramatic discrepancy may stem from imperfections in the ability of the MD model (see Methods) to reproduce accurate

equilibrium self-diffusivities at high PEO concentration. Though the PEO parameters previously yielded accurate conformational landscapes (24) and phase behavior in water,(45) it was not validated with equilibrium dynamics in mind.

3.7.5 MD simulations of solution structure

Though there are several metrics that report on the tetrahedrality of water in molecular simulations(85, 86, 88), many of these metrics can underestimate water coordination near surfaces(85, 86, 109, 110) due to the requirement of 4 neighboring waters. On the other hand, 3-body angle distributions do not require a given water molecule to be 4-coordinated, making them more robust to the geometric constraints introduced at interfaces.

Water's tetrahedral structure is closely related to its capacity to form approximately 4 hydrogen bonds per water molecule. In the present study, we note that both the number of water-water hydrogen bonds per water molecule, $N_{HB,wat-wat}/N_w$, and p_{tet} strongly correlate with the FFV of PEO-water mixtures [**Figure 3.10(a)-(b)**] and with each other [**Figure 3.10(c)**]. The overall decrease in $N_{HB,wat-wat}/N_w$ with increasing concentration in **Figure 3.10** is driven by the waters that directly hydrogen bond with PEO ether oxygens. The remaining hydration waters on the other hand exhibit enhanced tetrahedrality, with the net effect of an overall enhancement of the hydration layer tetrahedral structure.

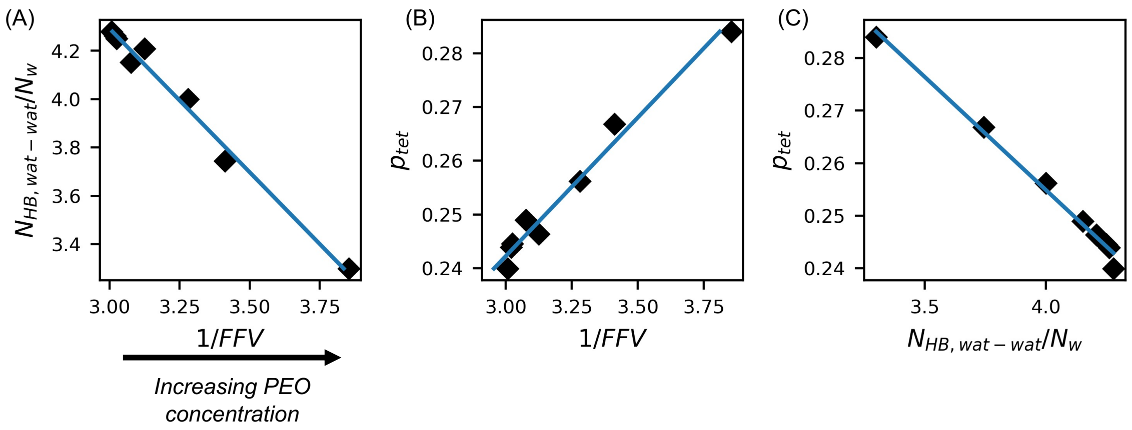


Figure 3.10. The MD-computed FFV exhibits strong correlations with water structural metrics such as (A) the average number of water-water hydrogen bonds per water molecule, $\boldsymbol{N_{HB,wat-wat}/N_w}$, and (B) the population of tetrahedrally coordinated waters, $\boldsymbol{p_{tet}}$. (C) $\boldsymbol{p_{tet}}$ correlates strongly with $\boldsymbol{N_{HB,wat-wat}/N_w}$.

3.7.6 Correlations between MD and ODNP water diffusivities and tetrahedral water population

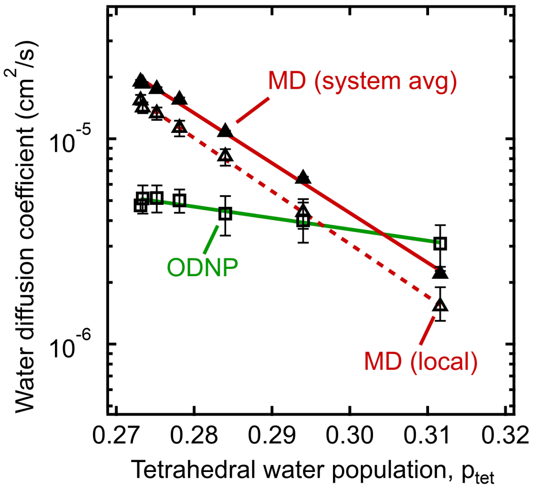


Figure 3.11. Comparing the correlation between population of tetrahedral waters p_{tet} and three estimates of the relative water diffusivity: system-averaged water self-diffusivity from MD D_{H_2O} (filled triangles), water self-diffusivity for waters within 0.8 nm of the radical oxygen from MD D_{local} (unfilled triangles), and ODNP-derived local water self-diffusivity using spin-labeled PEO D_{ODNP} (unfilled squares). Both MD estimates of water diffusivity show strong correlation with p_{tet} ($R^2 > 0.99$). While the ODNP results show a nominally favorable correlation with p_{tet} ($R^2 = 0.95$), the slope of the relationship is significantly smaller than the former two cases.

References

1. Landsman MR, Sujanani R, Brodfuehrer SH, Cooper CM, Darr AG, et al. 2020. Water Treatment: Are Membranes the Panacea? *Annual Review of Chemical and Biomolecular Engineering*. 11(1):559–85

2. Yasuda H, Lamaze CE, Peterlin A. 1971. Diffusive and hydraulic permeabilities of water in water-swollen polymer membranes. *Journal of Polymer Science Part A-2: Polymer Physics*. 9(6):1117–31

3. Paul DR. 1973. Relation between hydraulic permeability and diffusion in homogeneous swollen membranes. *Journal of Polymer Science: Polymer Physics Edition*. 11(2):289–96

4. Paul DR. 1974. Further comments on the relation between hydraulic permeation and diffusion. *Journal of Polymer Science: Polymer Physics Edition*. 12(6):1221–30

5. Paul DR. 1976. The Solution-Diffusion Model for Swollen Membranes. *Separation and Purification Methods*. 5(1):33–50

6. Hegde VH, Doherty MF, Squires TM. 2022. A two-phase model that unifies and extends the classical models of membrane transport. *Science*. 377(6602):186–91

7. Mehta A, Zydney AL. 2005. Permeability and selectivity analysis for ultrafiltration membranes. *Journal of Membrane Science*. 249(1):245–49

8. Geise GM, Lee H-S, Miller DJ, Freeman BD, McGrath JE, Paul DR. 2010. Water purification by membranes: The role of polymer science. *Journal of Polymer Science Part B: Polymer Physics*. 48(15):1685–1718

9. Yasuda H, Lamaze CE, Ikenberry LD. 1968. Permeability of solutes through hydrated polymer membranes. Part I. Diffusion of sodium chloride. *Die Makromolekulare Chemie*. 118(1):19–35

10. Mackie JS, Meares P. 1955. The diffusion of electrolytes in a cation-exchange resin membrane I. Theoretical. *Proc. R. Soc. Lond. A*. 232(1191):498–509

11. Geise GM, Paul DR, Freeman BD. 2014. Fundamental water and salt transport properties of polymeric materials. *Progress in Polymer Science*. 39(1):1–42

12. Masaro L, Zhu XX. 1999. Physical models of diffusion for polymer solutions, gels and solids. *Progress in Polymer Science*. 24(5):731–75

13. Monroe J, Barry M, DeStefano A, Aydogan Gokturk P, Jiao S, et al. 2020. Water Structure and Properties at Hydrophilic and Hydrophobic Surfaces. *Annual Review of Chemical and Biomolecular Engineering*. 11(1):523–57

14. Yoo H, Paranji R, Pollack GH. 2011. Impact of Hydrophilic Surfaces on Interfacial Water Dynamics Probed with NMR Spectroscopy. *J. Phys. Chem. Lett.* 2(6):532–36

15. Barnes AC, Bieze TWN, Enderby JE, Leyte JC. 1994. Dynamics of Water in the Poly(ethylene oxide) Hydration Shell: A Quasi Elastic Neutron-Scattering Study. *J. Phys. Chem.* 98(44):11527–32

16. Borodin O, Trouw F, Bedrov D, Smith GD. 2002. Temperature Dependence of Water Dynamics in Poly(Ethylene Oxide)/Water Solutions from Molecular Dynamics Simulations and Quasielastic Neutron Scattering Experiments. *J. Phys. Chem. B*. 106(20):5184–93

17. Borodin O, Bedrov D, Smith GD. 2002. Concentration Dependence of Water Dynamics in Poly(Ethylene Oxide)/Water Solutions from Molecular Dynamics Simulations. *J. Phys. Chem. B*. 106(20):5194–99

18. Tominaga T, Hishida M, Murakami D, Fujii Y, Tanaka M, Seto H. 2022. Experimental Evidence of Slow Mode Water in the Vicinity of Poly(ethylene oxide) at Physiological Temperature. *J. Phys. Chem. B*. 126(8):1758–67

19. Miller DJ, Dreyer DR, Bielawski CW, Paul DR, Freeman BD. 2017. Surface Modification of Water Purification Membranes. *Angewandte Chemie International Edition*. 56(17):4662–4711

20. Kang G, Cao Y. 2012. Development of antifouling reverse osmosis membranes for water treatment: A review. *Water Research*. 46(3):584–600

21. Padaki M, Surya Murali R, Abdullah MS, Misdan N, Moslehyani A, et al. 2015. Membrane technology enhancement in oil–water separation. A review. *Desalination*. 357:197–207

22. Chapman RG, Ostuni E, Liang MN, Meluleni G, Kim E, et al. 2001. Polymeric Thin Films That Resist the Adsorption of Proteins and the Adhesion of Bacteria. *Langmuir*. 17(4):1225–33

23. D'souza AA, Shegokar R. 2016. Polyethylene glycol (PEG): a versatile polymer for pharmaceutical applications. *Expert Opinion on Drug Delivery*. 13(9):1257–75

24. Sherck N, Webber T, Brown DR, Keller T, Barry M, et al. 2020. End-to-End Distance Probability Distributions of Dilute Poly(ethylene oxide) in Aqueous Solution. *J. Am. Chem. Soc.* 142(46):19631–41

25. Alessi ML, Norman AI, Knowlton SE, Ho DL, Greer SC. 2005. Helical and Coil Conformations of Poly(ethylene glycol) in Isobutyric Acid and Water. *Macromolecules*. 38(22):9333–40

26. Hager SL, Macrury TB. 1980. Investigation of phase behavior and water binding in poly(alkylene oxide) solutions. *Journal of Applied Polymer Science*. 25(8):1559–71

27. Huang L, Nishinari K. 2001. Interaction between poly(ethylene glycol) and water as studied by differential scanning calorimetry. *Journal of Polymer Science Part B: Polymer Physics*. 39(5):496–506

28. Hatakeyma T, Kasuga H, Tanaka M, Hatakeyama H. 2007. Cold crystallization of poly(ethylene glycol)–water systems. *Thermochimica Acta*. 465(1):59–66

29. Tyagi M, Murthy SSN. 2006. Dynamics of water in supercooled aqueous solutions of glucose and poly(ethylene glycol)s as studied by dielectric spectroscopy. *Carbohydrate Research*. 341(5):650–62

30. Shinyashiki N, Asaka N, Mashimo S, Yagihara S. 1990. Dielectric study on dynamics of water in polymer matrix using a frequency range 106–1010 Hz. *J. Chem. Phys.* 93(1):760–64

31. Sato T, Niwa H, Chiba A, Nozaki R. 1998. Dynamical structure of oligo(ethylene glycol)s-water solutions studied by time domain reflectometry. *J. Chem. Phys.* 108(10):4138–47

32. Murthy. 2000. Experimental Study of the Dynamics of Water and the Phase behavior of the Supercooled Aqueous Solutions of Propylene Glycol, Glycerol, Poly(ethylene glycol)s, and Poly(vinylpyrrolidone). *J. Phys. Chem. B*. 104(29):6955–62

33. Sudo S, Tsubotani S, Shimomura M, Shinyashiki N, Yagihara S. 2004. Dielectric study of the α and β processes in supercooled ethylene glycol oligomer–water mixtures. *J. Chem. Phys.* 121(15):7332–40

34. Aoki K, Hata R, Kaneyasu JF, Schwaab G, Shiraki K, Hattori T. 2018. Hydration of Aqueous Polymers Investigated by Terahertz Spectroscopy and Principal Component Analysis

35. Magazù S. 2000. NMR, static and dynamic light and neutron scattering investigations on polymeric aqueous solutions. *Journal of Molecular Structure*. 523(1):47–59

36. Li Y, Han Z, Ma C, Hong L, Ding Y, et al. 2022. Structure and dynamics of supercooled water in the hydration layer of poly(ethylene glycol). *Structural Dynamics*. 9(5):054901

37. Vrentas JS, Duda JL. 1977. Diffusion in polymer—solvent systems. I. Reexamination of the free-volume theory. *J. Polym. Sci. Polym. Phys. Ed.* 15(3):403–16

38. Amsden B. 2002. Modeling solute diffusion in aqueous polymer solutions. *Polymer*. 43(5):1623–30

39. Franck JM, Ding Y, Stone K, Qin PZ, Han S. 2015. Anomalously Rapid Hydration Water Diffusion Dynamics Near DNA Surfaces. *J. Am. Chem. Soc.* 137(37):12013–23

40. Biller JR, Barnes R, Han S. 2018. Perspective of Overhauser dynamic nuclear polarization for the study of soft materials. *Current Opinion in Colloid & Interface Science*. 33:72–85

41. Choo Y, Halat DM, Villaluenga I, Timachova K, Balsara NP. 2020. Diffusion and migration in polymer electrolytes. *Progress in Polymer Science*. 103:101220

42. Izadi S, Anandakrishnan R, Onufriev AV. 2014. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* 5(21):3863–71

43. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. 2004. Development and testing of a general amber force field. *Journal of Computational Chemistry*. 25(9):1157–74

44. He X, Man VH, Yang W, Lee T-S, Wang J. 2020. A fast and high-quality charge model for the next generation general AMBER force field. *J. Chem. Phys.* 153(11):114502

45. Sherck N, Shen K, Nguyen M, Yoo B, Köhler S, et al. 2021. Molecularly Informed Field Theories from Bottom-up Coarse-Graining. *ACS Macro Lett.* 10(5):576–83

46. Case DA. 2018. AMBER 2018

47. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, et al. 2016. Gaussian 16 Rev. B.01

48. Sezer D, Freed JH, Roux B. 2008. Parametrization, Molecular Dynamics Simulation, and Calculation of Electron Spin Resonance Spectra of a Nitroxide Spin Label on a Polyalanine α-Helix. *The Journal of Physical Chemistry B*. 112(18):5755–67

49. Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*. 98(12):10089–92

50. Eastman P, Pande V. 2010. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science & Engineering*. 12(4):34–39

51. Martínez L, Andrade R, Birgin EG, Martínez JM. 2009. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* 30(13):2157–64

52. Eastman P, Pande V. 2010. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science & Engineering*. 12(4):34–39

53. Trotzig C, Abrahmsén-Alami S, Maurer FHJ. 2007. Structure and mobility in water plasticized poly(ethylene oxide). *Polymer*. 48(11):3294–3305

54. Knauss R, Schiller J, Fleischer G, Kärger J, Arnold K. 1999. Self-diffusion of water in cartilage and cartilage components as studied by pulsed field gradient NMR. *Magnetic Resonance in Medicine*. 41(2):285–92

55. Foster KR, Cheever E, Leonard JB, Blum FD. 1984. Transport properties of polymer solutions. A comparative approach. *Biophys J*. 45(5):975–84

56. Abrahmsén-Alami S, Persson K, Stilbs P, Alami E. 1996. Effect of Temperature on NMR Self-Diffusion in Aqueous Associative Polymer Solutions. *J. Phys. Chem.* 100(11):4598–4605

57. Abrahmsén-Alami S, Stilbs P, Alami E. 1996. Water Self-Diffusion in Aqueous Associative Polymer Solutions. *J. Phys. Chem.* 100(16):6691–97

58. Gennes P-G de. 1979. *Scaling Concepts in Polymer Physics*. Cornell University Press. 336 pp.

59. Doi M, Edwards SF. 1988. *The Theory of Polymer Dynamics*. Clarendon Press. 420 pp.

60. Dubey V, Erimban S, Indra S, Daschakraborty S. 2019. Understanding the Origin of the Breakdown of the Stokes–Einstein Relation in Supercooled Water at Different Temperature–Pressure Conditions. *The Journal of Physical Chemistry B*. 123(47):10089–99

61. Bhanot C, Trivedi S, Gupta A, Pandey S, Pandey S. 2012. Dynamic viscosity versus probe-reported microviscosity of aqueous mixtures of poly(ethylene glycol). *The Journal of Chemical Thermodynamics*. 45(1):137–44

62. Lamanna R, Delmelle M, Cannistraro S. 1994. Solvent Stokes-Einstein violation in aqueous protein solutions. *Phys. Rev. E*. 49(6):5878–80

63. Meares P. 1954. The Diffusion of Gases Through Polyvinyl Acetate. *Journal of the American Chemical Society*. 76(13):3415–22

64. Cohen MH, Turnbull D. 1959. Molecular Transport in Liquids and Glasses. *The Journal of Chemical Physics*. 31(5):1164–69

65. Fujita H. 1961. Diffusion in polymer-diluent systems

66. Matsukawa S, Ando I. 1996. A Study of Self-Diffusion of Molecules in Polymer Gel by Pulsed-Gradient Spin−Echo 1H NMR. *Macromolecules*. 29(22):7136–40

67. Racko D, Chelli R, Cardini G, Bartos J, Califano S. 2005. Insights into positron annihilation lifetime spectroscopy by molecular dynamics simulations: Free-volume calculations for liquid and glassy glycerol. *Eur. Phys. J. D*. 32(3):289–97

68. Sherck N, Webber T, Brown DR, Keller T, Barry M, et al. 2020. End-to-End Distance Probability Distributions of Dilute Poly(ethylene oxide) in Aqueous Solution. *J. Am. Chem. Soc.* 142(46):19631–41

69. Cheng C-Y, Wang J-Y, Kausik R, Lee KYC, Han S. 2012. An ultrasensitive tool exploiting hydration dynamics to decipher weak lipid membrane–polymer interactions. *Journal of Magnetic Resonance*. 215:115–19

70. Schrader AM, Monroe JI, Sheil R, Dobbs HA, Keller TJ, et al. 2018. Surface chemical heterogeneity modulates silica surface hydration. *Proc Natl Acad Sci U S A*. 115(12):2890–95

71. Barnes R, Sun S, Fichou Y, Dahlquist FW, Heyden M, Han S. 2017. Spatially Heterogeneous Surface Water Diffusivity around Structured Protein Surfaces at Equilibrium. *J. Am. Chem. Soc.* 139(49):17890–901

72. Freed JH. 1978. Dynamic effects of pair correlation functions on spin relaxation by translational diffusion in liquids. II. Finite jumps and independent T_1 processes. *The Journal of Chemical Physics*. 68(9):4034–37

73. Hwang L, Freed JH. 1975. Dynamic effects of pair correlation functions on spin relaxation by translational diffusion in liquids. *J. Chem. Phys.* 63(9):4017–25

74. Overhauser AW. 1953. Polarization of Nuclei in Metals. *Phys. Rev.* 92(2):411–15

75. Carver TR, Slichter CP. 1953. Polarization of Nuclear Spins in Metals. *Phys. Rev.* 92(1):212–13

76. Armstrong BD, Han S. 2009. Overhauser Dynamic Nuclear Polarization To Study Local Water Dynamics. *Journal of the American Chemical Society*. 131(13):4641–47

77. Franck JM, Pavlova A, Scott JA, Han S. 2013. Quantitative cw Overhauser effect dynamic nuclear polarization for the analysis of local water dynamics. *Progress in Nuclear Magnetic Resonance Spectroscopy*. 74:33–56

78. Franck JM, Han S. 2019. Chapter Five - Overhauser Dynamic Nuclear Polarization for the Study of Hydration Dynamics, Explained. In *Methods in Enzymology*, ed AJ Wand. 615:131–75. Academic Press

79. González MA, Abascal JLF. 2011. A flexible model for water based on TIP4P/2005. *The Journal of Chemical Physics*. 135(22):224516

80. Zhu X, Riera M, Bull-Vulpe EF, Paesani F. 2023. MB-pol(2023): Sub-chemical Accuracy for Water Simulations from the Gas to the Liquid Phase. *J. Chem. Theory Comput.* 19(12):3551–66

81. Türke M-T, Bennati M. 2011. Saturation factor of nitroxide radicals in liquid DNP by pulsed ELDOR experiments. *Phys. Chem. Chem. Phys.* 13(9):3630

82. Armstrong BD, Han S. 2007. A new model for Overhauser enhanced nuclear magnetic resonance using nitroxide radicals. *The Journal of Chemical Physics*. 127(10):104508

83. Popp CA, Hyde JS. 1982. Electron-electron double resonance and saturation-recovery studies of nitroxide electron and nuclear spin-lattice relaxation times and Heisenberg exchange rates: lateral diffusion in dimyristoyl phosphatidylcholine. *Proc Natl Acad Sci U S A*. 79(8):2559–63

84. Enkin N, Liu G, Carmen Gimenez-Lopez M del, Porfyrakis K, Tkach I, Bennati M. 2015. A high saturation factor in Overhauser DNP with nitroxide derivatives: the role of 14 N nuclear spin relaxation. *Physical Chemistry Chemical Physics*. 17(17):11144–49

85. Chaimovich A, Shell MS. 2014. Tetrahedrality and structural order for hydrophobic interactions in a coarse-grained water model. *Physical Review E*. 89(2):

86. Monroe JI, Shell MS. 2019. Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *The Journal of Chemical Physics*. 151(9):094501

87. Stock P, Monroe JI, Utzig T, Smith DJ, Shell MS, Valtiner M. 2017. Unraveling Hydrophobic Interactions at the Molecular Scale Using Force Spectroscopy and Molecular Dynamics Simulations. *ACS Nano*. 11(3):2586–97

88. Errington JR, Debenedetti PG. 2001. Relationship between structural order and the anomalies of liquid water. *Nature*. 409(6818):318–21

89. Lum K, Chandler D, Weeks JD. 1999. Hydrophobicity at Small and Large Length Scales. *J. Phys. Chem. B*. 103(22):4570–77

90. Böhm F, Schwaab G, Havenith M. 2017. Mapping Hydration Water around Alcohol Chains by THz Calorimetry. *Angew. Chem. Int. Ed.* 56(33):9981–85

91. Jiao S, Rivera Mirabal DM, DeStefano AJ, Segalman RA, Han S, Shell MS. 2022. Sequence Modulates Polypeptoid Hydration Water Structure and Dynamics. *Biomacromolecules*. 23(4):1745–56

92. Ansaloni L, Minelli, Matteo, Baschetti, M. Giacinti, Sarti GC. 2014. Effect of relative humidity and temperature on gas transport in Matrimid(R): Experimental study and modeling. *Journal of Membrane Science*. 471:392–401

93. Moon JD, Galizia M, Borjigin H, Liu R, Riffle JS, et al. 2018. Water Vapor Sorption, Diffusion, and Dilation in Polybenzimidazoles. *Macromolecules*. 51(18):7197–7208

94. Moon JD, Galizia M, Borjigin H, Liu R, Riffle JS, et al. 2020. Modeling water diffusion in polybenzimidazole membranes using partial immobilization and free volume theory. *Polymer*. 189:122170

95. Lee WM. 1980. Selection of barrier materials from molecular structure. *Polymer Engineering & Science*. 20(1):65–69

96. Horn NR. 2016. A critical review of free volume and occupied volume calculation methods. *Journal of Membrane Science*. 518:289–94

97. Bondi A. 1964. van der Waals Volumes and Radii. *J. Phys. Chem.* 68(3):441–51

98. Park JY, Paul DR. 1997. Correlation and prediction of gas permeability in glassy polymer membrane materials via a modified free volume based group contribution method. *Journal of Membrane Science*. 125(1):23–39

99. Wu AX, Lin S, Rodriguez KM, Benedetti FM, Joo T, et al. 2021. Revisiting group contribution theory for estimating fractional free volume of microporous polymer membranes. *Journal of Membrane Science*, p. 119526

100. Edward JT. 1970. Molecular volumes and the Stokes-Einstein equation. *J. Chem. Educ.* 47(4):261

101. Moriguchi I, Kanada Y, Komatsu K. 1976. van der Waals Volume and the Related Parameters for Hydrophobicity in Structure-Activity Studies. *Chemical & Pharmaceutical Bulletin*. 24(8):1799–1806

102. Scherer JR, Bolton BA. 1985. Water in polymer membranes. 5. On the existence of pores and voids. *The Journal of Physical Chemistry*. 89(16):3535–40

103. Scherer JR. 1987. The partial molar volume of water in biological membranes. *Proc. Natl. Acad. Sci.* 84:7938–42

104. Bader RFW, Carroll MT, Cheeseman JR, Chang C. 1987. Properties of atoms in molecules: atomic volumes. *Journal of the American Chemical Society*. 109(26):7968–79

105. Trivedi S, Bhanot C, Pandey S. 2010. Densities of {poly(ethylene glycol)+water} over the temperature range (283.15 to 363.15)K. *The Journal of Chemical Thermodynamics*. 42(11):1367–71

106. Ayranci E, Sahin M. 2008. Interactions of polyethylene glycols with water studied by measurements of density and sound velocity. *The Journal of Chemical Thermodynamics*. 40(8):1200–1207

107. Volk A, Kähler CJ. 2018. Density model for aqueous glycerol solutions. *Exp Fluids*. 59(5):75

108. Tofts PS, Lloyd D, Clark CA, Barker GJ, Parker GJM, et al. 2000. Test liquids for quantitative MRI measurements of self-diffusion coefficient in vivo. *Magn. Reson. Med.* 43(3):368–74

109. Stock P, Monroe JI, Utzig T, Smith DJ, Shell MS, Valtiner M. 2017. Unraveling Hydrophobic Interactions at the Molecular Scale Using Force Spectroscopy and Molecular Dynamics Simulations. *ACS Nano*. 11(3):2586–97

110. Monroe J, Barry M, DeStefano A, Aydogan Gokturk P, Jiao S, et al. 2020. Water Structure and Properties at Hydrophilic and Hydrophobic Surfaces. *Annual Review of Chemical and Biomolecular Engineering*. 11(1):523–57

Chapter 4: Local water structures govern the mixing dynamics and thermodynamics of cryoprotectant mixtures

Adapted with permission from:

1. Debasish D. Mahanta[7], Dennis C. Robinson Brown, Simone Pezzotti, Songi Han Gerhard Schwaab, M. Scott Shell, Martina Havenith. *Local solvation structures govern the mixing thermodynamics of glycerol–water solutions. Chem Sci*. **2022**. DOI: 10.1039/D3SC00517H. Copyright 2022 Royal Society of Chemistry

2. Debasish D. Mahanta[8], Dennis C. Robinson Brown, Thomas R. Webber, Simone Pezzotti, Songi Han Gerhard Schwaab, M. Scott Shell, Martina Havenith. *Unraveling the Interplay between Structure, Dynamics, and Thermodynamics in Cryoprotectant Aqueous Solutions.* In Preparation

4.1 Introduction

As noted in **Chapter 2**, glycerol is a major cryoprotective agent (CPA)(1–4) used to promote protein stabilization upon exposure to supercooled temperatures as low as -196 C. Aqueous mixtures containing CPAs like glycerol mediate protein stabilization by inducing preferential solvation of the protein surface by water rather than glycerol(5). Other CPAs enhance protein stability in cryogenic environments via alternative mechanism such as in the case of DMSO which preferentially solvates the protein surface in contrast to glycerol. Often, heuristically designed multicomponent aqueous “recipes” containing DMSO, glycerol, and other additives are used improve the cryoprotective capacity of a CPA mixture. Incidentally, some CPA constituents like DMSO are cytotoxic at high concentration; hence, CPA mixtures must be tailored to maximize cryoprotection while minimizing toxicity. To date, the precise

---

[7] First-authorship shared between D.D.M. and D.C.R.B

[8] First-authorship shared between D.D.M and D.C.R.B

molecular-scale mechanisms underlying cryoprotection are poorly understood outside of mean field thermodynamic and kinematic approaches(6–8).

As glycerol-containing aqueous are nearly ubiquitous to the stabilization of biological materials *in vitro*, numerous experimental efforts have interrogated the glycerol-water structure such as dielectric spectroscopy(9–14), nuclear magnetic resonance (NMR)(15), x-ray photoelectron spectroscopy (XPS)(10), and infrared (IR) spectroscopy(4, 10, 16). These fundamental studies of glycerol-water mixtures provide insights into glycerol's influence on water's molecular structure. In one such example, Weeraratna et al. identified three distinct populations of water: (I) bulk water, (II) solvation waters hydrating glycerol, and (III) water molecules "confined" by glycerol using dielectric spectroscopy(10). Naturally, type I and II waters dominated at low glycerol concentration while type III waters dominated as glycerol concentration increases. However, the manner and magnitude by which these distinct molecular structural motifs impact the mixing thermodynamics of glycerol-water has not been previously investigated. Better understanding this structure-thermodynamics connection is an important precursor to interpreting the molecular mechanisms undergirding solute solvation thermodynamics in glycerol-water and rationally tailoring CPA mixture compositions for biomolecule cryoprotection.

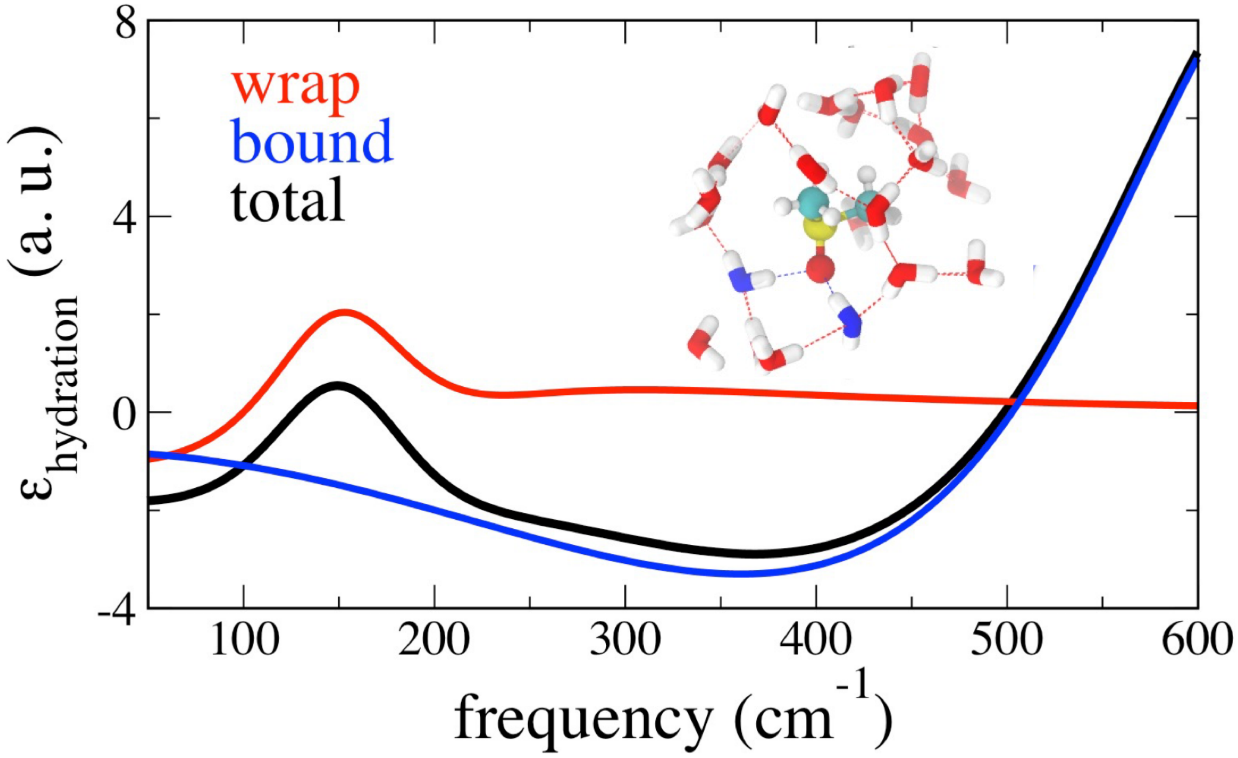


Figure 4.1 An idealized representation of a effective molar extinction coefficient spectra in DMSO-water is composed of the sum of wrap and bound water contributions. The wrap water contribution exhibits a peak in amplitude in the water-water HB-stretching region of the spectrum. The bound water contribution displays maximal enhancement in the THz frequency region corresponding to water librational modes (corresponding to greater steric hinderance owing to water-cosolvent hydrogen bonding). A simulation snapshot of a DMSO molecule illustrates the bound waters (in blue) and wrap waters (in red) local to the DMSO.

Unlike the above-mentioned experimental methods, Terahertz (THz) absorption spectroscopy directly probes the collective and solvation shell motions of water in aqueous mixtures without the need for model assumptions. The solvation shell of solute molecules were previously characterized by analyzing the difference THz spectra(10, 17–21)—subtracting the bulk water spectrum from the spectra of the glycerol-water mixtures. In calculating the difference spectrum, the prior works identified the spectroscopic features that probe distinct water structural motifs local to the hydrophobic and hydrophilic moieties of alcohols(17). I illustrate the water populations “wrapped” around hydrophilic moieties (“wrap” waters) and hydrogen bonded to the hydrophilic moiety ("bound" waters) of DMSO [**Figure 4.1**]. Specifically, the wrap and bound waters have THz-signatures in the 100-250 cm^{-1} and 350-600 cm^{-1} spectral regions, respectively. The population of wrap waters is proportional to the redshift in the water HB-stretching region (100-250 cm^{-1}) of the THz spectrum relative to bulk water—owing to weaker water-water hydrogen bonds in wrap waters(17, 19). On the other hand, increased water-solute hydrogen bonding—and hence, higher bound water population—leads to greater steric hinderance of water and a corresponding blueshift of the librational band of water (350-600 cm^{-1})(19). Notably, the balance of difference THz spectra signatures associated with hydration shell of hydrophobic moieties ("wrap" waters) and waters hydrogen bonded to hydrophilic moieties ("bound" waters) is predictive of alcohol solvation free energies(17, 19). This THz difference spectra data analysis protocol was given the apt name of "THz Calorimetry".

Terahertz (THz) absorption spectroscopy has the potential to address this challenge, especially when it is combined with the detailed molecular characterization offered by molecular dynamics (MD) simulations(10, 17). In particular, in a recent study on hydrated alcohols, we identified spectroscopic fingerprints in the THz range that directly probe the distinct structural motifs of the HB-network around the solute (17). We found that the water populations close to hydrophilic and hydrophobic moieties provide THz fingerprints that are well distinguishable from that of bulk water: "wrap" water molecules close to hydrophobic (CH_x) groups have a characteristic THz-signature in the 100-250 cm^{-1} region, where the translational motions of intermolecular stretching modes are active, while "bound" water molecules H-bonded to hydrophilic (OH) moieties have their imprint in the 350-600 cm^{-1} spectral range of water librations. Using THz-calorimetry to relate spectroscopic to thermodynamic quantities and MD simulations to independently quantify the distinct local contributions of the two identified populations to solvation entropy and enthalpy, we could experimentally show that the wrap population provides an entropic cost for (small) alcohol hydration, compensated by an enthalpic gain from the bound population(17, 22). These spectroscopic signatures of hydrophilic and hydrophobic hydration water populations are not specific of alcohol-water systems but have been recently observed for water hydrating a large variety of solutes and biomolecules(23–26).

In this chapter, we leverage the THz signatures or bound and wrap water structural motifs and atomistic molecular dynamics simulations to reveal the dependence of solvation thermodynamics on shifts in water's molecular structure in both glycerol-water and DMSO-water mixtures. In doing so, we establish a direct link between macroscopic thermodynamic metrics and sub-nanometer molecular structural details. This combined experimental-computational workflow suggests a path towards rational, physically-motivated design rules for CPA (or other stabilizing media) mixture compositions. Specifically, THz spectroscopy measurements directly probe the balance of local populations of bound and wrap water populations as a function of DMSO and glycerol concentration separately. Using atomistic molecular dynamics simulations, we perform detailed analyses of the water-water orientational ordering—via the relative population of tetrahedrally-coordinated waters p_{tet}–presented in different hydration water populations (e.g., bound and wrap waters). Further, both experimental and computational methods draw a direct connection between molecular structure and excess free energetic properties in glycerol-water and DMSO-water. Finally, we explore the relationship between molecular-level probes of water structure and ODNP-measured water dynamics to yield a similar structure-dynamics-thermodynamics connection to the one presented in **Chapter 2**.

4.2 MD Simulation Methodology

4.2.1 Glycerol-Water Simulation Protocol

We model glycerol-water mixtures using Bliek-Chelli parameters for glycerol(27, 28) and the OPC 4-site water model(29). This combination of models reliably reproduces the equilibrium thermophysical properties of glycerol-water under standard temperature and pressure conditions(29–31). Coulombic interactions are modelled using the particle-mesh Ewald summation scheme (PME)(32). We simulate glycerol-water mixtures with X_{gly} = 0, 0.012, 0.033, 0.05, 0.075, 0.10, 0.15, 0.20, 0.30, 0.40, 0.50, and 1.0 at 18°C and 1-bar using the GPU-optimized OpenMM package(33). All simulations are conducted in the NPT ensemble, with temperature and pressure held constant using a Langevin thermostat and Monte Carlo barostat(33), respectively. We generate initial configurations containing 729 glycerol molecules and vary the numbers of water molecules using Packmol software(34). We first minimize the energy of the initial configurations and equilibrate in the NPT ensemble for times ranging from 100 to 500 ns depending on the glycerol concentration. We conduct production runs for 100 ns with system coordinates saved every 100 ps. Following the simulations, we analyze the production run trajectories using the pytraj Python library(35) and in-house python code. For each simulation-computed quantity, we estimate the standard error of the mean using block averaging.

4.2.2 DMSO-Water Simulation Protocol

We model DMSO-water mixtures using the Fox-Kollman parameters(36) for DMSO and the TIP4P-EW 4-site water(37) model. Previous research has demonstrated that this specific combination of DMSO and water models reliably replicates standard temperature and pressure mixture properties such as density(36, 38). We model Coulombic interactions with the particle-mesh Ewald summation scheme (PME). To span the entire experimental composition range, we study the following exponentially spaced DMSO mole fraction, X_{DMSO} = 0, 0.01, 0.04, 0.06, 0.10, 0.16, 0.27, 0.45, 0.74, and 1.0 at 18°C and 1-bar using the GPU-optimized OpenMM package. Prior to conducting the MD simulations, we generate initial configurations containing 2000 water molecules with varying numbers of DMSO molecules using Packmol. Following system setup, we minimize energy minimize these initial configurations. Following energy minimization, we equilibrate the systems in the NPT ensemble for 100 ns by holding the temperature and pressure constant via a Langevin thermostat and Monte Carlo barostat, respectively. After setting up each system, we perform production simulations in two different ensembles: (1) NPT (using the same thermostat and barostat) and (2) NVE dynamics simulations. The NPT production runs are 100 ns long with system configurations recorded every 100 ps for later analysis. The NVE production runs are initiated from 5 separate, equally spaced NPT production run configurations, temperature equilibrated for 1 ns at 18°C, and finally 1 ns production runs. NVE production run coordinates are saved every 1 ps. All the analyses are conducted post-simulation using the pytraj.

4.2.3 MD Hydrogen Bonding Analyses

We define HBs via the widely-used geometric criteria of Luzar and Chandler(39), namely, O-H distance and $\angle O - OH$ angular cut-off values of 3.5 Å and 120° degrees, respectively. Tracking the waters participating in water-glycerol HBs, we directly quantify the mole fraction of bound waters (X_{bound}). We also calculate the mole fraction of wrap waters (X_{wrap}) by defining them as waters lying within the first hydration shell of glycerol molecules (~4.2 Å, as defined by the first minima of the 2D radial distribution function between glycerol heavy atoms and water oxygens), but not H-bonded to the glycerol molecule. Non-shell waters are simply all waters residing more than 4.2 Å from glycerol heavy atoms. We compute concentration of non-shell waters on a molar basis via $X_{non-shell} = (1 - X_{bound} - X_{wrap} - X_{gly})$.

4.2.4 MD Water Self-Diffusivity Analysis

Water self-diffusivity, D_{H_2O}, values are estimated from the results of the 5 separate NVE simulations via the slope of the mean-square displacement (MSD) curve

$$D_{H_2O} = \frac{1}{6} \lim_{t\to\infty} \frac{MSD(t)}{t} = \lim_{t\to\infty} \frac{\langle |r(t+\tau) - r(\tau)|^2 \rangle}{6t} \tag{4.1}$$

where r, t, and $\tau$are the position of a water oxygen, the current time step, and the initial time step. Here, $\langle\cdot\rangle$ is the ensemble average of a quantity across all water oxygen positions. We estimate the uncertainty in D_{H_2O} from the sample standard deviation.

4.2.5 MD Thermodynamics Analyses

Using the simulation-averaged molar energy $< E >$ and molar volume $< V >$ for each simulated glycerol-water mixture, we also estimate the molar energy of mixing $\Delta E_{mixing} =< E > - X_{gly} < E_{glycerol} > - \left(1 - X_{gly}\right) < E_{water} >$ and molar volume of mixing $\Delta V_{mixing} = < V > - X_{gly} < V_{glycerol} > - \left(1 - X_{gly}\right) < V_{water} >$. Here, $< Y_{water} >$ and $< Y_{glycerol} >$ correspond to the average molar value of Y for pure water and glycerol, respectively. The molar enthalpy of mixing ΔH_{mixing} hence stems from the relation $\Delta H_{mixing} = \Delta E_{mixing} + P\Delta V_{mixing}$ where P is system pressure (1 bar).

4.3 Characterizing Glycerol Hydration Layers

4.3.1 Analyzing THz Spectroscopic Signature of Bound Waters

The ensuing discussion of experimental results are adapted from our previous collaboration with Havenith et al(40); here, further details describing the experimental procedures and the theory underlying THz spectroscopy are given. To interrogate the spectroscopic signature of the glycerol-water hydration shell, we analyze the effective molar extinction coefficient, $\varepsilon_{hydration}(\nu)$ [depicted in **Figure 4.1**]. This hydration contribution to the molar extinction coefficient is given by

$$\varepsilon_{hydration}(\nu) = \varepsilon(\nu) - X_{gly}\varepsilon_{gly}(\nu) - X_{wat}\varepsilon_{wat}(\nu) \tag{4.2}$$

where ε_{gly} and ε_{wat} are the molar extinction coefficients of pure glycerol and water, respectively. Therefore, **Equation 4.3** is sensitive to the "excess" spectroscopic effect of mixing water and glycerol. By construction, $\varepsilon_{hydration}(\nu)$ identically approaches zero as $X_{gly} \rightarrow 1$ or 0. We depict the $\varepsilon_{hydration}(\nu)$ spectra for an experimental range from $X_{gly} =$ 0.01 to 0.8 in **Figure 4.2(a)**.

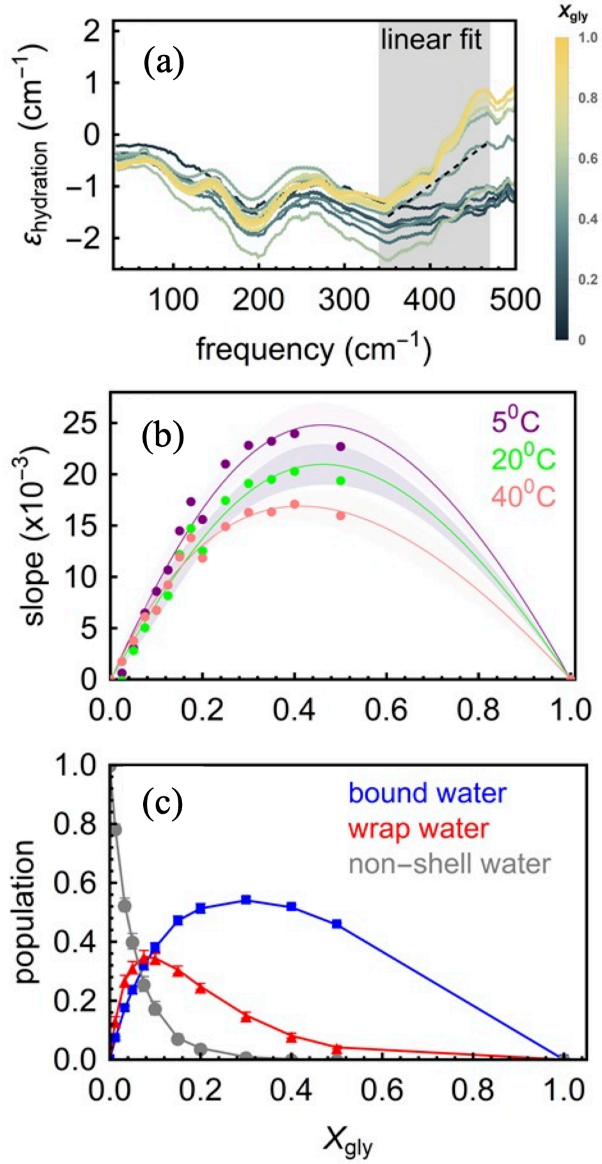


Figure 4.2 (a) Molar effective extinction coefficient ($\varepsilon_{hydration}$) at 20°C as calculated by subtracting the bulk water and bulk glycerol spectra from the glycerol-water mixtures spectra (equation 4). $\varepsilon_{hydration}$ increases with frequency for all the solutions in the 350-450 cm^{-1} spectral region (gray shaded area). The dotted black line is the linear fit in that region for glycerol mole fraction; $X_{gly} = 0.8$, as an example. The slope obtained from such fit is then used for interpreting the spectral trends. (b) The slopes derived from experimental $\varepsilon_{hydration}$ spectra are plotted for all X_{gly} at three different temperatures (5°C, 20°C and 40°C, respectively). The lines are guide to the eye and the error bars are indicated by the shaded regions. (c) The populations of various types of water in glycerol-water mixtures as predicted by MD simulations.

