New Jersey Institute of Technology Digital Commons @ NJIT

Dissertations

Theses and Dissertations

Fall 2017

Roles of cosolvents on protein stability

Zhaoqian Su New Jersey Institute of Technology

Follow this and additional works at: https://digitalcommons.njit.edu/dissertations Part of the <u>Other Physics Commons</u>

Recommended Citation

Su, Zhaoqian, "Roles of cosolvents on protein stability" (2017). *Dissertations*. 48. https://digitalcommons.njit.edu/dissertations/48

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Digital Commons @ NJIT. It has been accepted for inclusion in Dissertations by an authorized administrator of Digital Commons @ NJIT. For more information, please contact digitalcommons@njit.edu.

Copyright Warning & Restrictions

The copyright law of the United States (Title 17, United States Code) governs the making of photocopies or other reproductions of copyrighted material.

Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be "used for any purpose other than private study, scholarship, or research." If a, user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of "fair use" that user may be liable for copyright infringement,

This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve violation of copyright law.

Please Note: The author retains the copyright while the New Jersey Institute of Technology reserves the right to distribute this thesis or dissertation

Printing note: If you do not wish to print this page, then select "Pages from: first page # to: last page #" on the print dialog screen



The Van Houten library has removed some of the personal information and all signatures from the approval page and biographical sketches of theses and dissertations in order to protect the identity of NJIT graduates and faculty.

ABSTRACT

ROLES OF COSOLVENTS ON PROTEIN STABILITY

by Zhaoqian Su

The function of a protein is determined by its three-dimensional structure which emerges from the delicate balance of forces involving atoms of the protein and the solvent. This balance can be perturbed by changing temperature, pressure, pH and by adding organic molecules known as cosolvents to the solution. Despite the wide use of cosolvents to perturb protein structures in the lab and in living systems, their molecular mechanisms are still not well established. Understanding these mechanisms is a problem of substantial interest, with potential application to the design of new drugs to target proteins. In this dissertation, we probe the role of two major cosolvents, urea and trimethylamine N-oxide (TMAO) at atomic level.

Urea is widely used as a denaturant in the lab to destabilize native protein conformations. However, the atomic mechanism of this molecule remains a question of debate. To unravel its molecular mechanism, explicit all-atom molecular dynamics simulations of unrestrained and extended poly-alanine and poly-leucine dimers are performed. Consistent with experimental results, we find that the large non-polar side chain of leucine is affected by urea whereas backbone atoms and alanine's side chain are not. Urea is found to occupy positions between leucine's side chains that are not accessible to water. This accounts for extra Lennard-Jones bonds between urea and side chains that favors the unfolded state. These bonds compete with urea-solvent interactions that favor the folded state. The sum of these two energetic terms provides the enthalpic driving force for unfolding. It is shown here that this enthalpy correlates with the potential of mean force of poly-leucine dimers.

To provide insights into the stabilizing mechanisms TMAO on protein structures, microsecond all-atom molecular dynamics simulations of peptides and replica exchange molecular dynamics simulations (REMD) of the Trp-cage miniprotein are performed. Most previous studies have focused on the effect of this osmolyte on protein backbone. Our results are consistent with these studies as we show that TMAO induces the backbone to adopt compact conformations. However, it is shown that effects of TMAO on the backbone are not dominant. In particular, TMAO's effect on the backbone is overcompensated by its destabilizing effect on the hydrophobic core: non-polar peptides and residues forming the hydrophobic core of the Trp-cage protein adopt more extended conformations in solutions containing TMAO. It is found that a main interaction that can stabilize folded proteins are charge-charge interactions. In light of these results, we propose that competing effects of TMAO on hydrophobic and charge-charge interactions account for its net stabilizing role on proteins.

ROLES OF COSOLVENTS ON PROTEIN STABILITY

by Zhaoqian Su

A Dissertation Submitted to the Faculty of New Jersey Institute of Technology and Rutgers, The State University of New Jersey – Newark in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Applied Physics

Department of Physics

December 2017

Copyright O 2017 by Zhaoqian Su ALL RIGHTS RESERVED

APPROVAL PAGE

ROLES OF COSOLVENTS ON PROTEIN STABILITY

Zhaoqian Su

Dr. Cristiano L. Dias, Dissertation Advisor	
Assistant Professor of Physics, NJIT	

Dr. John F. Federici, Committee Member Distinguished Professor of Physics, NJIT

Dr. Gordon A. Thomas, Committee Member Professor of Physics, NJIT

Dr. Luis R. Cruz Cruz, Committee Member Associate Professor of Physics, Drexel University

Dr. Yong I. Kim, Committee Member Assistant Professor of Chemistry and Environmental Science, NJIT Date

Date

Date

Date

Date

BIOGRAPHICAL SKETCH

Author:	Zhaoqian Su
Degree:	Doctor of Philosophy
Date:	December 2017

Undergraduate and Graduate Education:

- Doctor of Philosophy in Applied Physics, New Jersey Institute of Technology, Newark, NJ, 2017
- Master of Science in Materials Science and Engineering, New Jersey Institute of Technology, Newark, NJ, 2012
- Bachelor degree of Science in Applied Physics, Taiyuan University of Technology, Taiyuan, P.R. China, 2008

Major: Applied Physics

Presentations and Publications:

- Z. Su, F. Mahmoudinobar and C. L. Dias, "Effects of TMAO on the conformation of peptides and its implications for small proteins," *Physical Review Letters*, 119:108102, 2017.
- Z. Su and C. L. Dias, "Molecular interactions accounting for protein denaturation by urea," *Journal of Molecular Liquids*, 228:168, 2017.
- Z. Su, F. Mahmoudinobar and C. L. Dias, "Roles of TMAO on hydrophobic and ionic interactions," *submitted for publication*, 2017.
- F. Mahmoudinobar, Z. Su, and C. L. Dias, "Thermodynamics of polar and non-polar amyloid fibrils," *submitted for publication*, 2017.
- S. R. Jampani, F. Mahmoudinobar, Z. Su and C. L. Dias, "Thermodynamics of A β 1621 dissociation from a fibril: Enthalpy, entropy, and volumetric properties," *Proteins: Structure, Function, and Bioinformatics*, 83:1963, 2015.
- Z. Su and C. L. Dias, "Driving β -strands into fibrils," The Journal of Physical Chemistry B, 118:10830, 2014.

- Z. Cheng, A. E. Delahoy, Z. Su, K. K. Chin, "Steady state minority carrier lifetime and defect level occupation in thin film CdTe solar cells," *Thin Solid Films*, 558:391, 2014.
- Z. Su, "Effects of TMAO on the conformation of poly-peptides and its applications for protein folding," *Biophysical Society 61st Annual Meeting*, New Orleans, Louisiana, 2017.
- Z. Su, "Effects of TMAO on the conformation of poly-peptides" *Dana Knox Research Showcase*, NJIT, Newark, NJ, 2017.
- Z. Su, "Effect of cosolvents on non-polar peptides," *American Physical Society* conference, Baltimore, MD, 2016.
- Z. Su, "Roles of urea and TMAO on extended non-polar peptides," *Dana Knox Research Showcase*, NJIT, Newark, NJ, 2016.
- Z. Su, "Roles of urea and TMAO on the interaction between extended non-polar peptides," *Biophysical Society 59th Annual Meeting*, Baltimore, MD, 2015.
- Z. Su, "Molecular interactions accounting for protein denaturation by urea," Johnson-Johnson 2015 Engineering Showcase, Johnson-Johnson World Headquarters, 2015.



I am not worthy of the least of all the mercies, and of all the truth, which thou hast shewed unto thy servant; for with my staff I passed over this Jordan; and now I am become two bands.

Genesis 32:10

ACKNOWLEDGMENT

First and foremost, I want to express my deepest gratitude to my advisor, Dr. Cristiano L. Dias, who advanced my knowledge in protein folding and sharpened my skills in MD simulation. Without his guidance and encouragement this thesis would not have been possible. He is an outstanding scientist, and I greatly admire his intellect, curiosity and passion for science. To me, he is more a mentor than an academic advisor. With his valuable help about everything I needed during the past years, I become a person who knows how to listen, speak, read, write and more importantly, solve problems and present solutions. It has been my great privilege to be his student, and I always cherish my time spent here.

I am grateful to Prof. Thomas, Prof. Federici, Prof. Cruz, and Prof. Kim for taking time out from their busy schedules to be on my doctoral committee and providing me with valuable comments and feedback.

I thank all the past and present members of the Dias group for being great colleagues during the past five years. Dr. Chitra Narayan tutored my first molecular dynamics simulation. Srinivas Jampani shared with us the advanced features of VMD and introduced Ubuntu to our daily life at work. I am also thankful to Farbod Mahmoudinobar. He contributed to the work in Chapter 4. My discussions with him were thoroughly enjoyable. I have had a great time interacting with other graduate and undergraduate students in the group. Working with Daniel Deboer, with whom I have overlapped the most, has been great fun. I wish them as well as Jiajie and Patrick the best in their graduate studies and beyond.

I thank Kevin Walsh, NJIT HPC administrator, for working ceaselessly to make sure our computing resources are running. Research would have been very hard without such support. I also thank Christine A. Oertel and Leslie M. Williams from the Department of Physics for taking care of my paperwork, travel reimbursements and making life convenient.

My stay at NJIT has been enriched by numerous friends here: Zimeng, Zhicheng, Yixuan, Tingting, Yi Wu, Zhewen, Hao Cui, Tao Lv, Chi Xu, Liyuan, Yuanyuan, Zhizhong and many other. Special mention must be made of Nengyi and Yuqian for countless tea breaks. We spent so much time together sharing our thoughts and happiness, comforting and encouraging each other. I could not imagine finishing my Ph.D. without the help and support from them.