Notably, $\varepsilon_{hydration}(\nu)$ displays no meaningful concentration-dependent enhancement in the THz region corresponding to non-hydrogen bonding ("wrap") water in the ~100-250 cm^{-1} spectral region. Further, water librational spectral region (~350-600 cm^{-1})—corresponding to water solute hydrogen bonds—displays a clear concentration dependence. A previous computational-THz study demonstrated that the slope of this librational is directly proportional to the number of solute-bonding waters per solute molecule, N_{bound}(19, 41). In conjunction, these trends (or lack thereof) in wrap and bound water signatures suggest that glycerol is well-integrated into the solution hydrogen bonding network with nearly all hydration waters participating in glycerol-water hydrogen bonds(11).

In **Figure 4.2(b)**, we examine the dependence of the librational region slope on glycerol mole fraction at three separate temperatures (5°, 20°, and 40° C). For the present discussion, we take the magnitude of the slope to be a proxy measurement of N_{bound}. At all experimental temperature, N_{bound} increases monotonically for $X_{gly} < 0.4$ and decreases monotonically for $X_{gly} > 0.4$. While N_{bound} must inevitably go to zero for pure glycerol, the concentration of maximum N_{bound} ($X_{glyc} = 0.4$) is non-obvious. The initial increase in N_{bound} stems from increasing glycerol-water interactions. At $X_{glyc} = 0.4$, the number of glycerol-water hydrogen bonds saturates, and glycerol-glycerol interactions progressively dominate glycerol-water interactions for $X_{glyc} > 0.4$. These findings coincide the previous THz and Dielectric spectroscopy studies(10, 11) which demonstrated that, upon increasing X_{glyc}, solution structure transitions from being dictated by (i) water-water interactions, to (ii) water-glycerol interactions (e.g., bound waters), and finally (iii) glycerol-glycerol interactions at approximately $X_{glyc} = 0.4$.

To further characterize these structural transformations, we performed classical MD simulations and evaluated the relative mole fractions of wrap and bound water populations, as well as of bulk-like water not involved in the hydration of glycerol solutes (denoted hereafter non-shell water) as a function of glycerol content. We identify these three populations in the simulations based on structural criteria, i.e., on their proximity with respect to OH and hydrophobic glycerol moieties and their H-bonding properties. The simulated variations in the three populations with X_{gly} are shown in **Figure 4.2c**, as obtained by direct counting of water molecules belonging to each identified population. Most notably, the bound water population follows most closely the trend observed experimentally for the slope, i.e., first increasing with X_{gly} for low glycerol content, saturating at around X_{gly} = (0.3-0.4), and then decreasing with further increasing the glycerol mole fraction. The simulated wrap water population shows a similar non-monotonic behavior but reaches saturation for much lower glycerol mole fractions (X_{gly} = 0.1). This is easily understood by considering that wrap water interacts much more weakly than bound water with glycerol (since it is not H-bonded), and therefore it is the easiest part of the hydration layer that is stripped off as soon as the glycerol molecules in the system become too many to be fully hydrated. The small population of wrap water above $X_{gly} > 0.1$ and the near bulk character of the wrap water signature region of the THz spectra [see **Figure 4.2**] prevents an unambiguous identification of this population in the THz-spectra. The non-shell water population shows a straightforward monotonic decrease with X_{gly} because increasing numbers of water molecules are involved in the solvation of glycerol molecules as their concentration increases.

4.4 Characterizing DMSO Hydration Layers

We now consider the structural details of DMSO-water mixtures. Unlike glycerol, DMSO contains hydrophobic (-CH_3) and hydrophilic (-S=O) moieties. Hence, we anticipate a substantial wrap water component from THz experiments à la previous investigations of hydration layers of alcohol-water mixtures(19,64). The experimental procedures will be described in detail in our soon to be submitted article[9]. For the present discussion, we consider principal component analysis (PCA) of the wrap (~100-250 cm^{-1}) and bound (~350-600 cm^{-1}) regions of the effective extinction spectra, $\varepsilon_{hydration}(\nu)$, for DMSO-water mixtures ranging from $X_{DMSO} = 0.01$ to 0.95. In **Figure 4.3(a)**, we depict the projections of extinction coefficient regions corresponding to wrap

$$P_{wrap} = proj\left[\varepsilon_{hydration}(100 < \nu < 250cm^{-1}), PC_{wrap}\right] \tag{4.3}$$

and bound

$$P_{bound} = proj\left[\varepsilon_{hydration}(350 < \nu < 600cm^{-1}), PC_{wrap}\right] \tag{4.4}$$

waters onto their first principal components P_{wrap} and P_{bound}, respectively. P_{wrap} increases with DMSO concentration until $X_{DMSO} = 0.35$ [**Figure 4.3(a)**]. Similarly, P_{bound} exhibits a maximum but at a higher DMSO concentration ($X_{DMSO} = 0.6$). Hence, upon increasing X_{DMSO} from 0 to 1, we fist observe the preferential hydration of the hydrophobic moieties (-CH_3) of DMSO for $X_{DMSO} < 0.35$. For $X_{DMSO} > 0.35$, the -CH_3 groups de-wet, and water begins to preferentially hydrate the hydrophilic moiety (-S=O). The competition between -CH_3 and -S=O hydration suggests a micro-heterogeneous structure of DMSO hydration waters.

We directly observe these distinct water structuring regions using analyses from the MD simulations [**Figure 4.6**]. Specifically, 2D radial distribution functions for water relative to the CH_3 [$g(r_{Ow-CH_3})$] and S=O [$g(r_{Ow-S=O})$] moieties display distinct signals. $g(r_{Ow-S=O})$ demonstrates the strong water-O_{DMSO} interaction with the first peak centered at ~0.28 nm. With its first peak centered at ~0.36 nm, $g(r_{Ow-CH_3})$ suggests that weaker water-CH_3 interactions. Further, water exists at higher local density near -S=O than -CH_3 as indicated by the higher

---

[9] This is manuscript #2 mentioned at the beginning of **Chapter 4**.

first peak amplitude of $g(r_{Ow-S=O})$ compared to $g(r_{Ow-CH_3})$ (between 25 and 300% higher). The magnitude of the first peak amplitude increases with X_{DMSO} for both RDFs.

To complement the THz measurements and accompanying PCA, we leverage classical MD simulation to directly compute the bound, wrap, and non-shell water mole fractions in water-DMSO as described the Methods section. We summarize the trends in these three mole fractions in **Figure 4.3(b)**. We note that the simulation computed wrap and bound water mole fractions exhibit qualitative agreement with the bound and wrap water signals indicated by the PCA scores of the experimental THz spectra [**Figure 4.3(a)**]. Specifically, X_{wrap} and X_{bound} saturate at DMSO mole fractions of $X_{DMSO} = 0.16$ and 0.45, respectively. Notably, the peak amplitude of X_{wrap} is 46% larger than in water-glycerol from our previous study(40). DMSO form fewer HBs with water than glycerol, which explains the relatively higher X_{wrap}. The wrap to bound water trade-off at $X_{DMSO} = 0.45$ stems from the near 1:1 ratio of water:DMSO paired with the preferential association of hydration of water near the HB-accepting sulfoxide oxygen. Naturally, as X_{DMSO} increases, larger portions of the water enter the hydration layer of DMSO molecules; hence monotonically decreasing the mole fraction of non-shell waters.

We note the quantitative difference between wrap and bound water signals derived from THz $\varepsilon_{hydration}$ spectra [**Figure 4.3(a)**] and MD simulations [**Figure 4.3(b)**]. This largely stems from the fundamental difference between the two measurements. For instance, normalizing X_{bound} and X_{wrap} by their sum (e.g., $\frac{X_{bound}}{X_{bound}+X_{wrap}}$) yields the fraction of the DMSO hydration shell occupied by bound and wrap waters, respectively. On the other hand, there is no obvious method of combining the bound and wrap principal component scores to reconstruct the composition of DMSO's hydration shell. To draw a more quantitative comparison between the principal components and MD hydration shell signatures, we rescale X_{bound} and X_{wrap} to new variables S_{bound} and S_{wrap}

$$S_i = \frac{X_i}{X_{bound}+X_{wrap}}\left(P_{bound} + P_{wrap}\right) \qquad (1)$$

where P_{bound} and P_{wrap} are the principal component scores corresponding to the bound and wrap water shown in **Figure 4.3(a)**. We depict these rescaled variables in **Figure 4.3(c)**. While the quantitative agreement between S_i and P_i is imperfect for both bound and wrap waters, we observe greater agreement between the peak positions of S_i and P_i than between X_i and P_i.

Specifically, we illustrate the rightward shift in the MD bound water signature peak location from $X_{DMSO} \approx 0.45$ into better agreement with the peak position of P_{bound} at $X_{DMSO} \approx 0.60$. Similarly, the MD wrap water signature peak position moves closer to the peak position of P_{wrap} by shifting from $X_{DMSO} = 0.16$ to 0.27.

Though the rescaled MD variables reproduce the essential trends of the THz principal component scores, we observe difference between the two in terms of signal amplitude and shape. In part these differences stem from the selected HB geometric criteria and assumed first hydration shell size. In **Figure 4.7**, we demonstrate the systematic effect of varying the hydration layer cut-off distance. Here, we find that increasing cut-off radii yields a systematic increase in the wrap water mole fraction X_{wrap}. On the other hand, X_{bound} is unaffected by the choice of hydration shell cutoff for radii larger than the HB distance criteria (e.g., >3.0 Å). Regardless of the chosen cut-off, the essential trends in X_{wrap} and X_{bound} remain unchanged. Additionally, while the DMSO force field parameters yield exceptionally accurate mixture densities at low concentrations ($X_{DMSO} < 0.27$), model performance marginally worsens at higher concentrations. This behavior is reflected in the indistinguishable model and experimental density for $X_{DMSO} < 0.27$ and the 2% larger model density for pure DMSO [**Figure 4.8**].

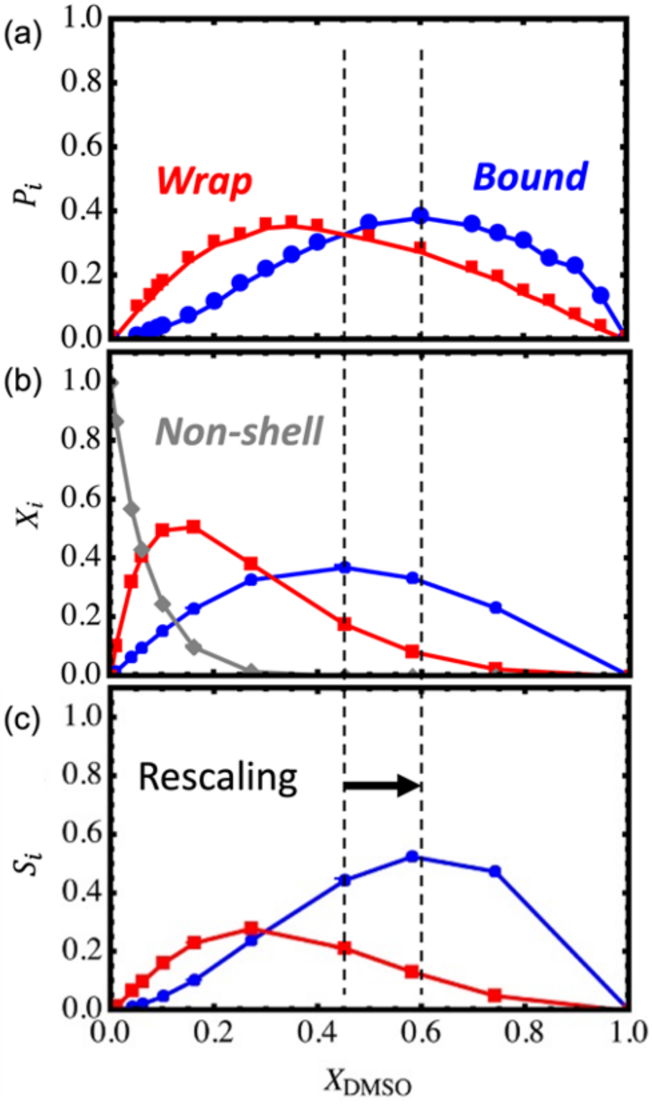


Figure 4.3 (a) The scores associated with the principal components of wrap, P_{wrap}, (red) and bound, P_{bound}, (blue) water display maxima at $X_{DMSO} \approx 0.35$ and 0.60, respectively. (b) Direct calculation of the mole fraction of wrap (X_{wrap}) and bound (X_{bound}) waters from the MD simulations demonstrate wrap and bound maxima at $X_{DMSO} \approx 0.16$ and 0.45, respectively. The non-shell waters are all waters not lying within the hydration shell of DMSO molecules. Hence, we define water mole fraction of non-shell waters by $X_{Non-shell} = X_{wat} - X_{wrap} - X_{bound}$. As X_{DMSO} increases, $X_{Non-shell}$ goes to zero because all waters are considered a part of DMSO hydration shells. (c) We rescale the MD computed X_{bound} and X_{wrap} to map to the THz-derived principal component scores. Both the peaks of the rescaled wrap, S_{wrap}, and bound, S_{wrap}, water signals shift closer to the respective wrap and bound water spectroscopic principal component scores. The rescaling procedure is described in detail in the main text.

4.5 Analyzing the Molecular Structure of Bound and Wrap Waters

Beyond reproducing the hydration structural information from THz experiments, we further analyze the MD trajectories to directly characterize the impact of increasing concentration on the hydration layer structure about DMSO. As detailed in our previous work(40) and the water literature more broadly(42–46), the orientational ordering of water in the vicinity of small molecules affects mixture thermodynamics. Specifically, we enumerate the water triplet angle configurations and analyze the resulting distribution, $P_{3b}(\theta)$. For the present analysis, we highlight angles corresponding to shifts in the tetrahedral ($\theta = 109.5°$) and simple fluid like ($\theta = 64°$) populations.

In **Figure 4.4(a)**, we present triplet angle distributions relative to the pure water distribution, $\Delta P_{3b}(\theta) = P_{3b}(\theta) - P_{3b,pure}(\theta)$, for bound and wrap water populations at X_{DMSO} from 0.01 to 0.74. Notably, the bound water tetrahedrality is less than that of pure water, $\Delta P_{3b,bound}(\theta = 109.5°) < 0$, while its icosahedrality is greater to compensate, $\Delta P_{3b,bound}(\theta = 64°) > 0$, for X_{DMSO}<0.27. The relative decrease in tetrahedrality for small X_{DMSO} stems from the formation of DMSO-water hydrogen bonds and thus decreased water-water orientational coordination. However, the relative population of tetrahedrally-coordinated bound waters increases systematically and eventually supersedes pure water tetrahedrality for X_{DMSO}>0.27. On the other hand, wrap waters exhibit enhanced tetrahedrality (and decreased icosahedrality) relative to pure water for the entire range of X_{DMSO}. Further, the tetrahedral enhancement of wrap waters residing near hydrophobic CH_3 groups coincides with previous findings(47–50) for water in the hydration shell of small (<1 nm in radius) hydrophobic moieties. We also note the presentation of a $\Delta P_{3b}(\theta)$ enhancement centered at $\theta \approx 130°$ in $\Delta P_{3b,bound}$. The enhancement at $\theta \approx 130°$ further distinguishes the structural features of water near the hydrophilic moieties of DMSO from the hydrophobic moieties. In **Figure 4.9**, we demonstrate that these high-θ ($\theta > 120°$) enhancements owe to monotonic increase in $P_{3b}(\theta > 120°)$ with increasing X_{DMSO}. The discrepant signatures of $\Delta P_{3b}(\theta)$ for bound and wrap waters further supports the presentation of nanoscopically distinct wetting (=O) and dewetting ($-CH_3$) regions of the DMSO hydration shell as X_{DMSO} increases.

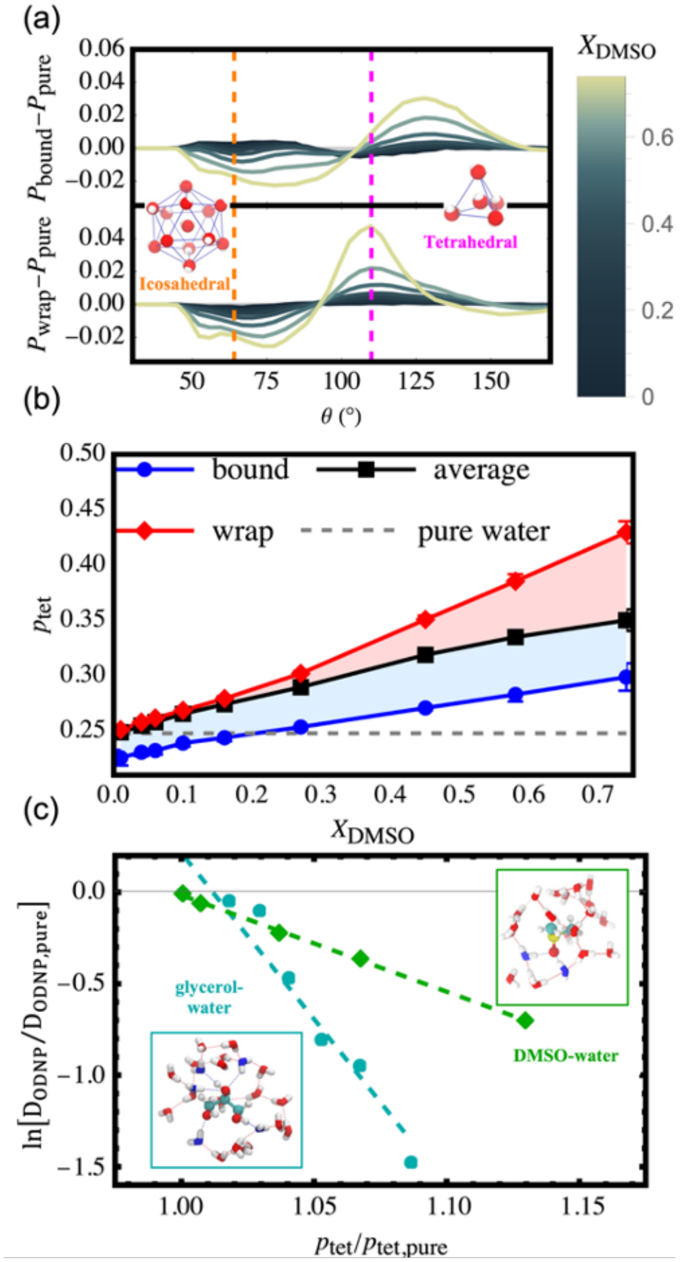


Figure 4.4 (a) We depict differential triplet angle distributions, $\Delta P_{3b,i}(\theta) = P_{3b,i}(\theta) - P_{pure}(\theta)$, for the bound and wrap water populations in DMSO-water. The wrap waters show a monotonic increase in tetrahedrality with increasing X_{DMSO}. On the other hand, bound water displays lower tetrahedrality relative to pure water for $X_{DMSO} < 0.27$ and higher tetrahedrality

relative to pure water for all $X_{DMSO} > 0.27$. These trends in tetrahedrality are numerically reflected by (b) the population of tetrahedrally coordinated waters, $p_{tet,i} = \int_{100°}^{120°} P_{3b,i}(\theta)\, d\theta$. Notably, $p_{tet,wrap}$ suggests that wrap waters become increasingly more tetrahedral than the system average over all water molecules (black) at higher X_{DMSO}. $p_{tet,bound}$ illustrates that bound waters are consistently ~10 to 15% less tetrahedrally coordinated than pure water. (c) The system averaged population of tetrahedral waters relative to pure water, $p_{tet}/p_{tet,pure}$, demonstrates remarkable corelation with the log-diffusivity of water in both glycerol-water (cyan) and DMSO-water (green) obtained from Overhauser Dynamic Nuclear Polarization (ODNP) experiments. The insets are simulation snapshots of the water environment within the hydration shell of DMSO (green box) and glycerol (cyan box) molecules at a water mole fraction of $X_{H_2O} = 0.9$. Here, blue and red oxygen atoms correspond to bound and wrap water molecules, respectively. Similarly, blue dashed lines illustrate hydrogen bonds between bound waters and the solute while red dashed lines illustrate hydrogen bonds between water molecules.

We directly quantify shifts in the tetrahedral population for bound and wrap water populations by integrating the corresponding distribution, $P_{3b}(\theta)$, over the (approximate) tetrahedral region, $100° < \theta < 120°$, to estimate the population of tetrahedrally coordinated waters, p_{tet} [**Figure 4.4(b)**]. At low X_{DMSO}, we find that the system-average (all water molecules included) p_{tet} is approximately the same magnitude of the wrap water tetrahedral population, $p_{tet,wrap}$. As observed in **Figure 4.4(b)**, the bound water tetrahedral population, $p_{tet,bound}$, lies below the system-average for $X_{DMSO} < 0.27$. However, X_{bound} is lower than X_{wrap} at low X_{DMSO}; hence, the dramatic reduction in $p_{tet,bound}$ does not strongly impact p_{tet} for small X_{DMSO} values. For larger values of X_{DMSO}, p_{tet} approaches $p_{tet,bound}$ as wrap waters are gradually replaced by bound waters [**Figure 4.3(b)**].

4.6 Developing the structure-dynamics connection

The experimental and computational analyses discussed above detail the structural motifs of DMSO-water but do not fully describe its molecular properties. Prior studies suggest that aqueous systems exhibit strong relationships between water structure and dynamics at the nanoscale(51–54). To directly probe water dynamics, we measure water diffusion coefficients, D_{ODNP}, in DMSO-water using Overhauser dynamic nuclear polarization (ODNP) experiments.

In **Figure 4.4(c)**, we depict the log-normalized ODNP water diffusion coefficient, $\ln\left[D_{ODNP}/D_{ODNP,pure}\right]$, against the normalized, MD-computed tetrahedral population, $p_{tet}/p_{tet,pure}$, for $X_{DMSO} = 0.01, 0.06, 0.10, 0.20$, and 0.30. Though ODNP and MD yield different absolute water self-diffusivities D_{H_2O}, both methods yield the same essential D_{H_2O} trend (with $R^2 \approx 0.95$) [**Figure 4.10**]. Notably, the log-diffusivity decreases linearly with $p_{tet}/p_{tet,pure}$. Hence, given that $p_{tet}/p_{tet,pure}$ increases nearly monotonically with X_{DMSO} [**Figure 4.4(b)**], $\ln\left[D_{ODNP}/D_{ODNP,pure}\right]$ decreases with the X_{DMSO} as well. By applying a linear fit to these data, we discover a remarkably well-defined relationship between molecular structure (via p_{tet}) and translational water dynamics (via D_{ODNP}). We coplot results for glycerol-water from our previous work,[27] finding a similarly strong structure-dynamics relationship. We note that the slope of the correlation differs based on the cosolvent, likely an effect of glycerol's three strong hydrogen bond donating groups. These correlations show how experimentally measured water diffusivity closely tracks the underlying structural changes in water.

Structure-dynamics relationships in aqueous environments have previously been explored in the context of supercooled water(55–59), hydration waters(25, 51, 53, 54, 60–62), and aqueous mixtures(20, 30, 39, 52, 63). In a recent computational study, we detailed emergent structure-diffusivity relations across a broad array of binary mixtures of varying temperature, composition, and cosolvent chemical identity(52). Further, we demonstrated that simple linear regression yields quantitative predictions of water self-diffusivities for 59 distinct simulations using only p_{tet} and two additional water structural metrics. The structure-dynamics connection shown in **Figure 4.4(c)** expand on these prior works by demonstrating the persistence of structure and dynamics even when derived computationally and experimentally, respectively. Further, the analyses pursued in the present work offer a workflow for co-interpreting experimental probes of molecular structure (e.g., THz) and dynamics (e.g., ODNP) by leveraging MD simulation calculations.

We performed a similar three-body angle distribution analysis for glycerol-water. The details for these calculations are found in recent publication in Chemical Science(40).

4.7 Developing the structure-dynamics connection

The structural changes detailed in **Section 4.5** are further anticipated to be affect the mixing thermodynamics of CPA-water mixtures. In the present discussion, we highlight the case of glycerol-water. Naturally, bound water population is connected to the extent of glycerol-water interactions and hence the enthalpy of mixing. Further, as previously demonstrated in THz calorimetry studies, the balance of wrap and bound THz signatures are suggestive of entropic and enthalpic trends, respectively(17, 64). Given that the THz spectrum only resolves trends in bound water, we analyze how THz signatures determine the enthalpy of mixing of glycerol-water in connection to MD simulations. However, for primary alcohols, which are more hydrophobic than glycerol, the global thermodynamic properties were found dominated by the wrap water population (i.e. by the hydrophobic solvation mechanism)(17, 64). **Figure 4.5(a)** demonstrates that this picture is reversed for water-glycerol mixtures. In the figure, we compare the spectroscopically measured slope with available mixing enthalpy (ΔH_{mixing}) data(65–67), as well as theoretical mixing enthalpy values from the present MD simulations, as a function of X_{gly}. The inverse of the slope is plotted since we expect bound water to contribute favorably to the mixing enthalpy (i.e. ΔH_{bound}<0 for each bound water on average). Within a 95% confidence interval, the slope and mixing enthalpy data correlate extremely well, suggesting that the changes in mixing enthalpy with X_{gly} can be quantitatively explained by considering the enthalpic term due to the variations in the bound water population, only. This is better demonstrated in **Figure 4.5(b)**, where ΔH_{mixing} is plotted against the slope, and a linear correlation is observed.

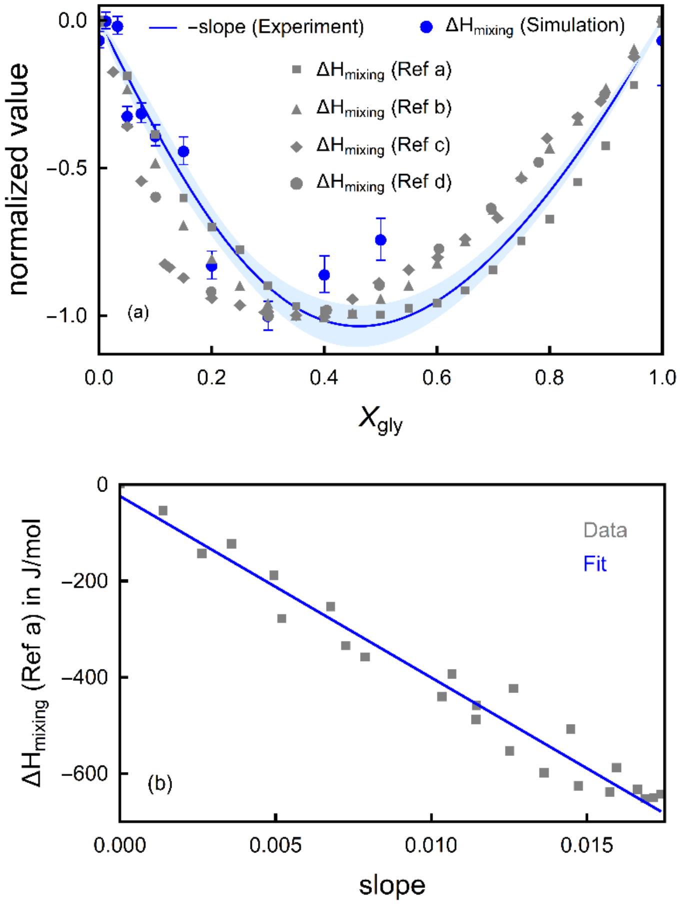


Figure 4.5. (a) Correlation of the experimentally determined slope (associated to the bound water population) with mixing enthalpies values (DH_{mixing}) as a function of glycerol content. DH_{mixing} values are taken from four literature data sets: Ref a; calculated (9), Ref b; extrapolated (9), Ref c (43), Ref d (42), as well as from the present MD simulations. The data sets are scaled (from 0 to -1) for better comparison. (b) Linear correlation between DH_{mixing} (from Ref a, as an example) and slope.

4.8 Conclusion

In summary the addition of CPAs to aqueous solutions alters the local molecular structures of hydration water and influences mixing thermodynamics and water dynamics. The basic understanding provided here on structure-thermodynamics and structure-dynamics relations offers a basis for a rational optimization of CPA-water mixture properties by tailoring the mixture composition. Our study highlights the importance of explicit inclusion of CPAs in simulation studies of proteins for the purpose of comparing computed conformational dynamics and interactions to experimental studies that may require CPAs as solution constituents, e.g., in the study of Double Electron-Electron Resonance (DEER) (61). By combining THz spectroscopy with classical MD simulations, we characterize way that collective molecular structure dictates the thermophysical properties of CPA-water solutions. We built our interpretation on the "wrap" and "bound" hydration water spectroscopic populations, i.e., well-defined THz markers that allow the direct monitoring of the changes in the hydrophobic and hydrophilic solvation mechanisms as a function of mixture composition, respectively. We propose that the mixing enthalpy between water and glycerol—a macroscopic property—is dominated by local contributions from the (hydrophilic) hydration water populations directly bound to glycerol. For glycerol, the bound water and not the wrap water population dictates mixing enthalpy since glycerol fits nicely into the water network and vice versa. By contrast, amphiphilic molecules like DMSO, present a more prominent entropically driven ("wrap") water population. Further, both DMSO- and glycerol-water mixtures display a strong structure-water dynamics connection—specifically between water tetrahedrality/self-diffusivity—as verified by complementary ODNP experiments. Though each mixture displays a distinct structure-dynamics relationship, they both exhibit a monotonic exponential relationship. The general picture emerging by combining these results is that the balance between local hydrophobic/wrap and hydrophilic/bound contributions dictates global mixture properties. The balance depends crucially on the size of the molecules that are mixed with water as well as on the number and distributions on polar groups. The findings detailed in this chapter offer exciting opportunities to tune in future studies the enthalpy and entropy of CPA-water mixtures by rationally modulating the nature of the mixture composition. For example, from the present results we can anticipate that the balance between the released water that is lost by the cell during freezing and the bound water population H-Bonded to a given cryoprotective solution (e.g., glycerol/DMSO/water) and biomolecules—a key parameter for cryoprotection—can be tailored by experimentally or computational exploiting the bound/wrap water signature and making direct connection to molecular properties of interest.

4.9 Appendix

4.9.1 Analysis of Distinct Water Environments Near Hydrophobic (-CH3) and Hydrophilic (=O)

The distinct low (wrap) and high (bound) THz signatures suggest the distinct "micro-heterogeneous" water structures near the polar (-S=O) and apolar (-CH_3) moieties of DMSO. We directly probe this "micro-heterogeneity" by analysing the 2D radial distribution functions (RDFs) of water near these polar and apolar moieties. At all water-DMSO concentrations, the water-(S=O) RDF $g(r_{Ow-S=O})$ exhibits "liquid-like" structuring. On the other hand, $g\left(r_{Ow-CH_3}\right)$ adopts a sigmoidal "de-wetting" behavior for all concentrations.

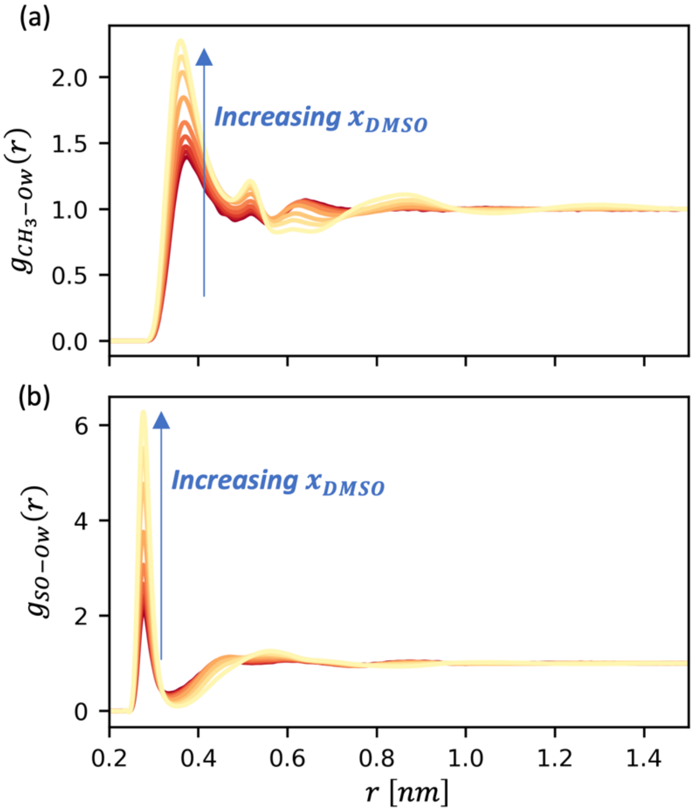


Figure 4.6. We depict the 2D radial distribution functions of water relative to the -CH3 [$g(r_{Ow-CH_3})$] and -S=O [$g(r_{Ow-S=O})$] moieties of DMSO. (a) $g(r_{Ow-CH_3})$] exhibits weaker water-CH3 interaction as suggested by the position of the first peak (~0.36 nm) compared to (b) the position of $g(r_{Ow-S=O})$'s first peak (~0.28 nm). Further, $g(r_{Ow-S=O})$ presents a higher first peak amplitude than $g(r_{Ow-CH_3})$, suggesting persistantly higher local water density near -S=O than -CH3. Here, the first peak amplitude difference varies between 25% ($x_{DMSO} = 0.01$) and 300% ($x_{DMSO} = 0.74$).

4.11.2 Sensitivity of Bound and Wrap Water Mole Fractions to Geometric Cutoffs

In **Figure 4.3** of the main text, we depict wrap and bound water mole fractions for fixed hydrogen bond criteria ($\angle O-OH = 150°$ and donor H-acceptor O distance 3 Å) and DMSO hydration layer cutoff distance (4.2 Å). Varying the values of these geometric criteria does not change the qualitative trends in bound x_{bound} and wrap x_{wrap} water mole fractions. Fixing the hydrogen bond angle criterion ("hbAng") and DMSO hydration layer cutoff ("cutoff") values of 150° and 4.2 Å while varying hydrogen bond distance criterion ("hbDist") between 3.0 and

4.0 Å, we observe a ~7% increase in the maximum x_{bound} at $x_{DMSO} = 0.45$ [**Figure 4.7(a)**]. Simultaneously, the maximum x_{wrap} decreases by ~3% [**Figure 4.7(d)**]. These effects stem from the relaxation of the definition of a DMSO-water hydrogen bonding pair upon increasing "hbDist". Fixing "hbDist" and "cutoff" values of 3.0 Å and 4.2 Å while varying "hbAng" between 120° and 170° we observe a ~70% decrease in the maximum x_{bound} at $x_{DMSO} = 0.45$ [**Figure 4.7(b)**]. Simultaneously, the maximum x_{wrap} increases by ~30% [**Figure 4.7(e)**]. Increasing "hbAng" further constrains the definition of a DMSO-water hydrogen bond and hence decreases x_{bound} and increasing x_{wrap}. Neither hydrogen bonding criterion appears to shift the peak location of x_{bound} towards the experimental observation [**Figure 4.3(a)**].

Fixing "hbDist" and "hbAng" values of 3.0 Å and 150° while varying "cutoff" between 3.0 and 6.0 Å we observe a near 2000% increase in the maximum x_{wrap} [**Figure 4.7(c)**]. x_{bound} remains unchanged because the hydrogen bonding criteria are fixed [**Figure 4.7(f)**]. Increasing "cutoff" relaxation the definition of a wrap water and hence increasing x_{wrap}. Decreasing the "cutoff" does shift the maximum x_{wrap} towards the observed experimental maximum. However, we do not obtain quantitative agreement between THz-derived and simulation-derived wrap-bound analyses in the absence of a rescaling like the one presented in the main text [**Figure 4.3(c)**].

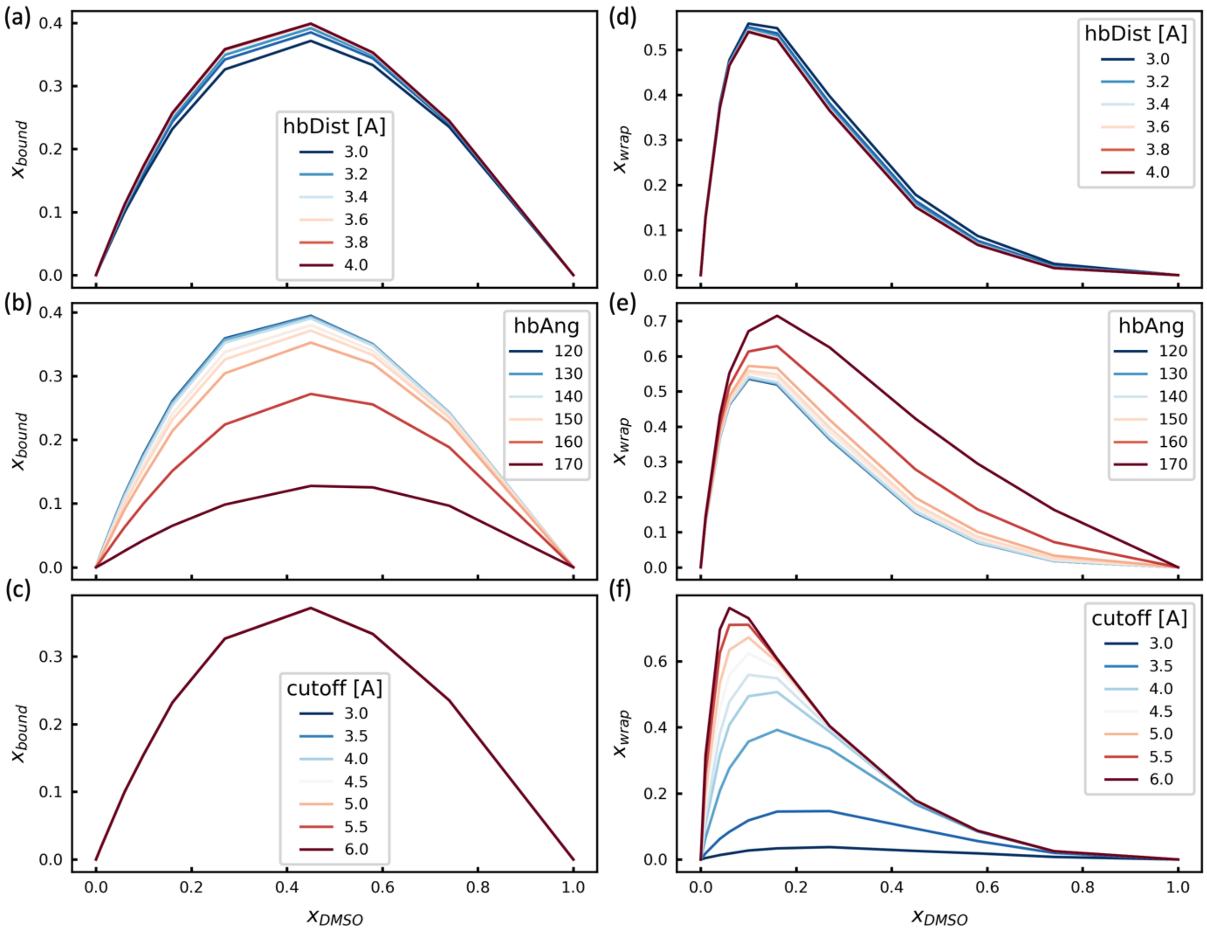


Figure 4.7. We depict bound water mole fractions x_{bound} as a function of DMSO mole fraction x_{DMSO} for (a) hydrogen bond angle criterion ("hbAng") and DMSO hydration layer cutoff ("cutoff") values of 150° and 4.2 Å with varying hydrogen bond distance criterion ("hbDist"), (b) "hbDist" and "cutoff" values of 3 Å and 4.2 Å with varying "hbAng", and (c) "hbDist" and "hbAng" values of 3 Å and 150° with varying "cutoff". We depict wrap water mole fractions x_{wrap} as a function of DMSO mole fraction x_{DMSO} for (d) "hbAng" and "cutoff" values of 150° and 4.2 Å with varying "hbDist", (e) "hbDist" and "cutoff" values of 3 Å and 4.2 Å with varying "hbAng", and (f) "hbDist" and "hbAng" values of 3 Å and 150° with varying "cutoff".

4.9.3 DMSO-Water Density Compared to Experimental Literature

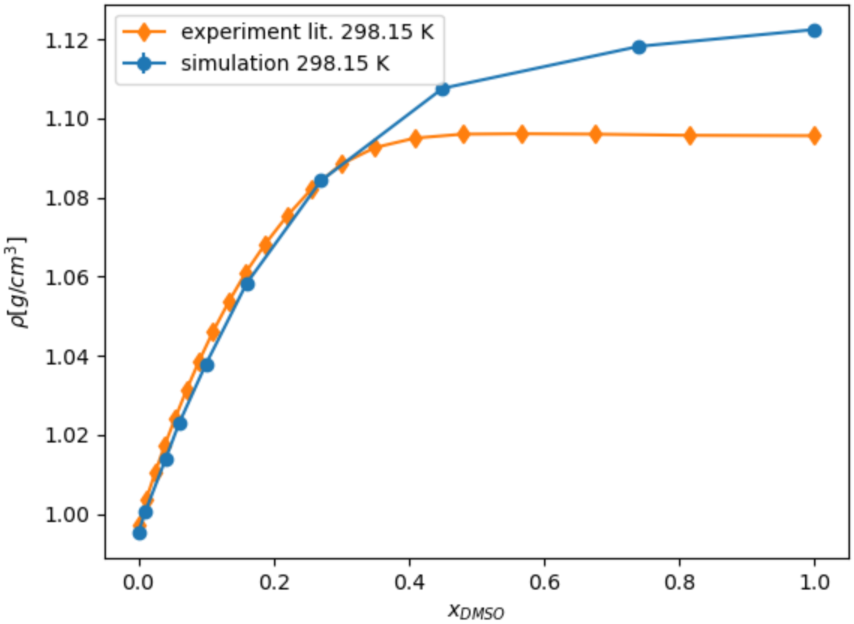


Figure 4.8. DMSO force field parameters yield reasonably accurate DMSO-water densities at low concentrations ($X_{DMSO} < 0.27$). Model accuracy decreases at higher concentrations. Specifically, experimental and MD densities are nearly indistinguishable for $X_{DMSO} < 0.27$. Simulations of pure DMSO yield 2% larger density compared to experiment.

4.9.4 Further Details on the Three-Body Angle Distributions

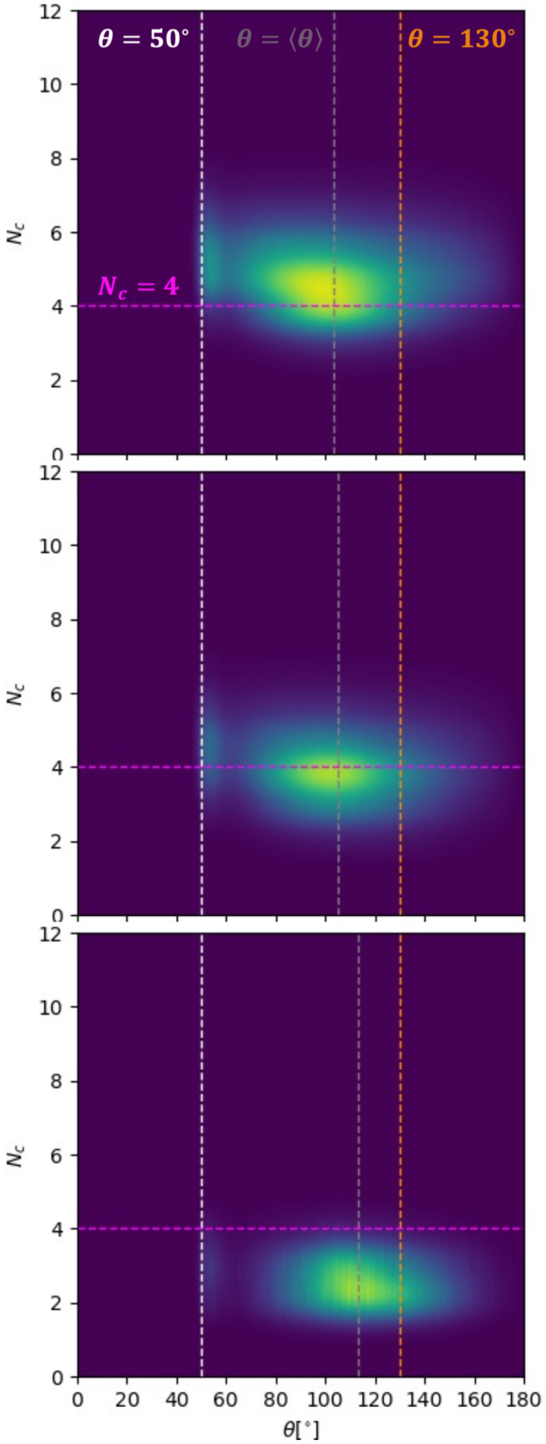


Figure 4.9. We depict 2D probability densities of water three body angles (x-axis) and the number of water molecules within a 0.34 nm cutoff of central water molecules (y-axis) for DMSO mole fractions of (top) 0.01, (middle) 0.16, and (bottom) 0.58. Here, the maximum probability density is represented in yellow and the minimum density (zero) is presented in dark blue. In each panel, vertical dashed white, grey, and orange lines correspond to $\theta = 50°$,

$\theta = \langle\theta\rangle$, and $\theta = 130°$, respectively. $\langle\theta\rangle$ is the mean three-body angle at the given concentration. The horizontal line indicates 4-coordinated waters ($N_c = 4$). As X_{DMSO} increases, the population shifts to lower N_c and higher θ. At low concentrations (e.g., $X_{DMSO} = 0.01$), there is a small population corresponding to overcoordinated waters centered at $\theta \approx 50°$ and $N_c \approx 5$. As X_{DMSO} increases, the "overcoordinated" population remains centered at $\theta \approx 50°$, but N_c decreases.