Furthermore, I would like to express special thanks to NJIT bible study club for providing me with such a warm welcoming home every Friday night.

Lastly but most importantly, I would like to thank my family for all their love and encouragement. For my parents, Yongqiang Su and Yaoxing Rui who raised me with a love of science and supported me in all my pursuits. And most of all for my loving, supportive, encouraging, and patient wife Lu whose faithful support during the final stages of this Ph.D. is so appreciated. Thank you.

TABLE OF CONTENTS

\mathbf{C}	hapt	\mathbf{er}		Page
1	INT	RODU	CTION	. 1
	1.1	Prote	in Structure	. 2
		1.1.1	Protein Primary Structure	. 3
		1.1.2	Protein Secondary Structure	. 3
		1.1.3	Protein Tertiary Structure	. 5
	1.2	Drivir	ng Force for Protein Folding	. 7
		1.2.1	Hydrogen Bonds	. 7
		1.2.2	Ionic Bonds (Salt Bridges)	. 8
		1.2.3	Van der Waals Bonds	. 9
		1.2.4	Hydrophobic Effect	. 10
		1.2.5	Configurational Entropy	. 12
	1.3	Molec	cular Dynamics Simulation	. 12
		1.3.1	Integrating the Equations of Motion	. 13
		1.3.2	Bonded and Non-bonded Interactions	. 15
		1.3.3	Simulating Rare Events	. 17
	1.4	Cosol	vent Effects on Protein Stability	. 22
		1.4.1	Protein Folding Equilibrium and Cosolvents	. 22
		1.4.2	Proposed Mechanisms for Urea-Induced Denaturation	. 24
		1.4.3	Proposed Mechanisms for TMAO Stabilization	. 26
2	DRI	VING	$\beta\text{-}\mathrm{STRANDS}$ INTO FIBRILS	. 29
	2.1	Intro	luction	. 29
	2.2	Metho	odology	. 32
	2.3	Resul	ts	. 35
	2.4	Concl	usion	. 41
	2.5	Assoc	iated Content	. 44

TABLE OF CONTENTS (Continued)

\mathbf{C}	hapt	er	Page
3 MOLECULAR INTERACTIONS ACCOUNTING FOR PROTEIN DENATURATION BY UREA		48	
	3.1	Introduction	48
	3.2	Methodology	51
	3.3	Results	55
	3.4	Conclusions	67
4	EFF I7	FECTS OF TMAO ON THE CONFORMATION OF PEPTIDES AND IS IMPLICATIONS FOR SMALL PROTEINS	. 69
	4.1	Introduction	69
	4.2	Methodology	71
	4.3	Results	77
	4.4	Conclusions	89
5	EFF	FECTS OF TMAO ON CHARGED INTERACTIONS	91
	5.1	Methodology	91
	5.2	Results	92
	5.3	Conclusions	98
6	SUN	MMARY AND FUTURE WORK	99
	6.1	Major Results of this Dissertation	99
	6.2	Foreseeing Future Work	101
B	IBLI(OGRAPHY	103

LIST OF TABLES

Tabl	Fable Pable	
3.1	System Details for Poly-leucine in Pure Water and Urea Solutions	54
3.2	Parameters of the AMBER-force-field for Urea	55
4.1	Details of the Simulations Performed for Poly-peptides and the Trp-cage Miniprotein	72
4.2	Thermodynamic Parameters for Trp-Cage at 298K in Pure Water	73
4.3	Thermodynamic Parameters for Trp-Cage at 298K in 5M TMAO Solution	74
4.4	Partial Charges and Lennard-Jones Parameters	76
4.5	Bond Parameters	76
4.6	Angle Parameters	77
4.7	Dihedral Angle Parameters	77
4.8	Average Fraction of Secondary Structures	82
4.9	Details of the Simulation Force Fields	90

LIST OF FIGURES

Figu	ire	Page
1.1	Schematics of amino acids and peptide bond formation.	2
1.2	Schematics of a polypeptide chain.	3
1.3	Secondary structures.	4
1.4	Backbone dihedral angles	5
1.5	Schematic Ramachandran plot showing allowed combinations of dihedral angles.	6
1.6	Visualization of tertiary structures	6
1.7	Schematic diagram showing the hydrogen bond pattern in an antiparrallel β -sheet	. 7
1.8	Peptide and Trp-cage protein with ionic bonds	8
1.9	Van der Waals forces and Leonard-Jones potential	9
1.10	Schematic representation of hydrophobic interaction	11
1.11	Principle of bond stretching and angle vibration.	16
1.12	Principle of improper dihedral angles.	16
1.13	Principle of proper dihedral angle	17
1.14	Schematic free energy profile	17
1.15	An example of 1-dimensional PMF	18
1.16	Schematic of a replica exchange molecular dynamics simulation	21
1.17	Schematic free energy landscape of a protein in water that exhibits two- state folding.	23
1.18	Atomic structure of urea (denaturant) and TMAO (protecting osmolyte)	. 24
2.1	Schematic representation of the reaction coordinate used to study the formation of peptide dimer trimer and tetramer.	. 33
2.2	Potential of Mean Force, number of hydrogen bonds, and potential energy to assemble β -strands in a hierarchical process.	36
2.3	Contribution of different chemical groups of the protein and solvent to tetramer formation.	. 39

LIST OF FIGURES (Continued)

Figu	ire	Page
2.4	Iso-surface of water distribution around the ground state of dimer, trimer and tetramer.	, . 41
2.5	Potential of mean force of two parallel and two antiparallel polyalanine peptides.	e . 44
2.6	Structure of two parallel polyalanine peptides	. 45
2.7	Dependence of the number of hydrogen bonds on the distance between centers-of-mass of peptides	n . 46
3.1	x-y components of the distance vector $\vec{\xi}$. 52
3.2	Histogram 38 windows in our simulaltions	. 53
3.3	Convergence of the PMF for poly-leucine in 7M urea solution	. 54
3.4	Distribution of the radius of gyration R_g of backbone atoms for poly- glycine, poly-alanine, and poly-leucine in aqueous urea solutions at 298 K and 1 atm	- 3 . 56
3.5	Potential of mean force (PMF) of poly-alanine and poly-leucine dimers in pure water and aqueous urea solutions.	n . 57
3.6	Potential of mean force (PMF) of poly-alanine and poly-leucine dimers in other concentrations.	n . 58
3.7	Preferential interaction $\Gamma(r)$ around dimers and monomeric peptides	. 60
3.8	Changes in energy terms when 7M urea is added to water	. 62
3.9	Volumes of poly-leucine in pure water and 7M urea solution	. 63
3.10	Lennard-Jones and electrostatic energies of peptide-solvent and solvent- solvent interactions	- . 65
4.1	Temperature dependence of RMSD distributions of Trp-cage miniprotein in pure water	n . 74
4.2	Fraction of folded replicas as a function of temperature for the Trp-cage miniprotein in pure water.	e . 75
4.3	Conformations of homopeptides in different aqueous solutions	. 78
4.4	SPCE water model and Netz TMAO model	. 80
4.5	Effects of TMAO on the conformation of non-polar peptides flanked by opposite charged residues	, 80

LIST OF FIGURES (Continued)

Figu	Ire	Page
4.6	Secondary structures of sample trajectories.	82
4.7	Probability distribution of the distance between charged hydrogen atoms of Lysine residue (K) and charged oxygen atoms of glutamic acid (E).	83
4.8	Free-energy landscape describing charge-charge interaction	83
4.9	Effects of TMAO on the Trp-cage miniprotein.	84
4.10	Preferential interaction $\Gamma(r)$ of poly-glycine, poly-leucine, $A\beta_{16-22}$ and Trp-cage.	87
4.11	Schematics of the most populated conformations and their allowed transitions.	88
5.1	PMF for the interaction between Na ⁺ and Cl ⁻ ions in pure water and aqueous TMAO solutions	93
5.2	Decomposition of PMFs of sodium chloride ion pair into enthalpies and entropies contribution	95
5.3	RDF of ions and different atoms of TMAO or water	97
6.1	Structure of CETP.	102

CHAPTER 1

INTRODUCTION

Proteins are the workhorse of the living cell actualizing the information blueprinted in the DNA [187]. These biomolecules perform and control almost all biological functions including catalysis (enzymes), transport (hemoglobin), transmission of information between specific cells and organs (hormones), activities in the immune system (antibodies), passage of molecules across cell membranes etc. [116]. To accomplish these tasks, proteins need to adopt a specific three-dimensional conformation, known as the native state [2]. In solution, proteins exist naturally in dynamic equilibrium between folded and unfolded states, and this equilibrium can be easily perturbed by changing temperature, pressure, pH and/or by adding small organic molecules known as co-solvents to the solution [26]. A main focus of this dissertation is to provide a molecular understanding of how model cosolvents affect the folding-unfolding equilibrium of proteins. In particular, we study effects of urea and TMAO, which are molecules used to destabilize and stabilize native protein structure, receptively. Understanding roles of these molecules can shed light on the new drug design to prevent protein conformational diseases

This dissertation is organized as follow. In Chapter 1, we provide an introduction to protein structure, molecular dynamics simulation and cosolvent effects on protein stability. In Chapter 2, we describe a framework to compute free-energies required to form peptide dimers, trimers and tetramers. In Chapter 3, we use this framework to determine roles of urea on the interaction between extended non-polar peptides. In Chapter 4, we provide insights into the stabilizing mechanisms of TMAO on peptides and protein structures. In Chapter 5, we probe in-depth understanding

of the stabilizing effects of TMAO on the distance-dependent charged interaction. In Chapter 6, we propose future research directions.



1.1 Protein Structure

Figure 1.1 Schematics of amino acids and peptide bond formation. (a) Schematics of an amino acid: Four groups of atoms that attach to a C_{α} atom consisting of an amino acid. (b) Peptide bond.

Structurally, proteins are heteropolymers consisting of the 20 naturally existing amino acids. Amino acids are made of four groups of atoms connected to a central carbon (C_{α}). These four groups are a single hydrogen atom (H), an amino group (NH₂), a carboxyl group (COOH) and a side chain (R) - see Figure 1.1a. What differentiates the amino acids from each other is the side chain, which are apolar, polar, acidic or basic. Amino acids can connect to each other via peptide bonds, in which amino group and the carboxyl group of two neighboring amino acids form a covalent bond with each other releasing a water molecule in this process - see Figure 1.1b. This leads to the formation of a long chain as shown in Figure 1.2. The repeating units $-N - C_{\alpha} - C - O$, which are connected to each other form the "backbone" of the protein.

1.1.1 Protein Primary Structure



Figure 1.2 Schematics of a polypeptide chain: A sequence of amino acids linked together via a peptide bond to form a polypeptide.

The primary structure is the simplest level of protein structure. It corresponds to the sequence in which the amino acids are linked together by peptide bonds. The protein sequence is determined by the nucleotide sequence of the segment of DNA containing the gene that codes for that protein. Each protein has a characteristic number of residues and amino acid sequence. Figure 1.2 shows the primary structure and how the peptide bond connects two adjacent amino acids starting from the N terminus to the C terminus.

1.1.2 Protein Secondary Structure

The next level of protein structure, i.e., secondary structure, refers to a local description of structure. Localized folding of a peptide chain are held together by weak bonds, e.g., hydrogen and Van der Waals bonds. Figure 1.3 shows the most common and important secondary structures which are called α -helices and β -sheets.

For α -helices, the carbonyl group (C = O) of amino acid *i* is hydrogen bonded to the amino group (N - H) of amino acid *i*+4 as shown in Figure 1.3a. This pattern



Figure 1.3 Secondary structures. Alpha-helix, antiparallel β -sheet, and parallel β -sheet. Dotted lines represent hydrogen bonds. Backbones and side chains are represented using a cartoon and van der walls representation, respectively

of bonding exerts force on the polypeptide chain into a helical structure that looks like a curled ribbon, containing 3.6 amino acids per turn. The R groups of the amino acids extend outward from the α -helix, where they are free to interact.

For β -sheets, two or more extended segments of a polypeptide chain line up next to each other, forming a sheet-like structure held together by hydrogen bonds. Hydrogen bonds are formed between carbonyl and amino groups of the backbone, whereas the side chain extend above and below the plane of the sheet. β -sheets can be parallel or antiparallel, depending on whether the β -strands run in the same or opposite directions, respectively. The direction is defined by the amino-carboxyl orientation of the chain as illustrated in Figure 1.3b,c.

The secondary structure can be approximately described by the backbone dihedral angles - see Figure 1.4. The dihedral angle around the C_{α} -C' bond is called psi (ψ) and the one around the N-C_{α} bond is called phi (ϕ). Because of steric constraints, not all the combinations of the dihedral angles are allowed. This can be visualized by plotting the ϕ and ψ angles against each other in a diagram called Ramachandran plot. Figure 1.5 shows various regions of the Ramachandran plot and the secondary structure of a protein to which it correspond. For example, the



Figure 1.4 Diagram showing backbone diheral angles: the angle of rotation around C_{α} -C' bond is called psi (ψ) and the one around the N-C_{α} bond is called phi (ϕ).

 $\phi \sim -160$ and $\psi \sim 160$ region corresponds roughly to β strands, $\phi \sim -60$ and $\psi \sim -60$ corresponds approximately to α helices.

1.1.3 Protein Tertiary Structure

The result of the secondary structure elements coming together in an energetically favorable way accounts for the tertiary structure, i.e., the three-dimensional shape of the protein molecule. For example, a tertiary structure can be formed by packing β -strands into one or many β -sheets with α -helices packed on top of the sheets. For an illustration of a tertiary structure, a cholesteryl ester transfer protein's (CETP) 3D structure is shown in Figure 1.6, where the protein is colored according to its secondary structure segments: β -strands are colored in yellow, whereas α -helices are shown in purple.

Although the three-dimensional shape of a protein may seem irregular and random, it is fashioned by many stabilizing and destabilizing forces between protein-



Figure 1.5 Schematic ramachandran plot indicating allowed combinations of dihedral angles defined in Figure 1.4. Red, blue and green areas correspond to regular β strands, right-handed α helices and left-handed α helices, respectively.



Example 1.6 Cholesteryl Ester Transfer Protein (CETP) **Figure 1.6** Visualization of tertiary structures. β -strands are colored in yellow, whereas the α -helices are shown in purple.

protein, protein-solvent, and solvent-solvent interactions. Anfinsen demonstrated that the primary sequence contains all the information to specify the three dimensional protein structure. This has given rise to the assumption that determing three dimensional structure of a protein from its primary sequence is possible.

1.2 Driving Force for Protein Folding

Native protein structures are typically only 5-10 kcal/mol more stable than unfolded states. This free-energy difference emerges from the sum of several weak interactions - like hydrogen bonds, ion pairs, van der Waals attractions, and hydrophobic interactions. These interactions are discussed in more detail below.

1.2.1 Hydrogen Bonds



Antiparallel β-sheet

Figure 1.7 Schematic diagram showing the hydrogen bond pattern in an antiparrallel β -sheet. Oxygen atoms are red; Nitrogen atoms are blue; The hydrogen atoms are white; The carbon atom in the main chain is cyan; Sidechains are not shown.

Hydrogen bonds play an important role in stabilizing protein structures. This weak bond is formed when a Hydrogen atom which is bonded to an electronegative atom (an atom that has relatively more tendency to attract bonding pair of electrons), is close to another electronegative atom that has a lone pair of electron. An example of hydrogen bond that is commonly seen in protein is between NH and O groups of the backbone. Figure 1.7 shows two peptides are aligned adjacent to each other such that hydrogen bonds can form between C=O groups of one peptide and NH groups on the other and vice versa. Compared to covalent bonds, hydrogen bonds are very weak (1 – 5 kcal/mol) and therefore can be easily destructed by heating up (1k_BT = 0.59 kcal/mol). Unfolding of tertiary and secondary structures of proteins happen as a result of this phenomenon. Hydrogen bonds are short-range and directional interactions, as they can only form at distances less than 3 Å and over a restricted range of angles.

1.2.2 Ionic Bonds (Salt Bridges)



Figure 1.8 (a) $A\beta_{16-22}$ peptide and (b) Trp-cage miniprotein.

Ionic bonds are formed between oppositely charged groups in proteins. Ionic bonds are short-range bonds and they are inversely dependent on the polarity of the environment around the two charges. For example, the attractive interaction between two charges buried in the nonpolar interior of a protein is approximately twenty times stronger than the attractive interaction between two charges in water. The dipoles of the water molecules surrounding each charge align in such a way that the charges are partially attenuated, decreasing the attraction between them. Because there is no water on the interior of a protein, and because most of the buried side chains are nonpolar, the interactions between side chains of opposite charge are highly favorable. We call ionic bonds in proteins salt bridges. Figure 1.8 shows the conformations of $A\beta_{16-22}$ peptide with charge-charge interaction and Trp-cage miniprotein with a key salt bridge. This indicates that charge-charge interaction between polar residues play an significant role in peptide and protein conformations. In Chapter 4, we demonstrate the addition of TMAO molecules into water is found to stabilize compact peptide structures by strengthening charge-charged interactions.





Figure 1.9 Van der Waals forces and Leonard-Jones potential. (a) van der Waal's interaction between two methane molecules. (b) The Leonard-Jones Potential energy between two atoms varies as a function of distance between the two atoms. Decomposition of the Leonard-Jones Potential energy into repulsive term (blue dashed line) and attractive term (red dashed line).

Van der Waals forces are caused by correlations in the fluctuating polarization of nearby atoms. As two atoms approach one another, the electron clouds become distorted so that there are transient, complementary dipoles induced in opposing surfaces. The favorable electrostatic interactions between these flickering dipoles are weak and act only at very short distances, but they add up so that molecules with lots of surface area interact strongly. As shown in Figure 1.9a, even though methane molecule has no net dipole, at an one instant its electron density may not be completely symmetrical, resulting in a temporary dipole. This can induce a temporary dipole in another molecule. The weak interaction of these temporary dipoles constitutes van der Waals forces. Moreover, the surface area of a molecule determines the strength of the van der Waals interactions between molecules. The larger the surface area, the larger the attractive force between two molecules, and the stronger the intermolecular forces.

The attractive part of Lennard-Jones potential is often used as an approximate model for van der Waal force. The curve in Figure 1.9b (Leonard-Jones Potential) shows how the potential energy between two atoms varies as a function of the distance between the atoms. The potential energy between two nonpolar groups in a molecule varies in the same way. When the nonpolar groups are at far distance, they do not interact (because they need to be close to distort one anothers electron clouds, producing those favorable flickering dipoles). As they approach more closely, the potential energy decreases (which is favorable) until a minimum is reached. Forcing the groups any closer causes a steep rise in energy (because the electron clouds start to overlap, which is unfavorable).