4.9.5 Comparison of ODNP and MD Water Self-Diffusivities

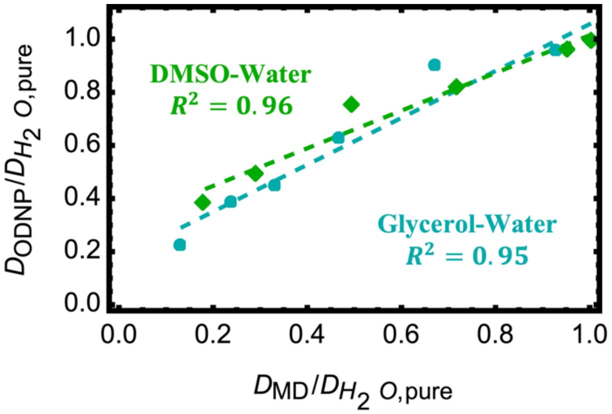


Figure 4.10. A direct comparison ODNP and MD-derived estimates of water self-diffusivity for glycerol-water(52) and DMSO-water. Here, ODNP (D_{ODNP}) and MD (D_{MD}) estimates are normalized by the pure water self-diffusivity from ODNP and MD, respectively. Though the two methods are not in perfect quantitative agreement, ODNP and MD yield similar qualitative trends for both DMSO-water ($R^2 = 0.96$) and glycerol-water ($R^2 = 0.95$)

Acknowledgements

M.H., G.S., S.P. and D.D.M. acknowledge financial support by European Research Council (ERC) Advanced Grant 695437 THz-Calorimetry. This study is funded by the Deutsche Forschungsgemeinschaft (DFG, German Re- search Foundation) under Germany's Excellence Strat- egy—EXC2033–390677874—RESOLV. We thank the Mercator Research Center Ruhr (MERCUR) for funding. M.S.S., and D.R.B. acknowledge support from the Center for Materials for Water and Energy Systems, an Energy Frontier Research Center funded by the US Department of Energy, Office of Science, Basic Energy Sciences under Award DE-SC0019272. Use was made of computational facilities purchased with funds from the National

Science Foundation (OAC-1925717) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. S.H. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG) im Rahmen der Exzellenzstrategie des Bundes und der Länder – EXC 2033 – Projektnummer 390677874 – RESOLVDFG-EXC-2033 project #390677874.

References

1. Weng L, Stott SL, Toner M. 2019. Exploring Dynamics and Structure of Biomolecules, Cryoprotectants, and Water Using Molecular Dynamics Simulations: Implications for Biostabilization and Biopreservation. *Annu. Rev. Biomed. Eng.* 21(1):1–31
2. Towey JJ, Soper AK, Dougan L. 2012. Molecular Insight Into the Hydrogen Bonding and Micro-Segregation of a Cryoprotectant Molecule. *The Journal of Physical Chemistry B*. 116(47):13898–904
3. Towey JJ, Soper AK, Dougan L. 2013. What happens to the structure of water in cryoprotectant solutions? *Faraday Discuss.* 167:159
4. Dashnau JL, Nucci NV, Sharp KA, Vanderkooi JM. 2006. Hydrogen Bonding and the Cryoprotective Properties of Glycerol/Water Mixtures. *J. Phys. Chem. B*. 110(27):13670–77
5. Gekko K, Timasheff SN. 1981. Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry*. 20(16):4667–76
6. Weng L, Chen C, Zuo J, Li W. 2011. Molecular Dynamics Study of Effects of Temperature and Concentration on Hydrogen-Bond Abilities of Ethylene Glycol and Glycerol: Implications for Cryopreservation. *J. Phys. Chem. A*. 115(18):4729–37
7. Toner M, Cravalho EG, Karel M. 1990. Thermodynamics and kinetics of intracellular ice formation during freezing of biological cells. *Journal of Applied Physics*. 67(3):1582–93
8. Karlsson JOM, Cravalho EG, Toner M. 1994. A model of diffusion-limited ice growth inside biological cells during freezing. *Journal of Applied Physics*. 75(9):4442–55
9. Hayashi Y, Puzenko A, Feldman Y. 2006. Slow and fast dynamics in glycerol–water mixtures. *Journal of Non-Crystalline Solids*. 352(42–49):4696–4703
10. Weeraratna C, Amarasinghe C, Lu W, Ahmed M. 2021. A Direct Probe of the Hydrogen Bond Network in Aqueous Glycerol Aerosols. *J. Phys. Chem. Lett.* 12(23):5503–11
11. Charkhesht A, Lou D, Sindle B, Wen C, Cheng S, Vinh NQ. 2019. Insights into Hydration Dynamics and Cooperative Interactions in Glycerol–Water Mixtures by Terahertz Dielectric Spectroscopy. *J. Phys. Chem. B*. 123(41):8791–99
12. Behrends R, Fuchs K, Kaatze U, Hayashi Y, Feldman Y. 2006. Dielectric properties of glycerol/water mixtures at temperatures between 10 and 50°C. *The Journal of Chemical Physics*. 124(14):144512

13. Puzenko A, Hayashi Y, Ryabov YE, Balin I, Feldman Y, et al. 2005. Relaxation Dynamics in Glycerol−Water Mixtures: I. Glycerol-Rich Mixtures. *J. Phys. Chem. B*. 109(12):6031–35
14. Hayashi Y, Puzenko A, Balin I, Ryabov YE, Feldman Y. 2005. Relaxation Dynamics in Glycerol−Water Mixtures. 2. Mesoscopic Feature in Water Rich Mixtures. *J. Phys. Chem. B*. 109(18):9174–77
15. Leavesley A, Wilson CB, Sherwin M, Han S. 2018. Effect of water/glycerol polymorphism on dynamic nuclear polarization. *Phys. Chem. Chem. Phys.* 20(15):9897–9903
16. Zelent B, Nucci NV, Vanderkooi JM. 2004. Liquid and Ice Water and Glycerol/Water Glasses Compared by Infrared Spectroscopy from 295 to 12 K. *J. Phys. Chem. A*. 108(50):11141–50
17. Conti Nibali V, Pezzotti S, Sebastiani F, Galimberti DR, Schwaab G, et al. 2020. Wrapping Up Hydrophobic Hydration: Locality Matters. *J. Phys. Chem. Lett.* 11(12):4809–16
18. Morawietz T, Marsalek O, Pattenaude SR, Streacker LM, Ben-Amotz D, Markland TE. 2018. The Interplay of Structure and Dynamics in the Raman Spectrum of Liquid Water over the Full Frequency and Temperature Range. *J. Phys. Chem. Lett.* 9(4):851–57
19. Pezzotti S, Sebastiani F, Van Dam EP, Ramos S, Conti Nibali V, et al. 2022. Spectroscopic Fingerprints of Cavity Formation and Solute Insertion as a Measure of Hydration Entropic Loss and Enthalpic Gain. *Angew Chem Int Ed*. 61(29):
20. Das Mahanta D, Patra A, Samanta N, Luong TQ, Mukherjee B, Mitra RK. 2016. Non-monotonic dynamics of water in its binary mixture with 1,2-dimethoxy ethane: A combined THz spectroscopic and MD simulation study. *The Journal of Chemical Physics*. 145(16):164501
21. Chakraborty S, Pyne P, Kumar Mitra R, Das Mahanta D. 2022. A subtle interplay between hydrophilic and hydrophobic hydration governs butanol (de)mixing in water. *Chemical Physics Letters*. 807:140080
22. Harrison GM, Remmelgas J, Leal LG. 1998. The dynamics of ultradilute polymer solutions in transient flow: Comparison of dumbbell-based theory and experiment. *Journal of Rheology*. 42(5):1039–58
23. Hua L, Zangi R, Berne BJ. 2009. Hydrophobic Interactions and Dewetting between Plates with Hydrophobic and Hydrophilic Domains. *J. Phys. Chem. C*. 113(13):5244–53
24. Patel AJ, Garde S. 2014. Efficient Method To Characterize the Context-Dependent Hydrophobicity of Proteins. *The Journal of Physical Chemistry B*. 118(6):1564–73
25. Monroe JI, Shell MS. 2018. Computational discovery of chemically patterned surfaces that effect unique hydration water dynamics. *Proceedings of the National Academy of Sciences*. 115(32):8093–98
26. Monroe JI, Jiao S, Davis RJ, Robinson Brown D, Katz LE, Shell MS. 2021. Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution. *Proc Natl Acad Sci USA*. 118(1):e2020205118
27. Chelli R, Procacci P, Cardini G, Della Valle RG, Califano S. 1999. Glycerol condensed phases Part I. A molecular dynamics study. *Phys. Chem. Chem. Phys.* 1(5):871–77
28. Blieck J, Affouard F, Bordat P, Lerbret A, Descamps M. 2005. Molecular dynamics simulations of glycerol glass-forming liquid. *Chemical Physics*. 317(2–3):253–57

29. Izadi S, Anandakrishnan R, Onufriev AV. 2014. Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* 5(21):3863–71

30. Chelli R, Procacci P, Cardini G, Califano S. 1999. Glycerol condensed phases Part II.A molecular dynamics study of the conformational structure and hydrogen bonding. *Physical Chemistry Chemical Physics*. 1(5):879–85

31. Jahn DA, Akinkunmi FO, Giovambattista N. 2014. Effects of Temperature on the Properties of Glycerol: A Computer Simulation Study of Five Different Force Fields. *The Journal of Physical Chemistry B*. 118(38):11284–94

32. Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*. 98(12):10089–92

33. Eastman P, Pande V. 2010. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science & Engineering*. 12(4):34–39

34. Martínez L, Andrade R, Birgin EG, Martínez JM. 2009. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* 30(13):2157–64

35. Roe DR, Cheatham TE. 2013. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* 9(7):3084–95

36. Fox T, Kollman PA. 1998. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B*. 102(41):8070–79

37. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, et al. 2004. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics*. 120(20):9665–78

38. Stachura SS, Malajczuk CJ, Mancera RL. 2018. Molecular dynamics simulations of a DMSO/water mixture using the AMBER force field. *J Mol Model*. 24(7):174

39. Luzar A, Chandler D. 1993. Structure and hydrogen bond dynamics of water–dimethyl sulfoxide mixtures by computer simulations. *The Journal of Chemical Physics*. 98(10):8160–73

40. Das Mahanta D, Brown DR, Pezzotti S, Han S, Schwaab G, et al. 2023. Local solvation structures govern the mixing thermodynamics of glycerol–water solutions. *Chem. Sci.* 14(26):7381–92

41. Ma CY, Pezzotti S, Schwaab G, Gebala M, Herschlag D, Havenith M. 2021. Cation enrichment in the ion atmosphere is promoted by local hydration of DNA. *Phys. Chem. Chem. Phys.* 23(40):23203–13

42. Starciuc T, Guinet Y, Hedoux A, Shalaev E. 2021. Water content thresholds in glycerol/water system: Low- and high-wavenumber Raman spectroscopy study. *Journal of Molecular Liquids*. 321:114678

43. Galamba N. 2013. Water's Structure around Hydrophobic Solutes and the Iceberg Model. *J. Phys. Chem. B*. 117(7):2153–59

44. Wu X, Lu W, Streacker LM, Ashbaugh HS, Ben-Amotz D. 2018. Temperature-Dependent Hydrophobic Crossover Length Scale and Water Tetrahedral Order. *The Journal of Physical Chemistry Letters*. 9(5):1012–17

45. Stock P, Monroe JI, Utzig T, Smith DJ, Shell MS, Valtiner M. 2017. Unraveling Hydrophobic Interactions at the Molecular Scale Using Force Spectroscopy and Molecular Dynamics Simulations. *ACS Nano*. 11(3):2586–97

46. Monroe JI, Shell MS. 2019. Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *The Journal of Chemical Physics*. 151(9):094501
47. Mandumpal JB, Kreck CA, Mancera RL. 2011. A molecular mechanism of solvent cryoprotection in aqueous DMSO solutions. *Phys. Chem. Chem. Phys.* 13(9):3839
48. Cheng C-Y, Song J, Pas J, Meijer LHH, Han S. 2015. DMSO Induces Dehydration near Lipid Membrane Surfaces. *Biophysical Journal*. 109(2):330–39
49. Yu Z-W, Quinn PJ. 1994. Dimethyl sulphoxide: A review of its applications in cell biology. *Bioscience Reports*. 14(6):259–81
50. Clever HL, Pigott SP. 1971. Enthalpies of mixing of dimethylsulfoxide with water and with several ketones at 298.15 K. *The Journal of Chemical Thermodynamics*. 3(2):221–25
51. Jiao S, Rivera Mirabal DM, DeStefano AJ, Segalman RA, Han S, Shell MS. 2022. Sequence Modulates Polypeptoid Hydration Water Structure and Dynamics. *Biomacromolecules*
52. Robinson Brown DC, Webber TR, Jiao S, Rivera Mirabal DM, Han S, Shell MS. 2023. Relationships between Molecular Structural Order Parameters and Equilibrium Water Dynamics in Aqueous Mixtures. *J. Phys. Chem. B*. acs.jpcb.3c00826
53. Monroe J, Barry M, DeStefano A, Aydogan Gokturk P, Jiao S, et al. 2020. Water Structure and Properties at Hydrophilic and Hydrophobic Surfaces. *Annual Review of Chemical and Biomolecular Engineering*. 11(1):523–57
54. Moon H, Collanton RP, Monroe JI, Casey TM, Shell MS, et al. 2022. Evidence for Entropically Controlled Interfacial Hydration in Mesoporous Organosilicas. *J. Am. Chem. Soc.* 144(4):1766–77
55. Duboué-Dijon E, Fogarty AC, Laage D. 2014. Temperature Dependence of Hydrophobic Hydration Dynamics: From Retardation to Acceleration. *J. Phys. Chem. B*. 118(6):1574–83
56. Chen S-H, Gallo P, Sciortino F, Tartaglia P. 1997. Molecular-dynamics study of incoherent quasielastic neutron-scattering spectra of supercooled water. *Physical Review E*. 56(4):4231–43
57. Dubey V, Erimban S, Indra S, Daschakraborty S. 2019. Understanding the Origin of the Breakdown of the Stokes–Einstein Relation in Supercooled Water at Different Temperature–Pressure Conditions. *The Journal of Physical Chemistry B*. 123(47):10089–99
58. Martiniano HFMC, Galamba N. 2013. Insights on Hydrogen-Bond Lifetimes in Liquid and Supercooled Water. *J. Phys. Chem. B*. 117(50):16188–95
59. Debenedetti PG, Stillinger FH. 2001. Supercooled liquids and the glass transition. *Nature*. 410(6825):259–67
60. Dahanayake JN, Mitchell-Koch KR. 2018. Entropy connects water structure and dynamics in protein hydration layer. *Physical Chemistry Chemical Physics*. 20(21):14765–77
61. M R, Ayappa KG. 2020. Dynamical Transitions of Supercooled Water in Graphene Oxide Nanopores: Influence of Surface Hydrophilicity. *The Journal of Physical Chemistry B*. 124(23):4805–20
62. Kundu A, Verma PK, Cho M. 2019. Water Structure and Dynamics in the Stern Layer of Micelles: Femtosecond Mid-Infrared Pump-Probe Spectroscopy Study. *The Journal of Physical Chemistry B*. 123(25):5238–45

63. Das Mahanta D, Rana D, Patra A, Mukherjee B, Mitra RK. 2018. Heterogeneous structure and solvation dynamics of DME/water binary mixtures: A combined spectroscopic and simulation investigation. *Chemical Physics Letters*. 700:50–56
64. Pezzotti S, Sebastiani F, Van Dam EP, Ramos S, Conti Nibali V, et al. 2022. Local thermodynamics of alcohols from THz-calorimetry: Spectroscopic fingerprints of entropic loss and enthalpic gain. Chemistry
65. Marcus Y. 2000. Some thermodynamic and structural aspects of mixtures of glycerol with water. *Phys. Chem. Chem. Phys.* 2(21):4891–96
66. Zehioua R, Coquelet C, Soo C-B, Richon D, Meniai A-H. 2009. Experimental and predicted excess molar enthalpies of some working pairs for absorption cycles. *Thermochimica Acta*. 495(1–2):72–80
67. Batov DV, Zaichikov AM, Slyusar VP, Korolev VP. 2001. Enthalpies of Mixing and State of Components in Aqueous-Organic Mixtures with Nets of Hydrogen Bonds. *Russian Journal of General Chemistry*. 71(8):1208–14

Chapter 5: Relationships Between Molecular Structural Order Parameters and Equilibrium Water Dynamics in Aqueous Mixtures

Reproduced with permission from:

Dennis C. Robinson Brown, Thomas R. Webber, Sally Jiao, Daniela M. Rivera Mirabal, Songi Han, M. Scott Shell. *Relationships Between Molecular Structural Order Parameters and Equilibrium Water Dynamics in Aqueous Mixtures. JPC B*. **2023**. DOI: 10.1021/acs.jpcb.3c00826. Copyright 2023 American Chemical Society

5.1 Introduction

The prediction of water transport and thermodynamic properties in solution is vital to an enormous range of applications spanning conventional chemical synthesis, water purification membranes, protein hydration and biological fouling, and nanotherapeutics, to name a few. However, directly estimating macroscopic and underlying molecular scale thermophysical properties can be challenging in complex mixtures and at chemically heterogeneous surfaces (e.g., proteins). Often the role of water's unique structure proves particularly important to making accurate predictions, in contrast to what is found for simpler liquids. For example, conventional approaches to capture diffusivity often employ continuum-based mechanistic models in dilute aqueous mixtures(1), water-polymer mixtures(2–4), and water purification membranes(5). However, continuum models cannot fully explicate water diffusivity in a wide range of solution environments(6–8); instead, shifts in molecular structure are a more precise signature of water mobility(6, 9–13).

Along these lines, decades of simulation efforts are now converging to a more unified view on the connections between water structure and its thermophysical properties(14–18). In their seminal work, Errington and Debenedetti(9) showed remarkably how water structure, captured by structural order parameters, is hierarchically connected to thermodynamic and dynamic property anomalies relative to simple liquids. More broadly, a wide range of studies have suggested manners by which water's molecular structure is fundamental to understanding water properties. Early efforts primarily probed water at interfaces with idealized solutes(15, 19, 20) and surfaces(12, 21, 22) that neglect the effects of topological and chemical heterogeneity. In one vital example, Lum, Chandler, and Weeks represented hydrophobic solutes/surfaces as hard spheres/walls(15). These simplified models provided important insight on distinct trends in water structure near small (<1 nm) and large (>1 nm) interfaces that persist in systems with greater atomistic detail(19, 23–25); namely, they supported the idea that water responds differently to small and large solutes, with unique thermodynamic driving forces in these length scale regimes. More recently, major efforts have now characterized the role of water density and interface fluctuations in governing water properties(26–30). Initial studies established these fluctuations as a key indicator of hydropathy at extended, flat surfaces. Remarkably, the works of Patel and coworkers(28, 31–33) suggest that density fluctuations can even be leveraged to map hydropathy at topologically and chemically complex protein-water interfaces(31).

Further, detailed structural analyses have become important tools for the rational design of materials(11, 34–36). In one recent example, Dallin et. al.(36) sought to leverage computationally-inexpensive water metrics to predict the hydration free energies at a variety of chemically and spatially heterogeneous self-assembled monolayer (SAM) surfaces. They found that characterizing only five structural metrics–pertaining to water's orientational order, density, and hydrogen bonding—the hydration free energies could be quite accurately predicted (< 4 $k_B T$). Such workflows may promise to circumvent expensive (or intractable) direct free energy calculations for large proteins and protein complexes.

While there has been substantial progress through simulations in identifying molecular order parameters that capture the relevance of water structure to emergent thermophysical properties, it has remained challenging to access experimental signatures of molecular-level water behavior which do the same. Water structure can be characterized by techniques like Terahertz(16, 37–39), Raman(40–43), IR(44–48), and vibrational sum-frequency generation spectroscopies(49–54). However, such efforts involve measurements that average over waters potentially experiencing a range of different molecular environments, blurring the structure-property relationship. A particularly promising experimental tool for probing *local* water behavior is Overhauser Dynamic Nuclear Polarization (ODNP)(55, 56). ODNP is an electron paramagnetic resonance (EPR)-enhanced NMR technique that directly measures the dynamic coupling between a radical spin probe molecule and the surrounding water protons within an approximate radius of one nanometer. The resultant NMR signal enhancements are used to directly measure local water translational self-diffusivity (< 1 nm from the spin probe) on the scale of tens of picoseconds to nanoseconds. ODNP enables the measurement of equilibrium water dynamics at room temperature in complex environments. Previous works have demonstrated that ODNP resolves local water diffusivities at chemically and topologically heterogeneous aqueous interfaces(57–61). These advances are certainly exciting, but it remains to be shown how such a locally measured diffusivity reflects underlying water structural behavior.

Statistical learning algorithms in conjunction with molecular simulations provide an opportunity to discover such relationships, and more generally to understand how water structure relates to functional thermodynamic and dynamic properties. Given a collection of simulations of water in varied environments, statistical learning approaches can identify important predictors (e.g., water structural order parameters) of quantities of interest (e.g., thermodynamic and dynamic property responses). In a recent example from Rego and coworkers, statistical learning algorithms—namely, artificial and convolutional neural networks—was leveraged to decipher the complex relationship between chemical patterning motifs at surfaces and surface hydropathy(35).

To illustrate the potential of these connections between water order parameters and properties, we briefly review results from our prior work(10, 62) suggesting a remarkably strong relationship between water's molecular structure—specifically tetrahedrality—and dynamics. **Figure 5.1** summarizes relationships found for several binary aqueous mixtures: (i) glycerol, (ii) poly(ethylene oxide) (PEO), and (iii) peptoids. Here, we characterize water's tetrahedrality by first enumerating the triplet angles, θ, between water oxygens and computing three-body angle distributions, $P_{3b}(\theta)$. We estimate the relative population of tetrahedral angles between triplets of water molecules, $p_{tet}/p_{tet,neat}$, by integrating over the tetrahedral region of $P_{3b}(\theta)$ $(100° < \theta < 120°)$ and normalizing p_{tet} by the tetrahedral population of neat water at 298K, $p_{tet,neat}$. We probe translational water dynamics via the logarithmic water self-diffusivity, $\ln\left[\frac{D_{H_2O}}{D_{H_2O,neat}}\right]$, relative to neat water at 298K. In all three systems, we observed a striking exponential $D_{H_2O} - p_{tet}$ relationship. Though the structure-dynamics correlation remains qualitatively similar between systems, the $\ln D_{H_2O} - p_{tet}$ slope is context dependent (i.e., specific to the particular solute or co-solvent). The distinction between PEO-water and glycerol-water likely stems from glycerol's capacity to both donate and accept hydrogen bonds, while PEO only accepts them. The more dramatic ln $D_{H_2O} - p_{tet}$ slope for the peptoid-water system—compared to PEO-water—is likely at least partially related to the fact that both D_{H_2O} and p_{tet} were computed for waters within the hydration layers of the peptoids as detailed by Jiao et. al(10) (versus throughout the entire mixture).

The strength of these correlations between tetrahedrality and diffusivity, spanning three chemically distinct systems, is particularly exciting as it suggests both the ability to predict diffusivity from water structure, but also the potential for ODNP-measurable local diffusivities to indirectly report on water structure. However, further investigation in a broader range of systems, and considering other measures of structural information, is clearly needed to more broadly understand the relationship between water structure and thermophysical properties like self-diffusivity.

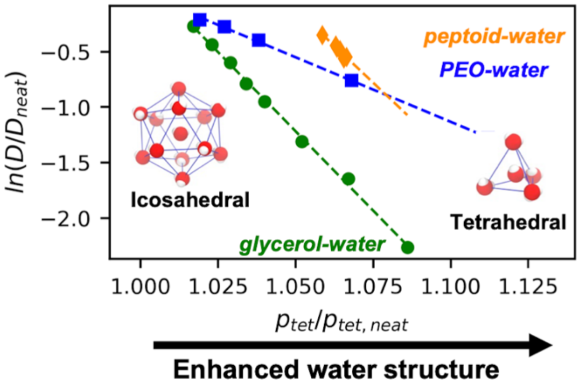


Figure 5.1. Water self-diffusivity, D_{H_2O}, strongly correlates with water's tetrahedrality, $p_{tet}/p_{tet,neat}$, for glycerol-water (green), PEO-water (blue), and peptoid-water (yellow) mixtures. Remarkably, an exponential relationship between dynamics and molecular structure persists across changes in mixture composition and the chemical identity of the additive. Water self-diffusivity, D_{H_2O}, is estimated via the slope of the mean-square displacement (MSD) curve. The population of tetrahedral water triplet angles, p_{tet}, is derived by integrating the triplet angle distribution, $P_{3b}(\theta)$, over the tetrahedral region ($100° < \theta < 120°$). The magnitude of the $D_{H_2O} - p_{tet}$ slope corresponds to the dynamic response to shifts in tetrahedrality. We hypothesize that glycerol-water shows the largest response owing to the ability of glycerol to accept and donate hydrogen bonds, while the peptoids and PEO only accept hydrogen bonds.

In the present work, we study a wide range of simulated aqueous mixtures and demonstrate with statistical learning an approach to learn structure-property connections. In **Section 5.2.3**, we highlight several water structural metrics and detail a strategy for drawing connections between these and macroscopic properties. **Section 5.3.1** describes the form and physical behavior of the structural metrics in the context of methanol-water at temperatures ranging from 273 to 373K. In **Section 5.3.2**, we implement principal component analysis (PCA) to show how subtle shifts in triplet angle distributions across a wide range of systems and conditions are captured by a small number of response degrees of freedom. **Section 5.3.3** then examines and interprets structure-dynamics and structure-thermodynamics correlations. In **Section 5.3.4**, we apply a feature selection procedure to identify relevant order parameters that accurately predict water self-diffusivity and excess hard sphere chemical potentials across numerous co-solvents and system conditions. Leveraging the particularly strong relationship between structure and dynamics, we propose a workflow for predicting the structural metrics and mechanisms underlying experimental probes of water properties in **Section 5.3.5**.

5.2 Methods

5.2.1. Molecular Dynamics Simulations

We consider the following binary aqueous mixtures: (i) glycerol-water at 291K at glycerol mole fractions of 0.01, 0.033, 0.05, 0.075, 0.10, 0.15, 0.20, and 0.30; (ii) PEO-water at 291K at PEO weight fractions of 0.005, 0.015, 0.05, 0.10, 0.20, 0.33, and 0.50; (iii) polypeptoid oligomers with variable hydrophobicity at 300K (see **Figure 5.2** for details); (iv) primary alcohol-water (methanol, ethanol, 1-propanol, and 1-butanol) at a weight fraction of 0.1 and temperatures of 273, 280, 291, 298, 309, 329, 350, and 373K; (v) DMSO-water mixtures at 291K of mole fractions 0.01, 0.04, 0.06, 0.10, 0.16, 0.27, and 0.45. These specific binary mixtures are chosen to span a wide variety of mixture conditions. In systems described above, we vary cosolvent chemistry (e.g., very hydrophilic glycerol vs. weakly hydrophilic butanol), mixture state (variations in temperature and composition), and cosolvent size (e.g., PEO vs. methanol).

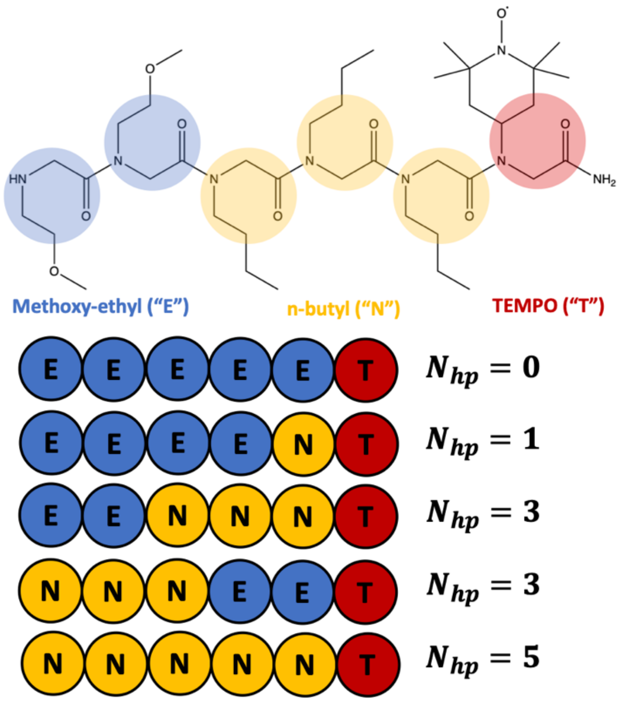


Figure 5.2. Schematic representation of polypeptoid oligomer chemistries from Jiao et. al.(10) where yellow and blue circles correspond to hydrophobic (n-butyl) and hydrophilic (methoxy-ethyl) moieties, respectively. The nominal oligomer hydrophobicity is modified by changing the number of hydrophobic monomers and their position.

The model parameters for the glycerol-water(62), PEO-water(63), and peptoid-water(10, 60) systems are described in previous works. Further, the simulations for the peptoid-water systems were are drawn from our past work(10, 60). We use DMSO force field parameters from Fox and Kollman(64) and the TIP4P-EW(65) 4-site water model. Finally, we consider methanol-water, ethanol-water, propanol-water, and butanol-water mixtures at alcohol weight fractions of 0.10 and at temperatures of 273, 280, 291, 298, 309, 329, 350, and 373K. As with the DMSO-water mixtures, we use the TIP4P-EW water model for the alcohol-water simulations. All non-electrostatic parameters for the alcohols are taken from the second-generation general AMBER force field (GAFF2)(66, 67). We scale partial charges using the scheme suggested by Fennell et. al.(68)

To prepare the glycerol-water, PEO-water, alcohol-water, and DMSO-water systems for simulation, we first generate initial atomic positions using PACKMOL(69) and energy minimize configurations using the OpenMM(70) molecular dynamics software. We equilibrate each system in the NPT ensemble via a Langevin thermostat(70) and Monte Carlo barostat(70) to maintain atmospheric pressure (1 bar). Following the equilibration period (between 50 and 200 ns dependent on the system), we conduct a production run in NPT for 100 ns, saving the atomic coordinates every 10 ps. To compute diffusivities, we draw 10 independent configurations from the NPT production run to generate starting configuration for 1 ns NVE production runs with coordinates saved every picosecond.

5.2.2 Water self-diffusivity

We compute water self-diffusivity by calculating the mean-squared displacement (MSD) of water oxygens as a function of time t

$$MSD(t) = \langle |r(t+\tau) - r(\tau)|^2 \rangle \qquad (5.1)$$

where $r(t)$ is the 3D position vector of a water oxygen and the brackets $\langle\rangle$ correspond to the ensemble average over several time windows and all water oxygens. We determine the self-diffusivity of water from the slope of the diffusive (linear) regime ($t > 5ps$) of the MSD curve. For the present study, we focus specifically on D_{H_2O} due to its ubiquity as a dynamic metric in research on aqueous solutions. Several studies have investigated the structure-dynamics correspondence using other dynamic metrics for translational(10, 11, 58), rotational(13, 71, 72), hydrogen bonding(12, 73), etc.; these are certainly of interest, but because translation and other water dynamics are frequently strongly correlated(72, 74–76), we focus on translational diffusivity here.

5.2.3 Water structure order parameters

While it's clear from **Figure 5.1** that p_{tet} displays remarkable correlation with diffusivity across several aqueous mixtures, the slope of these relationships varies widely depending on the cosolvent. Pursuant of more robust prediction of thermophysical properties from molecular structure, we consider a range of additional structural order parameters. Individual order parameters fall into one of two categories: orientational and translational. Translational order parameters measure the degree of spatial correlation between water molecules (e.g., local number densities). On the other hand, orientational order parameters probe angular correlations between water molecules (e.g., the population of tetrahedrally coordinated waters, p_{tet}). A full list of relevant order parameters for the present study are summarized in **Table 5.1**.

Tetrahedral order parameter. The three-body angle distribution and the related p_{tet} are not the only probes of tetrahedrality in aqueous environments. In particular, the tetrahedral order parameter, q, first investigated in detail by Errington and Debenedetti(9) has been widely studied as a metric of water structure since its introduction(8, 25, 77, 78). q is given by

$$q = 1 - \frac{3}{8}\sum_{j=1}^{3}\sum_{k=j+1}^{4}\left(\cos\psi_{jk} + \frac{1}{3}\right)^2 \qquad (5.2)$$

where ψ_{jk} is a three-body angle formed between a central water oxygen atom and two of its four nearest neighbors j and k. Here, we sum over all six three-body pairs between the central water and its four nearest neighbors. The ensemble average of q approaches 0 in fluids without interparticle correlations (e.g., an ideal gas) and is 1 for a perfectly tetrahedral crystal (e.g., ice at 0 K). For most dilute aqueous mixtures, q effectively quantifies the extent of a liquid's tetrahedrality as most waters retain four nearest water neighbors; however, its interpretation is more complex in concentrated aqueous mixtures, water under confinement, and in particular for water at surfaces because it requires locating four nearest water neighbors, which may not reside in the first coordination shell(79). When one or more of the nearest neighbors are outside of the central water's hydration shell, the more distant neighbors adopt more weakly-correlated orientations relative to the central water, hence decreasing q. While some prior efforts have proposed modified versions of q to account for undercoordination of water hydration shells(77, 80), we do not consider these alternative approaches here.

Specific volume and area. Favorable interactions between water and cosolvents like ethanol are known to produce negative volumes of mixing. To more precisely report atomic volumes at an individual molecule scale, we compute molecular volumes using the Voronoi approach, specifically through the quickhull SciPy implementation(81, 82). With this approach, we extract the approximate volume, v_{wat}, and surface area, a_{wat}, of every water molecule in each simulation snapshot for averaging. We also probe the asphericity parameter $\eta = \frac{a_{wat}^3}{36\pi v_{wat}^2}$ of water molecules. A perfect spherical Voronoi volume has $\eta = 1$ (e.g., ideal gas) while a polyhedral volume has $\eta > 1$, noting that a perfect tetrahedron has $\eta \approx 3.3$. Voronoi polyhedral calculations have been frequently used in a wide range of contexts from complex aqueous solutions(83) to supercooled fluids(84) to protein hydration layers(85).

Radial distribution functions. Local density correlations probed by 2D radial distribution functions (RDFs) report on shifts in the coordination environment near small molecules in aqueous mixtures. We quantify these shifts by integrating over the first hydration shell—as defined by the first minimum in an RDF—to obtain a coordination number. In the present work, we consider water-water ($n_{1,Ow-Ow}$) and solute-water ($n_{1,Sol-Ow}$) first hydration shell coordination numbers.

Via the water-water radial distribution function $g(r_{Ow-Ow})$, we also characterize the translational order of water by calculating the Errington-Debenedetti order parameter(9)

$$t = \frac{1}{\xi_c}\int_0^{\xi_c}|g(\xi) - 1|d\xi \qquad (5.3)$$

where $\xi = r_{Ow-Ow}\rho_{Ow-Ow}^{1/3}$. Here, r_{Ow-Ow} and ρ_{Ow-Ow} are the radial distance between water molecules and the bulk number density of water molecules, respectively. Further, $\xi_c = r_c\rho_{Ow-Ow}^{1/3}$ is the normalized cutoff distance defined by the first minimum of $g(r_{Ow-Ow})$. In the ideal gas limit, $g(r_{Ow-Ow}) \rightarrow 1$ and hence $t \rightarrow 0$. As short-range translational order develops (e.g., water upon cooling), t increases. t further increases in crystalline media with long-range translational order while the magnitude depends on the lattice structure.

Hard-sphere chemical potential and density fluctuations. While explicit calculation of hydrophobic hydration typically requires detailed and expensive free energy calculations(26, 28, 30, 31, 36), here we choose a simpler surrogate probe and estimate the excess chemical potential, μ_{HS}^{ex}, of solvating a hard sphere with radius 3.3-Å (ideal small hydrophobic molecule). We calculate μ_{HS}^{ex} via Widom insertion(86) in which hard sphere insertions are attempted at 10,000 random positions for 100 independent simulation snapshots and computing the negative log-probability of overlapping with zero atomic coordinates, $\mu_{HS}^{ex} = -k_BTlnP(N = 0)$. Naturally, lower probability of hard sphere insertion, $P(N = 0)$, increases μ_{HS}^{ex} and corresponds to a more hydrophilic local environment. μ_{HS}^{ex} has long been applied as a surrogate measurement hydropathy in aqueous environments from extended molecular surfaces(11, 23, 26, 27, 87) to protein hydration layers(27, 74, 87). In addition to the hard sphere chemical potential, we calculate the mean, $\langle N\rangle$, second moment, $\langle N^2\rangle$, and variance, $\langle \delta N^2\rangle = \langle (N - \langle N\rangle)^2\rangle$, of the hard sphere overlap distribution. These distribution statistics serve as additional order parameters that report on the local number density and corresponding density fluctuations.

Local structure index (LSI). The LSI is yet another method of quantifying the translational order of water. In contrast to t, the LSI measures fluctuations in the distance between the first and second hydration shells of water molecules(13, 88). The LSI for a central water molecule is defined as

$$LSI = \frac{1}{n}\sum_{i=1}^{n}(\Delta_i - \overline{\Delta})^2 \qquad (5.4)$$

where n is the index corresponding to the furthest nearest neighbor with radial distance $r_n < 3.7$Å, $\Delta_i = r_{i+1} - r_i$, and $\overline{\Delta} = \frac{1}{n}\sum_{i=1}^{n}\Delta_i$. Here, nearest neighbor radii are ordered by distance from the central atom $r_1 < r_2 < \cdots < r_n < 3.7$ Å. A molecule with randomly distributed environment (ideal gas) will have $LSI \approx 0$ Å^2. A tetrahedral solid like ice has $\langle LSI \rangle \approx 0.43$ Å^2. In aqueous solutions, $\langle LSI \rangle$ is intermediate between these two extremes with higher values corresponding to longer-range translational order. $\langle LSI \rangle$ has the potential to describe shifts in translational order due to the chemical identity of a given cosolvent. Some previous studies have even applied the LSI to quantify and characterize shifts in translational structure in protein hydration layers(89, 90).

Three-body angle distribution. As mentioned earlier, we construct distributions of triplet angles between each water molecule oxygen and its two nearest neighboring water oxygens. These three-body angle distributions, $P_{3b}(\theta)$, contain detailed information on water's fluctuating network structure. We specifically examine regions corresponding to tetrahedrally- ($\theta = 109.5°$) and icosahedrally structured ($\theta = 64°$) water populations. We then enumerate the population of tetrahedrally-coordinated waters, p_{tet}, by integrating over a region of $P_{3b}(\theta)$ approximately centered on the characteristic tetrahedral angle $\theta_{tet} = 109.5°$, from $100° < \theta < 120°$.

Beyond estimating the fraction of tetrahedrally-coordinated waters, we also extract several statistics from $P_{3b}(\theta)$: the mean $\cos\theta$, $\langle\cos\theta\rangle$, the variance of $\cos\theta$, $\langle\delta\cos^2\theta\rangle = \langle(cos\theta - \langle cos\theta\rangle)^2\rangle$, and the distribution entropy, $S_{3b} = -\sum_{i=1}^{N_{3b}} P_{3b}(\theta_i) ln P_{3b}(\theta_i)$. These additional metrics describe trends in $P_{3b}(\theta)$ beyond the tetrahedral region. Here, $\langle\cos\theta\rangle$ describes shifts away from $\cos(\theta \approx 109.5°) \approx 1/3$ observed in neat water at room temperature. On the other hand, increases in metrics like S_{3b} and $\langle\delta^2\cos\theta\rangle$ indicate higher probability of “non-tetrahedral” water configurations.

Table 5.1: A summary of the 20 structural order parameters considered in the feature selection procedure for predicting D_{H_2O} and μ_{HS}^{ex}

Relevant Distribution	**Order Parameter(s)**	**Formula**	**Notes**
water-water RDF: $g(r_{Ow-Ow})$	water-water coordination number: $n_{1,Ow-Ow}$	$n_{1,Ow-Ow} = \int_0^{r_c} 4\pi\rho_{wat}r_{Ow-Ow}^2 g(r_{Ow-Ow})\, dr_{Ow-Ow}$	• The bulk number density of water, ρ_{wat}, is system dependent. • The cutoff radius, r_c, is the first minimum of $g(r_{Ow-Ow})$.
	translational order parameter: t	See **Equation 3** for definition of t	
solute-water RDF $g(r_{Sol-Ow})$	solute-water coordination number: $n_{1,Sol-Ow}$	$n_{1,Sol-Ow} = \int_0^{r_c} 4\pi\rho_{wat}r_{Sol-Ow}^2 g(r_{Sol-Ow})\, dr_{Sol-Ow}$	The cutoff radius, r_c, is the first minimum of $g(r_{Sol-Ow})$.
three-body angle distribution: $P_{3b}(\theta)$	population of tetrahedrally-coordinated waters: p_{tet}	$p_{tet} = \int_{100°}^{120°} P_{3b}(\theta)d\theta$	N_{3b} is the total number of water triplet angles
	mean of $\cos\theta$: $\langle\cos\theta\rangle$	expected value of $\cos\theta$ from $P_{3b}(\theta)$	
	variance of $\cos\theta$: $\langle\delta\cos^2\theta\rangle$	$\langle\delta\cos^2\theta\rangle = \langle(\cos\theta - \langle\cos\theta\rangle)^2\rangle$	
	entropy of $P_{3b}(\theta)$: S_{3b}	$S_{3b} = -\sum_{i=1}^{N_{3b}} P_{3b}(\theta_i) ln P_{3b}(\theta_i)$	
tetrahedral order parameter distribution: $P(q)$	mean of q: $\langle q\rangle$	expected value of q from $P(q)$	See **Equation 5.2** for definition of q
	variance of q: $\langle\delta q^2\rangle$	$\langle\delta q^2\rangle = \langle(q - \langle q\rangle)^2\rangle$	
local structure index distribution: $P(LSI)$	mean of LSI: $\langle LSI\rangle$[Å^2]	expected value of LSI from $P(LSI)$	See **Equation 5.4** for definition of LSI
	variance of LSI: $\langle\delta(LSI)^2\rangle$ [Å^4]	$\langle\delta(LSI)^2\rangle = \langle(LSI - \langle LSI\rangle)^2\rangle$	
water specific volume distribution: $P(v_{wat})$	mean of v_{wat}: $\langle v_{wat}\rangle$ [Å^3]	expected value of v_{wat} from $P(v_{wat})$	See **Section 5.2.3** for a description of these calculations
	variance of v_{wat}: $\langle\delta v_{wat}^2\rangle$ [Å^6]	$\langle\delta v_{wat}^2\rangle = \langle(v_{wat} - \langle v_{wat}\rangle)^2\rangle$	
water specific area distribution: $P(a_{wat})$	mean of a_{wat}: $\langle a_{wat}\rangle$ [Å^2]	expected value of a_{wat} from $P(a_{wat})$	
	variance of a_{wat}: $\langle\delta a_{wat}^2\rangle$ [Å^4]	$\langle\delta a_{wat}^2\rangle = \langle(a_{wat} - \langle a_{wat}\rangle)^2\rangle$	
water asphericity distribution: $P(\eta)$	mean of η: $\langle\eta\rangle$	expected value of η from $P(\eta)$	
	variance of η: $\langle\delta\eta^2\rangle$	$\langle\delta\eta^2\rangle = \langle(\eta - \langle\eta\rangle)^2\rangle$	
Hard-sphere insertion distribution: $P(N)$	mean of N: $\langle N\rangle$	expected value of N from $P(N)$	See **Section 5.2.3** for a description of these calculations
	Squared mean of N: $\langle N^2\rangle$	expected value of N^2 from $P(N)$	
	variance of N: $\langle\delta N^2\rangle$	$\langle\delta N^2\rangle = \langle(N - \langle N\rangle)^2\rangle$	

5.2.4 Feature Selection

The evident persistent structure-dynamics relationship shown in **Figure 5.1** suggests that structure may be predictive of mixture thermophysical properties like D_{H_2O}. However, determining the relative importance of the numerous structural metrics described above for predicting D_{H_2O} (or other properties) is intractable by hand. For high-dimensionality datasets, mitigating the number of independent variables (features) in a statistical learning model is necessary to reduce model variance and improve extensibility. In the present study, feature selection methods aid in developing a molecular-physical intuition of the structural features underpinning shifts in water dynamics (D_{H_2O}) and solvation thermodynamics (μ_{HS}^{ex}). We primarily consider the sequential feature selection (SFS) procedure implemented via the python library MLxtend(91). Generally, SFS procedures search for the descriptive features in a "greedy" fashion—iteratively selecting the N features that best describe a dependent variable—here, diffusivity or chemical potential—for the training dataset. More specifically, we apply a forward SFS (fSFS) procedure.