1.2.4 Hydrophobic Effect

Hydrophobic interactions are very important driving force for protein folding. They can be described as the tendency for nonpolar surfaces to interact with each other rather than with water, i.e., they segregate. This leads to the burial of nonpolar side chains in the interior of proteins, which in turn leads to a collapse of the protein



Figure 1.10 Schematic representation of hydrophobic interaction. Hydrophobic solutes are colored in blue. Water molecules in the solvation shell are colored in red. Bulk water are shown in green. Arrow shows hydrophobic association and release of water to the bulk.

from an extended coil to a more compact, globular structure. This hydrophobic collapse happens not because van der Waals interactions are favorable, but because it is unfavorable for water molecules to organize around a hydrophobic surface. From the standpoint of the nonpolar molecule, it is at least as favorable to interact with water as it is to interact with another molecule like itself. From the standpoint of water, dipole-dipole interactions (between tow water molecules) are stronger than dipoleinduced dipole interactions (between water and a non-polar molecule). Thus, it is unfavorable for the hydrogen bonds between water molecules to be disrupted because a nonpolar molecule is in its way. To avoid this unfavorable situation, the water molecules become highly ordered around the nonpolar surface, forming structures that maximizing hydrogen bond formation as shown in Figure 1.10. In other words, water molecules around nonpolar surfaces are perturbed towards the solid phase, i.e., ice. This accounts for a reduction in the degree of order in the system which is entropically unfavorable. To avoid these unfavorable water structures, it is energetically favorable to bury non-polar surfaces away from water, i.e., non-polar side chains are buried in the protein interior. Consequently, the free energy of the system goes down.

1.2.5 Configurational Entropy

Conformational entropy is the major energy term that opposes folding of protein. When a protein is in unfolded state, it can form any arbitrary conformation of allowed ψ and ϕ angles. But when the protein folds into its native state, almost all those available conformations disappear and the protein can only adopt a reduced set of conformations. This reduction in available structures of protein is defined by loss of configurational entropy. As a result, when the folding reaction proceeds, the free energy eventually decreases as the loss of entropy starts to be compensated by the favorable hydrophobic and other interactions.

1.3 Molecular Dynamics Simulation

Molecular simulation is a very influential modeling toolbox, which enables us to follow and understand structure and dynamics with extreme detail, literally on scales where motion of individual atoms can be tracked [111]. We carry out computer simulations to understand the properties of assemblies of molecules in terms of their structure and the microscopic interactions between them.

Computer simulations act as a bridge between microscopic length and time scales and the macroscopic world of the laboratory: we provide a guess of the interactions between molecules, and obtain 'exact' predictions of bulk properties. The predictions are 'exact' in the sense that they can be made as accurate as we like, subject to the limitations imposed by our computer budget. Ultimately, bulk properties are compared with experimental measurements made on specific materials to validate initial guesses. From validated simulations, the hidden atomic level detail behind hulk measurements can be revealed [1, 142]. From validated simulations, the hidden atomic level detail behind bulk measurements can be revealed. We may also carry out simulations on the computer that are difficult or impossible in the laboratory (for example, working at extremes of temperature or pressure).

MD simulations fundamentally solve Newtons laws of motion for a set of classical interacting particles. This generates trajectories for the particles in the system that can be visualized and used to measure equilibrium and non-equilibrium properties such as transport coefficients or equilibrium thermodynamic ensemble averages. For example, to measure the mean value of a quantity X of interest (such as distance, contacts or binding energy between two molecules), one first records its value X_n at different times T_n , from which an average can be extracted:

$$\overline{X} = \frac{1}{K} \sum_{n=1}^{K} X_n \tag{1.1}$$

where K is the total number of measurements. \overline{X} is then the simulation estimate of the mean value of quantity X in the thermodynamic equilibrium.

1.3.1 Integrating the Equations of Motion

The particles in an MD system may represent for example connected atoms in a protein, a molecule of gas or some larger-scale coarse-grained entity. For a system of N atoms, the interactions between particles are specified by the interaction potential function $U(\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$, which is uniquely determined by their positions. The force \mathbf{F}_i on atom i can be written as the gradient of the potential energy $U(\mathbf{r}_i)$

$$\mathbf{F}_{\mathbf{i}} = -\frac{\partial U(r_i)}{\partial \mathbf{r}_{\mathbf{i}}},\tag{1.2}$$

and equation of motion of particle i is given by

$$M_i \frac{\partial^2 \mathbf{r_i}}{\partial t^2} = \mathbf{F_i}, i = 1...N.$$
(1.3)

In order to obtain the trajectory of the particles, we need to integrate the above equation of motion. A straightforward integration scheme that is widely used in MD simulation is called the Verlet algorithm. The simulation time is discretized into time steps of equal length Δt . Taylor expansion of the positions at times $(t + \Delta t)$ and $(t - \Delta t)$ gives:

$$r_i(t+\Delta t) = r_i(t) + v_i(t)\Delta t + \frac{f_i(t)}{2m}\Delta t^2 + \frac{\Delta t^3}{6}\ddot{r} + O(\Delta t^4)$$
(1.4)

$$r_i(t - \Delta t) = r_i(t) - v_i(t)\Delta t + \frac{f_i(t)}{2m}\Delta t^2 - \frac{\Delta t^3}{6}\ddot{r} + O(\Delta t^4)$$
(1.5)

Adding these two equations, we can obtain

$$r(t + \Delta t) = 2r_i(t) - r_i(t - \Delta t) + \frac{f_i(t)}{m} \Delta t^2 + O(\Delta t^4)$$
(1.6)

for the position update. Note that the expression for positions of particles are precise up to the $O(\Delta t^4)$, because the terms proportional to $O(\Delta t^3)$ canceled out. We further obtain velocity by substracting Equation (1.4) from Equation (1.5) and rearranging the terms:

$$v_i(t) = \frac{1}{2\Delta t} (r_i(t + \Delta t) - r_i(t - \Delta t)) + O(\Delta t^3)$$
(1.7)

The previous integration scheme is equivalently expressed in the so-called Leap-Frog algorithm:

$$r_i(t + \Delta t) = r_i(t) + v_i(t + \frac{\Delta t}{2})\Delta t$$
(1.8)

$$v_i(t + \frac{\Delta t}{2}) = v_i(t - \frac{\Delta t}{2}) + v_i(t + \frac{F_i}{m_i}\Delta t)$$

$$(1.9)$$

which provides the values for position and velocity for each time step. The appropriate choice of step size Δt will depend on the particular form of interaction potential between particles. Choosing Δt too large will lead to large differences in the forces at adjacent time-steps, and generate artifacts because the integrator does not converge. Typically, the steeper the potential or the higher the temperature the smaller the maximal time-step. On the other hand, picking too small a time-step will unnecessarily slow down the simulation, which could mean that the time trajectories are too short to generate meaningful time averages.

Despite its relative simplicity, the Verlet scheme remains a popular choice for MD simulations. Higher order integration schemes are useful for reproducing a single trajectory as accurately as possible for given initial condition, like when navigating a space probe through asteroid belt. But for sampling thermodynamic properties of a large molecular system, the Verlet integration has several important properties that make it a good algorithm. It only requires evaluation of the first derivative of the potential function, which makes it very efficient, as the force calculation is typically the most time-consuming part of simulation.

1.3.2 Bonded and Non-bonded Interactions

MD simulations commonly utilize pair-potentials, i.e., the potential energy is calculated from the sum of two-particle interactions. In order to calculate the potential energy for a given set of particle coordinates, one must first define the set of equations (known as force-filed) describing the different interactions that occur in the system. As an example, Equation (1.10)-(1.12) represent the interaction between atoms in the Amber force-field:

$$U = U_{bonded} + U_{non-bonded} \tag{1.10}$$

where U_{bonded} and $U_{non-bonded}$ represent covalent and weak interactions in the system of interest. Energies related to bond vibration, angle vibration, and movement around torsion angles of covalently linked atoms are described by:

$$U_{bonded} = \sum_{bonds} K_b (r - r_o)^2 + \sum_{angles} K_\theta (\theta - \theta_o)^2 + \sum_{dihedrals} K_\chi (1 + \cos(n\phi - \delta)) + \sum_{impropers} K_{imp} (\psi - \psi_o)^2$$

$$(1.11)$$

The first term in this equation accounts for bond stretching between two covalently bonded atoms i and j as shown in Figure 1.11. Bond-angle vibration between a triplet of atoms i-j-k is represented by a harmonic potential on the angle θ (second term of equation 1.11). Improper dihedrals are meant to keep certain groups of atoms in a plane (e.g. aromatic rings), or to prevent molecules from flipping over to their mirror images, see Figure 1.12.



Figure 1.11 Principle of bond stretching (left) and angle vibration (right).



Figure 1.12 Principle of improper dihedral angles. Out of plane bending for rings (left), substituents of rings (middle), out of tetrahedral (right). The improper dihedral angle ψ is defined as the angle between planes (i,j,k) and (j,k,l) in all cases.

The non-bonded potential consists of Lennard-Jones potential followed by the electrostatic potential:

$$U_{nonbonded} = \frac{1}{2} \sum_{i} \sum_{j \neq i} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \right] + \frac{q_i q_j}{4\pi\varepsilon r_{ij}}$$
(1.12)



Figure 1.13 Principle of proper dihedral angle.

Differences among the force fields used in MD simulation software are commonly not in the form of used potential energy function, but rather in the parameter of the force field. They are optimized by fitting the potential energy function, and functions derived from it, to target data in the form of experimental data or quantum mechanical data of model compounds.