We first shuffle and split the 59 data from various mixture compositions, cosolvent identity, and temperature into training and test sets composed of random4ly chosen subsets of 80 and 20% of the data, respectively. We then feed the training set into an N-feature fSFS algorithm with the following steps: (i) iterate through each structural metric, applying linear ordinary least squares (OLS) to each, (ii) select the feature with the highest R^2 for predicting the objective dependent variable, (iii) iterate through the remaining features, searching for the feature that most improves the R^2 for the dependent variable, and (iv) repeat the process (N-

2) additional times. In **Figure 5.1**, we demonstrate the linear relationship between $\frac{p_{tet}}{p_{tet,neat}}$ and $\ln[\frac{D_{H_2O}}{D_{H_2O,neat}}]$. This suggests that an OLS model simply and reasonably captures structure-dynamics relationships in binary aqueous mixtures. However, other regression methods such as symbolic regression may yield more robust structure-property relationships.

5.3 Results and Discussion

5.3.1 Methanol-water analysis

To illustrate the behavior of the order parameters discussed above and their relationship to translational water dynamics, we first consider the methanol-water series of simulations with variable temperature from 273 to 373 K at a fixed composition of 0.10 weight fraction of methanol. Similar analyses for aqueous binary mixtures with ethanol, propanol, butanol, glycerol, PEO, DMSO, and peptoid oligomers are detailed in **Section 5.5.1**.

Three-body angle distribution. We depict $P_{3b}(\theta)$ [**Figure 5.3(a)**] and the differential three-body angle distribution, $\Delta P_{3b}(\theta) = P_{3b}(\theta) - P_{3b,neat}(\theta)$, [**Figure 5.3(b)**] for methanol-water. Upon cooling methanol-water mixtures, we observe a systematic enhancement of the tetrahedral region ($\theta = 109.5°$) and corresponding reduction of the icosahedral ($\theta = 64°$) regions of $P_{3b}(\theta)$ and $\Delta P_{3b}(\theta)$. Further, we depict the degree of correlation between the various distribution statistics and water self-diffusivity. Notably, p_{tet} exhibits the greatest correlation with diffusivity ($R^2 \approx 1$), suggesting that this metric is a particularly predictive probe of equilibrium water dynamics in this system. As is discussed in further detail in **Section 5.3.3**, the $p_{tet} - D_{H_2O}$ relationship persists upon including simulation results from the other systems discussed in **Section 5.2.1**.

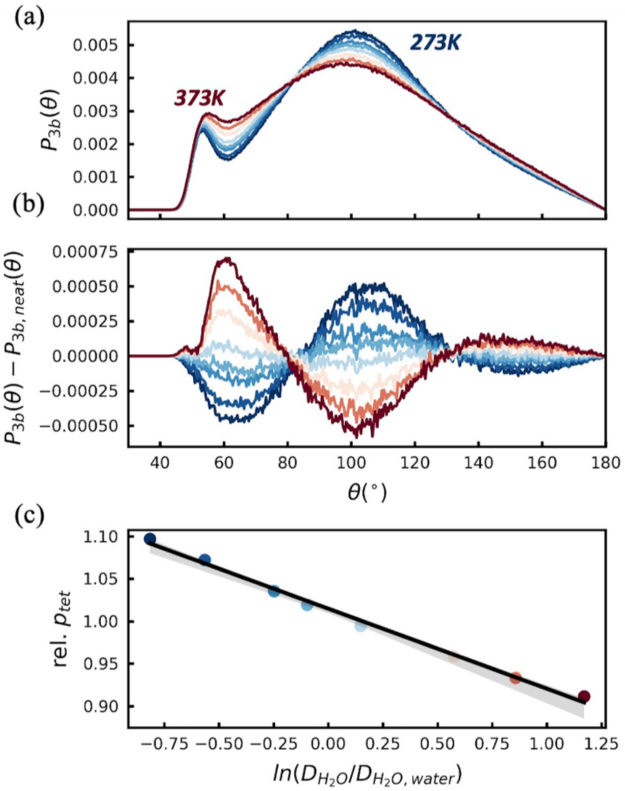


Figure 5.3. For methanol-water mixtures with temperatures increasing from 273 to 373 K, the (a) raw three-body angle distributions, $P_{3b}(\theta)$, and (b) differential three-body angle distributions, $P_{3b}(\theta) - P_{3b,neat}(\theta)$ exhibit enhancement and attenuation of the icosahedral ($\theta \approx 64^\circ$) and tetrahedral ($\theta \approx 109.5^\circ$) regions with temperature, respectively. (c) The population of tetrahedrally-coordinated waters, p_{tet}, strongly correlates with the log-diffusivity of water. A linear fit (black) and the corresponding 95% confidence interval (gray bands) are shown for the p_{tet}-D_{H_2O} relationship. Here, the p_{tet} (c) is presented relative to the value of the same order parameter in TIP4P-EW water at 298 K.

Other order parameters. To further our understanding of the connection between water structure and mixture thermophysical properties, we consider distributions of additional structural metrics (**Section 5.2.3**). **Figure 5.4(a)** depicts the distribution of the tetrahedral order parameter, $P(q)$, in methanol-water. For all considered temperatures, $P(q)$ appears bimodal, with peaks centered at $q \approx 0.45$ and $q \approx 0.8$. The population at $q \approx 0.45$ correspond to less tetrahedrally-coordinated, simple-fluid-like, waters while the population at $q \approx 0.8$ corresponds to tetrahedrally-coordinated water molecules. The bimodality of $P(q)$ is consistent with the original analysis of Errington and Debenedetti(9) and later works(78, 92). As we decrease the system temperature, $P(q)$ shifts from $q = 0.45$ to 0.8. Hence, the behavior of $P(q)$ is consistent with our observations of $P_{3b}(\theta)$ for methanol-water (**Figure 5.3**). Further, we note that $P(q)$ systematically broadens with increasing temperature. This suggests that, while tetrahedrality drops upon heating, water continues to display a substantial tetrahedral population.

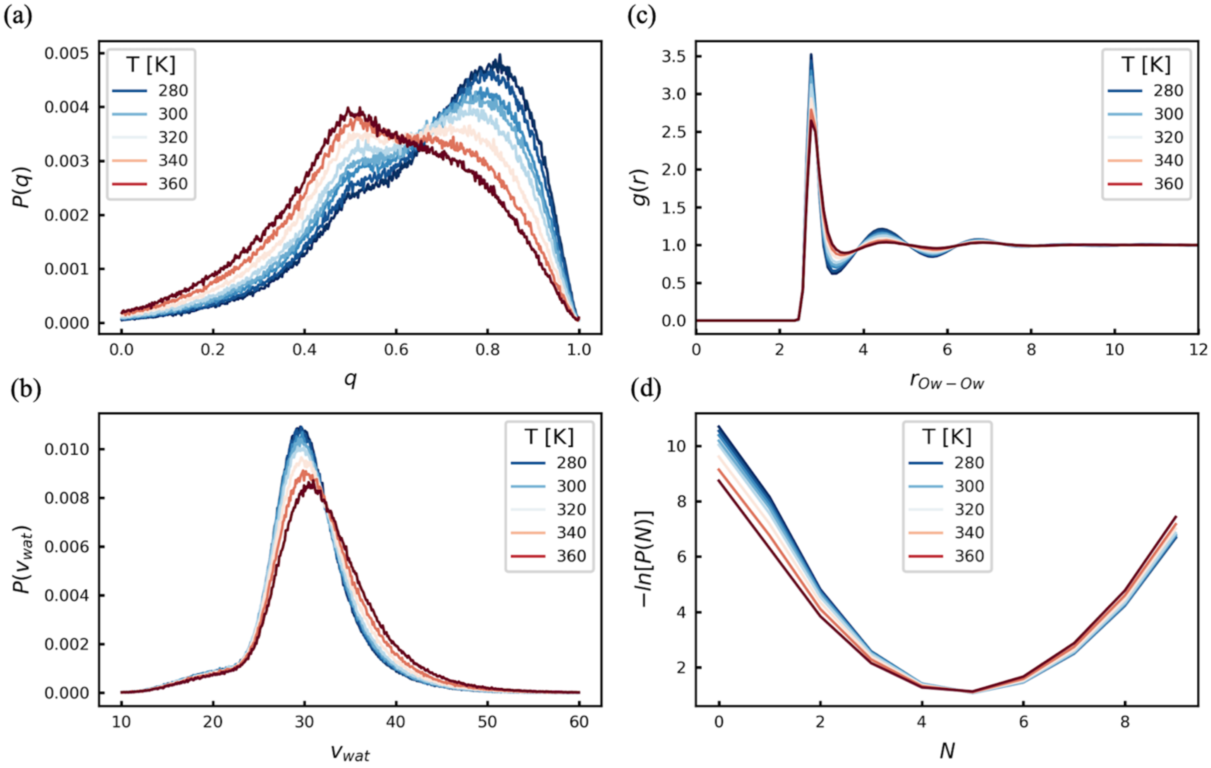


Figure 5.4. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2.3** for methanol-water at temperature of 273, 280, 291, 298, 309, 329, 350, and 373 K. (a) The distributions of the tetrahedral order parameter, $P(q)$, exhibit peaks at $q \approx 0.45$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.8$ upon cooling. (b) The distributions of water specific volumes, $P(v_{wat})$, broaden and exhibit larger mean specific volumes, $\langle v_{wat} \rangle$, as the temperature of the mixture increases. (c) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with decreasing temperature as demonstrated by the increasing peak and trough amplitudes. (d) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the systematic decrease in the affinity of an ideal hydrophobic molecule for methanol-water upon cooling. Simultaneously, decreased mixture temperature shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

We illustrate the effect of increasing temperature in methanol-water mixtures on the distribution of v_{wat} in **Figure 5.4(b)**. We note the systematic shift to lower average v_{wat} and a narrower distribution upon cooling methanol-water. The decrease in the distribution mean, $\langle v_{wat} \rangle$, with temperature is expected given the experimentally observed decrease in methanol-water density (and hence higher specific volume) upon heating(93). This decrease in molecular size is also reflected in the shift to smaller molecular surface area a_{wat} [**Figure 5.11**]. Additionally, we observe a systematic decrease in the mean of the $P(\eta)$ with increased temperature, suggesting that water molecular volumes adopt a more spherical shape as expected in ideal mixtures.

In **Figure 5.4(c)**, we show the $g(r_{Ow-Ow})$ peak amplitudes increase systematically with decreasing temperature, suggesting increasingly long-range translational order. We quantify this trend, observing systematic increases in $n_{1,Ow-Ow}$ and t upon cooling. Thus, we observe simultaneous enhancement of translational (e.g., $n_{1,Ow-Ow}$, and t) and orientational (e.g., $\langle q \rangle$ and p_{tet}). This strong coupling between translational and orientational order is consistent with the conclusions of the establishing work from Errington and Debenedetti(9) and more recent findings for liquid water(92) and other water-like fluids(78).

To visualize shifts in $P(N = 0)$ with varying temperature, we depict $-\ln[P(N)]$ in **Figure 5.4(d)**. We observe a general leftward shift (decreased $\langle N \rangle$) and broadening (increased $\langle \delta^2 N \rangle$) with increasing temperature. $P(N = 0)$ monotonically increases and hence μ_{HS}^{ex} decreases with increasing temperature. The increased solubility of an ideal hydrophobe with increasing temperature is anticipated with the decreasing density and lower cosolvent-water affinity (higher v_{wat}).

Finally, though we omit the LSI distribution from **Figure 5.4**, we observe consistent trends with the five distributions depicted above [**Figures 5.3(a) and 5.4**]. Specifically, as we decrease the temperature of methanol-water, the spacing between water molecules decrease (higher LSI); hence, $P(LSI)$ broadens and shifts towards higher $\langle LSI \rangle$ (more translationally structured water). This finding is consistent with the increased water-water coordination (increasing $n_{1,Ow-Ow}$) induced by decreased temperature. The distributions not directly depicted in **Figures 5.3 and 5.4** are shown in **Figure 5.11**.

At a single methanol-water concentration with varying temperature, we observe consistency between the above structural metrics and their physical interpretations. For example, the metrics derived from $P_{3b}(\theta)$ and $P(q)$ both suggest increasing tetrahedrality with decreasing temperature. For other cosolvents, as we vary the water content of a mixture, the interpretation of the structural metrics requires a subtler approach. For instance, p_{tet} monotonically increases while $\langle q \rangle$ monotonically decreases with increasing glycerol content. We discuss these subtleties in greater detail in **Section 3.3.**

5.3.2 Principal component analysis of the three-body angle distribution

In **Figure 5.3**, we show that the three-body angle distribution varies in predictable ways, along remarkably few degrees of freedom. Further, $\Delta P_{3b}(\theta)$ contains signatures of length scale-dependent hydration effects as noted in a previous computational study(79). In that work, Monroe and Shell applied principal component analysis (PCA) on $\Delta P_{3b}(\theta)$ to explicate the major variations in the three body angle distribution for hydration waters near colloids of varied size and attraction(79). This study demonstrated that the first (PC1) and second (PC2) principal components describe small (<1-nm) and large (>1-nm) length scale hydration effects, respectively. Transitions in water behavior from small to large length scales was first formally described by Lum, Chandler, and Weeks (LCW)(15). LCW theory describes how water molecules in vicinity of small (<1 nm radius) hydrophobic solutes (e.g., methane) exhibit persistent orientational ordering because they maintain the ability to hydrogen bond to one another. On the other hand, large (>1 nm radius) hydrophobic interfaces (e.g., oil-water interfaces) inhibit hydrogen bonding and induce dewetting of the interface. Discrepant water behavior at small and large length scale frequently emerges in both theoretical(14, 23, 33, 79, 94) and experimental(95) literature. While the transition from small to large length scale hydration behavior has typically considered water's hydrogen bonding capacity, Monroe and Shell's findings suggest that this behavior also manifests in the three body intermolecular water angle correlations(79), which is strongly reinforced by the results found here for a much broader range of chemically realistic systems.

To quantify the relevance of the three-body angle distribution more generally, we use the scikit-learn(96) package to perform PCA on the entire set of 59 differential three body angle distributions for all mixtures studied. The analysis reveals that approximately 97% of the variance in $\Delta P_{3b}(\theta)$—across the entire composition-temperature space—is describe by the first two principal components (PC1 and PC2) with a nontrivial 3% of the variance explained by the third component (PC3). While PCA alone does not reveal the physical origin of the PCs, comparing the PC signals to the trends in $\Delta P_{3b}(\theta)$ provides significant insight.

We note that PC1 exhibits peaks with positive and negative amplitude at $\theta \approx 64°$ (icosahedral) and $109.5°$ (tetrahedral), respectively [**Figure 5a**]. Hence, we interpret PC1 as the tradeoff between predominantly water-like (e.g., methanol-water at 273K) to simple fluid-like (e.g., methanol-water at 373K) environments. We characterize the tetrahedral-icosahedral tradeoff by computing projections of PC1 onto $\Delta P_{3b}(\theta)$ for each of the 59 simulations $p_1 = proj\big(PC_1, \Delta P_{3b}(\theta)\big)$. As p_1 measures the balance between tetrahedral and icosahedral water configurations, p_1 exhibits a nearly perfect correlation ($R \approx 1.0$) with p_{tet} with positive p_1 corresponding to higher tetrahedrality and negative p_1 corresponding to lower tetrahedrality than neat water [**Figure 5b**].

Notably, PC2 displays positive peaks at $\theta \approx 54°$ and 90° and negative peaks at $\approx 60°$ and $135°$. While most of these peaks are not easily ascribed to specific populational shifts as in the tetrahedral-icosahedral tradeoff of PC1, the PC2 peak at $\approx 90°$ stems from the increase in the population of ideal-gas-like water configurations observed at high temperature (e.g., methanol-water at 373K) and high cosolvent concentration (e.g. DMSO at $X_{DMSO} = 0.45$). We demonstrate that the projection of PC2 onto $\Delta P_{3b}(\theta)$, p_2, shows the most significant correlation ($R = -0.92$) with $\langle \cos\theta \rangle$, [**Figure 5.5(c)**] suggesting that p_2 probes changes in the mean of $P_{3b}(\theta)$. While it seems that p_2 must then correlate strongly with p_{tet} as well, this relationship breaks down at high concentrations (e.g., $X_{DMSO} = 0.45$) where significant rightward drift in the entire $P_{3b}(\theta)$ manifests (**Figure 5.17**).

In **Figure 5.5(a)**, we observe a pronounced peak in PC3 at $\theta \approx 50°$ that cannot be directly attributed to icosahedral or tetrahedral structures. Previous work has suggested that features at $\theta \approx 50°$ stem primarily from a small population of 5-coordinated water molecules(36, 79). These authors demonstrate higher $P_{3b}(50°)$ in 5-coordinated waters compared to 4-coordinated waters. Additionally, PC3 presents a diffuse, positive amplitude for $\theta > 120°$. However, this large-θ region of PC3 is difficult to attribute to specific phenomena due overlapping features of both PC1 and PC2. The large-θ effect manifests in the variable-concentration simulations in which increasing concentration yields a higher population of waters with $\theta > 120°$ and hence increasing p_3. The competition between the impact of high water-water coordination and large-θ effects is illustrated in **Figure 5.5(d)**. Here, the alcohol-water simulations, with variable temperature and fixed 10% weight fraction, exhibit a negative correlation between the increasing water-water coordination and decreasing p_{tet} with increasing temperature. On the other hand, the variable composition mixtures (e.g., DMSO-water) display the simultaneous enhancement of water tetrahedrality and high-θ water configurations.

We can further explain the PC trends by drawing comparisons to the prior work of Monroe and coworkers(34, 61, 79) in it was suggested that PC1 and PC2 correspond to water structural changes at small (<1.0 nm) and large length (>1.0 nm) scales, respectively. We consider previous $\Delta P_{3b}(\theta)$ results for an ideal small hydrophobic molecule, a methane molecule(79), and an ideal large hydrophobic surface, a methylated self-assembled monolayer (SAM)(11, 34), in **Figure 5.5(a)**. Notably, $\Delta P_{3b}(\theta)$ for methane resembles a superposition of PC1 and PC3 while $\Delta P_{3b}(\theta)$ for the SAM superposes signals from PC2 and PC3. Hence, we directly illustrate how PC1 and PC2 correspond to contributions from small and large length scale structural shifts in water. Both systems include contributions from PC3 variations because the probability of observing 5-coordinating waters decreases in the hydration layers at both small and large hydrophobic interfaces.

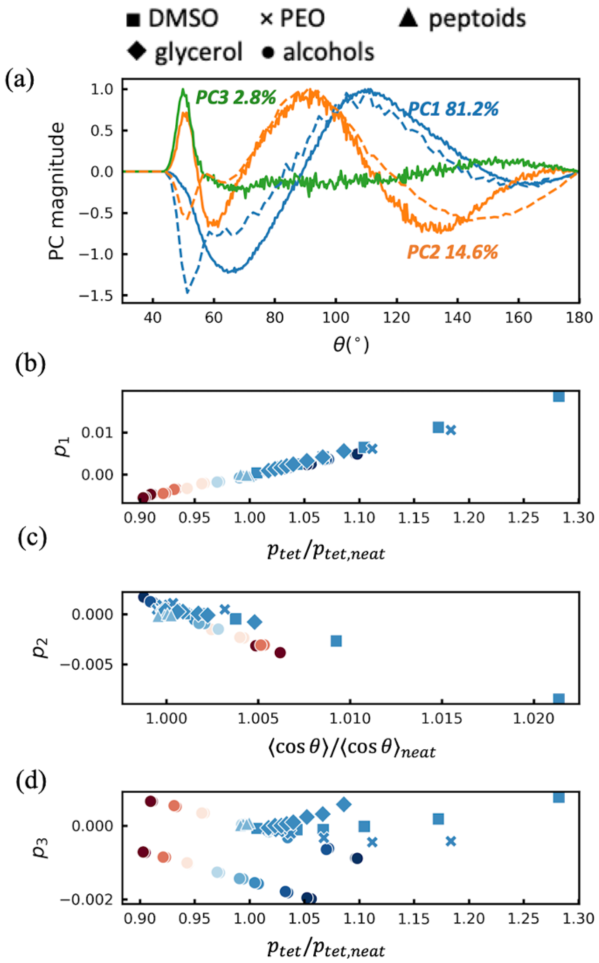


Figure 5.5. (a) Principal component analysis on the set of differential three-body angle distributions, $\Delta P_{3b}(\theta)$, for the 59 simulations described in **Section 5.2.1**. $\Delta P_{3b}(\theta)$ results for water populations in the first hydration shell of methane in water(79) and the hydration layer of a large hydrophobic surface, a methylated self-assembled monolayer (SAM)(34), are shown with dashed lines in blue and orange, respectively. The methane $\Delta P_{3b}(\theta)$ resembles a convolution of PC1 and PC3 while $\Delta P_{3b}(\theta)$ for the SAM convolves PC2 and PC3. (b) The projection of $\Delta P_{3b}(\theta)$ onto PC1 shows a strong positive correlation with the population of tetrahedrally coordinated $\frac{p_{tet}}{p_{tet,neat}}$ relative to neat water. (c) The projection of $\Delta P_{3b}(\theta)$ onto PC2 generally exhibits a negative correlation with $\frac{\langle\cos\theta\rangle}{\langle\cos\theta\rangle_{neat}}$. (d) The projection of $\Delta P_{3b}(\theta)$ onto PC3 shows system-dependent linear trends with $\frac{p_{tet}}{p_{tet,neat}}$. Here, $p_{tet,neat}$ and $\langle\cos\theta\rangle_{neat}$ are p_{tet} and $\langle\cos\theta\rangle$ for neat water at 298K.

5.3.3. Pearson correlation matrix

While the previous discussion of results for methanol-water systems and associated structural metrics provides initial interpretability, here we consider the behavior of a broader set of order parameters using the entire 59-system simulation dataset spanning multiple cosolvent chemistries, compositions, and temperatures. To understand the general relationships, we construct two separate Pearson correlation coefficient, R, matrices: one between the order parameters and D_{H_2O} and the other between the parameters and μ_{HS}^{ex}. For the sake of interpretability, we present only a subset of putative structural metrics in **Figures 5.6 and 5.7**. Full correlation matrices between all structural metrics water properties, D_{H_2O} and μ_{HS}^{ex}, are shown in **Figures S5.9 and S5.10**.

We first note several interesting correlations—or lack thereof—with D_{H_2O} [**Figure 5.6**]. As suggested by many findings in the literature(10–12, 58), water's translational mobility (e.g., D_{H_2O}) tends to be correlated ($|R| > 0.5$) to probes of translational order such as $n_{1,Ow-Ow}$, t, LSI, and $\langle N \rangle$. However, probes of the tetrahedral network structure including p_{tet}, S_{3b}, and p_1 exhibit even greater correlation with D_{H_2O} ($R > 0.9$) than the translation order parameters (apart from $n_{1,Ow-Ow}$). The most strongly-correlated order parameters with diffusivity (p_{tet}, S_{tet}, p_1, and $n_{1,Ow-Ow}$) are much more sensitive to the effects of temperature than less correlated parameters (t, LSI, and $\langle N \rangle$), which exhibit much smaller order parameter-diffusivity slopes for the alcohol series of simulations (see **Figure 5.19** for more details). Unlike the quantities derived from $P(\theta)$, q shows weaker correlation with D_{H_2O} when all 59 systems are considered. As mentioned in **Section 5.2.3**, water molecules can only be considered tetrahedral ($q = 1$) with all six nearest neighbors triplet angles being tetrahedral ($\psi_{jk} = 109.5°$). Further, lower water-water coordination number, $n_{1,Ow-Ow}$, accompanies increased additive concentration (e.g., glycerol-water, DMSO-water, and PEO-water series). Hence, the nearest neighbors outside of the first hydration shell have decreased orientational correlation with the central water molecule and thus lower tetrahedrality per q. On the other hand, we can eliminate the effect of cosolvent crowding by fixing mixture composition and decreasing temperature (e.g., alcohol-water series); in these cases, we observe a systematic increase in q upon cooling [**Figure 5.4(a)**]. For instance, we directly observe the opposite $D_{H_2O} - q$ relationships in the methanol-water series ($R < 0$) and glycerol-water series ($R > 0$) [**Figure 5.21**].

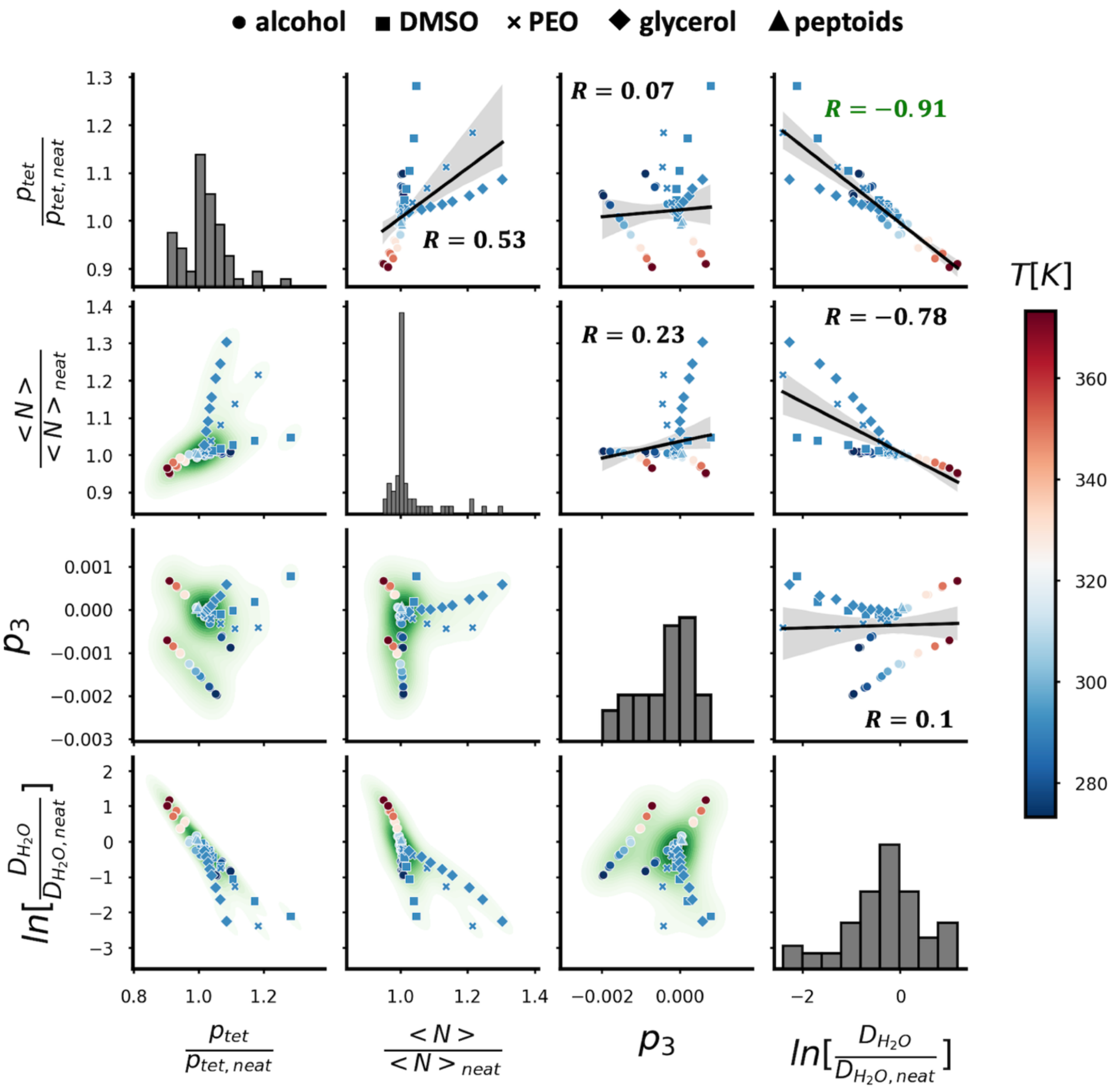


Figure 5.6. The upper off-diagonal tiles of the matrix depict the Pearson correlation between $\ln\left[\frac{D_{H_2O}}{D_{H_2O,neat}}\right]$ and a subset of relevant structural metrics. Here, linear fits between each pair of properties with 95% confidence bands shown in gray. Notably, we observe the largest correlation ($R = -0.91$) between $\ln\left[\frac{D_{H_2O}}{D_{H_2O,neat}}\right]$ and the relative p_{tet}, $\frac{p_{tet}}{p_{tet,neat}}$. Further, none of the three structural metrics presented ($\frac{p_{tet}}{p_{tet,neat}}$, $\frac{\langle N\rangle}{\langle N\rangle_{neat}}$, and p_3) are strongly correlated with the others. The diagonal tiles depict histograms of the simulation-observed D_{H_2O}, p_{tet}, $\langle N\rangle$, and p_3, separately. The lower off-diagonal tiles are estimates of the kernel density estimates (KDEs) of the bivariate distribution of each pair of properties in the matrix. The green shaded regions indicate the probability density with darker colors corresponding to higher probability density. In all off-diagonal tiles, the data points are colored according to the simulation temperature.

We separately consider μ_{HS}^{ex} as a surrogate measurement of the solvation thermodynamic environment for small, methane-like solutes. Unlike D_{H_2O}, μ_{HS}^{ex} exhibits a general lower degree of correlation with all order parameters considered [**Figure 5.7**]. Apart from $\langle v_{wat} \rangle$, $\langle a_{wat} \rangle$, and p_3, none of the order parameters have $|R| > 0.5$ correlations with μ_{HS}^{ex}. The lack of correlation is undoubtedly partially due to the subtle shifts in μ_{HS}^{ex} with both changes in temperature, solute chemistry, and composition. We physically interpret the correlation between the average molecular volume of a water molecule $\langle v_{wat} \rangle$ and μ_{HS}^{ex} as follows: increased molecular volume (decreasing mixture density) increases the favorability of solvating a hard sphere (lowers μ_{HS}^{ex}). A similar argument is made for the average molecular area of a water molecule $\langle a_{wat} \rangle$. The connection between μ_{HS}^{ex} and p_3, on the other hand, is less straightforward. As discussed in **Section 5.3.2**, we attribute PC3 primarily to the effect of a small population of 5-coordinated waters. Further, we note that Dallin et. al. found the population of 5-coordinated waters to predictive of solvation energetics—namely, the hydration free energy at self-assembled monolayers surfaces (SAMs)(36). The resulting projection of PC3 onto a given system's differential 3-body angle distribution quantifies the portion of 3-body angles between corresponding to a 5-coordinated water molecule. Hence, water molecules in hotter mixtures have a higher likelihood of 5-coordinating (p_3 increases) due to lower water-cosolvent affinity and simultaneously higher affinity for the small hydrophobe (μ_{HS}^{ex} increases).

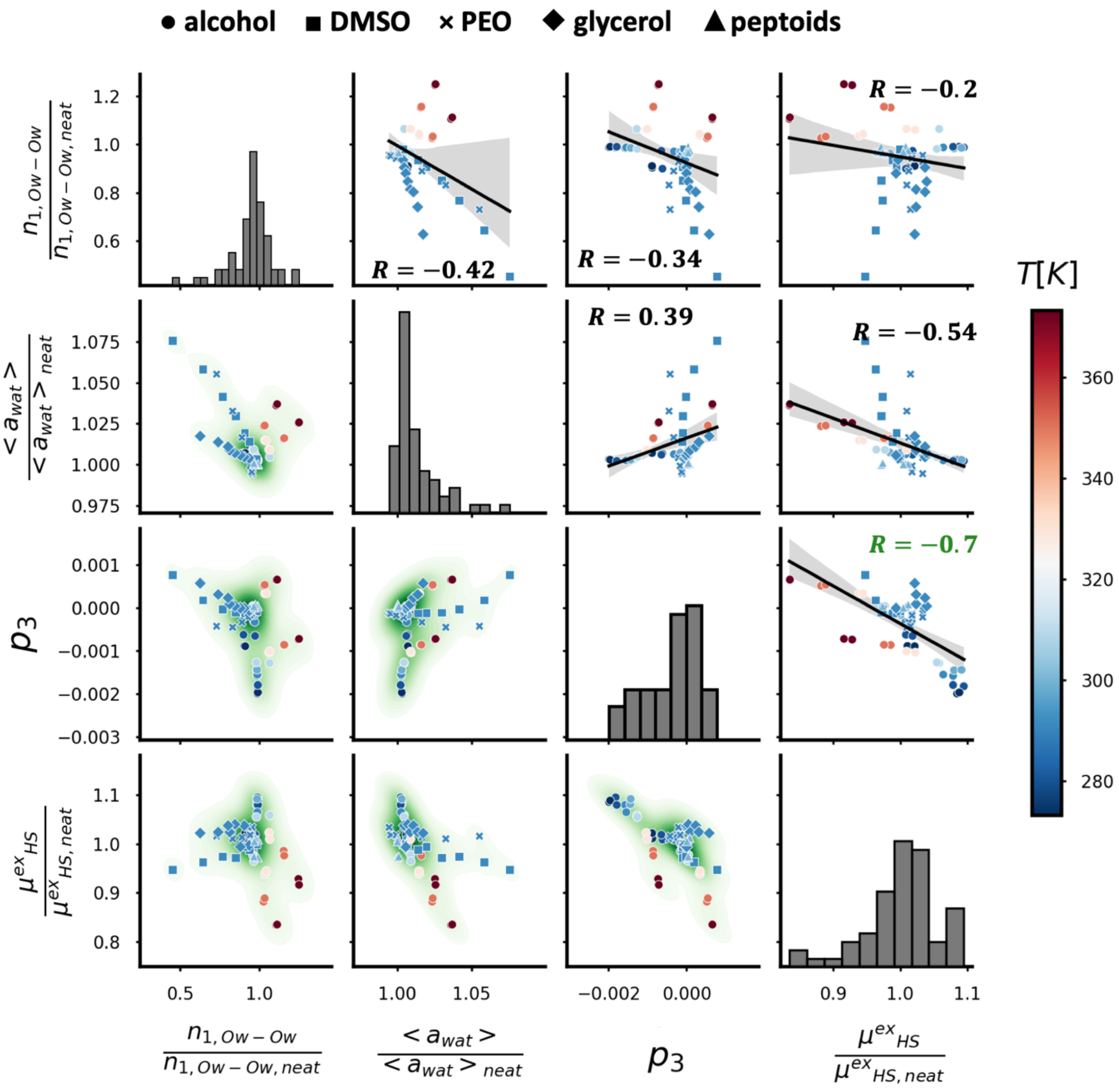


Figure 5.7. The upper off-diagonal tiles of the matrix depict the Pearson correlation between $\frac{\mu_{HS}^{ex}}{\mu_{HS,neat}^{ex}}$ and a subset of relevant structural metrics. Here, linear fits between each pair of properties with 95% confidence bands shown in gray. Notably, we observe the largest correlation ($R = -0.7$) between $\frac{\mu_{HS}^{ex}}{\mu_{HS,neat}^{ex}}$ and p_3. Further, none of the three structural metrics presented ($\frac{\langle a_{wat}\rangle}{\langle a_{wat}\rangle_{neat}}$,$\frac{n_{1,Ow-Ow}}{n_{1,Ow-Ow,neat}}$and p_3) are strongly Pearson-correlated with the others. The diagonal tiles depict histograms of simulation-observed μ_{HS}^{ex}, $\langle a_{wat}\rangle$, $n_{1,Ow-Ow}$, and p_3, separately. The lower off-diagonal tiles are estimates of the kernel density estimates (KDEs) of the bivariate distribution of each pair of properties in the matrix. The green shaded regions indicate the probability density with darker colors corresponding to higher density. In all off-diagonal tiles, the data points are colored according to the simulation temperature.

5.3.4. Statistical learning approach to property prediction

Though atomistic molecular simulations provide the detailed configurational information necessary to resolve the numerous water structural metrics described in **Section 5.2.3**, experimental techniques that probe molecular scale features are typically limited to a single type of measurement. For example, while an experimental technique like ODNP yields spatially resolved water diffusivity measurements, one cannot determine the molecular driving forces underlying the dynamic shifts without molecular simulation or other experimental methods. The correlations described above offer the opportunity of developing a deeper and more general understanding of the structural means and magnitudes through which observed dynamic or thermodynamic shifts emerge.

To quantify the relationship between a collection of water structural metrics and observable properties, here we apply statistical learning algorithms to systematically select the metrics that best describe and predict changes in equilibrium dynamics (diffusivity) and solvation thermodynamics (hard sphere excess chemical potential). In statistical learning applications, researchers frequently seek to reduce redundant or uninformative features from models to improve computational efficiency and improve the robustness of predictions(36, 97, 98). Due to the potentially large numbers of putative features (possibly thousands), these efforts tend to leverage low intervention dimensionality-reduction strategies(99–101). Hence, dimensionality-reduction algorithms like principal component analysis(16, 79, 102) and feature selection strategies(36) are often applied.

Due to the high computational cost of direct atomistic simulation, generating large datasets for statistical learning methods can be challenging. Fortunately, we have access to numerous simulations from our previous works(10, 62) and the additional simulations described in **Section 5.2.1**. We leverage the structural metrics detailed in **Section 5.2.3** paired with feature selection to develop a general connection between molecular scale structure, dynamics, and solvation thermodynamics. More specifically, using the fSFS procedure described in the **Section 5.2.4**, we systematically rank 20 structural order parameters [**Table 1**] and three principal component projections (23 order parameters in total) in order of their importance to predicting the log self-diffusivity of water $\ln\left(\frac{D_{H_2O}}{D_{H_2O,neat}}\right)$ and, separately, the excess hard sphere chemical potential $\frac{\mu_{HS}^{ex}}{\mu_{HS,neat}^{ex}}$. Here for predictive models, we normalize each quantity by its value in "neat" water at 298K and 1bar.

Applying fSFS, we find that the three most important features for describing the variation in $\ln\left(\frac{D_{H_2O}}{D_{H_2O,neat}}\right)$ to be p_{tet}, $\langle N\rangle$, and p_3 in order of decreasing importance. Further, we develop three multiple linear regression (MLR) models considering increasing numbers of independent variables: (1) $\frac{p_{tet}}{p_{tet,neat}}$, (2) $\frac{p_{tet}}{p_{tet,neat}}$ and $\frac{\langle N\rangle}{\langle N\rangle_{neat}}$, and (3) $\frac{p_{tet}}{p_{tet,neat}}$, $\frac{\langle N\rangle}{\langle N\rangle_{neat}}$, p_3. Here, $\frac{p_{tet}}{p_{tet,neat}}$ is the greatest predictor of D_{H_2O}, with $R = -0.91$ [**Figure 5.6**]. In **Figure 5.8**, we depict the performance of each of the three models compared to direct measurement of D_{H_2O}. Remarkably, even for the simplest model, (1), we find high predictive capacity ($R_{test}^2 = 0.83$). Here, we observe the poorest model agreement for the systems with variable composition (e.g., glycerol-water, DMSO-water, and PEO-water). For example, as suggested by **Figure 5.1**, depending on the solute identity the slope of the diffusivity-tetrahedrality relationship changes dramatically.

Upon including only one additional feature $\langle N\rangle$, we find significantly better model performance ($R_{test}^2 = 0.96$). We hypothesize that p_{tet} captures shifts in orientational structure while $\langle N \rangle$ corrects the model to account for composition-induced changes in translational order. At first glance, **Figure 5.19** suggests that $n_{1,Ow-Ow}$ is superior to $\langle N \rangle$ as a translational order parameter ($R = 0.9$ rather than $R = -0.8$). However, $n_{1,Ow-Ow}$ is colinear with p_{tet} ($R = -0.9$). This collinearity results in a similar predictive power for both $n_{1,Ow-Ow}$ and p_{tet}.

The importance of quantifying both orientational and translational order was previously noted by Errington and Debenedetti(9). In this work, they discovered that pair of order parameters q and t signaled the emergence of anomalous thermophysical properties in supercooled water. Though here we highlight different structural metrics, the parallel suggests that the interplay between orientational and translational order are natural descriptors, which here are identified using automated feature selection. Further, adding a third feature, p_3, does not dramatically increase the predictive capacity of the MLR model ($R^2_{test} = 0.99$). This may suggest that water mobility is primarily determined by water-water translational and orientational correlations even in environments heavily crowded by cosolvent molecules.

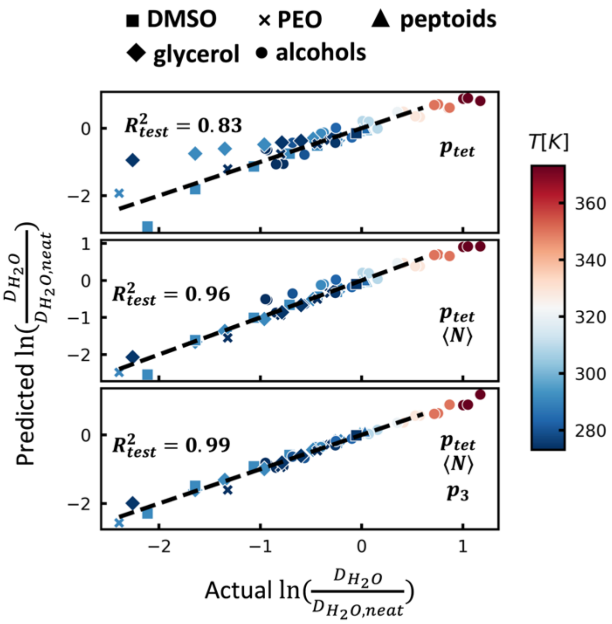


Figure 5.8. From top to bottom, performance of model (1) with dependent variable $\frac{p_{tet}}{p_{tet,neat}}$, (2) with dependent variables $\frac{p_{tet}}{p_{tet,neat}}$ and $\frac{\langle N \rangle}{\langle N \rangle_{neat}}$, and (3) with dependent variables $\frac{p_{tet}}{p_{tet,neat}}$, $\frac{\langle N \rangle}{\langle N \rangle_{neat}}$, and p_3 to predict D_{H_2O}. Model (1) does not perfectly capture the different D_{H_2O}-p_{tet} relationships across distinct systems as shown in **Figure 5.1**. Model (2) significantly improves upon the accuracy. Here, the inclusion of a metric for translational order, $\langle N \rangle$, largely corrects for the poor performance for high concentration glycerol-water and DMSO-water mixtures. While adding the independent variable, p_3, in model (3) marginally increases the accuracy of D_{H_2O} prediction, the physical interpretation of this improvement is not entirely clear. Including higher numbers of independent variables greater than three does not dramatically improve performance [**Figure 5.22**].

We also seek the top three features that best describe trends in solvation thermodynamics. Implementing the same fSFS procedure that we apply to water self-diffusivity, we construct the following MLR models (4) p_3, (5) p_3 and $\frac{n_{1,Ow-Ow}}{n_{1,Ow-Ow,neat}}$, and (6) p_3, $\frac{n_{1,Ow-Ow}}{n_{1,Ow-Ow,neat}}$, and $\frac{\langle a_{wat} \rangle}{\langle a_{wat} \rangle_{neat}}$ to predict for $\frac{\mu^{ex}_{HS}}{\mu^{ex}_{HS,neat}}$. In **Figure 5.9**, we compare the test set performance of models (4)-(6) to computed values of μ^{ex}_{HS}. Unlike the results shown in **Figure 5.8**, the single variable model (4) poorly predicts ($R^2 = 0.37$) trends in μ^{ex}_{HS}. The underperformance of model (4) may initially seem surprising given the high Pearson correlation ($|R| = 0.7$) between μ^{ex}_{HS} and p_3. However, we note that the $p_3 - \mu^{ex}_{HS}$ relationship exhibits trends with opposing slopes for variable concentration and variable temperature simulations as further discussed in **Section 5.3.3.**

With the inclusion of $n_{1,Ow-Ow}$, model (5) significantly improves performance on the test set ($R^2_{test} = 0.67$). While model (2) leverages both translational and orientational metrics to predict D_{H_2O}, model (5) relies on two metrics that quantify water-water radial correlations to predict μ^{ex}_{HS}. Here, $n_{1,Ow-Ow}$ improves model performance by mitigating the deviation from the linear μ^{ex}_{HS}-structure relationship observed in the variable concentration simulations. We improve μ^{ex}_{HS} prediction in model (6) by accounting for the impact of water's increasing specific molecular area, $\langle a_{wat}\rangle$, with decreasing μ^{ex}_{HS}. Upon including $\langle a_{wat}\rangle$, we observe comparable test set performance ($R^2_{test} = 0.88$) to models (1) and (2). We also note that similar performance to model (6) is attained by swapping $\langle a_{wat}\rangle$ for $\langle v_{wat}\rangle$.