1.3.3 Simulating Rare Events



Figure 1.14 Schematic free energy profile as a function of reaction coordinate. Crossing from state A to state B is slowed down by the barrier.
If free-energy barriers are much larger than k_BT (the typical energy scale in molecular simulations), then only a very rare fluctuation will allow the simulated system to cross the barrier, leading to sampling problems. But if you have some knowledge of the barrier, then progress can be made to accelerate the simulation. Several such techniques are described below. But first, to characterize a barrier, choose a reaction coordinate ξ and plot the free energy $F(\xi)$, as shown in Figure 1.14. The energy minima separated by a barrier can represent for example different conformational states of a molecule, or a system with unbound and bound ligand respectively. The reaction coordinate ξ characterizing the transition can be for example torsional angle or distance between a receptor and a ligand. It is very important to pick a suitable order parameter to describe the desired transition.



Figure 1.15 An example of 1-dimensional free-energy landscape (shown in red). Biasing umbrella sampling potential (green) is shown as well as the resulting biased free-energy landscape (blue).

Umbrella Sampling Method Umbrella sampling is one of the most straightforward method to overcome high barriers. This method introduces a biasing potential to the Hamiltonian of a system, which depends on the reaction coordinate ξ and is designed to raise probability of unlikely events. Figure 1.15 shows how umbrella potential (green) changes the shape of the underlying free-energy landscape (red) such that high-energy structures are sampled more frequently.

The resulting ensemble of modified Hamiltonian, shown in Equation 1.13, includes more conformations around region what are rarely sampled by traditional MD simulations. However, the statistics calculated from an umbrella sampling simulation correspond to the H_{bias} Hamiltonian in Equation 1.13, where the physical system actually obeys the H_{orig} Hamiltonian.

$$H_{bias}(r) = H_{orig}(r) + \frac{1}{2}k(\xi - \xi_o)$$
(1.13)

 H_{orig} is the original, unbiased Hamiltonian, k is the strength of the harmonic bias, ξ is the reaction coordinate, ξ_0 is reference point. Since the exact biasing potential is known, and the sampling provides information about the total biased Hamiltonian, we can use the information to infer the original Hamiltonian along the PMF that our simulation has successfully characterized through sampling. However, since the umbrella potential increases monotonically on the umbrella center, configurations far away from that center will be sampled insufficiently, leading to poor convergence in those regions. To overcome this issue, a series of umbrella sampling simulations are conducted in intervals along the reaction coordinate (called windows), which are used to assemble pieces of the PMF near the center of the respective umbrellas. These pieces are then stitched together to approximate the total, unbiased PMF. The free energy of the biased potential along the PMF is related to the probability density function at that point according to

$$P_{bias}(r_o) = \frac{\int exp(-\beta H_{bias}(r))\delta(r-r_o))dr}{exp(-\beta A_{bias}(r_o))}$$
(1.14)

where P is the probability distribution function, δ is the Dirac delta function that to extract only those ensemble members that correspond to the specific point r_o on the PMF, and A is the free energy along the PMF at that value. The unbiased probability distribution, which is directly related to the unbiased free energy up to an arbitrary constant, can be estimated according to

$$P_{unbias}(r) = exp(-\beta(A_{bias} - A_{unbias}))exp\left[\beta(\frac{1}{2}k(\xi - \xi_o))\right]P_{bias}(r)$$
(1.15)

where A is the Helmholtz free energy along the PMF. The unbiased probability distribution function is estimated for each window, and must be recombined to calculate the full PMF.

Replica Exchanges Molecular Dynamics In traditional MD simulations, one trajectory was generated under one condition or some conditions that were set, one big problem about such simulation is the lack of efficiency in sampling the canonical distribution due to the trapping into a large number of local minima, which slow this sampling of phase space. Several attempts have been made to overcome this problem. One of them is Replica Exchange Molecular Dynamics (REMD). In this method as shown in Figure 1.16, a series of simulations are conducted parallelly at different temperatures, and exchanges of configurations are attempted periodically.

Since the time needed to travel across the barrier is $\propto exp(\Delta F/k_BT)$, it will be less difficult to cross the barrier at higher temperature. Parallel tempering utilize this principle by coupling simulations at multiple temperatures. The higher temperature is beneficial for the system to escape local free-energy minima, while simulations at the lower temperature T of interest produce the correct ensemble averages.

In this scheme, n simulations (MD or MC) are run parallelly at temperatures T1 < T2 < T3 < ... Tn, respectively. After the certain steps, two neighboring temperatures $T_a < T_{a+1}$ are chosen and configurations are swapped with desired



Figure 1.16 Schematic of a replica exchange molecular dynamics simulation. Replica are represented as red, blue and silver arrows at their corresponding temperature. The exchange attempts are shown between adjacent replicas connected by orange arrows.

exchange probability that meets detailed balance condition

$$p = min\left(1, exp\left[\left(\frac{1}{k_B T_a} - \frac{1}{k_B T_{a+1}}\right)\left(V\left(r_a^N\right) - V\left(r_{a+1}^N\right)\right)\right]\right)$$
(1.16)

where $V(r_a^N)$ and $V(r_{a+1}^N)$ represent the potential energy evaluated for the configuration of particles in a-th and a+1-th replica, respectively. In the variant of parallel tempering for MD, called replica exchange molecular dynamics (REMD) [182], it is further necessary to rescale velocities of particles by a factor $\sqrt{\frac{T_a}{T_{a+1}}}$ for configuration which was at T_{a+1} prior to accepting the exchange, and by an inverse factor for velocities of configuration originally at T_a

Since the probability of accepting a configuration swap will drop fast to zero if T_a and T_{a+1} are excessively far apart, adding more intermediate temperatures can be more efficient than running just two replicas at T1 and Tn. A key input parameter for a replica exchange simulation is the number of replicas and range of temperatures

to be covered by those replicas. Even though more replicas are more expensive, but closer temperatures result in larger acceptance rates for replica swaps. So the number of replicas needs to be optimized to attain the most efficient simulation.

1.4 Cosolvent Effects on Protein Stability

In solution, folded and unfolded states of proteins exist in a dynamic equilibrium. This equilibrium can be easily perturbed by the addition of small organic molecules known as co-solvents to the solution[26]. The underlying molecular mechanism of these cosolvents is a question of debate [173]. Here, we review current ideas on various aspects of cosolvents on protein denaturation and stabilization.

1.4.1 Protein Folding Equilibrium and Cosolvents

As for any chemical equilibria, protein folding and unfolding transitions can be characterized by thermodynamic parameters. In particular, the difference in Gibbs free energy between folded and unfolded states quantifies the stability of the native state for proteins that fold without intermediate structure. For these two states, the folding equilibrium can be written as:

$$U \frac{k_f}{k_u} F, \tag{1.17}$$

where U and F denote unfolded and folded states, respectively. k_f and k_u are the folding and unfolding rates, respectively. These quantities determine the Gibb free energy of folding, as well as the concentration of folded and unfolded proteins according to the following equations:

$$\Delta G = -RT \ln \frac{k_f}{k_u},\tag{1.18}$$

$$\frac{[F]}{[U]} = \frac{k_f}{k_u},\tag{1.19}$$

where [F] and [U] are the concentrations of folded and unfolded protein, respectively. The equilibrium concentrations of the folded and the unfolded species are reached when they match the rates of conversion from one species to the other.



Figure 1.17 Schematic free energy landscape of a protein in water that exhibits two-state folding. The reaction coordinate ξ represents the averaged folding/unfolding pathway.

Proteins that exhibit two-state folding transitions have a relatively smooth free energy landscape with two minima along the folding pathway. The free energy barrier that separates the two minima corresponds to the transition state (TS) [104]. A schematic free energy landscape of a protein that exhibits twostate folding is displayed in Figure 1.17. The equilibrium between folded and unfolded protein structures is governed by the interplay between protein-protein interactions, solvent-solvent interactions, protein-solvent interactions and the configuration entropy. Folding of a protein increases the number of protein-protein contacts and solvent-solvent contacts. Unfolding, on the other hand, increases the number of protein-solvent contacts and the configurational entropy of the system. The stability of a folded protein is affected by the addition of cosolvents to the solution. Cosolvents that shift the equilibrium toward the unfolded ensemble are termed denaturants, e.g., urea, whereas those that favor the folded ensemble, e.g. are known as protecting osmolytes. Thus, the equilibrium can be tuned in either direction, depending on the identity of the cosolvent. Below we discuss effects of urea and TMAO on the stability of the native state.



Figure 1.18 Atomic structure of urea (denaturant) and TMAO (protecting osmolyte). Carbon is shown cyan, nitrogen in blue,oxygen in red, and hydrogen in white.

1.4.2 Proposed Mechanisms for Urea-Induced Denaturation

Urea molecule is widely used as a denaturant in the lab whereby it perturbs native protein conformations to favor the unfolded state [78, 26]. However, urea's atomic mechanism has been a question of debate since its effect was quantify in the 60's [173]—although important insights have been obtained recently [76, 26, 81, 58, 80]. Urea may impede the formation of native structures either by forming strong bonds with the protein (direct mechanism) [28, 178, 83, 81, 109, 156, 123, 105, 130, 46] or by modifying the structure of water accounting for different solvation properties (indirect mechanism) [67, 15, 88, 44, 63]. Moreover, it is still not clear whether urea affects mainly the backbone or side chains of the protein. Answers to these open questions can be obtained using simplified computational frameworks that can quantify interaction strengths in the solvent and protein at the atomic level.

In recent years, several studies have emerged showing that, in solution, urea can replace water without changing significantly its structural properties since both molecules are highly polar [177, 173, 95, 155, 102, 83, 9, 90]. This has been a strong argument against urea's indirect mechanism which assumes that the formation of water structures is disrupted by this molecule [67, 129]. Urea's direct mechanism is, therefore, gaining increased acceptance in the scientific community. In particular, it is supported by explicit all-atom computer simulations [28, 178] wherein urea is shown to form strong non-bonded interactions with the protein providing the driving force for unfolding. Some studies indicate that these protein-urea interactions emerge mainly from hydrogen bonding with the backbone [18, 109, 57, 123, 9] whereas others point to dispersion interactions with non-polar groups of the proteins [159, 206, 102, 158, 80]. Separating contributions from backbone and side chain atoms is not an easy task even in computer simulations that have access to atomic detail [28, 81, 18].

Traditionally, model compounds have been used to estimate interaction strengths between different chemical groups in proteins [177, 156, 140, 170, 193, 179]. In particular, free-energies to transfer non-polar amino acids from pure water to urea solutions were shown to be negative implying that urea solutions can accommodate non-polar amino acids better than pure water [140]. As a result, hydrophobic interactions are weaker in urea solutions and this could account for protein denaturation. Small non-polar compounds (e.g., methane and alanine) as well as glycine were not found to be significantly affected by urea [140]. Computer simulations of model compounds point to the same conclusion [170, 193] but also highlight strong hydrogen bonding of urea with small charged solutes [57]. Different conclusions can be drawn from studies of simple model compounds because they do not account simultaneously for contributions from backbone and side chain atoms in the same proportion as in real proteins. Moreover, solvation of side chains are affected by the backbone as reported in a recent study [192, 77]. These limitations of model compounds can be overcome by studying homopeptides [81, 179, 102, 130]. Recently, Horinek and Netz computed free-energies τ to transfer homopeptides from pure water to aqueous solutions containing urea using all-atom molecular dynamics simulations and different force-fields [81]. Direct and indirect effects of urea were singled out by differentiating τ with respect to urea concentration. This provided important insights into effects of force-fields on urea's mechanism [81].

In Chapter 3 of this dissertation, we study effects of urea on non-restrained poly-glycine, poly-alanine, and poly-leucine monomers as well as on the interaction of extended poly-alanine and poly-leucine dimers. For the extended dimers, we compute the free-energy required to bring two peptides that are initially non-interacting all the way to a distance ξ , i.e., we compute potentials of mean force (PMF). Our choice of peptides and approach allows us to probe the effect of urea on two types of interactions that are essential to account for protein folding, i.e., hydrophobic interactions between side chains of alanine and leucine residues and interactions between backbone atoms. Notice that, in addition to non-polar residues, urea has also been shown to bind favorably to polar and charged residues [177, 178].

1.4.3 Proposed Mechanisms for TMAO Stabilization

The stabilizing effect of TMAO on native structures is often explained by its exclusion from the vicinity of the protein surface [41, 110, 171] due to favorable TMAO-water interactions [86, 16, 211, 132, 162, 99, 167]. Exclusion may be more pronounced close to the main chain of the protein, i.e., the backbone [167, 21, 7, 8, 6]. This explains the adoption of more compact conformations by poly-glycine (which is commonly used as a model of the protein backone) in aqueous TMAO solution compared to pure water [82, 181]. Despite these results, TMAO exclusion from the protein surface as a mechanism to protect the native state has been challenged by recent studies. In particular, computer simulations and experiments have shown that TMAO stabilizes compact conformations of some non-polar polymers through direct interactions [157, 131]. Other studies have suggested that protecting effects of TMAO emerge because this molecule acts as a crowding agent [120] and/or by weakening the strength of hydrogen bonds between the protein and water molecules [120]. Effects of TMAO were also reported to emerge from favorable interactions of this osmolyte with the heterogeneous protein surface that is generated upon folding [108]. In later studies, TMAO was not excluded from air-water or polypeptide-water interfaces which is a result supported by both experiments and computer simulations [64, 162].

Previous studies have shown that TMAO has little effect on the strength of hydrophobic interactions between small hydrophobic compounds, e.g., methane (CH₄) [4, 5], whereas this osmolyte has a non-zero effect on larger non-polar molecules, e.g., neopentane (C₄H₁₂) [68, 157]. Effects on the latter depends on the force-field used to simulate TMAO. In particular, studies using a five-site model for neopentane reported that the penalty for transferring this non-polar molecule to water was reduced and increases when simulations were performed using Kast and Netz force-fields for TMAO, respectively [68]. This implies that hydrophobic interactions can be made stronger or weaker with the addition of TMAO to the solution depending on simulation parameters. Effects of Kast TMAO on the interaction between a pair of neopentane molecules modeled as a simple Lennard-Jones sphere have also been studied showing that this osmolyte weakens hydrophobic interactions [144]. Recently, Kast model was shown to reproduce experimental properties of binary water-TMAO solutions [124]. However, Netz model may provide a more realist description of proteins in water-TMAO solutions as it was calibrated to reproduce experimental transfer free-energies of poly-glycine from pure water to TMAO solutions [167].

In Chapter 4, we highlight the effects of TMAO on the molecular forces stabilizing native protein structures by studying small peptides and the Trp-cage miniprotein. Consistent with other studies [82, 37], we find that TMAO favors compact conformations of a peptide model of the protein backbone, i.e., poly-glycine. However, the addition of even the smallest non-polar side chain (i.e., $-CH_3$ group of alanine) to the backbone counteract this effect while larger non-polar side chains account for peptide swelling. This suggests that TMAO can destabilize the hydrophobic core of proteins. We also study conformations of non-polar peptides with charged terminal residues. We find that these peptides become more compact when TMAO is added to water due to stronger interactions between charged residues. In light of these results, we hypothesize that competition of TMAO's effect on hydrophobic and charged interactions accounts for its net stabilizing role in proteins. Extensive replica exchange molecular dynamics (REMD) simulations of the small Trp-Cage protein in pure water and TMAO solution are performed to test this hypothesis. Accordingly, we find that residues that form the hydrophobic core of Trp-cage sample more extended conformations while distances between its charged residues decrease when TMAO is added to water.

CHAPTER 2

DRIVING β -STRANDS INTO FIBRILS

2.1 Introduction

The hierarchical organization of building blocks into complex supra-structures is a recurrent theme in biology. At the molecular level α -helices and β -strands are building units that pack into globular shapes forming the native state of proteins [146, 145]. The length of these units and their packing topology accounts for the diversity of known protein folds. The formation of protein cross- β or fibril structures is a simpler case of hierarchical organization in which β -strands building blocks are identical [56, 71, 137, 148]. These conformations have received a lot of attention as they are linked to diseases like Alzheimer's and Parkinson's [55]. Despite these studies, it is still not clear what are the forces driving peptides into β -sheets and fibrils, and how they differ from the ones driving residues into globular proteins [166]. This is of fundamental importance to understand the pathology of diseases and it is the focus of the present work.

Under appropriate experimental conditions, X-Ray patterns corresponding to cross- β structures have been identified for proteins with seemingly unrelated amino acid sequences [60, 59, 164, 194]. This ubiquity suggests that backbone properties (which are common to all proteins) could be responsible for fibrils [60]. Accordingly, energetic models of the Protein Data Bank (PDB) have shown that the free-energy landscape for misfolding is dominated by interactions involving backbone atoms [65, 33]. In this framework, the role of sidechains is to modulate the propensity of fibril formation [36]. In particular, experiments have shown that mutations accounting for an increase in the concentration of non-polar residues in the fibril core have a higher fibrillization rate [141, 45, 70] while an increase in the net peptide charge has the opposite effect [35]. This has lead to the formulation of the amyloid self-organization principle according to which fibril stability is enhanced by maximizing the number of hydrophobic and favorable electrostatic interactions (including salt bridges and hydrogen bonds) [190, 189, 175]. Despite these insights, the question of how backbone and sidechain atoms contribute to fibrillization is still open.

Thermodynamics provides a quantitative framework to study the forces driving conformational changes in proteins [2]. These phenomena are often described by a two-state reaction equation with an equilibrium constant, K, that can be measured and used to compute differences in Gibbs free energy, $\Delta G = -\text{RT} \ln(K)$ [150]. These free energy differences result usually from the sum of large opposing terms. For example, in protein folding at ambient conditions the entropic energy $(-T\Delta S)$ favors the unfolded state while enthalpy (ΔH) favors the native state—each term contributing ~ 100 - 200 kcal/mol and resulting in $\Delta G \sim 10$ kcal/mol [23, 24, 150]. Main contributions to $-T\Delta S$ comes from mainchain $(-T\Delta S_{\text{mainchain}})$ and water molecules around the protein $(-T\Delta S_{\text{water}})$ which favor, respectively, unfolded and native states. ΔH emerges from changes in energy due to covalent-bonds (ΔH_{CB}) and non-covalent-bonds (ΔH_{NCB}) . Thus,

$$\Delta G = -T\Delta S_{\text{mainchain}} - T\Delta S_{\text{water}} + \Delta H_{\text{CB}} + \Delta H_{\text{NCB}}.$$
(2.1)

The four terms in the right-hand-side of Eq. 2.1 are not directly accessible experimentally. This has led to controversies in molecular interpretations of measured $-T\Delta S$ and ΔH . For example, in protein folding at ambient conditions, the entropic energy of water molecules around non-polar groups, which gives rise to hydrophobic interactions [66], is believed to be the main force driving protein folding [98, 53, 107]. However, $-T\Delta S$, which is the entropic quantity that can be measured experimentally, appears unfavorable to the folded state. This is because it is dominated by $-T\Delta S_{\text{mainchain}}$. Thus, studies aiming to describe protein folding have shifted back and forth between sidechain and backbone centered views [118, 161, 21].