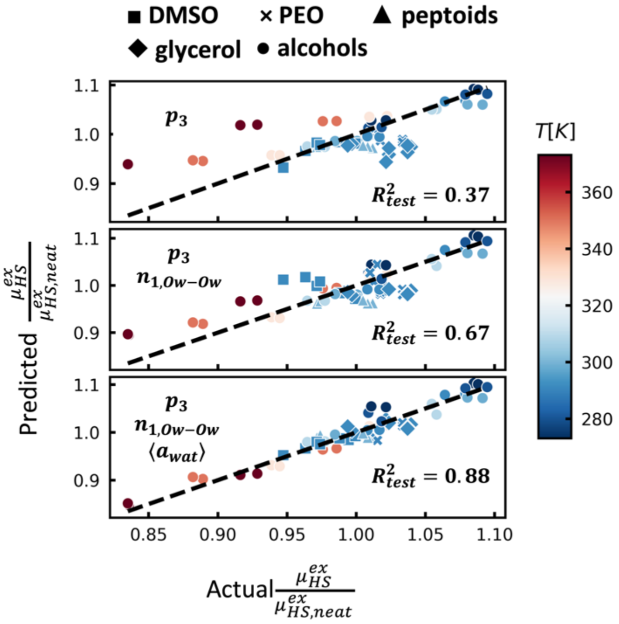


Figure 5.9. Performance of models (4) (with dependent variable p_3), (5) (with dependent variables p_3 and $\frac{n_{1,OwOw}}{n_{1,OwOw,neat}}$), and (6) (with dependent variables p_3, $\frac{n_{1,OwOw}}{n_{1,OwOw,neat}}$, and $\frac{\langle a_{wat}\rangle}{\langle a_{wat}\rangle_{neat}}$) to predict μ^{ex}_{HS} on the test data set. We observe weak model performance including the single independent variable, p_3, in model (4). Here, the

relationship partially owes to the increased likelihood of observing supra-tetrahedral (5-coordinated) water coordination environments with increasing temperature in alcohol-water. In model (5), we observe a substantial increase in model accuracy by accounting for the mean change in the water-water coordination number, $n_{1,Ow-Ow}$. Only upon including water's average molecular surface area, $\langle a_{wat} \rangle$, does structure reliably predict μ_{HS}^{ex}. Including higher numbers of independent variables greater than three does not dramatically improve performance [**Figure 5.23**].

We further validate the fSFS-selected features by applying alternative feature selection strategies—namely—LASSO regression. In **Figures S5.14 and S5.15,** we depict the results of a LASSO regression procedure for estimating the relative importance of structural features for diffusivity-prediction. LASSO regression indicates that six structural metrics are predictive of D_{H_2O}. As with the fSFS results, the LASSO procedure includes p_{tet}, p_3, and $\langle N \rangle$ among the most descriptive structural metrics. We further note that fSFS is a greedy algorithm—choosing only the most descriptive structural metric with each successive iteration. Alternatively, even upon tuning the LASSO learning rate with 4-fold cross-validation, LASSO regression yields a greater number of descriptive structural metrics. However, depending on the choice of test and training datasets, LASSO produces different feature rankings. Notably, in the first case [**Figure 5.24**] we observe the highest importance score for $\langle N \rangle$ while in a second case [**Figure 5.25**] we find that p_{tet} is most important. This strong dependency on the choice of training data suggests that, even with variable standardization, LASSO-based feature selection cannot yield an absolute ranking of features in the present case. We provide details on these analyses in the **Section 5.5.4**.

5.3.5 Inferring structure metrics from dynamics

The diffusivity-structure relationships above provide a potential framework for predicting the structural mechanisms driving shifts in dynamics with application in more complex mixtures and at chemically heterogeneous surfaces such as proteins-water interfaces. While this fSFS-MLR procedure selects the most important structural metrics, it does not inherently quantify the magnitude of these metrics' effect on dynamics. For example, though glycerol-water at mole fraction 0.05 is accurately predicted by model (2), we cannot determine the extent to which this agreement is driven by translational and orientational structural effects.

We consider model (3) for predicting water self-diffusivity using only two order parameters: $\frac{\langle N\rangle}{\langle N\rangle_{neat}}$ and $\frac{p_{tet}}{p_{tet,neat}}$. As noted in **Section 5.3.4**, including more than three independent parameters in the MLR model does not meaningfully improve performance. Further, selecting only one translational ($\frac{\langle N\rangle}{\langle N\rangle_{neat}}$) and one orientational ($\frac{p_{tet}}{p_{tet,neat}}$) structural metric improves the interpretability of the predictions. Invoking this simple two parameter model, we employ a Bayesian linear regression framework to estimate the posterior distribution of fitting parameter values, $p(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{y})$

$$p(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{y}) \propto p(\boldsymbol{\beta})p(\boldsymbol{X},\boldsymbol{y}|\boldsymbol{\beta}) \qquad (5.5)$$

where $\boldsymbol{\beta}$ is the vector, $[\beta_{p_{tet}}, \beta_{\langle N\rangle}, \beta_0]$. Here, $\beta_{p_{tet}}$ and $\beta_{\langle N\rangle}$ are slope parameters corresponding to $\frac{p_{tet}}{p_{tet,neat}}$ and $\frac{\langle N\rangle}{\langle N\rangle_{neat}}$, respectively, and β_0 is the intercept. $\boldsymbol{X}$ is a 59x3 array with each row corresponding to a simulation observation, column one corresponds to an observation of $\frac{p_{tet}}{p_{tet,neat}}$, column two corresponds to $\frac{\langle N\rangle}{\langle N\rangle_{neat}}$, and column three is unity for every element to account for the intercept. $\boldsymbol{y}$ is a 59-element column vector where each value is a simulation estimate of D_{H_2O}. We note that the entire set of dependent and independent variables are transformed to unit-variance, zero-mean representation. $p(\boldsymbol{\beta})$ is the prior distribution of $\boldsymbol{\beta}$. $p(\boldsymbol{X},\boldsymbol{y}|\boldsymbol{\beta})$ is the likelihood of observing $\boldsymbol{X}$ and $\boldsymbol{y}$ conditioned on $\boldsymbol{\beta}$. Here, we take $p(\boldsymbol{\beta})$ to be a bivariate normal distribution, $\mathcal{N}(0,\boldsymbol{\Sigma})$, with zero-mean and unit covariance $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. We detail the impact of this prior assumption in further detail in the **Section 5.5.6**.

More interestingly, we leverage $p(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{y})$ to generate new predictions of diffusivity via the posterior predictive distribution

$$p(y^{new}|\boldsymbol{x^{new}},\boldsymbol{X},\boldsymbol{y}) \propto \int p(\boldsymbol{\beta}|\boldsymbol{X},\boldsymbol{y})p(y^{new}|\boldsymbol{x^{new}},\boldsymbol{\beta})d\boldsymbol{\beta} \qquad (5.6)$$

where $\boldsymbol{x^{new}}$ is a proposed new vector of structural metrics, $\boldsymbol{x^{new}} = \left[\frac{p_{tet}^{new}}{p_{tet,neat}} \frac{\langle N\rangle^{new}}{\langle N\rangle_{neat}}\right]$. y^{new} is the new D_{H_2O} prediction given $\boldsymbol{x^{new}}$. With one additional application of Bayes' Theorem, we infer a posterior distribution of structural metrics

$$p(\boldsymbol{x^{new}}|y^{new},\boldsymbol{X},\boldsymbol{y}) \propto p(\boldsymbol{x^{new}})p(y^{new}|\boldsymbol{x^{new}},\boldsymbol{X},\boldsymbol{y}) \qquad (5.7)$$

where $\boldsymbol{x^{new}}$ is a predicted observation vector $[p_{tet}^{new}\ \langle N\rangle^{new}]$ given a value of $y^{new} = D_{H_2O}^{new}$. We assume the prior distribution, $p(\boldsymbol{x^{new}})$, is a bivariate normal distribution, $\mathcal{N}(0,\boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. Further interrogation of this prior assumption is discussed in **Section 5.5.6**. To numerically sample $p(\boldsymbol{x^{new}}|y^{new},\boldsymbol{X},\boldsymbol{y})$, we randomly draw a vector of structural metrics, $\boldsymbol{x^{new}}$, from $p(\boldsymbol{x^{new}})$ and draw 20,000 estimates of y^{new} conditioned on the training data ($\boldsymbol{X}$ and $\boldsymbol{y}$) and $\boldsymbol{x^{new}}$. We repeat this sampling procedure 20,000 times. We then estimate the posterior distribution, $p(\boldsymbol{x^{new}}|y^{new},\boldsymbol{X},\boldsymbol{y})$, as a discrete histogram of y^{new} and $\boldsymbol{x^{new}}$. Finally, we approximate $p(\boldsymbol{x_1^{new}}|y^{new} = y_1^{new},\boldsymbol{X},\boldsymbol{y})$ for a given $y^{new} = y_1^{new}$ by selecting the y^{new}-bin value closest to y_1^{new}.

As a proof of concept, we test this Bayesian framework for two log-diffusivity values corresponding approximately to (i) methanol-water at 298K with methanol weight fraction of 0.10 and (ii) DMSO-water at 291 K with X_{DMSO} =0.45. We analyze posterior predict distributions on $p_{tet}/p_{tet,neat}$ [**Figures 5.10a** and **b**] and $\langle N\rangle/\langle N\rangle_{neat}$ [**Figures 5.10c** and **d**]. For $p_{tet}/p_{tet,neat}$, the predicted means are 0.5% larger and 9% smaller than the direct simulation observations for systems (i) and (ii), respectively. Simultaneously, the $\langle N\rangle/\langle N\rangle_{neat}$ mean predictions are 3% larger and 8% larger than the direct observation for systems (i) and (ii), respectively. The decreased accuracy of the prediction for prediction (ii) stems from the corresponding poorer performance of MLR model (2) to predict the behavior of DMSO at $X_{DMSO} = 0.45$ [**Figure 5.8**]. Though the performance of these structure predictions is imperfect, this approach yields the qualitatively correct behavior in both structural metrics.

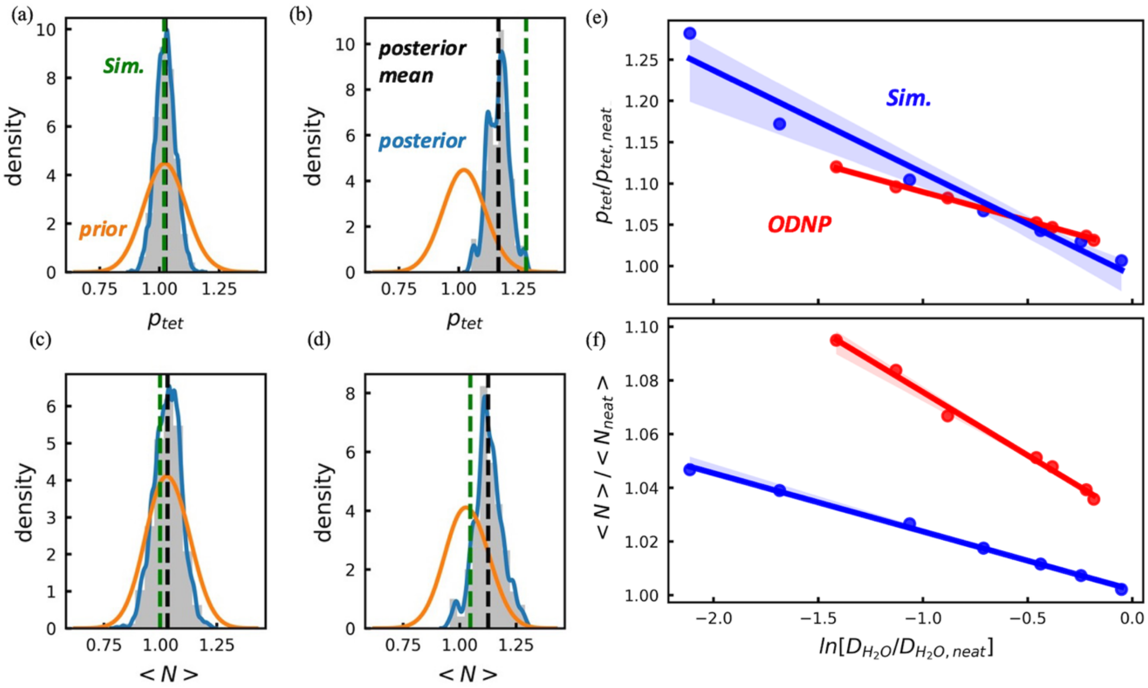


Figure 5.10. In panels (a)-(d), we depict prior (orange) and posterior predictive (blue) distributions generated by the Bayesian regression model trained on p_{tet}, $\langle N\rangle$, and p_3 to predict D_{H_2O}. The p_{tet} predictive distributions for (a) methanol-water at 298K with methanol weight fraction of 0.10 and (b) DMSO-water with $X_{DMSO} = 0.45$ and 291K yield mean values $\left\langle \frac{p_{tet}}{p_{tet,neat}} \right\rangle_{posterior}$ (black vertical line). The $\langle N\rangle$ predictive distributions for (c) methanol-water at 298K and (d) DMSO-water at $X_{DMSO} = 0.45$ and 291K yield mean values $\left\langle \frac{\langle N\rangle}{\langle N\rangle_{neat}} \right\rangle_{posterior}$ (black vertical line). Simulation means are depicted with green vertical lines. Model predictions are closer to the

simulation measurement for methanol-water at 298K compared to the DMSO-water case. Predictions of (e) $\frac{p_{tet}}{p_{tet,neat}}$ and (f) $\frac{\langle N \rangle}{\langle N \rangle_{neat}}$ generated from ODNP-derived D_{H_2O} values for a range of DMSO-water mixtures from $X_{DMSO} = 0$ to 0.5 (red) are compared to simulation-derived properties (blue). In panels (e) and (f), the lines represent linear least squares fits between the log-diffusivity and a given structure metric. Here, the shaded bands corresponding to 95% confidence intervals on the fit.

We extend this line of reasoning towards inferring water structure from experimental measurements of water self-diffusivity. Specifically, we measure D_{H_2O} by performing ODNP experiments for DMSO-water at compositions spanning $X_{DMSO} = 0$ to 0.5 (see **Section 5.2**). We first note that experimental D_{H_2O} values are not numerically identical to the corresponding simulation-estimated diffusivities. This stems partly from the ODNP assumptions used to compute D_{H_2O} via the so-called force-free hard sphere model (FFHS)(57, 103, 104). Further discrepancies may stem from the DMSO and water model force field accuracy. However, experimental and computational estimates of D_{H_2O} exhibit the same qualitative behavior for DMSO-water mixtures.

With experimental D_{H_2O} measurements in hand, we perform the same structure-prediction procedure as in case (ii) above to several DMSO-water mixtures. In **Figures 5.10e** and **5.10f**, we observe monotonic decreases in both $p_{tet}/p_{tet,neat}$ and $\langle N \rangle / \langle N \rangle_{neat}$ with increasing X_{DMSO} (decreasing D_{H_2O}). The diffusivity-structure relationships predicted from experimental results and direct computation are qualitatively similar. Further, just as with example case (ii), there is greater disparity between the inferred $\frac{\langle N \rangle}{\langle N \rangle_{neat}}$ and the direct simulation estimate than between the inferred $\frac{p_{tet}}{p_{tet,neat}}$ and its direct computation. We note that the accuracy of the $\frac{\langle N \rangle}{\langle N \rangle_{neat}}$ prediction decreases with increasing X_{DMSO}. This decreasing performance is expected given the decreasing performance of the MLR model with increasing X_{DMSO} [**Figure 5.8**]. Though not in quantitative accordance, the qualitative agreement between direct calculation and the Bayesian predictions of water structure suggests a new, systematic method of understanding the molecular scale modes, magnitudes, and mechanisms underlying changes in thermophysical properties like D_{H_2O}.

5.4 Conclusions

We utilize molecular simulations to probe water's molecular structure and water properties in a broad range of binary mixtures. We generate a library of 23 structural metrics for a set of 59 binary aqueous mixtures of varying temperature, composition, and cosolvent identity. We leverage PCA to quantify and characterize the subtler trends in water's triplet angle distribution, $\Delta P_{3b}(\theta)$, across temperature-composition space. The vast majority (97%) of the variance in $\Delta P_{3b}(\theta)$ is captured by the first two principal components (PC1 and PC2). Further, we posit that PC1 and PC2 capture shifts in water's tetrahedral network structure on small (<1 nm) and large (> 1 nm) length scales, respectively. In keeping with several previous works(14, 15, 17, 105–107), we observe enhanced water structure in the presence of small hydrophobic solutes—in our case hydrophilic solutes as well—and a disruption in water's tetrahedral network in the vicinity of large hydrophobic surfaces. This PCA framework enables new methods to characterize shifts in hydration water properties from small ideal hydrophobic molecules (e.g., methane) to solutes with intermediate hydropathy (e.g., PEO) to macromolecules and molecular surfaces.

Aggregating 20 structural metrics and the projections of each $\Delta P_{3b}(\theta)$ onto the principal components, we analyze the connections between the combined 23 structural metrics, water dynamics (D_{H_2O}), and solvation thermodynamics (μ_{HS}^{ex}). Using forward sequential feature selection (fSFS), we discover that D_{H_2O} is predicted with remarkably high accuracy ($R^2 \approx 0.96$) via multiple linear regression (MLR) with two independent variables: p_{tet} and $\langle N \rangle$. Here, p_{tet} and $\langle N \rangle$ track water's tetrahedral and translational ordering, respectively. The apparent translation-orientation dependence of D_{H_2O} parallels Errington and Debenedetti's seminal work(9) demonstrating the importance of both orientational (q) and translational (t) order in signaling anomalous properties in supercooled water. The compatibility of this prior work and the present work implies the importance of the translation-orientation pairing to describe—and potentially predict—thermophysical properties in aqueous environments. μ_{HS}^{ex}, on the other hand, is predicted with lower accuracy ($R^2 \approx 0.88$) via MLR with three independent variables: p_3, $n_{1,Ow-Ow}$, and $\langle a_{wat} \rangle$. Hence, solvation thermodynamics, as probed by μ_{HS}^{ex}, appears to depend more so on probes of local spatial correlations and water density than water's tetrahedrality. However, other aqueous systems exhibit clear relationships between solvation thermodynamic properties and water's orientational order. For example, previous studies point to significant correlation between solvation thermodynamics in colloidal hydration shells[81], extended surfaces, and protein hydration layers[109].

This feature selection workflow should be generalizable to additional water properties and aqueous environments. As an extension of this procedure, future studies should consider local hydration layer properties (e.g., protein-water interfaces) as opposed to system-average properties as in the current work. This would offer a more systematic method of understanding the structural and dynamic features underpinning hydration layer properties. Finally, we propose a method of leveraging relatively inexpensive molecular simulations to approximate water's molecular structure from ODNP-derived diffusivity data. Though the predictions are not one-to-one with direct calculations from MD simulations, we capture the qualitative structural trends observed in simulations. Hence, this work also offers the prospect of informed, but low intervention, forecasting of properties not directly measured by experiment.

5.5 Appendix

5.5.1 Further Analysis of Solution Structure

In **Figure 5.11-18**, we depict distributions of varied structural metrics for each of the 8 binary aqueous mixture types (methanol [**Figure 5.11**], ethanol [**Figure 5.12**], propanol [**Figure 5.13**], butanol [**Figure 5.14**], glycerol [**Figure 5.15**], PEO [**Figure 5.16**], DMSO [**Figure 5.17**], and peptoid [**Figure 5.18**]) and all 59 individual simulations. The first four **Figures 5.11-14** are color-coded by the simulation temperature with a blue-red color scale where the blue and red correspond to cold (e.g., 273.15K) and hot (e.g., 373.15K) temperatures, respectively. In **Figures 5.15-17**, the composition is represented by a yellow-green color scale with yellow corresponding to dilute aqueous mixtures and green corresponding to higher cosolvent concentrations. The differences the system-average structural distributions between the peptoid sequences depicted in **Figure 5.2** are not perceptible when plotted in **Figure 5.18**.

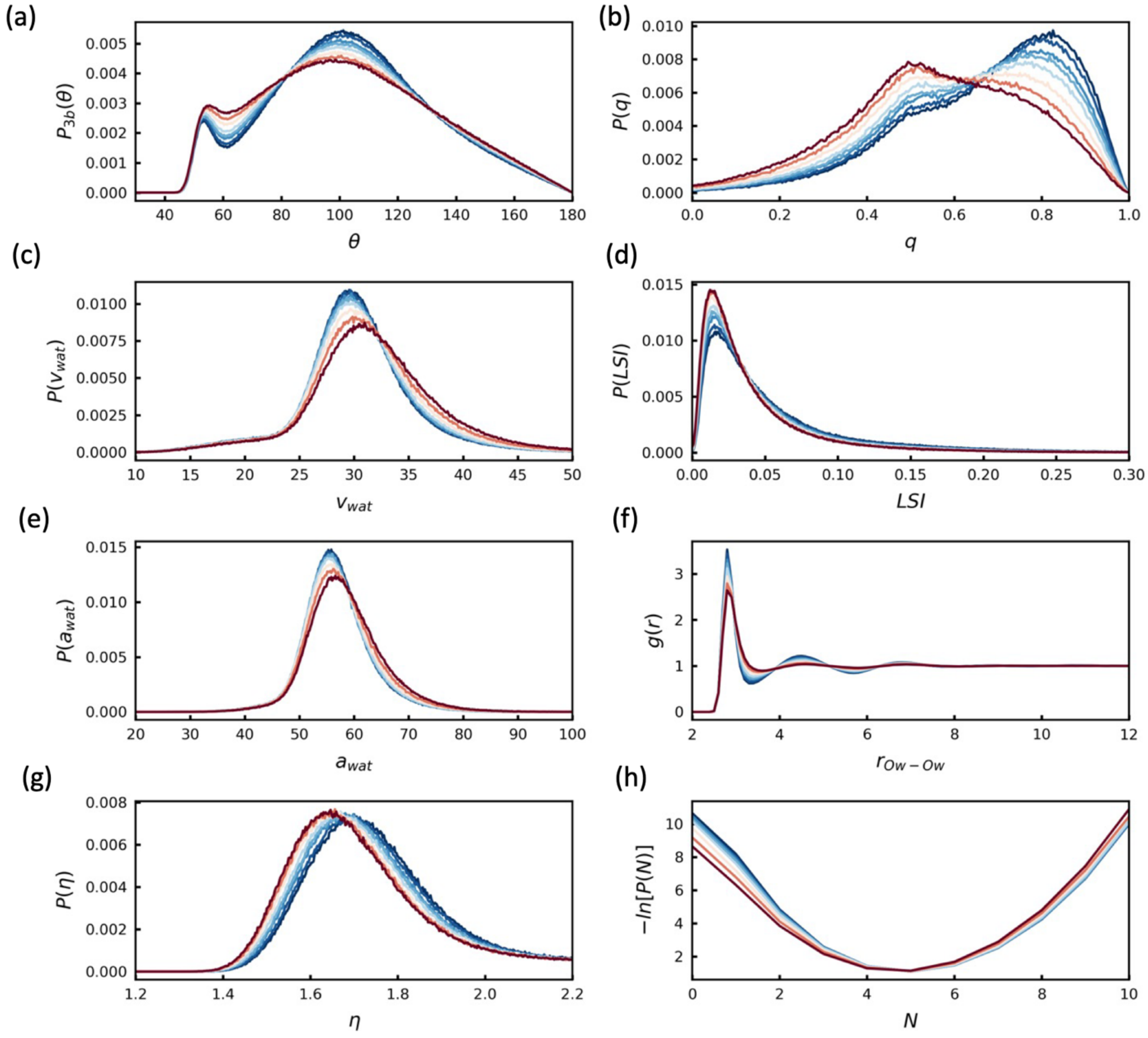


Figure 5.11. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2** for methanol-water at temperature of 273.15, 280.15, 291.15, 298.15, 309.15, 329.15, 350.15, and 373.15K. Here, red corresponds to higher temperature and blue to colder temperature. (a) The distributions of the tetrahedral order parameter, $P_{3b}(\theta)$, exhibit a systematic decreases in the icosahedral, $\theta = 64°$, and increase in the tetrahedral, $\theta = 109.5°$, regions with decreasing temperature. (b) The tetrahedral order parameter distribution, $P(q)$, exhibit peaks at $q \approx 0.4$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.8$ upon cooling. (c) The distributions of water specific volumes, $P(v_{wat})$, broaden and exhibit larger mean specific volumes, $\langle v_{wat} \rangle$, as the temperature of the mixture increases. (d) The distributions of LSI, $P(LSI)$, broaden and exhibit larger means, $\langle LSI \rangle$, as the temperature of the mixture decreases. (e) The distributions of water specific areas, $P(a_{wat})$, broaden and exhibit larger mean specific volumes, $\langle a_{wat} \rangle$, as the temperature of the mixture increases. (f) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with decreasing temperature as demonstrated by the increasing peak and trough amplitudes. (g) The distributions of water asphericity, $P(\eta)$, exhibit smaller (more spherical) mean asphericity, $\langle \eta \rangle$, as the temperature of the mixture increases. (h) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the systematic decrease in the affinity of an ideal hydrophobic molecule for methanol-water upon cooling. Simultaneously, decreased mixture temperature shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

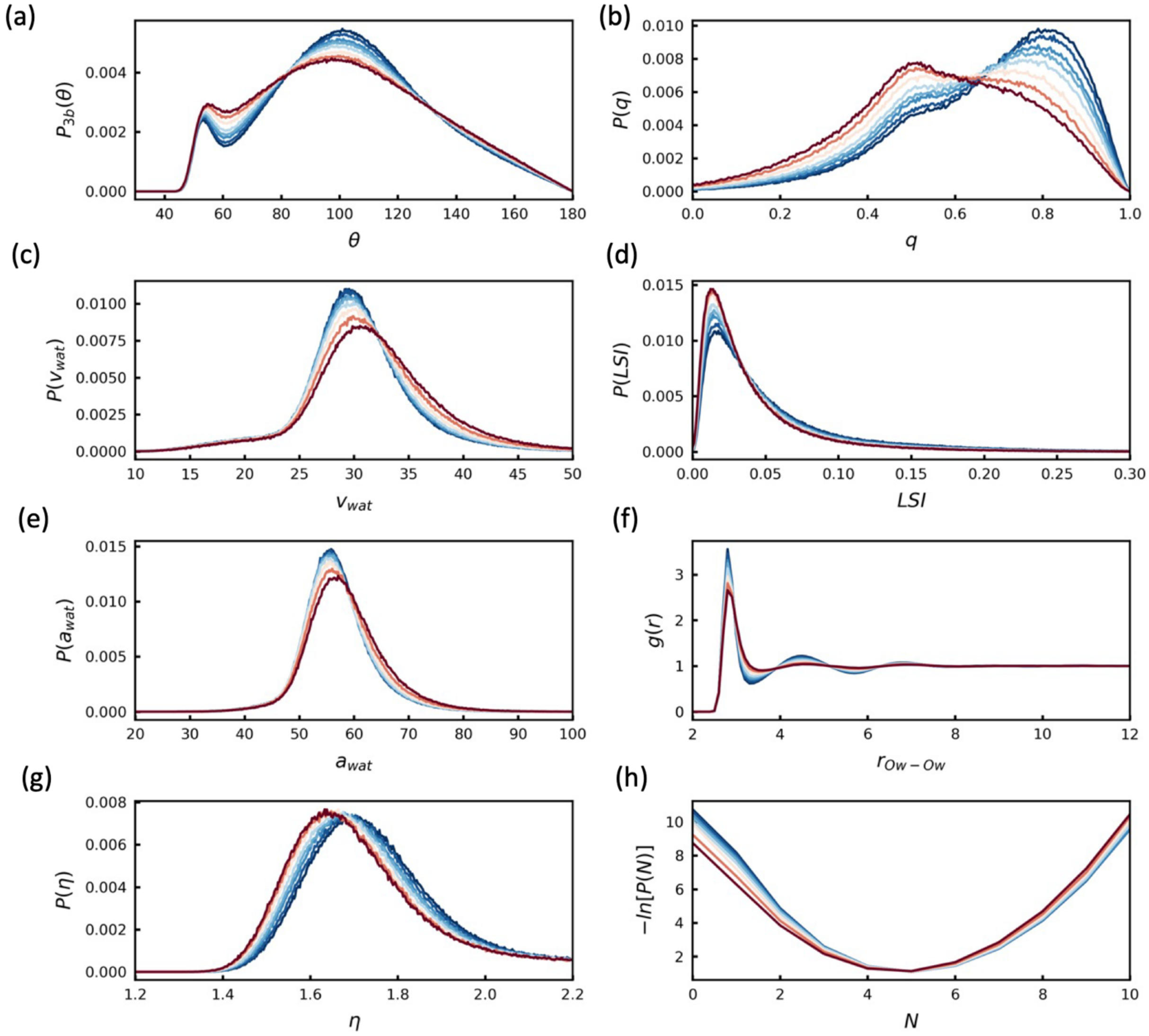


Figure 5.12. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2** for ethanol-water at temperature of 273.15, 280.15, 291.15, 298.15, 309.15, 329.15, 350.15, and 373.15K. Here, red corresponds to higher temperature and blue to colder temperature. (a) The distributions of the tetrahedral order parameter, $P_{3b}(\theta)$, exhibit a systematic decrease in the icosahedral, $\theta = 64°$, and increase in the tetrahedral, $\theta = 109.5°$, regions with decreasing temperature. (b) The tetrahedral order parameter distribution, $P(q)$, exhibit peaks at $q \approx 0.4$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.8$ upon cooling. (c) The distributions of water specific volumes, $P(v_{wat})$, broaden and exhibit larger mean specific volumes, $\langle v_{wat} \rangle$, as the temperature of the mixture increases. (d) The distributions of LSI, $P(LSI)$, broaden and exhibit larger means, $\langle LSI \rangle$, as the temperature of the mixture decreases. (e) The distributions of water specific areas, $P(a_{wat})$, broaden and exhibit larger mean specific volumes, $\langle a_{wat} \rangle$, as the temperature of the mixture increases. (f) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with decreasing temperature as demonstrated by the increasing peak and trough amplitudes. (g) The distributions of water asphericity, $P(\eta)$, exhibit smaller (more spherical) mean asphericity, $\langle \eta \rangle$, as the temperature of the mixture increases. (h) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the systematic decrease in the affinity of an ideal hydrophobic molecule for ethanol-water upon cooling. Simultaneously, decreased mixture temperature shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

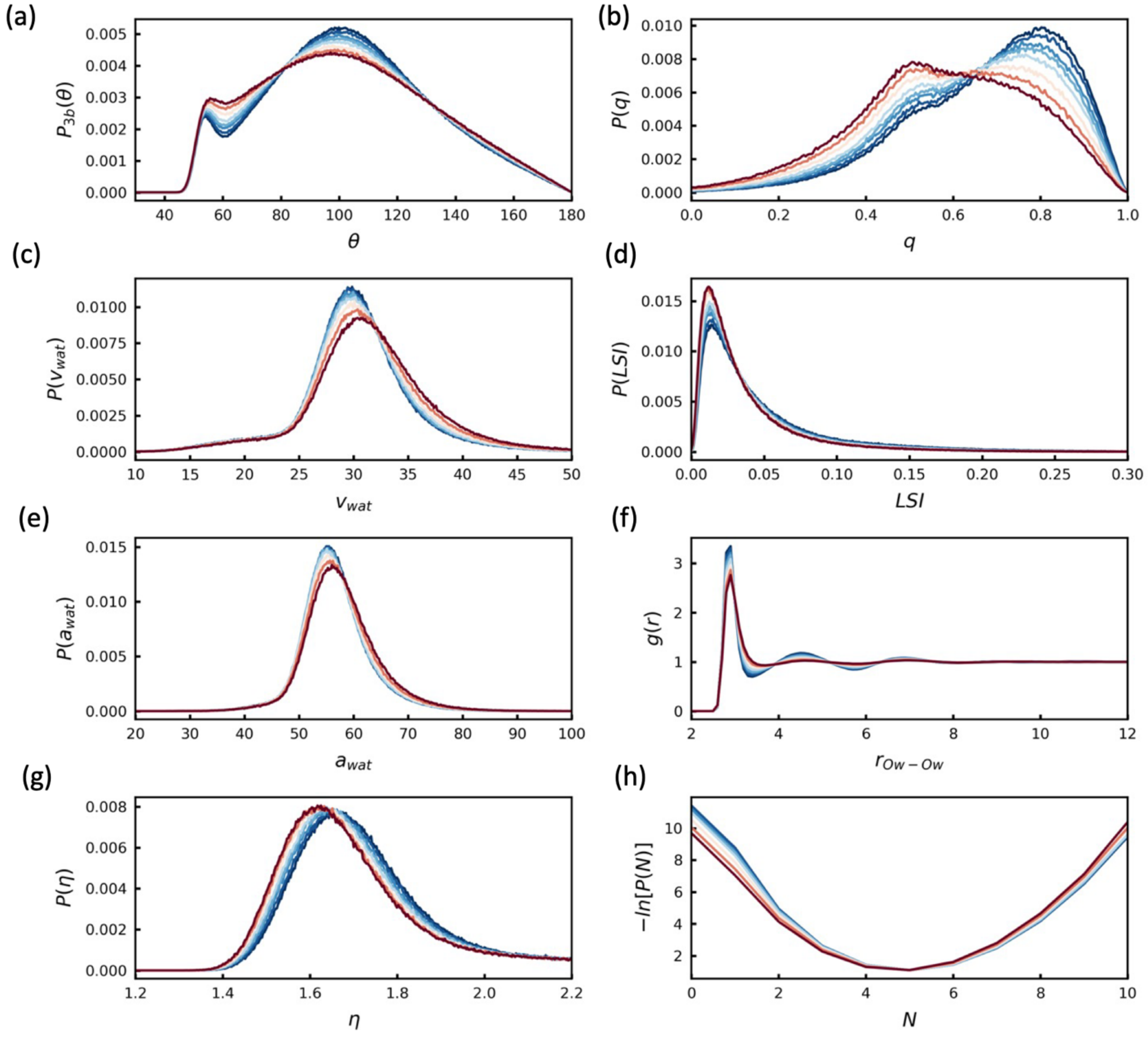


Figure 5.13. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2** for propanol-water at temperature of 273.15, 280.15, 291.15, 298.15, 309.15, 329.15, 350.15, and 373.15K. Here, red corresponds to higher temperature and blue to colder temperature. (a) The distributions of the tetrahedral order parameter, $P_{3b}(\theta)$, exhibit a systematic decrease in the icosahedral, $\theta = 64°$, and increase in the tetrahedral, $\theta = 109.5°$, regions with decreasing temperature. (b) The tetrahedral order parameter distribution, $P(q)$, exhibit peaks at $q \approx 0.4$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.8$ upon cooling. (c) The distributions of water specific volumes, $P(v_{wat})$, broaden and exhibit larger mean specific volumes, $\langle v_{wat} \rangle$, as the temperature of the mixture increases. (d) The distributions of LSI, $P(LSI)$, broaden and exhibit larger means, $\langle LSI \rangle$, as the temperature of the mixture decreases. (e) The distributions of water specific areas, $P(a_{wat})$, broaden and exhibit larger mean specific volumes, $\langle a_{wat} \rangle$, as the temperature of the mixture increases. (f) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with decreasing temperature as demonstrated by the increasing peak and trough amplitudes. (g) The distributions of water asphericity, $P(\eta)$, exhibit smaller (more spherical) mean asphericity, $\langle \eta \rangle$, as the temperature of the mixture increases. (h) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the systematic decrease in the affinity of an ideal hydrophobic molecule for propanol-water upon cooling. Simultaneously, decreased mixture temperature shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

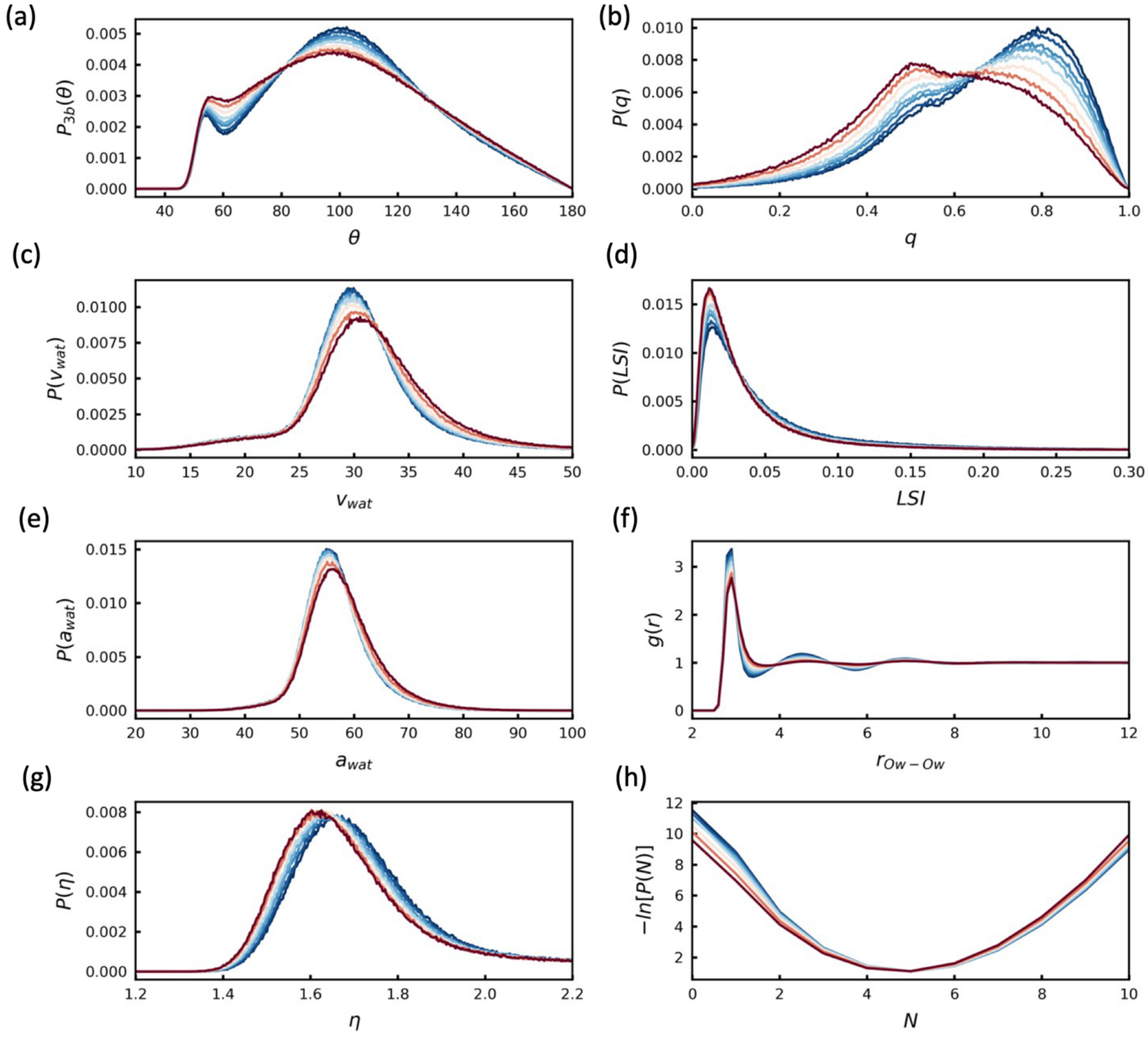


Figure 5.14. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2** for butanol-water at temperature of 273.15, 280.15, 291.15, 298.15, 309.15, 329.15, 350.15, and 373.15K. Here, red corresponds to higher temperature and blue to colder temperature. (a) The distributions of the tetrahedral order parameter, $P_{3b}(\theta)$, exhibit a systematic decrease in the icosahedral, $\theta = 64°$, and increase in the tetrahedral, $\theta = 109.5°$, regions with decreasing temperature. (b) The tetrahedral order parameter distribution, $P(q)$, exhibit peaks at $q \approx 0.4$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.8$ upon cooling. (c) The distributions of water specific volumes, $P(v_{wat})$, broaden and exhibit larger mean specific volumes, $\langle v_{wat} \rangle$, as the temperature of the mixture increases. (d) The distributions of LSI, $P(LSI)$, broaden and exhibit larger means, $\langle LSI \rangle$, as the temperature of the mixture decreases. (e) The distributions of water specific areas, $P(a_{wat})$, broaden and exhibit larger mean specific volumes, $\langle a_{wat} \rangle$, as the temperature of the mixture increases. (f) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with decreasing temperature as demonstrated by the increasing peak and trough amplitudes. (g) The distributions of water asphericity, $P(\eta)$, exhibit smaller (more spherical) mean asphericity, $\langle \eta \rangle$, as the temperature of the mixture increases. (h) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the systematic decrease in the affinity of an ideal hydrophobic molecule for butanol-water upon cooling. Simultaneously, decreased mixture temperature shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

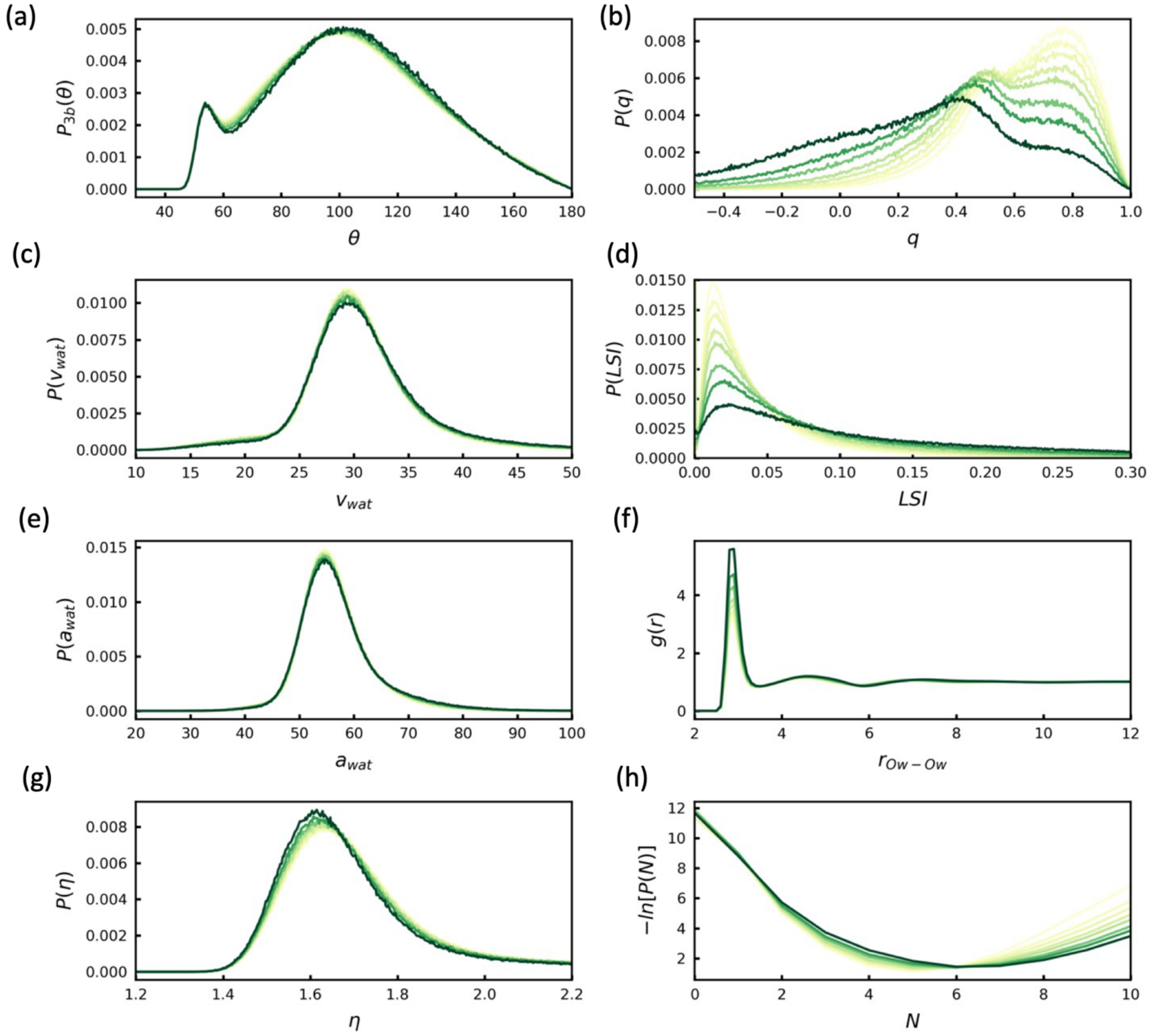


Figure 5.15. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2** for glycerol-water at X_{glyc} of 0.01, 0.033, 0.05, 0.075, 0.10, 0.15, 0.20, and 0.30. Here, green corresponds to higher concentration and yellow to lower concentration. (a) The distributions of the tetrahedral order parameter, $P_{3b}(\theta)$, exhibit a systematic decrease in the icosahedral, $\theta = 64°$, and increase in the tetrahedral, $\theta = 109.5°$, regions with decreasing temperature. (b) The tetrahedral order parameter distribution, $P(q)$, exhibit peaks at $q \approx 0.4$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.4$ upon increasing glycerol concentration. (c) The distributions of water specific volumes, $P(v_{wat})$, exhibit slightly larger mean specific volumes, $\langle v_{wat} \rangle$, as the concentration of the mixture increases. (d) The distributions of LSI, $P(LSI)$, broaden and exhibit larger means, $\langle LSI \rangle$, as the concentration of the mixture increases. (e) The distributions of water specific areas, $P(a_{wat})$, exhibit slightly larger mean specific volumes, $\langle a_{wat} \rangle$, as the concentration of the mixture increases. (f) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with increasing concentration as demonstrated by the increasing peak and trough amplitudes. (g) The distributions of water asphericity, $P(\eta)$, exhibit smaller (more spherical) mean asphericity, $\langle \eta \rangle$, as the concentration of the mixture increases. (h) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the similarity of the affinity of an ideal hydrophobic molecule for glycerol-water upon crowding. Simultaneously, decreased mixture concentration shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

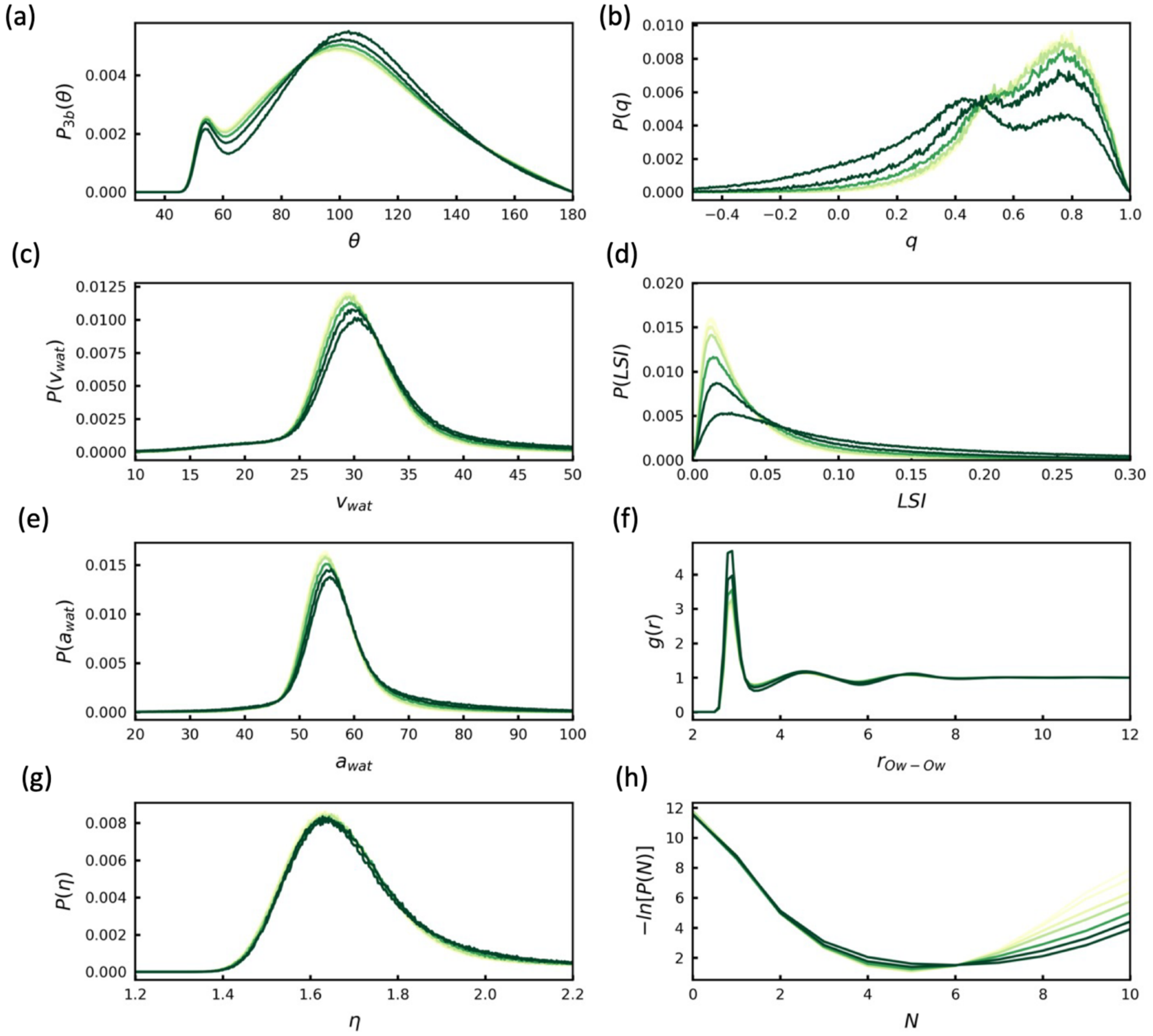


Figure 5.16. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2** for PEO-water at W_{PEO} of 0.005, 0.015, 0.05, 0.10, 0.20, 0.33, and 0.50. Here, green corresponds to higher concentration and yellow to lower concentration. (a) The distributions of the tetrahedral order parameter, $P_{3b}(\theta)$, exhibit a systematic decrease in the icosahedral, $\theta = 64°$, and increase in the tetrahedral, $\theta = 109.5°$, regions with decreasing temperature. (b) The tetrahedral order parameter distribution, $P(q)$, exhibit peaks at $q \approx 0.4$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.4$ upon increasing glycerol concentration. (c) The distributions of water specific volumes, $P(v_{wat})$, exhibit larger mean specific volumes, $\langle v_{wat} \rangle$, as the concentration of the mixture increases. (d) The distributions of LSI, $P(LSI)$, broaden and exhibit larger means, $\langle LSI \rangle$, as the concentration of the mixture increases. (e) The distributions of water specific areas, $P(a_{wat})$, exhibit larger mean specific volumes, $\langle a_{wat} \rangle$, as the concentration of the mixture increases. (f) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with increasing concentration as demonstrated by the increasing peak and trough amplitudes. (g) The distributions of water asphericity, $P(\eta)$, exhibit smaller (more spherical) mean asphericity, $\langle \eta \rangle$, as the concentration of the mixture increases. (h) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the similarity of the affinity of an ideal hydrophobic molecule for PEO-water upon crowding. Simultaneously, decreased mixture concentration shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

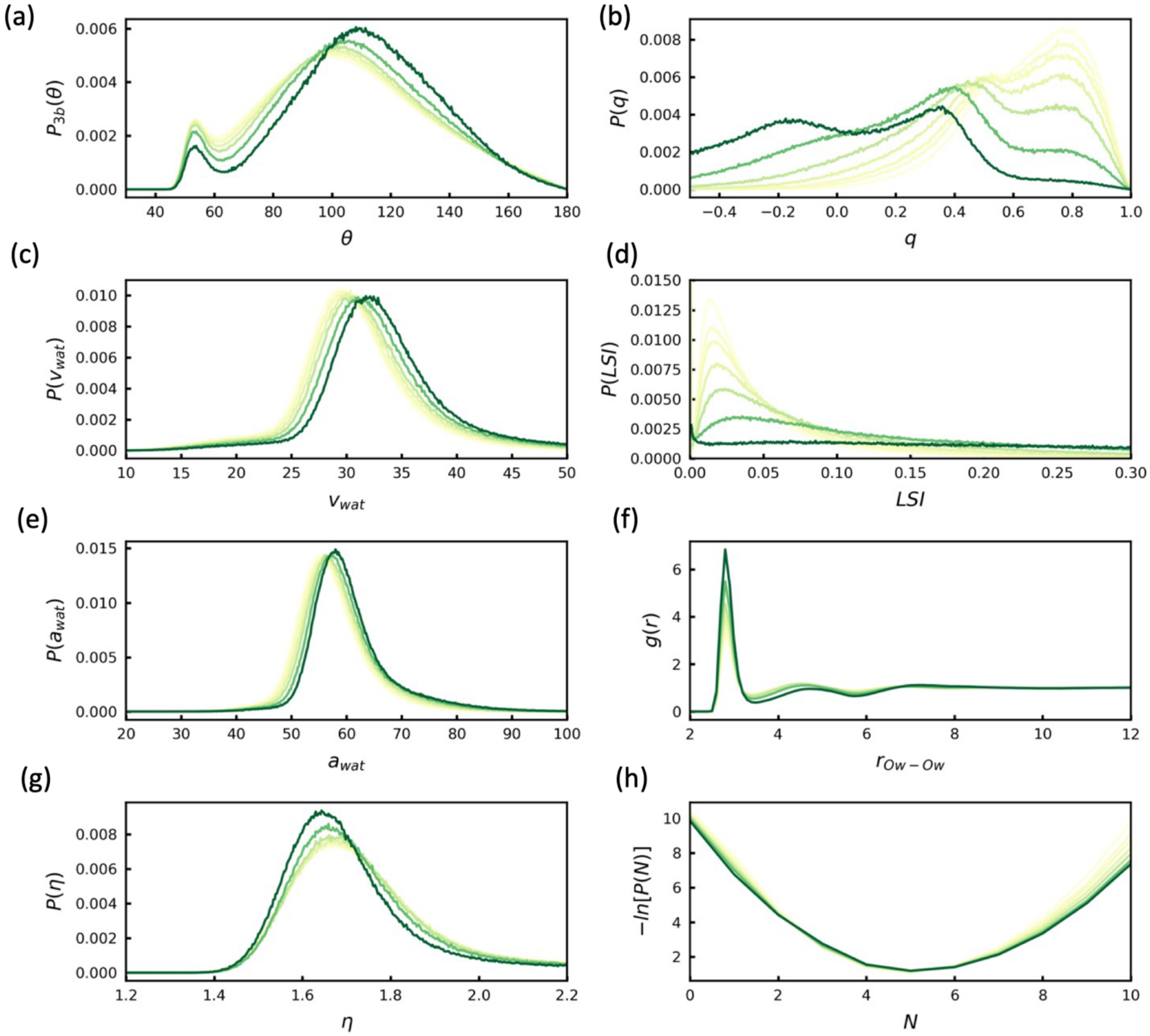


Figure 5.17. We depict probability distributions for a subset of structural metrics discussed in **Section 5.2** for DMSO-water at X_{DMSO} of 0.01, 0.04, 0.06, 0.10, 0.16, 0.27, and 0.45. Here, green corresponds to higher concentration and yellow to lower concentration. (a) The distributions of the tetrahedral order parameter, $P_{3b}(\theta)$, exhibit a systematic decrease in the icosahedral, $\theta = 64°$, and increase in the tetrahedral, $\theta = 109.5°$, regions with decreasing temperature. (b) The tetrahedral order parameter distribution, $P(q)$, exhibit peaks at $q \approx 0.4$ (less tetrahedral) and $q \approx 0.8$ (more tetrahedral). The distribution systematically shifts towards $q = 0.4$ upon increasing glycerol concentration. (c) The distributions of water specific volumes, $P(v_{wat})$, exhibit larger mean specific volumes, $\langle v_{wat} \rangle$, as the concentration of the mixture increases. (d) The distributions of LSI, $P(LSI)$, broaden and exhibit larger means, $\langle LSI \rangle$, as the concentration of the mixture increases. We note that $\langle LSI \rangle$ at the highest concentration $X_{DMSO} = 0.45$ approaches 0 due to the increasing space between water molecules and their nearest neighbors. (e) The distributions of water specific areas, $P(a_{wat})$, exhibit larger mean specific volumes, $\langle a_{wat} \rangle$, as the concentration of the mixture increases. (f) 2D water-water radial distribution function (RDFs), $g(r)$, show enhanced translational order in water with increasing concentration as demonstrated by the increasing peak and trough amplitudes. (g) The distributions of water asphericity, $P(\eta)$, exhibit smaller (more spherical) mean asphericity, $\langle \eta \rangle$, as the concentration of the mixture increases. (h) The negative log-probability of hard sphere insertion distribution, $-\ln[P(N)]$, demonstrates the similarity of the affinity of an ideal hydrophobic molecule for DMSO-water upon crowding. Simultaneously, decreased mixture concentration shifts the distribution towards lower mean overlap numbers, $\langle N \rangle$.

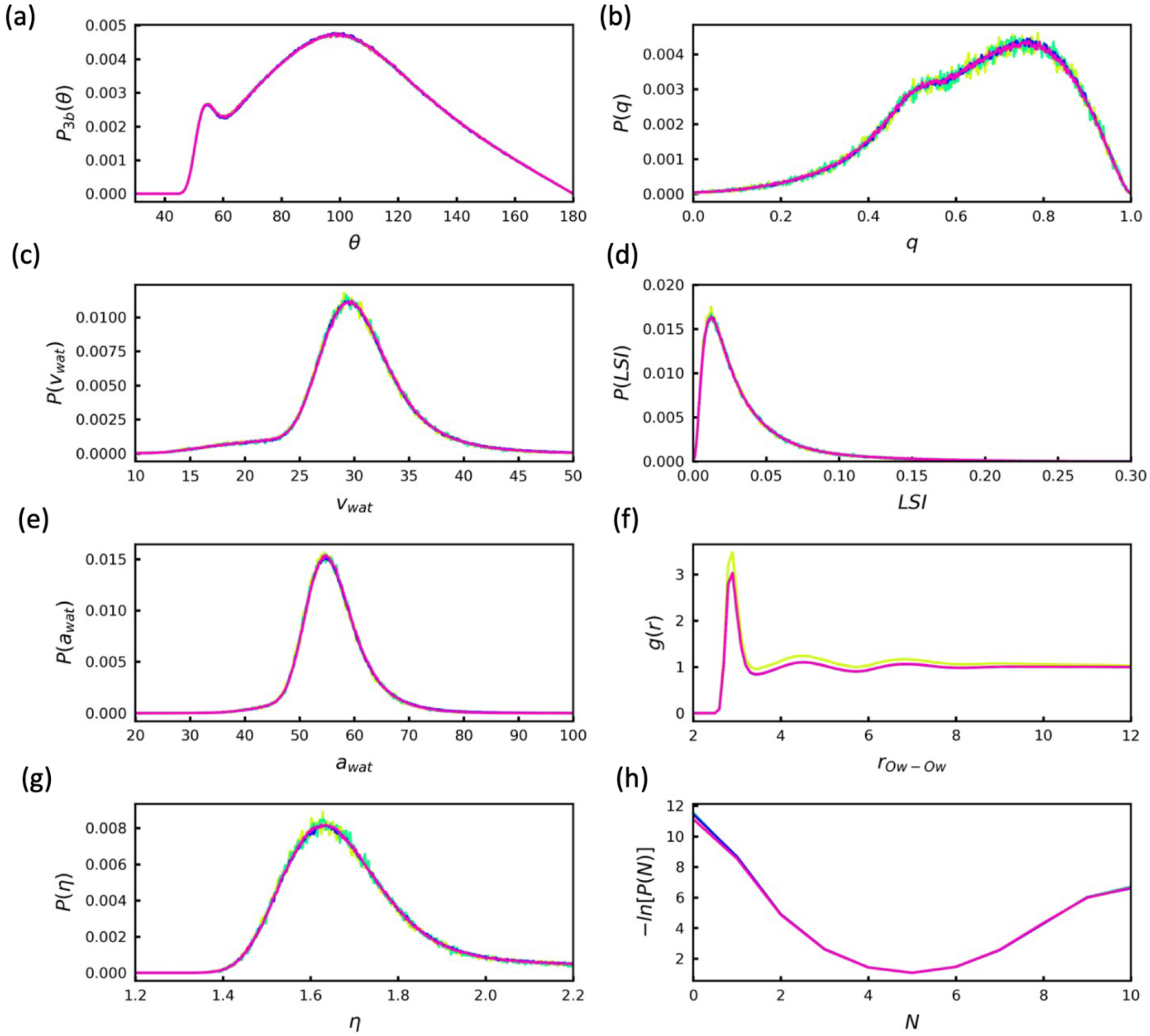


Figure 5.18. We depict probability distributions for a subset of structural metrics discussed in **Section 2** for the variable-hydropathy peptoids shown in **Figure 5.2**. Here, each color corresponds to a different peptoids chemistry at infinite dilution in water. For the peptoid systems, the differences between distributions are not visible. Here, we show (a) $P_{3b}(\theta)$, (b) $P(q)$, (c) $P(v_{wat})$, (d) $P(LSI)$, (e) $P(a_{wat})$, (f) $g(r)$, (g) $P(\eta)$, and (h) $-\ln[P(N)]$.

5.5.2 Structure-dynamics and structure-thermodynamics correlations

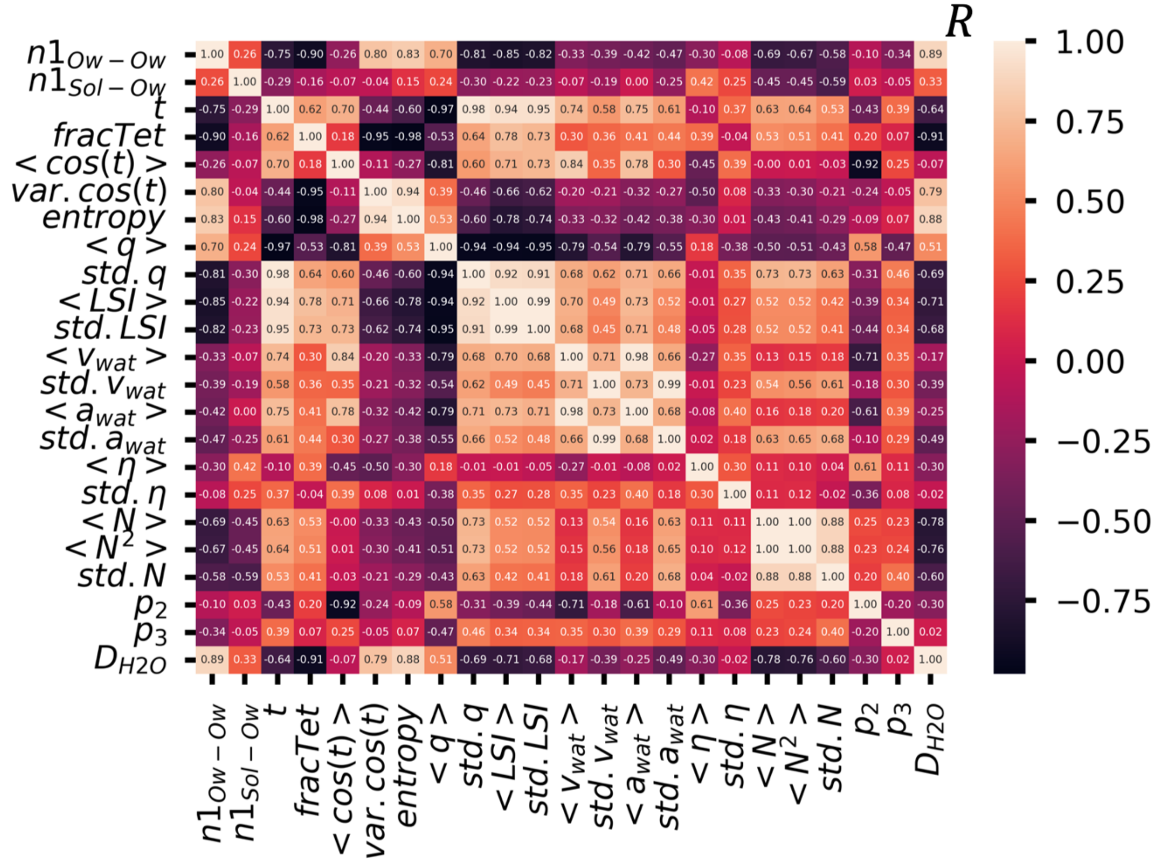
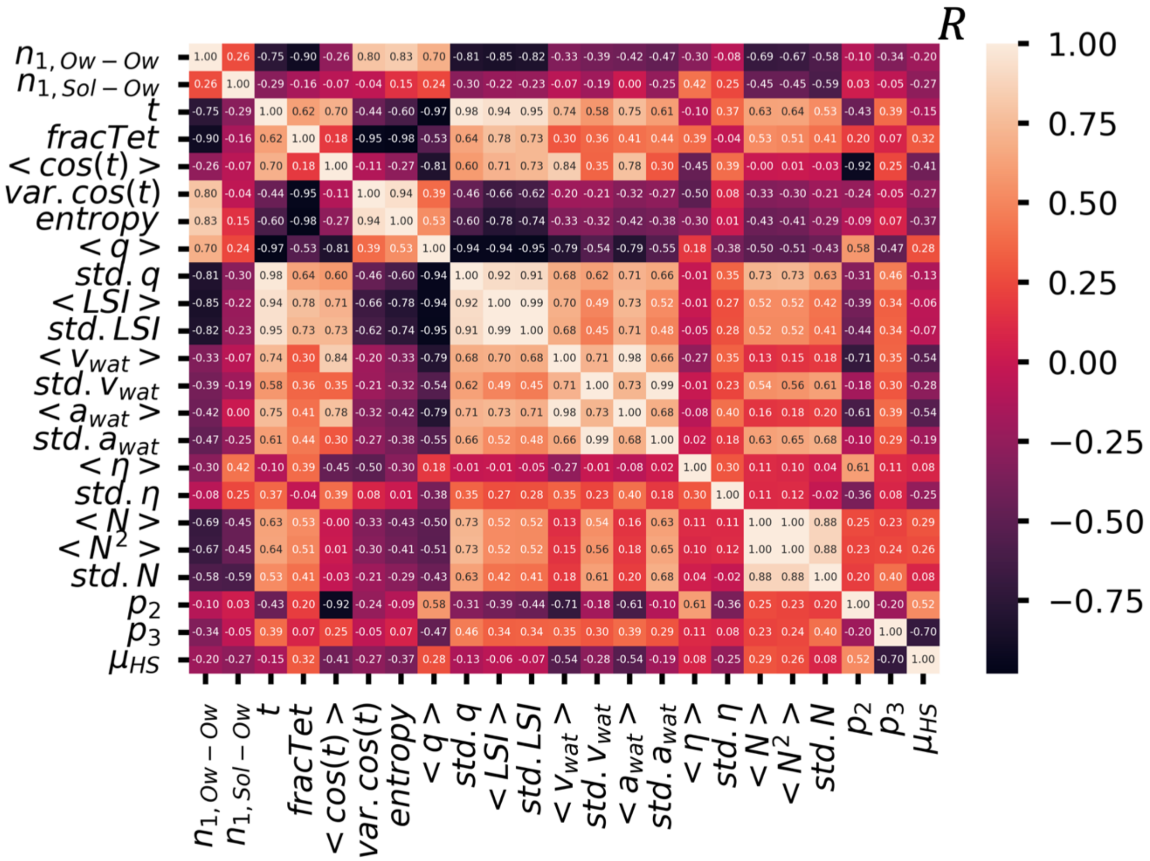


Figure 5.19. Pearson correlation matrix of structural metrics and water self-diffusivity (last column/row).



	$n_{1,Ow-Ow}$	$n_{1,Sol-Ow}$	t	$fracTet$	$<cos(t)>$	$var.cos(t)$	$entropy$	$<q>$	$std.q$	$<LSI>$	$std.LSI$	$<v_{wat}>$	$std.v_{wat}$	$<a_{wat}>$	$std.a_{wat}$	$<\eta>$	$std.\eta$	$<N>$	$<N^2>$	$std.N$	p_2	p_3	μ_{HS}
$n_{1,Ow-Ow}$	1.00	0.26	-0.75	-0.90	-0.26	0.80	0.83	0.70	-0.81	-0.85	-0.82	-0.33	-0.39	-0.42	-0.47	-0.30	-0.08	-0.69	-0.67	-0.58	-0.10	-0.34	-0.20
$n_{1,Sol-Ow}$	0.26	1.00	-0.29	-0.16	-0.07	-0.04	0.15	0.24	-0.30	-0.22	-0.23	-0.07	-0.19	0.00	-0.25	0.42	0.25	-0.45	-0.45	-0.59	0.03	-0.05	-0.27
t	-0.75	-0.29	1.00	0.62	0.70	-0.44	-0.60	-0.97	0.98	0.94	0.95	0.74	0.58	0.75	0.61	-0.10	0.37	0.63	0.64	0.53	-0.43	0.39	-0.15
$fracTet$	-0.90	-0.16	0.62	1.00	0.18	-0.95	-0.98	-0.53	0.64	0.78	0.73	0.30	0.36	0.41	0.44	0.39	-0.04	0.53	0.51	0.41	0.20	0.07	0.32
$<cos(t)>$	-0.26	-0.07	0.70	0.18	1.00	-0.11	-0.27	-0.81	0.60	0.71	0.73	0.84	0.35	0.78	0.30	-0.45	0.39	-0.00	0.01	-0.03	-0.92	0.25	-0.41
$var.cos(t)$	0.80	-0.04	-0.44	-0.95	-0.11	1.00	0.94	0.39	-0.46	-0.66	-0.62	-0.20	-0.21	-0.32	-0.27	-0.50	0.08	-0.33	-0.30	-0.21	-0.24	-0.05	-0.27
$entropy$	0.83	0.15	-0.60	-0.98	-0.27	0.94	1.00	0.53	-0.60	-0.78	-0.74	-0.33	-0.32	-0.42	-0.38	-0.30	0.01	-0.43	-0.41	-0.29	-0.09	0.07	-0.37
$<q>$	0.70	0.24	-0.97	-0.53	-0.81	0.39	0.53	1.00	-0.94	-0.94	-0.95	-0.79	-0.54	-0.79	-0.55	0.18	-0.38	-0.50	-0.51	-0.43	0.58	-0.47	0.28
$std.q$	-0.81	-0.30	0.98	0.64	0.60	-0.46	-0.60	-0.94	1.00	0.92	0.91	0.68	0.62	0.71	0.66	-0.01	0.35	0.73	0.73	0.63	-0.31	0.46	-0.13
$<LSI>$	-0.85	-0.22	0.94	0.78	0.71	-0.66	-0.78	-0.94	0.92	1.00	0.99	0.70	0.49	0.73	0.52	-0.01	0.27	0.52	0.52	0.42	-0.39	0.34	-0.06
$std.LSI$	-0.82	-0.23	0.95	0.73	0.73	-0.62	-0.74	-0.95	0.91	0.99	1.00	0.68	0.45	0.71	0.48	-0.05	0.28	0.52	0.52	0.41	-0.44	0.34	-0.07
$<v_{wat}>$	-0.33	-0.07	0.74	0.30	0.84	-0.20	-0.33	-0.79	0.68	0.70	0.68	1.00	0.71	0.98	0.66	-0.27	0.35	0.13	0.15	0.18	-0.71	0.35	-0.54
$std.v_{wat}$	-0.39	-0.19	0.58	0.36	0.35	-0.21	-0.32	-0.54	0.62	0.49	0.45	0.71	1.00	0.73	0.99	-0.01	0.23	0.54	0.56	0.61	-0.18	0.30	-0.28
$<a_{wat}>$	-0.42	0.00	0.75	0.41	0.78	-0.32	-0.42	-0.79	0.71	0.73	0.71	0.98	0.73	1.00	0.68	-0.08	0.40	0.16	0.18	0.20	-0.61	0.39	-0.54
$std.a_{wat}$	-0.47	-0.25	0.61	0.44	0.30	-0.27	-0.38	-0.55	0.66	0.52	0.48	0.66	0.99	0.68	1.00	0.02	0.18	0.63	0.65	0.68	-0.10	0.29	-0.19
$<\eta>$	-0.30	0.42	-0.10	0.39	-0.45	-0.50	-0.30	0.18	-0.01	-0.01	-0.05	-0.27	-0.01	-0.08	0.02	1.00	0.30	0.11	0.10	0.04	0.61	0.11	0.08
$std.\eta$	-0.08	0.25	0.37	-0.04	0.39	0.08	0.01	-0.38	0.35	0.27	0.28	0.35	0.23	0.40	0.18	0.30	1.00	0.11	0.12	-0.02	-0.36	0.08	-0.25
$<N>$	-0.69	-0.45	0.63	0.53	-0.00	-0.33	-0.43	-0.50	0.73	0.52	0.52	0.13	0.54	0.16	0.63	0.11	0.11	1.00	1.00	0.88	0.25	0.23	0.29
$<N^2>$	-0.67	-0.45	0.64	0.51	0.01	-0.30	-0.41	-0.51	0.73	0.52	0.52	0.15	0.56	0.18	0.65	0.10	0.12	1.00	1.00	0.88	0.23	0.24	0.26
$std.N$	-0.58	-0.59	0.53	0.41	-0.03	-0.21	-0.29	-0.43	0.63	0.42	0.41	0.18	0.61	0.20	0.68	0.04	-0.02	0.88	0.88	1.00	0.20	0.40	0.08
p_2	-0.10	0.03	-0.43	0.20	-0.92	-0.24	-0.09	0.58	-0.31	-0.39	-0.44	-0.71	-0.18	-0.61	-0.10	0.61	-0.36	0.25	0.23	0.20	1.00	-0.20	0.52
p_3	-0.34	-0.05	0.39	0.07	0.25	-0.05	0.07	-0.47	0.46	0.34	0.34	0.35	0.30	0.39	0.29	0.11	0.08	0.23	0.24	0.40	-0.20	1.00	-0.70
μ_{HS}	-0.20	-0.27	-0.15	0.32	-0.41	-0.27	-0.37	0.28	-0.13	-0.06	-0.07	-0.54	-0.28	-0.54	-0.19	0.08	-0.25	0.29	0.26	0.08	0.52	-0.70	1.00

Figure 5.20. Pearson correlation matrix of structural metrics and excess hard-sphere chemical potential (last column/row).

5.5.3 Further details on structure-dynamics correlation

In the main text, we depict correlations between a limited subset of structural metrics and water self-diffusivity [**Figure 5.6**]. We note the weaker correlations between translational order parameters like t, $\langle q \rangle$, and $\langle LSI \rangle$ compared to those shown in **Figure 5.6**. This derives primarily from the opposite diffusivity-structure relationships for the alcohol-water mixtures versus the remaining solutes as depicted in **Figure 5.21**. For example, the ensemble average tetrahedral order parameter, $\langle q \rangle$, monotonically increases with decreasing water self-diffusivity, D_{H_2O}, in alcohol-water while $\langle q \rangle$ monotonically decreases with decreasing D_{H_2O} in the variable composition simulations.

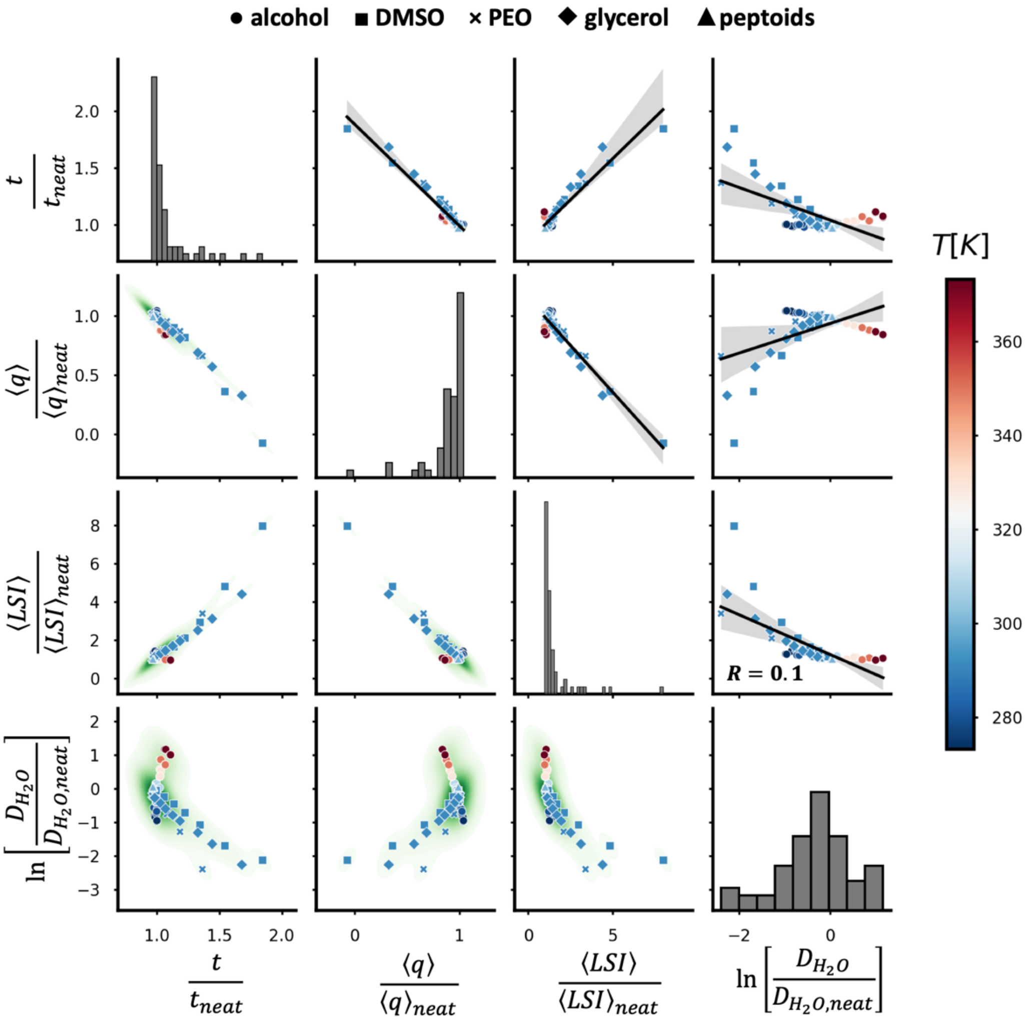


Figure 5.21: We depict the relatively poor correlation between water self-diffusivity and several structural metrics ($\langle q \rangle$, t, and $\langle LSI \rangle$). We note the near perfect correlation between the local structure index $\langle LSI \rangle$, translational order parameter, t, and tetrahedral order parameter, $\langle q \rangle$. However, we observe opposite D_{H_2O}-structure relationships for the primary alcohol mixtures compared to the other systems. Here, the symbols are assigned as in the text.

5.5.4 Model performance with an increasing number of features

In **Figure 5.22**, we depict the effect of increasing the number of independent structural metrics, n, on the performance of MLR models to predict D_{H_2O}. Though the training set performance (via R^2 and $RMSE$) increases monotonically—but with diminishing returns for more than three features—with increasing feature inclusion, test set performance slightly decreases for $n > 6$.

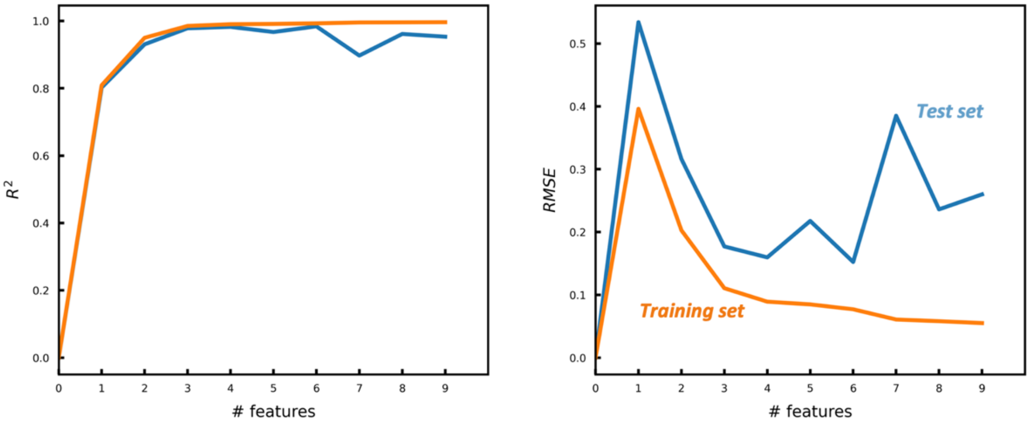


Figure 5.22: Training and test set performance of MLR models to predict D_{H_2O} with increasing numbers of independent structural variables. We observe diminishing improvement of MLR predictions for greater than three features. In fact, test set performance worsens when we include more than six features. Here, the test set is composed of a randomly-selected ~20% share of the dataset (12 simulations) and the training set is composed of the remaining data (47 simulations). We also perform 4-fold cross-validation on the training set to choose model parameters and evaluate the training set performance.

Similarly, we show the effect of n on the performance of MLR models to predict μ_{HS}^{ex} [**Figure 5.23**]. Unlike predicting D_{H_2O}, the performance of the chemical potential prediction models (via R^2 and $RMSE$) increases monotonically for both test and training sets. However, model performance similarly plateaus for $n > 3$.

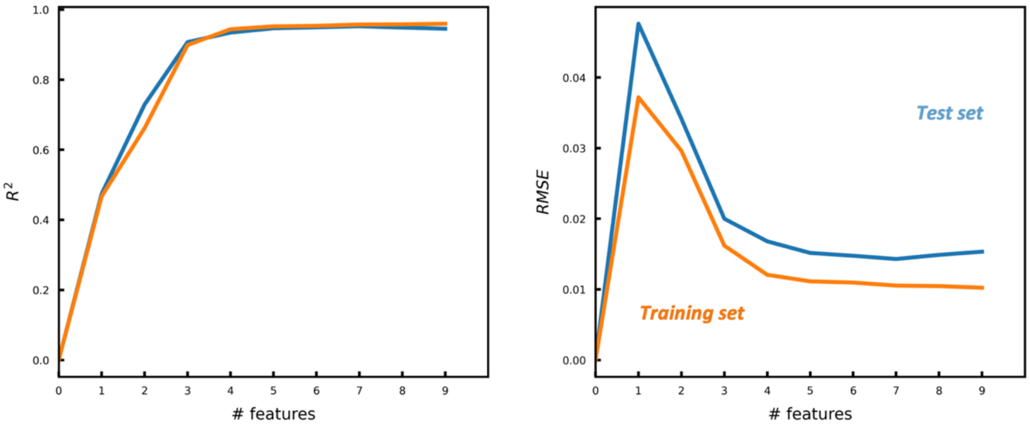


Figure 5.23: Training and test set performance of MLR models to predict μ_{HS}^{ex} with increasing numbers of independent structural variables. We observe diminishing improvement of MLR predictions for greater than three features. Here, the test set is composed of a randomly-selected ~20% share of the dataset (12 simulations) and the training set is composed of the remaining data (47 simulations). We also perform 4-fold cross-validation on the training set to choose model parameters and evaluate the training set performance.

5.5.5 LASSO REGRESSION

As described in **Section 5.2.4** of the main text, we develop a LASSO regression model on a randomly selected subset of data that is 80% of the size of the entire dataset. We implement LASSO feature selection using the tools provided scikit-learn(96). To select the learning rate, α, of the LASSO model, we apply a hyperparameter tuning algorithm built into scikit-learn with 4-fold cross-validation.

The six features with non-zero LASSO importance are shown in **Figure 5.24.** Here, the top three features selected by forward sequential feature selection (fSFS)—namely, p_{tet}, $\langle N \rangle$, and p_3—are also predicted to be important by LASSO. However, we note that LASSO regression predicts that S_{3b} and $\langle N \rangle$ explain the training set D_{H_2O} trend better than p_{tet} in contrast to the fSFS results presented in the main body. While a linear regression model trained only on S_{3b} yields similar performance to one trained solely on p_{tet}, a model trained on $\langle N \rangle$ alone performs much more poorly (see the $\langle N \rangle$-D_{H_2O} relationship in **Figure 5.6** with $R = -0.78$). This apparent inconsistency is in part explained by our manual selection of p_{tet} as the most important feature. However, this does not necessarily explicate the higher LASSO importance of $\langle N \rangle$ compared to p_{tet}.

To fully explain this discrepancy, we perform the LASSO feature selection once more using a different randomly selected training set. In **Figure 5.25**, we demonstrate that our LASSO procedure now yields 8 important features that still include p_{tet}, $\langle N \rangle$, and p_3. However, p_{tet} is not the most important structural metric for predicting D_{H_2O}. Hence, LASSO feature selection strongly depends on the training set, and Lasso regression alone should not be used to rank feature importance.

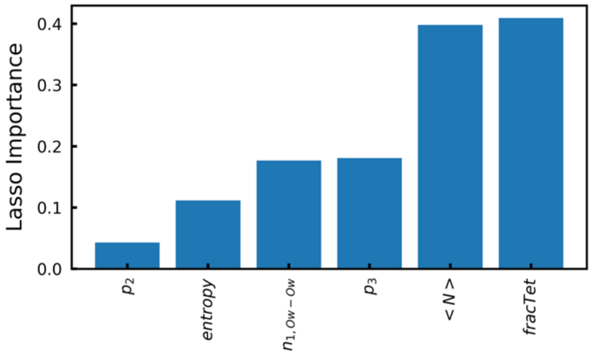


Figure 5.24. First test to determine rank diffusivity-predictive features by importance (larger values correspond to higher predictive capacity) using LASSO regression. This set of features includes the top three features from sequential feature selection (**Figure 5.7**). All features not shown have importance values of zero.

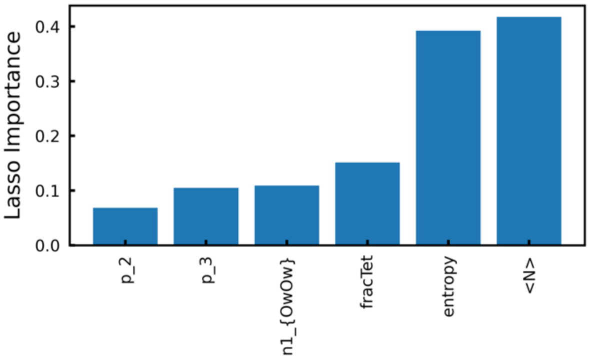


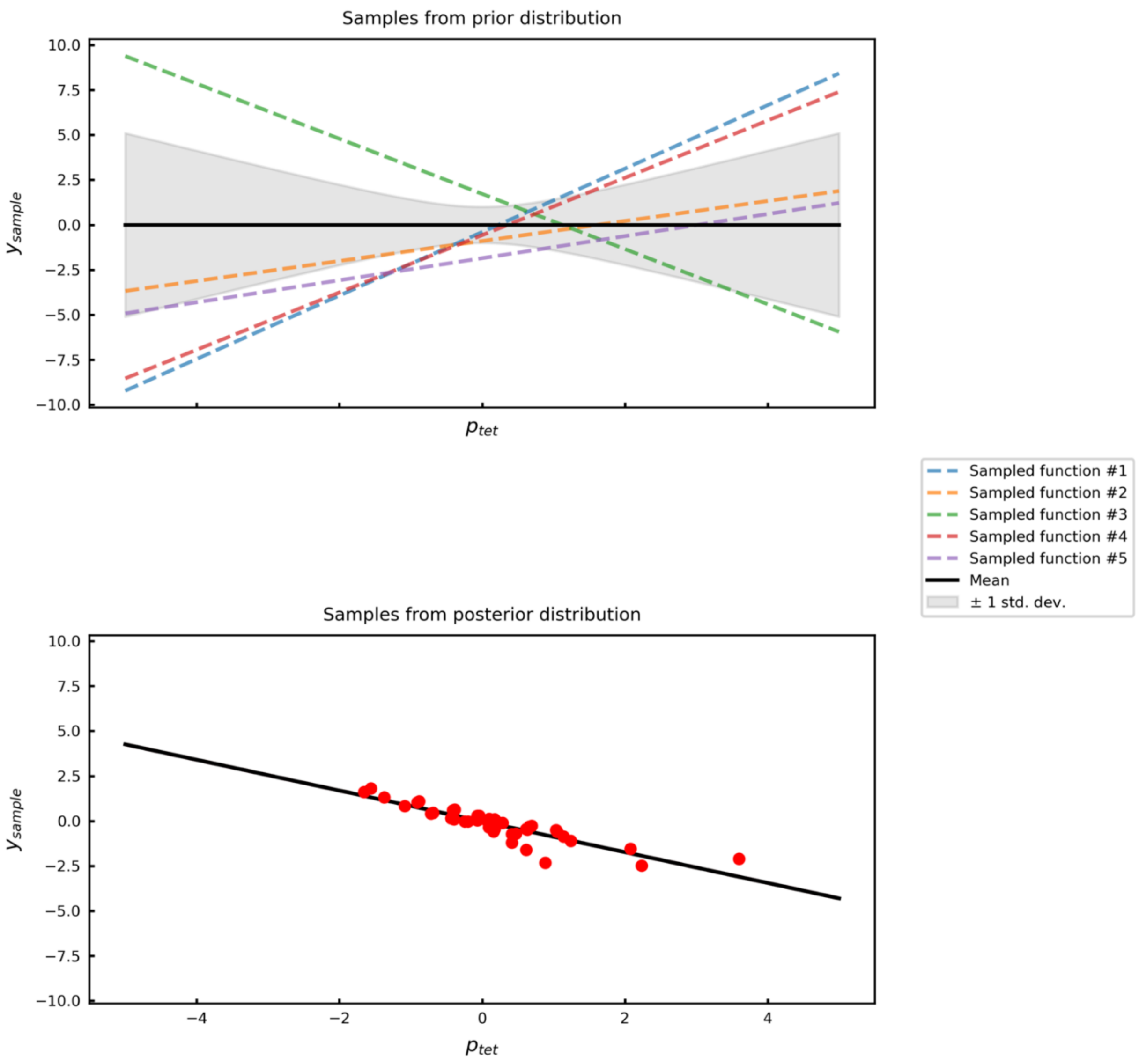
Figure 5.25. Second test to determine rank diffusivity-predictive features by importance (larger values correspond to higher predictive capacity) using LASSO regression. This set of features includes the top three features from sequential feature selection (**Figure 5.7**). This test case contains 6 fewer features than test case #1 (**Figure 5.24**). All features not shown have importance values of zero.