In protein folding, the importance of hydrophobic interactions can be inferred from the positive curvature of ΔG with respect to temperature [150, 98, 51, 13] which is characteristic of non-polar solvation and accounts for heat and cold denaturations of proteins [50, 49, 52]. Also, the diversity of native folds can only be encoded in the amino acid sequence, suggesting that sidechain properties and, in particular, the burial of non-polar sidechains in the dry protein core has to be responsible for folding. Currently, it is still a question of debate how intra- and inter-peptide hydrogen bonds contribute to secondary-structure formation [12]. It is being argued that hydrogen bonding are stabilizing [69, 160] and destabilizing [119, 205]. This question is of fundamental importance to understand fibril formation as inter-peptide hydrogen bonds are maximized in these structures.

Here we study fibril formation using extended poly-alanine peptides as unit blocks for aggregation and an umbrella sampling protocol to compute free energies to form peptide dimer, trimer, and tetramer. The latter can be considered the smallest repeating unit of a fibril. We find that Lennard-Jones and electrostatic energies of chemical groups in the protein and solvent are one order of magnitude larger than the overall enthalpy of the system. Thus, small errors in modeling these interactions can account for large errors in the total enthalpy of the system—highlighting the need for accurate models. We show that the total average number of hydrogen bonds in the system does not change during aggregation as the result of a compensatory mechanism where the formation of one inter-peptide hydrogen bond accounts for the rupture of two peptide-water bonds and the release of water molecules from the neighborhood of the peptide leading to the formation of one extra water-water bond [98, 62, 61, 185, 94]. A consequence of this compensation is that electrostatic energies related to hydrogen bonds are not minimized during fibril formation. Thus, hydrogen bonds do not drive fibril formation. However, we argue that inter-peptide hydrogen bonds play an important role in fibril formation since aggregation without the formation of these bonds is energetically prohibitive. Both mainchain and sidechain atoms contribute actively to minimize Lennard-Jones interactions during fibril formation. This knowledge of how different chemical groups of the protein contribute to minimize the energy of the system is of fundamental importance to develop strategies to inhibit fibrilization related to diseases and to develop better coarse-grain models of proteins.

2.2 Methodology

The system studied in this work consists of up to four poly-alanine peptides, i.e., ALA_{10} , immersed in a periodic box containing 5,500 TIP3P water molecules (0.03 M peptide). Poly-alanine peptides have been shown experimentally to aggregate and to form fibrils at conditions of 10 μ M peptide, pH 7, 0.1 M salts at 25°C [17]—see also [59]. In our simulations, peptides are made "infinite" through the use of periodic boundary conditions by attaching the carbonyl-group of residue 1 to the amine-group of residue 10 in the z-direction. The use of "infinite" peptides eliminates effects from chain ends causing all residues to be equivalent and to resemble amino acids in the middle of a β -strand. A pressure of 38 bar is applied along the z-direction to keep the box from collapsing. The magnitude of this pressure is chosen to ensure an average peptide length of ~ 3.5 nm. Therefore, peptides are stretched in our simulations. Notice that the main energetic term opposing aggregation in Eq. 2.1 is the reduction in the entropy of the mainchain. This "dissociation" force is not taken into account in our simulations since mainchain entropies of stretched peptides are essentially the same in all states implying that $-T\Delta S_{\text{mainchain}} = 0$ in Eq. 2.1. Our simulation setup is designed to describe "aggregations" forces, i.e., changes in the enthalpy of peptide/water and entropy of water ($\Delta H_{\rm CB}$, $\Delta H_{\rm NCB}$, and $-\Delta S_{\rm water}$ in Eq. 2.1) when solvated amino acids are brought to interact with each other.



Figure 2.1 Schematic representation of the reaction coordinate used to study the formation of peptide dimer (a), trimer (b), and tetramer (c).

A pressure of 1 atm is applied along x and y directions to account for water density at ambient pressure. Simulations are carried out using GROMACS and the Amber99-sd-ildn forcefield [113]. Temperature (298 K) and pressure are controlled using the v-rescale thermostat ($\tau_{\rm T} = 1$ ps) and the Parrinello–Rahman barostat ($\tau_{\rm P} =$ 1 ps), respectively. A time step of 2 fs is used and the neighbor list is updated every 10 steps. Electrostatics is treated using the Smooth Particle Mesh Ewald method with a grid spacing of 0.13 nm and a 1.3 nm real-space cutoff [20].

To determine the free-energy landscape of peptide tetramer formation, we perform three sets of simulations containing two, three, and four peptides, respectively. In the first set, peptides are arranged in an anti-parallel orientation and an umbrella sampling protocol is used to sample the space given by the distance between centers-of-mass of peptides (ξ_2 in Figure 2.1*a*) in the range 0.4 nm to 2.0 nm. Different windows in which peptides are restrained by a spring that has a constant of 5,000 kJ mol⁻¹ nm⁻² are simulated. Equilibrium distances of springs in neighboring windows differ in steps of 0.05 nm. Each window is simulated for 100 ns and the Potential of Mean Force (PMF) to form a dimer is computed using the Weighted Histogram Analysis Method (WHAM) [84]. In coarse-grained simulations, fibrillization was found to start with the formation of anti-parallel structures follow by a transition to parallel conformations after enough peptides are added to β -sheets [10, 139]. This suggests that parallel β -structures are less stable than anti-parallel ones when sheets are formed by a small number of peptides while the opposite is expected for large sheets. For poly-valine, the number of peptides required to stabilize parallel β -sheet was found to be 14 [10].

To perform the second set of simulations, a peptide is added to the simulation box of the dimer. In these simulations, a spring is used to restrain centers-of-mass of the dimer at a distance corresponding to the minimum of its PMF. Umbrella sampling simulations are repeated for this new system to sample the distance between centers-of-mass of a reference peptide in the dimer and the additional chain (ξ_3 in Figure 2.1*b*). Configurations from these simulations are used to compute the PMF of trimer formation using WHAM. At last, the trimer is restrain to the configuration corresponding to the minimium in its PMF using two springs connecting a reference peptide to the other two chains. A peptide is added to the simulation box which now comprises two sets of anti-parallel peptides. Umbrella sampling simulations are performed using the distance between the reference peptide in the trimer and the added chain as the new reaction coordinate (ξ_4 in Figure 2.1*c*). The PMF to form a tetramer is computed using WHAM.

To define hydrogen bonds, we employ a commonly used geometrical definition in which these bonds are formed when the distance between hydrogen (H) donor (D) and acceptor (A) is smaller than 0.4 nm and the angle H-D-A is smaller than 30°. In the calculation of quantities involving solute–solvent and solvent–solvent atoms, all solvent (water) molecules were taken into account. In the calculation of spatial distribution functions of water—see Figure 2.4—the simulation box was divided in bins of length 0.02 nm. Spatial distribution functions are given in units of the ratio of the density of water in the simulation and the density of water of an ideal fluid for each bin.

2.3 Results

In Figure 2.2*a*, we show PMF(s) of peptide dimer, trimer, and tetramer. These three PMF(s) are computed separately as described in the methodology section and the continuous reaction coordinate ξ is obtained by concatenating inverted order parameters of peptide dimer ($\xi_2^{\text{new}} = 2.0 - \xi_2$), trimer ($\xi_3^{\text{new}} = 2.0 - \xi_3$), and tetramer ($\xi_4^{\text{new}} = 2.0 - \xi_4$). In this procedure, the PMF of the peptide trimer at $\xi_3^{\text{new}} = 0.0$ is shifted to match the minimum in the PMF of the peptide dimer. Similarly for the tetramer: its PMF at $\xi_4^{\text{new}} = 0.0$ is shifted to match the free-energy minimum of the trimeric system. This concatenation procedure describes the assembly of fibrils through monomer addition [38, 79] where peptide dimer is formed first followed by docking of a third peptide and subsequent tetramer formation.

In Figure 2.2b we show how numbers of hydrogen bonds in the system change during peptide dimer, trimer, and tetramer formation. The number of hydrogen bonds for isolated peptides is used as a reference, i.e., $N_{Hbond}(\xi = 0) \equiv 0$. For all values of ξ , the total (or net) number of hydrogen bonds does not change significantly. A similar behavior was reported recently for β -hairpin formation of GB1 peptide [134] and β -sheet formation of model peptides made of glycine, alanine, value, and leucine residues [133]. To understand this result in more detail, we decompose the total number of hydrogen bonds into contributions from peptide-peptide, peptide-water, and water-water bonds. We observe an almost perfect compensatory mechanism [98, 62, 61, 185, 94] where the formation of one peptide-peptide hydrogen bond is preceded by the rupture of two peptide-water bonds accounting for water release into the bulk and the formation of one additional water-water bond. Notice that during peptide trimer formation (at $\xi\sim 2.8~{\rm nm})$ the third peptide docks onto the dimer (i.e., β -sheet) without forming inter-peptide hydrogen bonds—see Figure 2.1. This process involves displacement of water molecules from the space between sidechains in the β -sheet (see "HB" configuration in Figure 2.4*a*) to the bulk during docking. In this



Figure 2.2 Potential of Mean Force (a), number of hydrogen bonds (b), and potential energy (c) to assemble β -strands in a hierarchical process. These quantities were computed with respect to their values at non-interacting peptide conformations ($\xi = 0$). Number of hydrogen bonds was divided into contributions from peptide-peptide (pp), water-water (ww), and peptide-water (pw). The net number of hydrogen bonds is also shown. The potential energy was divided into electrostatic and van der Waals contributions. Dashed lines separate simulations of dimer (left), trimer (middle), and tetramer (right). Error bars were estimated from block averages by dividing the simulation in five blocks.

case, rupture of peptide-water hydrogen bonds during docking of the third peptide is compensated by newly formed water-water bonds. The generality of the observed hydrogen bond compensation for peptides with different amino acid sequences and situations might be due to the small size and polarity of water which can penetrates small cavities to saturate non-satisfied hydrogen bonds.