5.5.6 Bayesian Methods

Structure Inference Methodology. To illustrate the procedure for predicting structure from dynamics described in **Section 5.3.4**, we present a simplified example for the case of a simple linear regression model

$$y = \ln\left[\frac{D_{H_2O}}{D_{H_2O,neat}}\right] = \beta\left[\frac{p_{tet}}{p_{tet,neat}}\right] + \beta_0$$

where β and β_0 are the slope and intercept, respectively. We visualize the consequences of the standard normal prior distributions $p(\beta)$ and $p(\beta_0)$ by drawing five sample linear regression functions (**Figure 5.26**). In the absence of training data, the five samples are completely random and produce a uniform $y = 0$ for all $\frac{p_{tet}}{p_{tet,neat}}$ values. Upon including the training dataset, we draw an additional five fitting functions from the posterior distribution, $p(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{y})$. These new sample functions are indistinguishable.


Samples from prior distribution

Samples from posterior distribution

Figure 5.26. We depict five sample fitting functions from the prior distribution, $p(\boldsymbol{\beta})$ (top) and the posterior distribution, $p(\boldsymbol{\beta}|\boldsymbol{X}, \boldsymbol{y})$ (bottom). In the bottom figure, the red points are the training dataset. Here, the shaded bands corresponding to standard deviation of the sample fits.

Testing the Effect of Priors on Structure Prediction. To extract meaningful predictions from Bayesian inference methods, we must carefully consider the form of the prior. In the main text, we propose standard normal distributions (zero-mean and unit variance) for new values of the independent variables (e.g., $p(X_i^{new}) = \mathcal{N}(0, \Sigma)$ with $\Sigma = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$). In **Figure 5.10**, we show qualitative agreement between direct simulation estimates and predictions using standard normal priors on p_{tet} and $\langle N \rangle$. These predictions are promising, but how will changing the prior affect these predictions?

We test the robustness of our predictions by instead considering uniform distribution priors spanning $X_i^{new} = -3$ to 3. In other words, we propose that all X_i^{new} values (p_{tet} or $\langle N \rangle$) lie between ±3 standard deviations of the mean value (here, $\langle X_i^{new} \rangle = 0$). We observe the impact of this new assumption inf **Figure 5.27**. Notably, the behavior of the predictions is qualitatively similar to the results shown in **Figure 5.10**. However, we observe larger variance in the predictions of $\langle N \rangle$ (**Figure 5.27(f)**).

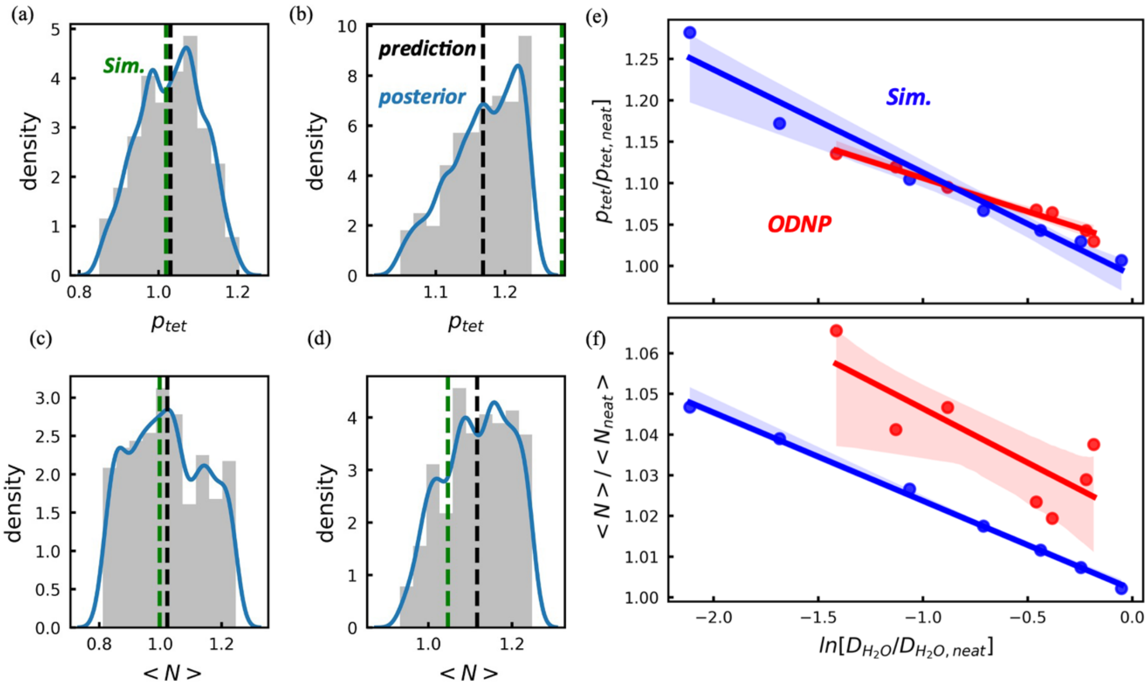


Figure 5.27. In panels (a)-(d), we depict the posterior predictive (blue) distribution generated by the Bayesian regression model trained on p_{tet}, $\langle N \rangle$, and p_3 to predict D_{H_2O}. The p_{tet} predictive distributions for (a) methanol-water at 298.15K with methanol weight fraction of 0.10 and (b) DMSO-water at $X_{DMSO} = 0.45$ and 291.15K yield $p_{tet}/p_{tet,neat}$ (black vertical line) predictions. The $\langle N \rangle$ predictive distributions for (c) methanol-water at 298.15K with methanol weight fraction of 0.10 and (d) DMSO-water at $X_{DMSO} = 0.45$ and 291.15K yield $\langle N \rangle / \langle N \rangle_{neat}$ (black vertical line) predictions. Simulation means are depicted with green vertical lines. Model predictions are closer to the simulation measurement for methanol-water at 298.15K compared to the DMSO-water case. Predictions of (e) $\frac{p_{tet}}{p_{tet,neat}}$ and (f) $\frac{\langle N \rangle}{\langle N \rangle_{neat}}$ generated from ODNP-derived D_{H_2O} values for a range of DMSO-water mixtures from $X_{DMSO} = 0$ to 0.5 (red) are compared to direct simulation estimates (blue). In

panels (e) and (f), the lines represent linear least squares fits between the log-diffusivity and a given structure metric. Here, the shaded bands corresponding to 95% confidence intervals on the fit.

ASSOCIATED CONTENT

Supporting Information.

Additional structural details not provided in the main text; full Pearson correlation matrices between structural, dynamic, and thermodynamic properties; notes on sequential feature selection model performance; an alternative feature selection methodology; further details on the Bayesian methodology presented in the main text. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*M. Scott Shell. Email: shell@ucsb.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

This work was supported as part of the Center for Materials for Water and Energy Systems (M-WET), an Energy Frontier Research Center funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences under Award #DE-SC0019272. Use was made of computational facilities purchased with funds from the National Science Foundation (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara. S.J. acknowledges support by the National Science Foundation Graduate Research Fellowship (DGE 1650114). D.R.M. acknowledges support from a "Future Leaders of Advanced Materials" internship sponsored by the MRSEC Program of the National Science Foundation under Award No. DMR 1720256.

References

1. Edward JT. 1970. Molecular volumes and the Stokes-Einstein equation. *J. Chem. Educ.* 47(4):261
2. Yasuda H, Lamaze CE, Ikenberry LD. 1968. Permeability of solutes through hydrated polymer membranes. Part I. Diffusion of sodium chloride. *Die Makromolekulare Chemie*. 118(1):19–35
3. Yasuda H, Peterlin A, Colton CK, Smith KA, Merrill EW. 1969. Permeability of solutes through hydrated polymer membranes. Part III. Theoretical background for the selectivity of dialysis membranes. *Die Makromolekulare Chemie*. 126(1):177–86
4. Yasuda H, Lamaze CE, Peterlin A. 1971. Diffusive and hydraulic permeabilities of water in water-swollen polymer membranes. *Journal of Polymer Science Part A-2: Polymer Physics*. 9(6):1117–31
5. Mackie JS, Meares P. 1955. The diffusion of electrolytes in a cation-exchange resin membrane I. Theoretical. *Proc. R. Soc. Lond. A*. 232(1191):498–509

6. Monroe J, Barry M, DeStefano A, Aydogan Gokturk P, Jiao S, et al. 2020. Water Structure and Properties at Hydrophilic and Hydrophobic Surfaces. *Annual Review of Chemical and Biomolecular Engineering*. 11(1):523–57

7. Sharma M, Yashonath S. 2006. Breakdown of the Stokes−Einstein Relationship: Role of Interactions in the Size Dependence of Self-Diffusivity. *J. Phys. Chem. B*. 110(34):17207–11

8. Dubey V, Erimban S, Indra S, Daschakraborty S. 2019. Understanding the Origin of the Breakdown of the Stokes–Einstein Relation in Supercooled Water at Different Temperature–Pressure Conditions. *The Journal of Physical Chemistry B*. 123(47):10089–99

9. Errington JR, Debenedetti PG. 2001. Relationship between structural order and the anomalies of liquid water. *Nature*. 409(6818):318–21

10. Jiao S, Rivera Mirabal DM, DeStefano AJ, Segalman RA, Han S, Shell MS. 2022. Sequence Modulates Polypeptoid Hydration Water Structure and Dynamics. *Biomacromolecules*. 23(4):1745–56

11. Monroe JI, Shell MS. 2018. Computational discovery of chemically patterned surfaces that effect unique hydration water dynamics. *Proceedings of the National Academy of Sciences*. 115(32):8093–98

12. Stirnemann G, Castrillón SR-V, Hynes JT, Rossky PJ, Debenedetti PG, Laage D. 2011. Non-monotonic dependence of water reorientation dynamics on surface hydrophilicity: competing effects of the hydration structure and hydrogen-bond strength. *Physical Chemistry Chemical Physics*. 13(44):19911

13. Duboué-Dijon E, Laage D. 2015. Characterization of the Local Structure in Liquid Water by Various Order Parameters. *J. Phys. Chem. B*. 119(26):8406–18

14. Chandler D. 2005. Interfaces and the driving force of hydrophobic assembly. *Nature*. 437(7059):640–47

15. Lum K, Chandler D, Weeks JD. 1999. Hydrophobicity at Small and Large Length Scales. *The Journal of Physical Chemistry B*. 103(22):4570–77

16. Wirtz H, Schäfer S, Hoberg C, Havenith M. 2018. Differences in Hydration Structure Around Hydrophobic and Hydrophilic Model Peptides Probed by THz Spectroscopy. *J Infrared Milli Terahz Waves*. 39(9):816–27

17. Head-Gordon T, Sorenson JM, Pertsemlidis A, Glaeser RM. 1997. Differences in hydration structure near hydrophobic and hydrophilic amino acids. *Biophysical Journal*. 73(4):2106–15

18. Johnson ME, Malardier-Jugroot C, Head-Gordon T. 2010. Effects of co-solvents on peptide hydration water structure and dynamics. *Phys. Chem. Chem. Phys.* 12(2):393–405

19. Athawale MV, Goel G, Ghosh T, Truskett TM, Garde S. 2007. Effects of lengthscales and attractions on the collapse of hydrophobic polymers in water. *PNAS*. 104(3):733–38

20. Lynden-bell RM, Head-gordon T. 2006. Solvation in modified water models: towards understanding hydrophobic effects. *Molecular Physics*. 104(22–24):3593–3605

21. Romero-Vargas Castrillón S, Giovambattista N, Aksay IA, Debenedetti PG. 2009. Effect of Surface Polarity on the Structure and Dynamics of Water in Nanoscale Confinement. *The Journal of Physical Chemistry B*. 113(5):1438–46

22. Giovambattista N, Debenedetti PG, Rossky PJ. 2007. Hydration Behavior under Confinement by Nanoscale Surfaces with Patterned Hydrophobicity and Hydrophilicity. *J. Phys. Chem. C*. 111(3):1323–32

23. Patel AJ, Varilly P, Jamadagni SN, Acharya H, Garde S, Chandler D. 2011. Extended surfaces modulate hydrophobic interactions of neighboring solutes. *Proceedings of the National Academy of Sciences*. 108(43):17678–83
24. Mittal J, Hummer G. 2008. Static and dynamic correlations in water at hydrophobic interfaces. *Proceedings of the National Academy of Sciences*. 105(51):20130–35
25. Wu X, Lu W, Streacker LM, Ashbaugh HS, Ben-Amotz D. 2018. Temperature-Dependent Hydrophobic Crossover Length Scale and Water Tetrahedral Order. *The Journal of Physical Chemistry Letters*. 9(5):1012–17
26. Godawat R, Jamadagni SN, Garde S. 2009. Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proceedings of the National Academy of Sciences*. 106(36):15119–24
27. Jamadagni SN, Godawat R, Garde S. 2011. Hydrophobicity of Proteins and Interfaces: Insights from Density Fluctuations. *Annual Review of Chemical and Biomolecular Engineering*. 2(1):147–71
28. Patel AJ, Varilly P, Chandler D. 2010. Fluctuations of Water near Extended Hydrophobic and Hydrophilic Surfaces. *J. Phys. Chem. B*. 114(4):1632–37
29. Willard AP, Chandler D. 2008. The role of solvent fluctuations in hydrophobic assembly. *Journal of Physical Chemistry B*. 112(19):6187–92
30. Cui D, Ou S-C, Patel S. 2015. Protein Denaturants at Aqueous–Hydrophobic Interfaces: Self-Consistent Correlation between Induced Interfacial Fluctuations and Denaturant Stability at the Interface. *J. Phys. Chem. B*. 119(1):164–78
31. Rego NB, Xi E, Patel AJ. 2021. Identifying hydrophobic protein patches to inform protein interaction interfaces. *Proc Natl Acad Sci USA*. 118(6):e2018234118
32. Patel AJ, Garde S. 2014. Efficient Method To Characterize the Context-Dependent Hydrophobicity of Proteins. *The Journal of Physical Chemistry B*. 118(6):1564–73
33. Xi E, Patel AJ. 2016. The hydrophobic effect, and fluctuations: The long and the short of it. *Proc Natl Acad Sci USA*. 113(17):4549–51
34. Monroe JI, Jiao S, Davis RJ, Robinson Brown D, Katz LE, Shell MS. 2021. Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution. *Proc Natl Acad Sci USA*. 118(1):e2020205118
35. Rego NB, Ferguson AL, Patel AJ. 2022. Learning the relationship between nanoscale chemical patterning and hydrophobicity. *Proc. Natl. Acad. Sci. U.S.A.* 119(48):e2200018119
36. Dallin BC, Kelkar AS, Van Lehn RC. 2023. Structural features of interfacial water predict the hydrophobicity of chemically heterogeneous surfaces. *Chem. Sci.* 14(5):1308–19
37. Hoshina H, Iwasaki Y, Katahira E, Okamoto M, Otani C. 2018. Structure and dynamics of bound water in poly(ethylene-vinylalcohol) copolymers studied by terahertz spectroscopy. *Polymer*. 148:49–60
38. Pal S, Samanta N, Das Mahanta D, Mitra RK, Chattopadhyay A. 2018. Effect of Phospholipid Headgroup Charge on the Structure and Dynamics of Water at the Membrane Interface: A Terahertz Spectroscopic Study. *J. Phys. Chem. B*. 122(19):5066–74
39. Penkov N, Shvirst N, Yashin V, Fesenko E, Fesenko E. 2015. Terahertz Spectroscopy Applied for Investigation of Water Structure. *J. Phys. Chem. B*. 119(39):12664–70
40. Cross PC, Burnham J, Leighton PA. 1937. The Raman Spectrum and the Structure of Water. *J. Am. Chem. Soc.* 59(6):1134–47

41. Perera P, Wyche M, Loethen Y, Ben-Amotz D. 2008. Solute-Induced Perturbations of Solvent-Shell Molecules Observed Using Multivariate Raman Curve Resolution. *J. Am. Chem. Soc.* 130(14):4576–77

42. Sovago M, Kramer Campen R, Bakker HJ, Bonn M. 2009. Hydrogen bonding strength of interfacial water determined with surface sum-frequency generation. *Chemical Physics Letters*. 470(1):7–12

43. Waluyo I, Huang C, Nordlund D, Bergmann U, Weiss TM, et al. 2011. The structure of water in the hydration shell of cations from x-ray Raman and small angle x-ray scattering measurements. *The Journal of Chemical Physics*. 134(6):064513

44. Medders GR, Paesani F. 2015. Infrared and Raman Spectroscopy of Liquid Water through “First-Principles” Many-Body Molecular Dynamics. *J. Chem. Theory Comput.* 11(3):1145–54

45. Rankin BM, Ben-Amotz D, van der Post ST, Bakker HJ. 2015. Contacts Between Alcohols in Water Are Random Rather than Hydrophobic. *J. Phys. Chem. Lett.* 6(4):688–92

46. Persson RAX, Pattni V, Singh A, Kast SM, Heyden M. 2017. Signatures of Solvation Thermodynamics in Spectra of Intermolecular Vibrations. *J. Chem. Theory Comput.* 13(9):4467–81

47. Sun Y, Petersen PB. 2017. Solvation Shell Structure of Small Molecules and Proteins by IR-MCR Spectroscopy. *J. Phys. Chem. Lett.* 8(3):611–14

48. Kundu A, Verma PK, Cho M. 2019. Water Structure and Dynamics in the Stern Layer of Micelles: Femtosecond Mid-Infrared Pump-Probe Spectroscopy Study. *The Journal of Physical Chemistry B*. 123(25):5238–45

49. Ji N, Ostroverkhov V, Tian CS, Shen YR. 2008. Characterization of Vibrational Resonances of Water-Vapor Interfaces by Phase-Sensitive Sum-Frequency Spectroscopy. *Phys. Rev. Lett.* 100(9):096102

50. Ostroverkhov V, Waychunas GA, Shen YR. 2005. New Information on Water Interfacial Structure Revealed by Phase-Sensitive Surface Spectroscopy. *Phys. Rev. Lett.* 94(4):046102

51. Sovago M, Campen RK, Wurpel GWH, Müller M, Bakker HJ, Bonn M. 2008. Vibrational Response of Hydrogen-Bonded Interfacial Water is Dominated by Intramolecular Coupling. *Phys. Rev. Lett.* 100(17):173901

52. Myalitsin A, Urashima S, Nihonyanagi S, Yamaguchi S, Tahara T. 2016. Water Structure at the Buried Silica/Aqueous Interface Studied by Heterodyne-Detected Vibrational Sum-Frequency Generation. *J. Phys. Chem. C*. 120(17):9357–63

53. Cyran JD, Backus EHG, Nagata Y, Bonn M. 2018. Structure from Dynamics: Vibrational Dynamics of Interfacial Water as a Probe of Aqueous Heterogeneity. *J. Phys. Chem. B*. 122(14):3667–79

54. Zhang Y, de Aguiar HB, Hynes JT, Laage D. 2020. Water Structure, Dynamics, and Sum-Frequency Generation Spectra at Electrified Graphene Interfaces. *The Journal of Physical Chemistry Letters*. 11(3):624–31

55. Franck JM, Han S. 2019. Chapter Five - Overhauser Dynamic Nuclear Polarization for the Study of Hydration Dynamics, Explained. In *Methods in Enzymology*, ed AJ Wand. 615:131–75. Academic Press

56. Armstrong BD, Han S. 2009. Overhauser Dynamic Nuclear Polarization To Study Local Water Dynamics. *Journal of the American Chemical Society*. 131(13):4641–47

57. Franck JM, Ding Y, Stone K, Qin PZ, Han S. 2015. Anomalously Rapid Hydration Water Diffusion Dynamics Near DNA Surfaces. *J. Am. Chem. Soc.* 137(37):12013–23
58. Barnes R, Sun S, Fichou Y, Dahlquist FW, Heyden M, Han S. 2017. Spatially Heterogeneous Surface Water Diffusivity around Structured Protein Surfaces at Equilibrium. *J. Am. Chem. Soc.* 139(49):17890–901
59. Ortony JH, Qiao B, Newcomb CJ, Keller TJ, Palmer LC, et al. 2017. Water Dynamics from the Surface to the Interior of a Supramolecular Nanostructure. *J. Am. Chem. Soc.* 139(26):8915–21
60. Jiao S, DeStefano A, Monroe JI, Barry M, Sherck N, et al. 2021. Quantifying Polypeptoid Conformational Landscapes through Integrated Experiment and Simulation. *Macromolecules*. 54(11):5011–21
61. Schrader AM, Monroe JI, Sheil R, Dobbs HA, Keller TJ, et al. 2018. Surface chemical heterogeneity modulates silica surface hydration. *Proceedings of the National Academy of Sciences*. 115(12):2890–95
62. Das Mahanta D, Robinson Brown D, Pezzotti S, Han S, Schwaab G, et al. 2023. Local Water Structures Govern the Mixing Thermodynamics of Glycerol-Water Solutions. ChemRxiv
63. Sherck N, Webber T, Brown DR, Keller T, Barry M, et al. 2020. End-to-End Distance Probability Distributions of Dilute Poly(ethylene oxide) in Aqueous Solution. *J. Am. Chem. Soc.* 142(46):19631–41
64. Fox T, Kollman PA. 1998. Application of the RESP Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B*. 102(41):8070–79
65. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, et al. 2004. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. *The Journal of Chemical Physics*. 120(20):9665–78
66. Vassetti D, Pagliai M, Procacci P. 2019. Assessment of GAFF2 and OPLS-AA General Force Fields in Combination with the Water Models TIP3P, SPCE, and OPC3 for the Solvation Free Energy of Druglike Organic Molecules. *Journal of Chemical Theory and Computation*. 15(3):1983–95
67. Sprenger KG, Jaeger VW, Pfaendtner J. 2015. The General AMBER Force Field (GAFF) Can Accurately Predict Thermodynamic and Transport Properties of Many Ionic Liquids. *The Journal of Physical Chemistry B*. 119(18):5882–95
68. Fennell CJ, Wymer KL, Mobley DL. 2014. A Fixed-Charge Model for Alcohol Polarization in the Condensed Phase, and Its Role in Small Molecule Hydration. *J. Phys. Chem. B*. 118(24):6438–46
69. Martínez L, Andrade R, Birgin EG, Martínez JM. 2009. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* 30(13):2157–64
70. Eastman P, Pande V. 2010. OpenMM: A Hardware-Independent Framework for Molecular Simulations. *Computing in Science & Engineering*. 12(4):34–39
71. Friant-Michel P, Wax J-F, Meyer N, Xu H, Millot C. 2019. Translational and Rotational Diffusion in Liquid Water at Very High Pressure: A Simulation Study. *J. Phys. Chem. B*. 123(47):10025–35
72. Tan P, Huang J, Mamontov E, García Sakai V, Merzel F, et al. 2020. Decoupling between the translation and rotation of water in the proximity of a protein molecule. *Phys. Chem. Chem. Phys.* 22(32):18132–40

73. Martiniano HFMC, Galamba N. 2013. Insights on Hydrogen-Bond Lifetimes in Liquid and Supercooled Water. *J. Phys. Chem. B*. 117(50):16188–95

74. Heyden M. 2019. Heterogeneity of water structure and dynamics at the protein-water interface. *J. Chem. Phys.* 150(9):094701

75. Laage D. 2006. A Molecular Jump Mechanism of Water Reorientation. *Science*. 311(5762):832–35

76. Chopra R, Truskett TM, Errington JR. 2010. On the Use of Excess Entropy Scaling to Describe the Dynamic Properties of Water. *J. Phys. Chem. B*. 114(32):10558–66

77. Nayar D, Agarwal M, Chakravarty C. 2011. Comparison of Tetrahedral Order, Liquid State Anomalies, and Hydration Behavior of mTIP3P and TIP4P Water Models. *J. Chem. Theory Comput.* 7(10):3354–67

78. Jabes BS, Agarwal M, Chakravarty C. 2010. Tetrahedral order, pair correlation entropy, and waterlike liquid state anomalies: Comparison of GeO2 with BeF2, SiO2, and H2O. *The Journal of Chemical Physics*. 132(23):234507

79. Monroe JI, Shell MS. 2019. Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *The Journal of Chemical Physics*. 151(9):094501

80. Baksi A, Biswas R. 2020. Does Confinement Modify Preferential Solvation and H-Bond Fluctuation Dynamics? A Molecular Level Investigation through Simulations of a Bulk and Confined Three-Component Mixture. *J. Phys. Chem. B*. 124(51):11718–29

81. SciPy 1.0 Contributors, Virtanen P, Gommers R, Oliphant TE, Haberland M, et al. 2020. Author Correction: SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat Methods*. 17(3):352–352

82. Barber CB, Dobkin DP, Huhdanpaa H. 1996. The quickhull algorithm for convex hulls. *ACM Trans. Math. Softw.* 22(4):469–83

83. Idrissi A, Damay P, Yukichi K, Jedlovszky P. 2008. Self-association of urea in aqueous solutions: A Voronoi polyhedron analysis study. *The Journal of Chemical Physics*. 129(16):164512

84. Yeh Y, Mou C-Y. 1999. Orientational Relaxation Dynamics of Liquid Water Studied by Molecular Dynamics Simulation. *J. Phys. Chem. B*. 103(18):3699–3705

85. Persson F, Söderhjelm P, Halle B. 2018. The geometry of protein hydration. *The Journal of Chemical Physics*. 148(21):215101

86. Widom B. 1963. Some Topics in the Theory of Fluids. *The Journal of Chemical Physics*. 39(11):2808–12

87. Acharya H, Vembanur S, Jamadagni SN, Garde S. 2010. Mapping hydrophobicity at the nanoscale: Applications to heterogeneous surfaces and proteins. *Faraday Discussions*. 146:353–65

88. Shiratani E, Sasai M. 1996. Growth and collapse of structural patterns in the hydrogen bond network in liquid water. *The Journal of Chemical Physics*. 104(19):7671–80

89. Kuffel A, Zielkiewicz J. 2012. Why the Solvation Water around Proteins Is More Dense than Bulk Water. *J. Phys. Chem. B*. 116(40):12113–24

90. Kuffel A, Czapiewski D, Zielkiewicz J. 2014. Unusual structural properties of water within the hydration shell of hyperactive antifreeze protein. *The Journal of Chemical Physics*. 141(5):055103

91. Raschka S. 2018. MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack. *JOSS*. 3(24):638

92. Kumar P, Buldyrev SV, Stanley HE. 2009. A tetrahedral entropy for water. *Proc. Natl. Acad. Sci. U.S.A.* 106(52):22130–34

93. Perry RH, Green DW, eds. 2008. *Perry's chemical engineers' handbook*. New York: McGraw-Hill. 1 p. 8th ed ed.

94. Athawale MV, Sarupria S, Garde S. 2008. Enthalpy−Entropy Contributions to Salt and Osmolyte Effects on Molecular-Scale Hydrophobic Hydration and Interactions. *J. Phys. Chem. B*. 112(18):5661–70

95. Di W, Gao X, Huang W, Sun Y, Lei H, et al. 2019. Direct Measurement of Length Scale Dependence of the Hydrophobic Free Energy of a Single Collapsed Polymer Nanosphere. *Physical Review Letters*. 122(4):047801

96. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, et al. 2012. Scikit-learn: Machine Learning in Python

97. Glielmo A, Husic BE, Rodriguez A, Clementi C, Noé F, Laio A. 2021. Unsupervised Learning Methods for Molecular Simulation Data. *Chem. Rev.* 121(16):9722–58

98. Zebari R, Abdulazeez A, Zeebaree D, Zebari D, Saeed J. 2020. A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction. *JASTT*. 1(2):56–70

99. Statt A, Kleeblatt DC, Reinhart WF. 2021. Unsupervised learning of sequence-specific aggregation behavior for a model copolymer. *Soft Matter*. 17(33):7697–7707

100. Löpez CA, Vesselinov VV, Gnanakaran S, Alexandrov BS. 2019. Unsupervised Machine Learning for Analysis of Phase Separation in Ternary Lipid Mixture. *J. Chem. Theory Comput.* 15(11):6343–57

101. Rodriguez-Nieva JF, Scheurer MS. 2019. Identifying topological order through unsupervised machine learning. *Nature Physics*, pp. 790–95

102. Altis A, Nguyen PH, Hegger R, Stock G. 2007. Dihedral angle principal component analysis of molecular dynamics simulations. *The Journal of Chemical Physics*. 126(24):244111

103. Freed JH. 1978. Dynamic effects of pair correlation functions on spin relaxation by translational diffusion in liquids. II. Finite jumps and independent T_1 processes. *The Journal of Chemical Physics*. 68(9):4034–37

104. Hwang L-P, Freed JH. 1975. Dynamic effects of pair correlation functions on spin relaxation by translational diffusion in liquids. *The Journal of Chemical Physics*. 63(9):4017

105. Bowron DT, Filipponi A, Roberts MA, Finney JL. 1998. Hydrophobic Hydration and the Formation of a Clathrate Hydrate. *Phys. Rev. Lett.* 81(19):4164–67

106. Davis JG, Gierszal KP, Wang P, Ben-Amotz D. 2012. Water structural transformation at molecular hydrophobic interfaces. *Nature*. 491(7425):582–85

107. Funke S, Sebastiani F, Schwaab G, Havenith M. 2019. Spectroscopic fingerprints in the low frequency spectrum of ice (Ih), clathrate hydrates, supercooled water, and hydrophobic hydration reveal similarities in the hydrogen bond network motifs. *J. Chem. Phys.* 150(22):224505

Chapter 6: Molecular-scale driving forces of solute-surface interactions

6.1 Introduction

As discussed in **Chapter 1**, current generation water purification membranes exhibit a nearly linear tradeoff between selectivity for desired constituents and water permeability. A key plank of next generation membrane material development is to overcome the selectivity-permeability tradeoff. The design of antifouling membranes is complimentary to overcoming this selectivity-permeability tradeoff. Specifically, anti-fouling membranes reduce the accumulation of undesired wastewater constituents and hence improve water permeability through membranes over time. Heuristically, antifouling membranes are (i) smooth, (ii) charge neutral, and (iii) hydrophilic(1–3). A net membrane surface charge induces the accumulation of charged foulants driven by electrostatics(3). Microscopically rough membrane surfaces yield increased foulant adhesion within the troughs of the rugged surface(1, 3). Surface hydrophilization is hypothesized to induce a bound water layer that further reduces fouling by organic species(2, 3). A variety of surface chemistry modifications(4–9) have been implemented finely tune solute surface interactions to further improve membrane antifouling. However, experimental methods and continuum fluid models cannot directly probe the molecular mechanisms underpinning the antifouling property. Instead, synergistic studies using experiment and detailed atomistic molecular simulations are required to understand the fundamental molecular driving forces of antifouling in different chemical contexts.

Most of the discussion in the previous five chapters pertains to aqueous mixtures rather than macromolecular interfaces found in Membrane Science. Still, much of the physical insight and even some of the results highlighted in the previous chapters aide in our understanding of water properties at membrane-water interfaces. Specifically, local aqueous environments in OEO/water mixtures serve as illustrative proxies for hydrated polymer membranes. However, it is necessary to explicitly investigate fouling mechanisms at more realistic polymer surfaces. Direct simulation of water purification membranes with atomistic resolution is intractable, but it is tractable to construct representative model systems to replicate the chemical and spatial arrangement of polymer surfaces inspired by antifouling materials utilized in reverse osmosis (RO) and nanofiltration (NF) membranes.

Experimentally quantifying and characterizing surface fouling at the molecular level remains an open challenge. Though methods like Surface Plasmon Resonance (SPR) can directly probe membrane foulants of interest to for water treatment such as proteins(10–13), directly measuring the chemical state and composition of the fouling layer remains challenging. Ambient-pressure X-ray Photoelectron Spectroscopy (AP-XPS) is an interesting alternative approach to probe fouling mechanisms and the surface features imbuing antifouling character. Traditional ultra-high vacuum (UHV) XPS spectroscopy probes both the atomic composition and the abundance of specific chemical functionalities within the first ~10 nm of surfaces with compositional accuracy up to ~0.1%. AP-XPS techniques leverage the same underlying physics as UHV XPS but near ambient conditions. Hence, AP-XPS enables the study of the abundance of specific chemical species sorbed within a hydration layer compared to the vapor phase.

In this chapter, we explain our efforts to develop an experimental-simulation workflow leveraging vacuum solute-surface affinity calculations and ambient-pressure x-ray photoelectron spectroscopy (APXPS) measurements. Following this computational-experimental study, we describe a representative antifouling OEO brush surface and leverage umbrella sampling calculation to characterize solute-surface interactions for a subset of the solutes from the vacuum study. Further, we leverage the principal component analysis from **Chapter 5** to explore the relationship between solution structure sorption thermodynamics.

6.2 Methodology

6.2.1 Molecular Dynamics Simulation Preparation

For the vacuum study, we use the ideal hydrophobic (methyl-terminated) and hydrophilic (hydroxyl-terminated) alkylsiloxane SAM surfaces described in the prior works by Monroe and coworkers(14, 15). For the hydrated antifouling study, we study oligo (ethylene oxide) (OEO) brush surfaces by grafting methyl-terminated ethylene oxide (EO) 6-mers on each side of grid of sulfur atoms arranged on a gold hexagonally close-packed (Au-hcp) lattice as described by Ismail and coworkers(16). We apply EO parameters developed in our prior investigation of the PEO configurational landscape in water(17). Grafting chains with Au-hcp spacing and on an 8x9 grid, we obtain a surface with dimensions $3.897\ nm \times 4.000\ nm$. Sulfur atom positions are fixed during simulation. We add OPC waters to the OEO system using the Packmol(18) software with an initial box length of $12\ nm$. To initialize simulations with a solute, we place the solute above the OEO surface again using Packmol.

6.2.2 Molecular Dynamics Simulation Procedure

We utilize OpenMM package(19) to conduct all simulations described in this chapter. We utilized particle-mesh Ewald (PME)(20) to compute electrostatic interactions. Non-bonded interactions are treated with nonbonded Lennard Jones (LJ) potentials cut off at 10Å. Per standard OpenMM procedure, we constrain all intermolecular hydrogen bonds the

SHAKE(21) algorithm and maintain water rigidity using SETTLE(22). For all simulations, we maintain a constant temperature of 298 K using a Langevin thermostat(19) with an integration timestep of 2 fs, friction coefficient of 0.1 ps^{-1}. For all simulations in the NPT ensemble, we utilize a Monte Carlo barostat—described in the OpenMM documentation(19)—with Monte Carlo barostat moves every 200 fs and a setpoint pressure of 1 bar. Prior to conducting umbrella sampling calculations, we equilibrate the hydrated OEO surface for 5-ns in the NVT ensemble at 500 K to relax the OEO chain configurations more rapidly. We follow this melting simulation with a 5-ns NPT ensemble simulation at 298.15 K and 1 bar to cool the system to the desired production temperature. Finally, we run a production simulation in the NPT ensemble at 298 K and 1 bar for 100-ns, saving the system coordinates every 10-ps. We generate solute-free structural information such as the density profile and three-body angle distributions using this trajectory.

Using Packmol(18), we insert a solute molecule at a random position 4-nm above the sulfur atom z-coordinate of the equilibrated and hydrated OEO system (without a solute). We equilibrate the solute at this farthest position for 2-ns in the NVT ensemble followed by 10-ns in the NPT ensemble. We apply an umbrella sampling procedure to compute potential of mean force (PMF) profiles between a given surface-solute pair. Specifically, we apply a harmonic centroid potential between the centroid of the brush sulfur atoms and the centroid of the given solute molecule with a force constant of $k = 1000 kJ/mol/nm^2$. We perform umbrella simulations at 30 umbrella states evenly spaced in the interval $[1.0nm, 4.0nm]$. We briefly equilibrate each of these states in the NVT ensemble for 100-ps followed by the NPT ensemble for 500-ps. Finally, we perform each umbrella state production run for 12-ns and save the system coordinates every 1-ps. The first 2-ns of each state production run is discarded. Using

MBAR(23, 24), we compute PMFs in the absence of a solute restraining potential. We estimate the uncertainties in the free energies by: (1) computing the PMFs for five independent blocks of production data and (2) calculating the standard error from the five block PMFs.

The system equilibration and umbrella sampling procedures for the vacuum SAM systems are essentially identical to that of the OEO surface described above. The primary difference is the shorter equilibration and production times utilized. In **Figure 6.1**, we depict an example of a PMF profile for inserting a methane molecule at a specified solute-surface distance from an ideal hydrophobic (methylated) SAM surface in vacuum. The PMF profile plateaus to a "bulk" vacuum value at solute-surface distances greater than ~2.5 nm to $PMF \approx 2k_BT$. PMF values lower than $2k_BT$ correspond to more favorable methane-surface distances than in vacuum. The depth of free energy minimum ($\sim 2k_BT$) at ~1.5 nm determines the solute-surface binding affinity.

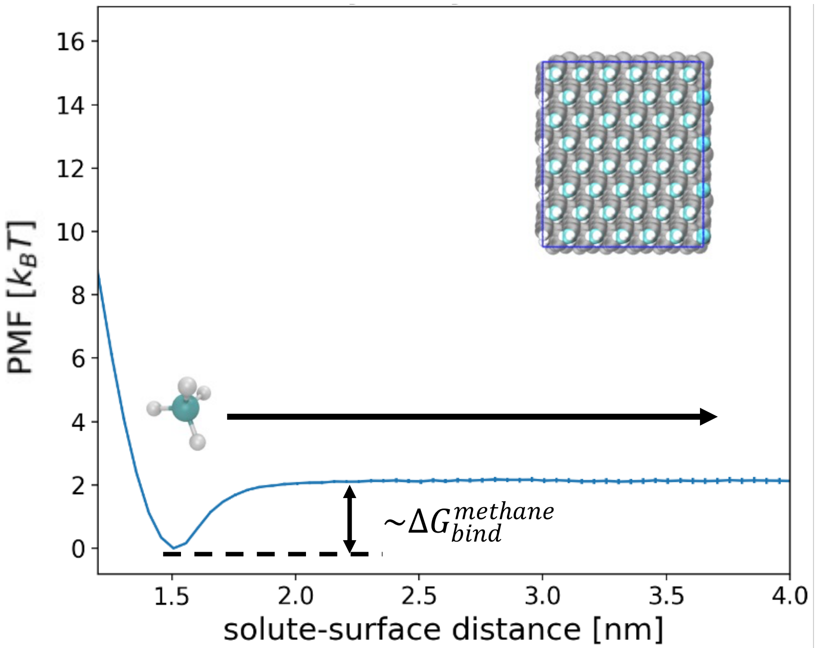


Figure 6.1 An illustrative example of an umbrella sampling calculation: The potential of mean force (PMF) of a single methane molecule as a function of distance from a methylated SAM surface indicates methane's preference for a given solute-surface distance. PMF values less than bulk (PMF lower than about $2k_BT$) suggest greater methane affinity relative to bulk water while the opposite is true for PMF values greater than bulk. The

free energy minimum at ~1.5 nm gives one measure of the binding free energy of methane for the SAM $\Delta G_{bind}^{methane}$.

6.3 Characterizing Solute-Surface Interactions in Vacuum

While the primary focus of the previous chapters are aqueous systems, the focus of the current section lies in the vacuum domain. We elect to probe solute-surface interactions in the absence water due to the present lack of an unambiguous complimentary experimental probe of solute-surface interactions in the infinite dilution limit. In vacuum, we can draw direct connections between atomistic molecular simulations and ambient pressure x-ray photoelectron spectroscopy (APXPS). As discussed in the introduction to this chapter, XPS enables the study of surface composition with high resolution (at approximately 0.1% atom fraction) and surface sensitivity (~10 nm) under ultra-high vacuum ($<10^{-8}$ Torr) conditions. Our collaborators' APXPS experimental apparatus enables XPS measurements near ambient pressure and hence at hydrated surfaces. As an incremental step in developing a complimentary simulation-APXPS workflow, we investigate solute-surface interactions for a suite of solute chemistries in the absence of water at ideal hydrophilic and hydrophobic SAM surfaces previously studied in work by the Shell group(14) and others(25–28).

6.3.1 Calculation of Solute Affinities in Vacuum

For the present discussion, we consider small hydrocarbon solutes due to their relevance to wastewater streams from oil and gas production. Further, small molecule solutes are more computationally tractable for the free energy calculations that we consider in this chapter. Specifically, we study alkanes (methane, ethane, propane, butane) and alcohols (methanol, ethanol, propanol, isopropanol, butanol) to systematically investigate the effect of solute hydrophobicity. We also consider water (TIP4P-EW) and acetic acid solute-surface interactions. Based on water solubility, the solutes with the highest (hypothetical) hydrophilicity contain higher ratios of hydrophilic heavy atoms (e.g., O and N) to hydrophobic heavy atoms (e.g., C). For instance, butanol is more hydrophobic than methanol. However, a previous work by Monroe and coworkers(14) demonstrated that interactions between solutes and hydrated SAM surfaces are poorly predicted by simple solubility arguments. Thus, detailed experimental or computational methods are required to properly describe solute-surface affinities.

In **Figure 6.2**, we depict the PMF profiles for several of the above-mentioned solutes at the ideal hydrophobic (methylated) and hydrophilic (hydroxylated) surfaces described in **Section 6.2**. Here, the plateau values at solute-surface separation—defined as the distance between the fixed sulfur atoms and the solute center of mass—greater than 2.0 nm give an estimate of solute-surface binding affinity with larger values corresponding to higher affinity. Notably, each solute exhibits higher affinity for the hydroxylated surface compared to the methylated surface. We find the largest contrast between hydrogen bonding (HB) solutes. Each alkane presents ~k_BT higher affinity for the hydroxylated surface compared to the methylated surface. On the other hand, hydrogen bonding solutes show considerably higher affinity (>$7k_BT$)

for the hydroxylated SAM than the methylated SAM.

The massive enhancement of sorption of hydrogen bonding solutes at hydroxylated surface clearly stems from solute-surface hydrogen bonds. In **Figure 6.7(a)**, we apply that a simple linear model: $\Delta G_{model}^{OH} = N_{HB} \Delta G_{HB}$ where N_{HB} is the average number of solute-surface hydrogen bonds and ΔG_{HB} is a fitting parameter. This model yields moderate agreement with the observed ΔG_{bind}^{OH} $[R^2 = 0.50]$. However, N_{HB} alone cannot forecast solute-surface affinity because neither non-hydrogen bonding (non-HB) solutes nor the hydrophobic surface participate in hydrogen bonds. In addition to solute-surface hydrogen bonding, solute molecular weight is a major determinant of surface binding affinity. We apply a second linear model: $\Delta G^{i}_{model} = N_{heavy}\Delta G_{VDW}$ where N_{heavy} is the number of solute heavy atoms and ΔG_{VDW} is a fitting parameter. This model yields reasonable predictions of $\Delta G^{CH_3}_{bind}$ ($R^2 = 0.80$) [**Figure 6.7(a)**] but fails to forecast ΔG^{OH}_{bind} ($R^2 \approx 0$) [**Figure 6.7(b)**]. N_{heavy} fails to capture the underlying trend in ΔG^{OH}_{bind} due to the large discrepancy in OH-SAM affinity between non-HB and HB solutes. We interpret the relationship between N_{heavy} and ΔG_{bind} as owing to the attractive van der Waals interactions between the solute and surface.

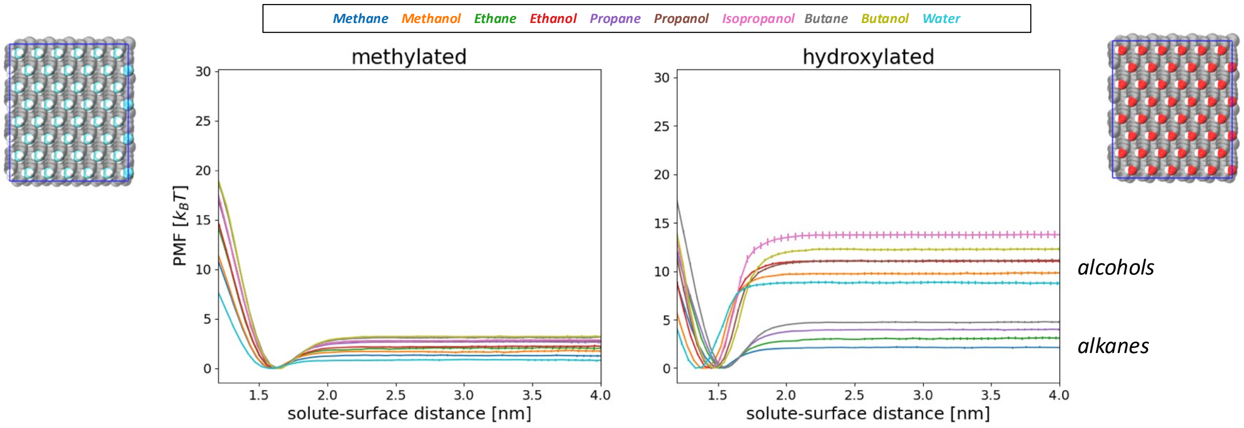


Figure 6.2 Comparing PMFs for a range of solutes for model hydrophobic and hydrophilic SAMs. The (hydrophobic) methylated surface shows a monotonic increase in well-depth with solute molecular weight with the lowest and greatest affinities demonstrated by TIP4P-EW water and butanol, respectively. A similar molecular weight dependence is observed for the (hydrophilic) hydroxylated surface. However, the alkane solutes display a markedly lower solute-surface affinity compared to the nominally hydrophilic solutes (alcohols and TIP4P-EW water).

6.3.2 Structural Drivers of Solute-Surface Interactions

In **Section 6.3.1**, we briefly note the solute features that correlate with solute-surface binding affinity. Leveraging the two solute properties (N_{HB} and N_{heavy}), we can construct a simple model to predict binding affinity between the OH-SAM and the library of solutes considered here. Specifically, we propose a bilinear model to predict the solute binding free energy ΔG_{model}

$$\Delta G_{model}^{OH} = N_{heavy}\Delta G_{VDW}^{OH} + N_{HB}\Delta G_{HB}^{OH}$$

where ΔG_{VDW}^{OH} and ΔG_{HB}^{OH} are linear least squares fitting parameters corresponding to the van der Waals and hydrogen bonding contributions to the binding free energy, respectively. Additionally, we consider the simpler model for the CH_3-SAM: $\Delta G_{model}^{CH_3} = N_{heavy}\Delta G_{VDW}^{CH_3}$.

The resulting model predictions depicted in **Figure 6.3** are imperfect but capture the underlying trends in ΔG_{bind} for both surfaces. In **Figure 6.3(a)**, we depict semi-quantitative prediction of ΔG_{bind}^{OH} from ΔG_{model}^{OH}. While the primary alcohols and alkanes follow ΔG_{model}^{OH}, isopropanol and acetic acid (syn- and anti- isomers) stray from the predictions of ΔG_{model}^{OH}. We now detail how these three solutes deviate from ΔG_{model}^{OH}. Though isopropanol exhibits more negative ΔG_{bind}^{OH} than the other alcohols, it simultaneously has the same N_{heavy} and N_{HB} (within error) as 1-propanol. The difference isopropanol's molecular geometry compared to 1-propanol dramatically affects solute-surface interaction. Hence, additional solute or surface properties must be considered to predict isopropanol's affinity for the OH-SAM more accurately. The two acetic acid (AA) isomers have the same N_{heavy} but present substantially different N_{HB} (~10% larger for anti-AA). The difference between ΔG_{bind}^{OH} and ΔG_{model}^{OH} for anti-AA and syn-AA are $-5\ k_BT$ and $2.5\ k_BT$, respectively. At the same time, the two water models present an even larger difference in N_{HB} (~13% larger for TIP4P-EW compared to OPC) but are nearly quantitatively predicted by the ΔG_{model}^{OH}. Though both water models present relatively distinct solute-surface hydrogen bonding, these poses the same underlying molecular geometry (rigid bonds and angles) and thus similar ΔG_{bind}^{OH} (less than $k_B T$ difference). On the other hand, anti- and syn-AA have the exact opposite dihedral configurations. Just as with isopropanol, additional molecular details are required construct a robust model for solute-surface affinity.

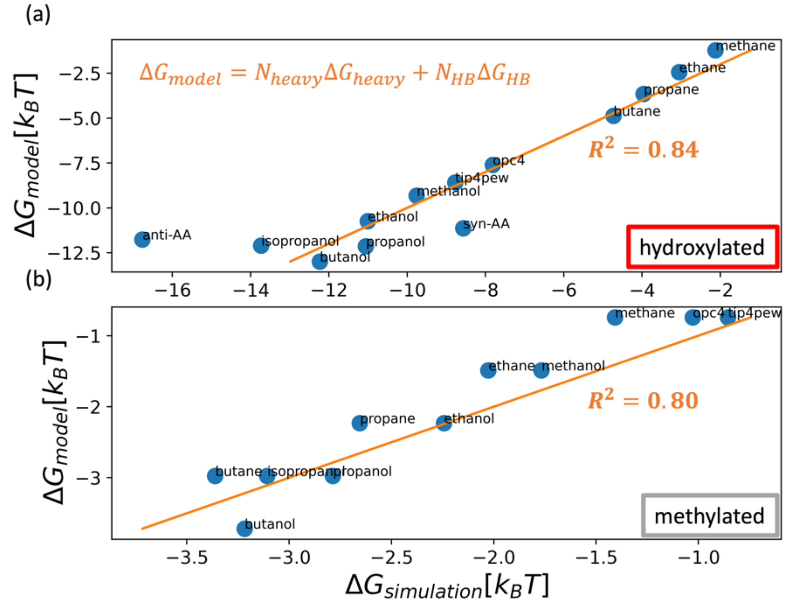


Figure 6.3 Simple two-parameter bilinear models ΔG_{model}^{i} take the number of solute heavy atoms N_{HB} and the average number of solute-surface hydrogen bonds N_{HB} as inputs to predict the simulated solute-surface binding affinity $\Delta G_{simulation}$ for both the (a) hydroxylated and (b) methylated SAM surface. The fitting parameters ΔG_{heavy} and ΔG_{HB} correspond to the free energy components corresponding to van der Waals interactions and hydrogen bonding, respectively. (a) Solute-surface affinities at the hydroxylated surface are best captured for alkanes and primary alcohols. The syn-acetic acid (AA_syn) and anti-acetic acid (AA_anti) isomers fall far off the trend line (orange) in opposite directions. This stems from the dramatic shifts the hydrogen bonding potential of acetic acid upon changing its conformation while leaving N_{heavy} unchanged. (b) Solute-surface affinities at the methylated surface are broadly captured by ΔG_{model}^{CH3}. Here, the methylated surface cannot hydrogen bond to the solutes and hence $N_{HB} = 0$; hence, reducing the model to $\Delta G_{model}^{CH3} = N_{heavy}\Delta G_{heavy}$.

6.3.3 Drawing Connections to APXPS Measurements

While our simulation studies provide some insight into the molecular scale drivers of solute-surface interactions, experimental methods are ultimately required to verify and contextualize the computational results. To draw a direct experimental connection, we consider APXPS measurements by our collaborators at the Advanced Light Source (ALS) at Lawrence Berkeley National Lab (LBNL)(29). The APXPS measurements considered equivalent methylated and hydroxylated SAM surface to the simulation study discussed above. Utilizing DFT simulations under APXPS conditions, one can determine the binding energy shifts in XPS spectra induced by adsorption of gaseous species onto a substrate.

For the present study, hydroxyl-containing solutes adsorbing to a SAM, the O 1s peak of an XPS spectra gives relative concentrations of oxygen-containing species owing to the underlying SAM material, vapor, and adsorbed to the surface. Our collaborators' experimental findings for both the methylated and hydroxylated SAM surfaces in the presence of water, butanol (BuOH), and acetic acid (AcOH) are depicted in **Figure 4-3** of the Mikayla Barry's dissertation(29). Experimentally, no adsorption characteristic peaks emerge for the hydrophobic surface and the hydrophilic surface only exhibits acetic acid adsorption. On the other hand, the simulation results suggest that every solute adsorbs to both SAM surfaces to varying degrees. However, we must consider a major caveat before discarding our findings. The magnitude of the hydrophobic SAM-solute binding free energies suggests sub-monolayer—or even single molecule—surface coverage of the methylated SAM within the pressure limits of the APXPS measurements. Therefore, the adsorbed solute layer may not be within the detection limits of APXPS. Even the hydrophilic solutes (except for acetic acid) may insufficiently adsorb to the hydroxylated SAM to detect with APXPS. Further, the observation that the hydroxylated SAM presents (detectable) acetic acid sorption using APXPS is in qualitative agreement with the substantially higher acetic acid binding free energy compared to the other hydrophilic solutes (~$6k_BT$ higher than butanol).

Ultimately, further iteration on both the experimental and computational methodology is required to draw quantitative connections. For example, more appreciable solute sorption may be observed at the CH_3 SAM surface if a pressure at or above the surface saturation pressure (for a given solute) is utilized for APXPS measurements. However, higher experimental operating pressures will mandate higher simulated pressures to quantitatively connect experiment to simulations; and hence, grand canonical ensemble simulations like those briefly mentioned in **Section 6.3.1** would provide a more favorable method.

6.4 Solute-Surface Interactions at a Hydrated Antifouling Surface

In this section, we return to the aqueous systems by probing solute-surface interactions at hydrated oligo(ethylene) oxide (OEO) brush surfaces to mimic antifouling layers in water purification membranes. However, our use of a 2D model system does build upon the idealized hydroxylated SAM surface discussed in **Section 6.3**. Unlike the model SAM systems, the OEO present a more molecularly rough surface and water permeable inner surface. These features increase the complexity of the simulations and post-production analyses but serve as an important building block to leveraging molecular simulations for developing realistic design motifs.

6.4.1 Solute Affinities in Solution

In **Figure 6.5**, we summarize the binding free energetics for a subset of the solutes considered in **Section 6.3**. Specifically, we exclude the alkane solutes for the present case. We also add the ideal hydrophilic solute ammonia as a contrast to ideal hydrophobe-surface interactions (here represented by methane). In **Figure 6.4(a)**, we depict the position-dependent affinity of the suite of solutes for the OEO brush surface. Notably, each solute displays either a free energy minimum or a weak free energy barrier for the solute at distances greater than the solute-surface interface as determined by the mean z-position of the instantaneous water-brush interface at $z = z_{int} \approx 1.9\ nm$. We determine the instantaneous interface using a method developed by Willard and Chandler(30, 31). The mean instantaneous interface is indicated with a vertical dashed line in **Figure 6.4(a)**.

To quantify the solute-surface interactions, we estimate the binding free energy ΔG_{bind} from the solute-surface distance at the net PMF minimum. In **Figure 6.4(c)**, we display the ΔG_{bind} values of each solute for the OEO brush surface and previous findings for ideal hydrophobic and hydrophilic surfaces [top-down surface depictions shown in **Figure 6.4(b)**]—the same surfaces that we discuss in **Section 6.3**—from a study by Monroe and Shell(14). In accordance with the heuristic fouling resistance of hydrophilic functionalizations (including EO-based coatings), the OEO brush surface exhibits reduced solute binding affinity relative to the ideal hydrophobic SAM surface and binding affinities within ~$k_B T$ of the ideal hydrophilic SAM. Specifically, the larger organic molecules—namely, benzene and phenol—display the most favorable (negative) ΔG_{bind} for each surface under consideration. On the other hand, molecules ammonia and boric acid are weakly attracted to the OEO brush ($0 < \Delta G_{bind} < k_B T$). Physical intuition may aide in the interpretation of the binding affinity trends shown in **Figure 6.4(c)**. For instance, one may expect hydrophilic surfaces to develop a strong “bound” water layer owing primarily to water-surface hydrogen bonding; hence, we encounter an entropic penalty to disrupting the surface-water hydrogen bonds to bind any solute to the surface. However, this explanation relies on the same heuristic reasoning that we aim to avoid in this study.

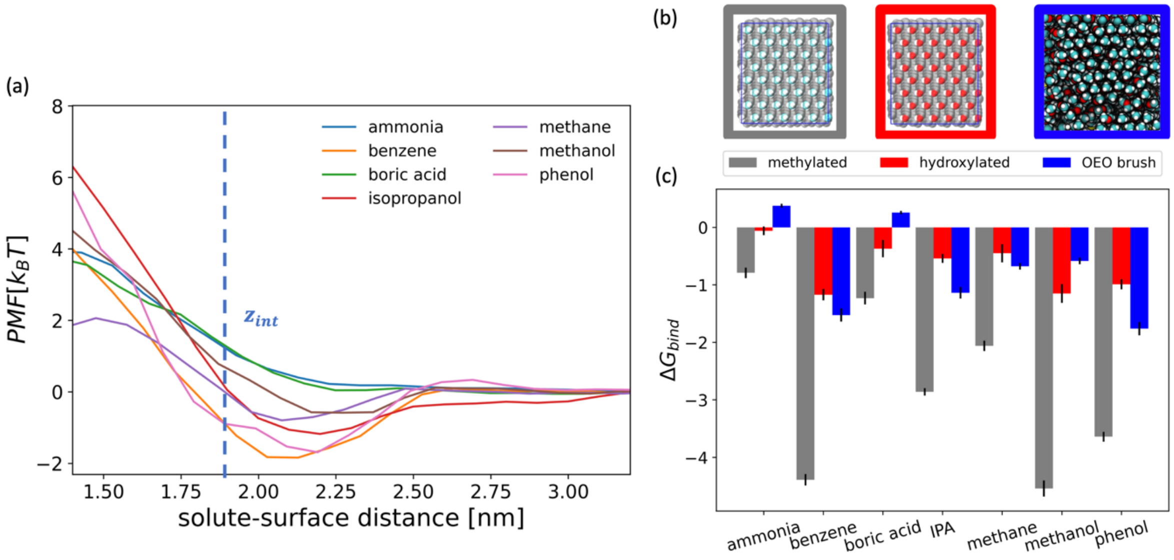


Figure 6.4 (a) OEO surface-solute PMF profiles suggest solute-surface interactions ranging from strong attraction (e.g., phenol) to weak repulsion (e.g., ammonia and boric acid) for the entire range of solutes. (b) Top-down representations of the OEO brush surface in comparison to ideal hydrophobic (methylated) and hydrophilic (hydroxylated) SAM surfaces studied by Monroe and Shell(14). (c) OEO surface-solute binding free energies ΔG_{bind} mimic the behavior of the magnitude and sign of the solute-surface affinities of the ideal hydrophilic (hydroxylated) SAM surface for all solutes excluding ammonia and boric acid.

Instead, we may perform several decompositions of the solute-surface PMFs. For instance, Monroe and coworkers demonstrated PMF decomposition schemes to access the van der Waals, electrostatic, entropic, and energetic components(14, 32). For the present discussion, we consider the following entropy-energetics decomposition scheme:

$$\Delta G_{bind} = \Delta\langle U_{SW}\rangle_{bind} + \Delta S_{res,bind} \qquad (6.1)$$

where $\Delta\langle U_{sw}\rangle_{bind}$ and $\Delta S_{res,bind}$ are the direct solute-surface interaction and water-surface interface restructuring components of the binding free energy. ΔG_{bind} is defined by the minimum of the solute-surface PMF as above. We compute $\Delta\langle U_{sw}\rangle_{bind}$ from a separate solute-surface direct interaction PMF, $\Delta\langle U_{sw}\rangle(z)$ derived by computing solute-environment interactions as a function of position z from the surface. We define $\Delta\langle U_{sw}\rangle_{bind}$ from the value of $\Delta\langle U_{sw}\rangle(z)$ at the solute-surface distance at minimum distance of the net PMF z_{bind}.

We calculate the position dependent entropy of restructuring PMF $\Delta S_{res}(z)$ from **Equation 6.1** and hence $\Delta S_{res,bind} = \Delta S_{res}(z = z_{bind})$. Contrary to the proposed interpretation of the binding free energy, the OEO brush surface exhibits favorable restructuring entropy [**Figure 6.5(c)**] and unfavorable direct interaction [**Figure 6.5(b)**] terms in a qualitatively similar manner to the model hydrophobic SAM. Specifically, while both the methylated and OEO surfaces entropically favor solute sorption, magnitudes of the $\Delta S_{res,bind}$ and $\Delta\langle U_{sw}\rangle_{bind}$ contributions ensure reduced solute sorption compared to the methylated SAM surface. Therefore, we cannot simply predict the thermodynamics driving forces behind solute-surface interactions using heuristics and intuition.

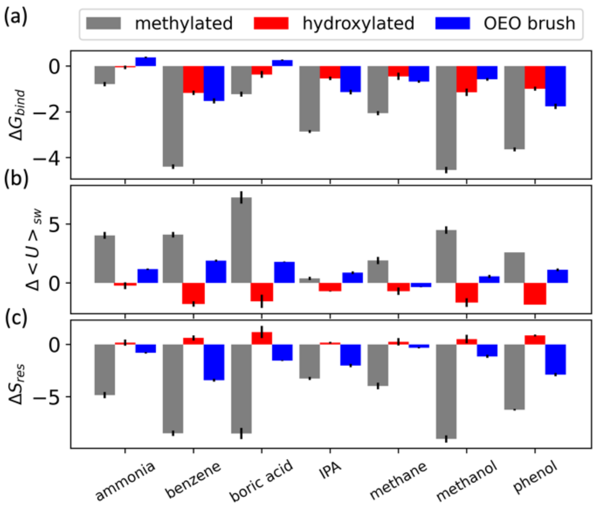


Figure 6.5 (a) The binding free energies ΔG_{bind} presented in **Figure 6.4c** are replicated here as a point of reference when interrogating the direct energetic $\Delta\langle U\rangle_{sw}$ and restructuring entropy ΔS_{res} components of ΔG_{bind}. (b) The direct energetic component of the binding free energy $\Delta\langle U\rangle_{sw}$ accounts for the direct solute-environment interactions at the solute-surface interface. Here, the OEO brush surface free energy displays a positive directionality more akin to the ideal hydrophobic (methylated) surface. (c) The entropic component of the binding free energy $\Delta S_{res,bind}$ accounts for the free energetic cost of restructuring the solution environment at the solute-surface interface. Here, the restructuring entropic contribution of the OEO surface more closely resembles the negative directionality of the ideal hydrophobic surface.

6.4.2 Building Structure-Affinity Connections

We discover persistent relationships between water properties and solution structure in **Chapter 5**. The three-body angle distribution of water $P_{3b}(\theta)$ presents a particularly detailed picture of water structure. Remarkably, by performing a principal component (PC) analysis on $P_{3b}(\theta)$ for 59—ostensibly simple—binary mixtures we recover signatures of water structure near small molecules (with a smallest dimension < 1 nm) but also signatures of the water dewetting at an extended hydrophobic surface [see **Figure 5.5(a)**]. To extend this analysis the OEO brush surface, we compute the difference three-body angle distributions

$$\Delta P_{3b}^{NS}(\theta; \Delta z) = P_{3b}^{NS}(\theta; \Delta z) - P_{3b,pure}(\theta)$$

where $P_{3b}^{NS}(\theta; \Delta z)$ is the three-body angle distributions of waters (in the absence of a solute) within a 5 Å-thick slab of water molecules at a position Δz above the water-brush interface. Once again, $P_{3b,pure}(\theta)$ is the three-body angle distribution of pure OPC water [see **Figure 6.6(a)**]. Inspecting $\Delta P_{3b}^{NS}(\theta; \Delta z)$ at seven Δz (-0.12, 0.13, 0.47, 0.81, 1.15, 1.49, and 1.83 nm), we observe a simultaneous enhancement of water's tetrahedral character and attenuation of water's icosahedral (simple fluid) character relative to pure water closer to the brush surface (smaller Δz). The same tetrahedral-icosahedral tradeoff was noted by Monroe and Shell in their investigation of the OH- and CH_3-SAM surfaces(15). At distances far enough from the OEO-water interface ($\Delta z = 1.83$ nm), $\Delta P_{3b}^{NS}(\theta; \Delta z) \rightarrow 0$ as water becomes bulk-like. We also consider a similar difference three-body angle distribution with the additional condition that the water molecules lie within a 7 Å radial cutoff of the heavy atoms of a phenol molecule centered at the same seven Δz values $\Delta P_{3b}^{S}(\theta; \Delta z)$ [see **Figure 6.6(b)**]. Like $\Delta P_{3b}^{NS}(\theta; \Delta z)$, $\Delta P_{3b}^{S}(\theta; \Delta z)$ presents an increasing tradeoff between water's tetrahedral and icosahedral character as Δz decreases. However, $\Delta P_{3b}^{S}(\theta; \Delta z)$ does not approach 0 in the bulk ($\Delta z = 1.83$ nm). However, this is not surprising given that water's tetrahedral character is typically enhanced in the vicinity of small molecules as shown in **Chapter 3** and many other studies(33–36).

While the apparent linear enhancement of water structure with decreasing Δz may appear to contradict findings from **Section 6.4.1**. Namely, that OEO brush-solute affinity is driven by a free energetic gain in the solution restructuring entropy $\Delta S_{res}(\Delta z = 0) < 0$. However, solution restructuring depends also on the molecular structure of the OEO brush as the solute approaches the interface. Additionally, solution structure is not solely determined by the orientational ordering of water—as measured by three-body angle distributions—but translational ordering as well. However, we can extract additional physical intuition from these two variants of $\Delta P_{3b}(\theta)$ by considering the three principal components detailed in **Chapter 5**. Specifically, by taking the projections of the $\Delta P_{3b}(\theta)$'s onto the three most important PC components—$p_i = proj(\Delta P_{3b}(\theta), PC_i)$—we extract position-dependent measures of the phenomena underlying the principal components. Recall that PC1 measures the tradeoff between tetrahedral and icosahedral waters and resembles $\Delta P_{3b}(\theta)$ for water molecules within the hydration layer an ideal small hydrophobic molecule (methane). PC2 probes the shift in $\Delta P_{3b}(\theta)$ that accompany transitions going from a hydrated small molecule to an extended hydrophobic interface; and hence, PC2 strongly resembles the $\Delta P_{3b}(\theta)$ at a water-methylated SAM interface. PC3 probes the relative abundance of 5-coordinated water molecules as discussed in previous studies(25, 32, 36).

In **Figure 6.6(c)**, we overlay the principal component projections for water in the absence of a solute p_i^{NS} and within the hydration shell of phenol p_i^S as a function of Δz. Both p_i increase monotonically with decreasing Δz as implied by **Figure 6.6(a) and (b)**. The PC1 projections exhibit distinct behavior for all but one Δz with $p_1^S > p_1^{NS}$. Notably, p_1^S approaches a finite value in bulk because $\Delta P_{3b}^S(\theta; \Delta z) \neq 0$ at large Δz as expected. Like the p_1 data, both p_2 series display a monotonic increase with decreasing Δz. $\Delta P_{3b}(\theta; \Delta z)$ do not display obvious signs of a transition to water structure near a hydrophobic surface such as the CH3-SAM, but both $\Delta P_{3b}^{NS}(\theta;\Delta z)$ and $\Delta P_{3b}^{S}(\theta;\Delta z)$ data sets present a small shift in the mean with varying Δz. Here, both p_2^{NS} and p_2^{S} approach 0 because the structure water near a phenol at Δz far from 0 does not closely resemble structure near a hydrophobic surface (such as the CH3-SAM). Finally, both p_3 series appear identical relative to the magnitude of the PC1 and PC2 projections. The p_3 decrease as Δz decrease because 5-coordinated waters become less favorable as density increases near the brush-water interface. We have made initial attempts to connect the PC projections more quantitatively to sorption thermodynamics (e.g., ΔS_{res}), but our efforts suggest that additional structural features may be required to develop more predictive models for solute-surface affinity.

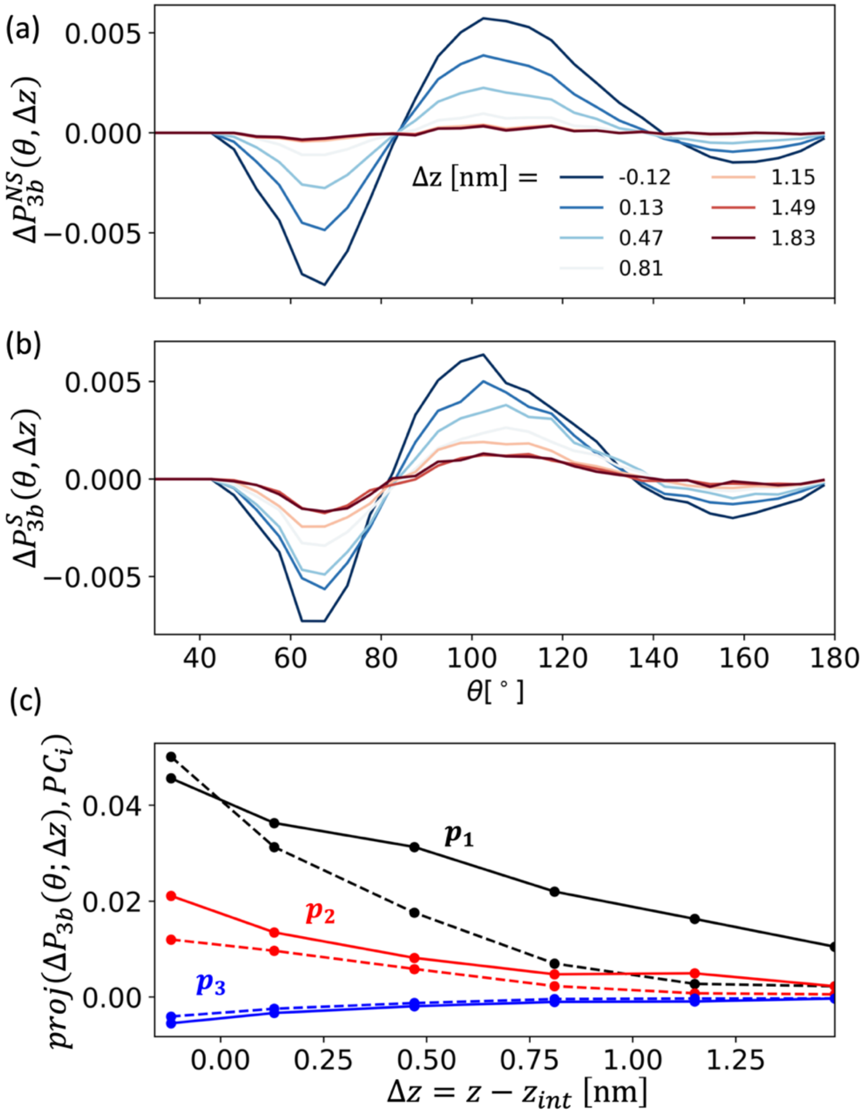


Figure 6.6 (a) We depict the non-solute ("NS") difference three body angle distribution $\Delta P_{3b}^{NS}(\theta;\Delta z)$ for waters within a 5 Å-thick slab of water molecules at a position Δz above the water-brush interface. (b) Similarly, we show the solute ("S") difference three body angle distribution $\Delta P_{3b}^{S}(\theta;\Delta z)$ for waters within a 7 Å radial cutoff of the heavy atoms of a phenol molecule centered at the same Δz positions. Here, we consider seven Δz values: -0.12, 0.13, 0.47, 0.81, 1.15, 1.49, and 1.83 nm. Both populations of water exhibit a monotonic increase in tetrahedral coordination and simultaneous decrease in icosahedral coordination. (c) We depict the projection of the difference three body angle distributions onto the three principal components (computed in **Chapter 5**) for "NS" $p_i^{NS} = proj(\Delta P_{3b}^{NS}(\theta;\Delta z), PC_i)$ and "S" the: $p_i^{NS} = proj(\Delta P_{3b}^{NS}(\theta;\Delta z), PC_i)$ waters. "NS" and "S" projections are differentiated by dashed and solid lines, respectively. p_1^{NS}, p_1^{S}, p_2^{NS}, and p_2^{S} monotonically decrease far from the OEO brush-water interface, but p_1^{S} approaches a finite bulk value while the other approach 0. The trends in p_1^{NS} and p_1^{S} quantify the tetrahedral-icosahedral tradeoff implied by the difference three body distributions as Δz decreases. The increase in p_2^{NS} and p_2^{S} as Δz decreases suggest the emergence of large length scale water structuring nearer to the brush surface. p_3^{NS} and p_3^{S} are nearly indistinguishable and monotonically increase towards 0 in the bulk as the frequency of 5-coordinated waters increases.

6.5 Conclusions

In the preceding discussion, we detail our progress in developing molecular intuition for anti-fouling at membrane-relevant materials. In **Section 6.3**, we advance efforts to draw connections between experimentally accessible and direct simulation-derived measures of foulant sorption in vacuum. We detail our molecular-scale analysis for the determinants of the sorption of small molecule solutes of interest for produced water treatment. For a broad set of neutral solute chemistries, we demonstrate semi-quantitative prediction of solute-surface affinity ΔG_{bind} leveraging only the solute molecular weight and its degree of hydrogen bonding (average number of hydrogen bonds formed) to a hydroxylated-SAM. We attempt to draw a connection between our simulation and AP-XPS measurements from our collaborators. The practical limitations of AP-XPS experimental conditions and sensitivity may hinder our ability to directly compare methodologies for the time-being. However, there is another AP-XPS apparatus at the LBL Advanced Light Source (ALS) that operates at higher pressures than the procedure mentioned above. Once personnel are available to interrogate this matter further, direct simulation-XPS comparisons may become more tractable.

In **Section 6.4**, we highlight computational advances on the study of antifouling at a membrane relevant OEO brush system by computing the potential-of-mean-force (PMF) for a subset of solutes from the vacuum study. We discover comparable solute-surface binding affinities observed at a model hydrophilic surface (OH-SAM) for the entire range of solutes. We decompose the PMFs into direct energetic $\Delta\langle U_{sw}\rangle$ (solute interactions with its environment) and solution restructuring entropy ΔS_{res} (entropic penalty to restructure the solution). Interestingly, opposite to the ideal hydrophilic surface (OH-SAM), the solute-OEO sorption is driven by favorable $\Delta S_{res} < 0$ despite the similar solute binding affinities for the two surfaces. This discrepancy likely stems from the water-SAM hydrogen bond network formed by the OH-SAM and not the OEO brush surface. We then leverage the PCA of the three-body angle distributions from **Chapter 5** to examine the position dependence of water behavior in connection with the binding affinity of phenol. While this is a *de novo* approach to analyzing the structural determinants of local water properties, further study is necessary to directly predict properties (as done in **Chapter 5**) like solute-surface affinities.

6.6 Appendix

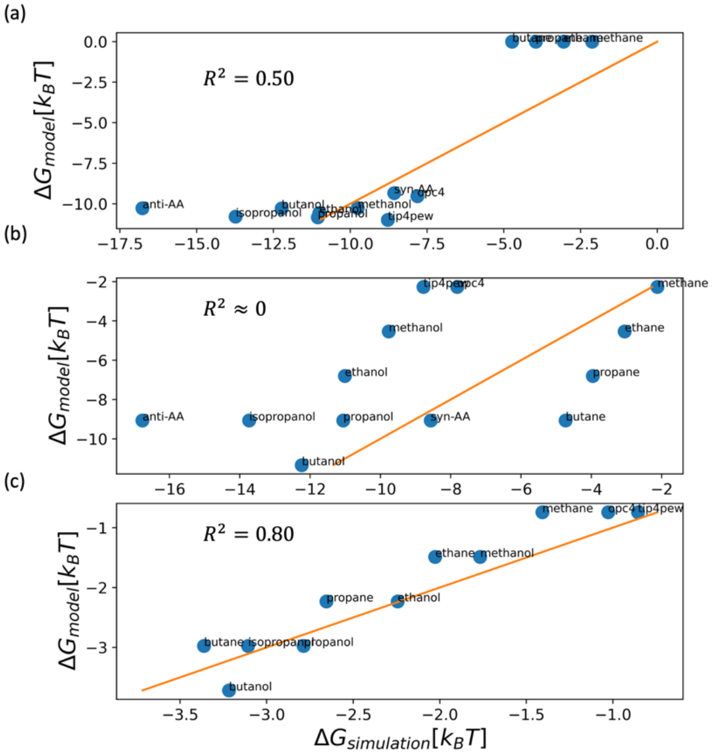


Figure 6.7 (a) Using the model $\Delta G^{OH}_{model} = N_{HB}\Delta G_{HB}$, we partially capture the variation in solute-OH SAM binding affinities. However, the average number of hydrogen bonds N_{HB} alone does not predict the behavior of non-hydrogen bonding ("non-HB") solutes. (b) Using the model $\Delta G^{OH}_{model} = N_{heavy}\Delta G_{VDW}$, we do not predict solute-OH SAM binding affinities. If we separately fit the trends for "non-HB" and hydrogen bonding ("HB") solutes, the two trends would reasonably capture the variation in ΔG^{OH}_{bind}. (c) Using the model $\Delta G^{CH_3}_{model} = N_{heavy}\Delta G_{VDW}$, we predict solute-CH$_3$ SAM binding affinities much better than the model presented in panel (a).

References

1. Zhong Z, Li D, Zhang B, Xing W. 2012. Membrane surface roughness characterization and its influence on ultrafine particle adhesion. *Separation and Purification Technology*. 90:140–46
2. Jiang S, Li Y, Ladewig BP. 2017. A review of reverse osmosis membrane fouling and control strategies. *Science of The Total Environment*. 595:567–83
3. Miller DJ, Dreyer DR, Bielawski CW, Paul DR, Freeman BD. 2017. Surface Modification of Water Purification Membranes. *Angewandte Chemie International Edition*. 56(17):4662–4711
4. Gu M, Kilduff JE, Belfort G. 2012. High throughput atmospheric pressure plasma-induced graft polymerization for identifying protein-resistant surfaces. *Biomaterials*. 33(5):1261–70
5. Zhou M, Liu H, Kilduff JE, Langer R, Anderson DG, Belfort G. 2009. High-Throughput Membrane Surface Modification to Control NOM Fouling. *Environmental Science & Technology*. 43(10):3865–71
6. Rana D, Matsuura T, Narbaitz RM, Feng C. 2005. Development and characterization of novel hydrophilic surface modifying macromolecule for polymeric membranes. *Journal of Membrane Science*. 249(1–2):103–12
7. Rana D, Matsuura T, Narbaitz R. 2006. Novel hydrophilic surface modifying macromolecules for polymeric membranes: Polyurethane ends capped by hydroxy group. *Journal of Membrane Science*. 282(1–2):205–16
8. Wang Y, Wang T, Su Y, Peng F, Wu H, Jiang Z. 2005. Remarkable Reduction of Irreversible Fouling and Improvement of the Permeation Properties of Poly(ether sulfone) Ultrafiltration Membranes by Blending with Pluronic F127. *Langmuir*. 21(25):11856–62
9. Chen W, Peng J, Su Y, Zheng L, Wang L, Jiang Z. 2009. Separation of oil/water emulsion using Pluronic F127 modified polyethersulfone ultrafiltration membranes. *Separation and Purification Technology*. 66(3):591–97
10. Chapman RG, Ostuni E, Takayama S, Holmlin RE, Yan L, Whitesides GM. 2000. Surveying for Surfaces that Resist the Adsorption of Proteins. *Journal of the American Chemical Society*. 122(34):8303–4
11. Chapman RG, Ostuni E, Liang MN, Meluleni G, Kim E, et al. 2001. Polymeric Thin Films That Resist the Adsorption of Proteins and the Adhesion of Bacteria. *Langmuir*. 17(4):1225–33
12. Chaudhury MK, Whitesides GM. 1992. Correlation Between Surface Free Energy and Surface Constitution. *Science*. 255(5049):1230–32
13. Harder P, Grunze M, Dahint R, Whitesides GM, Laibinis PE. 1998. Molecular Conformation in Oligo(ethylene glycol)-Terminated Self-Assembled Monolayers on Gold and Silver Surfaces Determines Their Ability To Resist Protein Adsorption. *The Journal of Physical Chemistry B*. 102(2):426–36
14. Monroe JI, Jiao S, Davis RJ, Robinson Brown D, Katz LE, Shell MS. 2021. Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution. *Proc Natl Acad Sci USA*. 118(1):e2020205118
15. Monroe JI, Shell MS. 2018. Computational discovery of chemically patterned surfaces that effect unique hydration water dynamics. *Proceedings of the National Academy of Sciences*. 115(32):8093–98

16. Ismail AE, Grest GS, Stevens MJ. 2007. Structure and Dynamics of Water near the Interface with Oligo(ethylene oxide) Self-Assembled Monolayers. *Langmuir*. 23(16):8508–14

17. Sherck N, Webber T, Brown DR, Keller T, Barry M, et al. 2020. End-to-End Distance Probability Distributions of Dilute Poly(ethylene oxide) in Aqueous Solution. *J. Am. Chem. Soc.* jacs.0c08709

18. Martínez L, Andrade R, Birgin EG, Martínez JM. 2009. PACKMOL: A package for building initial configurations for molecular dynamics simulations. *J. Comput. Chem.* 30(13):2157–64

19. Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, et al. 2017. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS Comput Biol*. 13(7):e1005659

20. Darden T, York D, Pedersen L. 1993. Particle mesh Ewald: An $N \cdot \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics*. 98(12):10089–92

21. Andersen HC. 1983. Rattle: A "velocity" version of the shake algorithm for molecular dynamics calculations. *Journal of Computational Physics*. 52(1):24–34

22. Miyamoto S, Kollman PA. 1992. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *J. Comput. Chem.* 13(8):952–62

23. Shirts MR, Chodera JD. 2008. Statistically optimal analysis of samples from multiple equilibrium states. *The Journal of Chemical Physics*. 129(12):124105

24. Chodera JD, Swope WC, Pitera JW, Seok C, Dill KA. 2007. Use of the Weighted Histogram Analysis Method for the Analysis of Simulated and Parallel Tempering Simulations. *J. Chem. Theory Comput.* 3(1):26–41

25. Dallin BC, Kelkar AS, Van Lehn RC. 2023. Structural features of interfacial water predict the hydrophobicity of chemically heterogeneous surfaces. *Chem. Sci.* 14(5):1308–19

26. Acharya H, Vembanur S, Jamadagni SN, Garde S. 2010. Mapping hydrophobicity at the nanoscale: Applications to heterogeneous surfaces and proteins. *Faraday Discussions*. 146:353–65

27. Patel AJ, Garde S. 2014. Efficient Method To Characterize the Context-Dependent Hydrophobicity of Proteins. *The Journal of Physical Chemistry B*. 118(6):1564–73

28. Godawat R, Jamadagni SN, Garde S. 2009. Characterizing hydrophobicity of interfaces by using cavity formation, solute binding, and water correlations. *Proceedings of the National Academy of Sciences*. 106(36):15119–24

29. Barry ME. 2021. *The effects of surface chemistry on polymer hydration, adsorption, and fouling*. UCSB

30. Willard AP, Chandler D. 2009. Coarse-grained modeling of the interface between water and heterogeneous surfaces. *Faraday Discuss.* 141:209–20

31. Willard AP, Chandler D. 2010. Instantaneous Liquid Interfaces. *J. Phys. Chem. B*. 114(5):1954–58

32. Monroe JI, Shell MS. 2019. Decoding signatures of structure, bulk thermodynamics, and solvation in three-body angle distributions of rigid water models. *The Journal of Chemical Physics*. 151(9):094501

33. Jiao S, Rivera Mirabal DM, DeStefano AJ, Segalman RA, Han S, Shell MS. 2022. Sequence Modulates Polypeptoid Hydration Water Structure and Dynamics. *Biomacromolecules*. 23(4):1745–56

34. Mahanta DD, Brown DR, Pezzotti S, Han S, Schwaab G, et al. 2023. Local solvation structures govern the mixing thermodynamics of glycerol–water solutions. *Chem. Sci.*
35. Stock P, Monroe JI, Utzig T, Smith DJ, Shell MS, Valtiner M. 2017. Unraveling Hydrophobic Interactions at the Molecular Scale Using Force Spectroscopy and Molecular Dynamics Simulations. *ACS Nano*. 11(3):2586–97
36. Bowron DT, Filipponi A, Roberts MA, Finney JL. 1998. Hydrophobic Hydration and the Formation of a Clathrate Hydrate. *Phys. Rev. Lett.* 81(19):4164–67

Chapter 7: Conclusions and Outlook

7.1 Summary

Here, we briefly summarize the findings presented in **Chapters 2-6**. In **Chapter 2,** we present a first-of-its-kind quantitative comparison between water dynamic metrics in glycerol-water as probed by molecular dynamics simulations and Overhauser Dynamic Nuclear Polarization (ODNP) experiments. This molecular-scale investigation appears to confirm the *structure-dynamics-thermodynamics hypothesis* for glycerol-water. **Chapter 3** highlights our further interrogation of the water *structure-dynamics* relationship in PEO-water solutions—a system relevant to membrane materials development—in close collaboration with experimental water dynamics probes from nanoscopic (ODNP) and microscopic (Pulsed-Field Gradient NMR) perspectives. In **Chapter 4**, we return small molecule solutes (DMSO and glycerol) to more closely inspect the hydration layer structure. Specifically, we decompose the solute hydration layer into solute hydrogen bonding ("bound") and non-hydrogen bonding ("wrap") populations and investigate their molecular structural features. We synthesize our findings from **Chapters 2-4** in **Chapter 5** to predict functional dynamics and thermodynamics properties from molecular structural metrics. Finally, we briefly discuss efforts to provide a molecular-scale picture solute-surface interactions at an extended oligoethylene (oxide) surface in **Chapter 6**.

7.2 Conclusions and Future Work

Through the cumulative efforts presented in **Chapters 2-5**, we demonstrate that *water structure is predictive of system properties and function*. However, the present efforts center highly idealized systems such as binary aqueous mixtures. To directly apply the methods and structure-property relationships elucidated above to industrially relevant systems, we must substantially extend our efforts to chemically and topographically heterogeneous interfaces such as those between proteins and water. For instance, the interrelationship between water structure, dynamics, and hydropathy presented at model extended surfaces(1) does not simply extend to the protein-water interfaces(2). Instead, the precise 3D-arrangement of chemical groups—with variable hydropathy and charge—strongly influences functional properties at protein-water interfaces(2–5). The vast and dynamic feature space at hydrated protein surfaces makes conventional analysis methods for predicting interfacial properties from water structure intractable.

In **Chapter 5**, we achieve a high degree of success leveraging very simple statistical learning methodologies to predict functional mixture properties from water molecular structural metrics. Our approach is reminiscent of a study by Dallin and coworkers(4) in which they apply a simple feature selection procedure and linear regression to accurately predict the hydration free energy (HFE)—a measure of hydropathy—at chemically heterogeneous self-assembled monolayer (SAM) surfaces. While we do detect structural signatures of solute-surface sorption at the OEO brush surface, we cannot simply predict sorption free energies from water structure using regression [**Chapter 6**]. Hence, we will likely be unable to predict a property like water self-diffusivity at the more chemically and topologically diverse surfaces presented at protein-water interfaces.

We note, however, that the principal component analysis (PCA) detailed in **Chapter 5** contains signatures of collective water network structures at both small (effective hydrated radius < 1 nm)—PC1—and large—PC2—hydrophobic interfaces. We hypothesize that these signatures may strongly indicate hydrophobic regions at the protein-water interface. For instance, a water near a large hydrophobic region of a surface is likely to resemble PC2 more closely than PC1. As detailed in numerous previous studies at heterogeneous molecular surfaces(4, 6–10), the tendency for a surface (or surface sub-region) to de-wet is highly suggestive of function. In the context of protein engineering, active sites typically de-wet more readily to promote interactions with ligand molecules(9). A purely structure-based probe of local hydration environments is highly desirable due to the high, and potentially insurmountable, computational cost of direct HFE calculation.

References

1. Monroe JI, Shell MS. 2018. Computational discovery of chemically patterned surfaces that effect unique hydration water dynamics. *Proceedings of the National Academy of Sciences*. 115(32):8093–98
2. Barnes R, Sun S, Fichou Y, Dahlquist FW, Heyden M, Han S. 2017. Spatially Heterogeneous Surface Water Diffusivity around Structured Protein Surfaces at Equilibrium. *J. Am. Chem. Soc.* 139(49):17890–901
3. Heyden M. 2019. Heterogeneity of water structure and dynamics at the protein-water interface. *J. Chem. Phys.* 150(9):094701
4. Dallin BC, Kelkar AS, Van Lehn RC. 2023. Structural features of interfacial water predict the hydrophobicity of chemically heterogeneous surfaces. *Chem. Sci.* 14(5):1308–19
5. Monroe JI, Jiao S, Davis RJ, Robinson Brown D, Katz LE, Shell MS. 2021. Affinity of small-molecule solutes to hydrophobic, hydrophilic, and chemically patterned interfaces in aqueous solution. *Proc Natl Acad Sci USA*. 118(1):e2020205118
6. Patel AJ, Garde S. 2014. Efficient Method To Characterize the Context-Dependent Hydrophobicity of Proteins. *The Journal of Physical Chemistry B*. 118(6):1564–73
7. Patel AJ, Varilly P, Chandler D. 2010. Fluctuations of Water near Extended Hydrophobic and Hydrophilic Surfaces. *J. Phys. Chem. B*. 114(4):1632–37
8. Patel AJ, Varilly P, Jamadagni SN, Acharya H, Garde S, Chandler D. 2011. Extended surfaces modulate hydrophobic interactions of neighboring solutes. *Proceedings of the National Academy of Sciences*. 108(43):17678–83

9. Rego NB, Xi E, Patel AJ. 2021. Identifying hydrophobic protein patches to inform protein interaction interfaces. *Proc Natl Acad Sci USA*. 118(6):e2018234118

10. Dallin BC, Van Lehn RC. 2019. Spatially Heterogeneous Water Properties at Disordered Surfaces Decrease the Hydrophobicity of Nonpolar Self-Assembled Monolayers. *The Journal of Physical Chemistry Letters*, pp. 3991–97

11. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*. 596(7873):583–89

12. Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, et al. 2022. AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Research*. 50(D1):D439–44