In all-atom models, hydrogen bonds emerge from electrostatic interactions involving X–H···Y chemical groups, where X and Y are electronegative atoms and H is hydrogen. Thus, if hydrogen bonds are a main force driving aggregation, the electrostatic energy is expected to correlate with the PMF, i.e., it should decrease whenever the PMF become a minimum. In Figure 2.2*c*, we show the dependence of the electrostatic energy of the system on the reaction coordinate ξ . It does not correlate with the PMF being mostly indifferent to tetramer formation. In contrast, the energy due to Lennard-Jones interactions is favorable to peptide aggregation. As a result, the sum of Lennard-Jones and electrostatic energy, i.e., non-bonded interactions, is favorable to aggregation.

Figure 2.3 quantifies contributions of different chemical groups to changes in electrostatic and Lennard-Jones energies. Panel *a* shows a compensatory mechanism for electrostatic interactions where unfavorable mainchain-water interactions are balanced by favorable water-water and mainchain-mainchain interactions. Since Nitrogen and Oxygen are mainchain atoms contributing the most to the peptide's electrostatic energy, the observed compensatory mechanism for the electrostatic energy can be mapped to the formation of one inter-peptide hydrogen bond which requires breakage of two peptide-water hydrogen bonds and subsequent formation of one water-water hydrogen bond. Moreover, the total change in electrostatic energy emerging from these contributions is approximately zero—as shown in Figure 2.2*c*. Hence, it is not expected to drive aggregation. However, a hypothetical process in which mainchain NH– and CO– groups are buried away from water without forming inter- or intra-peptide hydrogen bonds would increase the electrostatic energy by a prohibitive large amount, i.e., $E_{Sol-Main}^{elect.} + E_{Sol-Sol}^{elect.} \sim 9.4 \text{ kJ/mol/residue}$. This suggests that a main role of inter- and/or intra-peptide hydrogen bonds is to penalize structures for which the overall number of hydrogen bonds in the system is not optimized.

Notice that in implicit water models the formation of mainchain hydrogen bonds is taken into account by a decrease in the energy of the system [174]. This favors peptide conformations with optimized secondary structures. This approach was successful in predicting the structure of various non-amyloid [200, 31, 126, 125, 163] and amyloid [30, 54] peptides. But, this description of hydrogen bonds does not mimic the energetics of explicit water simulations as described above. Thus, there is a trade-off between being accounting for the entropic contribution of the backbone using implicit water coarse-grained models and describing with greater accuracy non-covalent interactions using all-atom models. A successful strategy could be to combine these two approaches [135].

Figure 2.3*b* shows how burial of sidechain and mainchain atoms away from water during aggregation affects Lennard-Jones interactions. In this process, water molecules are transferred to the bulk accounting for a modest decrease in Lennard-Jones energy. This change is comparable to the reduction in Lennard-Jones energy due to sidechain-sidechain interactions. Main contributions to changes in Lennard-Jones energy are due to water-mainchain, water-sidechain, mainchain-mainchain, and mainchain-sidechain interactions. Breaking bonds between water and peptide atoms, i.e., water-mainchain and water-sidechain bonds, accounts for a large increase in Lennard-Jones energy. However, this process is overcompensated by the favorable formations of new bonds between atoms of the peptide, i.e., mainchain-mainchain and mainchain-sidechain. This leads to a net Lennard-Jones energy that is favorable to aggregation—as shown in Figure 2.2*c*.



Figure 2.3 Contribution of different chemical groups of the protein and solvent to tetramer formation. Electrostatic (a), Lennard-Jones (b), and non-bonded-interactions (c) are decomposed into solvent-solvent (black), solvent-mainchain (red), solvent-sidechain (green), mainchain-mainchain (blue), mainchain-sidechain (orange), sidechain-sidechain (brown) contributions.

Figure 2.3c shows the sum of Lennard-Jones and electrostatic energies, i.e., non-bonded energies, for interactions between the different chemical groups of the system. The overall non-bonded energy is favorable to aggregation (as displayed in Figure 2.2c) and Figure 2.3c shows that it emerges from large contributing terms that have opposite signs. Contributing terms are one order of magnitude larger than the overall non-bonded energy. This highlights the importance of using accurate models since small errors in modeling the strength of one bond can produce large errors in the total enthalpy of the system.

In Figure 2.4 we show the spatial distribution of water around ground states of peptide dimer (panel a), trimer (b), and tetramer (c). First column corresponds to a cross-sections of the peptide's main axis. In an anti-parallel β -sheet, residues along a strand can either be hydrogen-bonded to the neighboring strand or non-hydrogenbonded [32, 87]. Sidechains of hydrogen-bonded residues all face the same direction which is called the hydrogen-bonded face (HB) of the β -sheet. Similarly defined is the non-hydrogen-bonded face (NHB). A detailed analysis of water distribution around a β -sheet for different non-polar amino acids was reported recently [196]. It showed that water distribution is different at HB and NHB faces. This result was particularly striking for leucine and value residues. For alanine, polar groups of the mainchain are more exposed to water at the NHB face compared to HB. Thus, water molecules were shown to penetrate deeper in the space between sidechains at the NHB face. This result is visible in the cross-section view of dimers and tetramers where iso-surfaces are more pronounced and penetrate deeper between sidechains at the NHB face. At the HB face, distributions of water in addition of being perpendicular to the axis of the peptide also have a parallel component which enhances electrostatic interactions between water and polar groups of the mainchain that are partially buried between C_{α} atoms. The lateral view (last column in Figure 2.4) shows a more pronounced distribution of water facing NH– groups of the mainchain compared to CO– groups.



Figure 2.4 Iso-surface of water distribution around the ground state of dimer (a), trimer (b), and tetramer (c). Columns correspond to cross-section of the peptide's axis, non-hydrogen-bonded face (NHB), hydrogen-bonded face (HB), and lateral view. Isovalues of dimer, trimer and tetramer are 5.04278, 5.04278, and 10.5715.

This result was studied in detail in reference [196] and it was related to the hydrogenreceiver nature of the CO– group which allows hydrogen-bonding with water with a greater angular freedom.

2.4 Conclusion

In summary, to investigate the energetics of fibril formation, we performed extensive molecular dynamics simulations of poly-alanine in water. To understand which chemical groups of the peptide favor fibril formation, we decomposed non-bonded interactions into contributions from mainchain and sidechain atoms. We show that changes in the energy of the system due to mainchain atoms play a major role in this process. We used a geometrical definition to compute the number of hydrogen bonds in the system. We show that due to a compensatory mechanism, the total number of hydrogen bonds in the system does not change significantly during fibrillization [62, 98, 94, 185]. Furthermore, main changes in the electrostatic energy of the system are related to hydrogen bonds and, as a result of the compensatory mechanism, it does not decrease during fibrillization. Therefore, while fibril formation accounts for an increase in the number inter-peptide hydrogen bonds there is no apparent energy gain in the formation of these bonds. This leads to the question of what is the energetic role of inter-peptide hydrogen bonds [12]

We argue that peptide aggregation without the formation of inter-peptide hydrogen bonds produces a large electrostatic penalty. Thus, in all-atom simulations secondary structures do not form to minimize energetic terms associated with hydrogen bonds but to avoid the energetic penalty of having non-satisfied polar groups pointing towards the dry core of the protein. This implies that disordered configurations with polar groups exposed to water and secondary-structure configurations might not be very different with respect to the energy of hydrogen bonds. Therefore, the main role of hydrogen bonds is to reduce the number of peptide conformations and, in particular, the number of compact peptide structures. This result could have important implications for the development of coarse-grained models.

Moreover, if we assume that energies to form sidechain-sidechain and mainchainmainchain hydrogen bonds are similar then our results suggest that burial of nonpolar sidechains in the protein core without the formations of hydrogen bonds could be subjected to large penalties. Thus, conformational changes of proteins after the formation of a dry core would involve little change in the number of both sidechain and backbone hydrogen bonds since this requires overcoming large energy barriers. This is consistent with reported results from coarse grained simulations in which transitions between β -barrel structures and fibrils for the polar NHVTLSQ peptide occurred with little variation in the number of hydrogen bonds [172]. In addition, the penalty of having non-saturated hydrogen bonds in the protein core could be responsible for faster fibrillization rates of polar sequences compared to non-polar ones [117], since the formation of sidechain hydrogen bonds would occur promptly after the creation of a dry core in the case of polar sequences while a larger (and more frustrated) phase space has to be sampled before in-registry fibrils can form in non-polar peptides. In the case of trans-membrane proteins, it has been reported that hydrogen bonds between polar sidechains and the backbone play an important role in the dynamics and stability of α -helical structures [165]. This is also consistent with our results which associates an energetic penalty to polar sidechains exposed to membranes that can, however, be minimized through sidechain-mainchain hydrogen bonding.

Despite these conceptual and quantitative contributions, limitations of the current work should also be noted. While all-atom models have been optimized over the years to account for folding of several proteins [112, 169], there are still uncertainties regarding force-field parameters. In particular, all-atom models do not account for context dependent interactions whereas experiments are suggesting that the strength of hydrogen bonds could be dependent on the environment [48, 47, 93, 92, 69]. Moreover, the setup used in this work only considers stretched peptide structures which is a geometry that facilitates the observed compensatory mechanism for hydrogen bonds. In contrast, when peptides are not interacting they can assume conformations for which the total number of hydrogen bonds in the system is not optimized. While results on GB1 model peptide are showing that the compensatory mechanism is also valid for unconstrained structures [134], this requires further investigation.



Figure 2.5 Potential of mean force of two parallel and two antiparallel polyalanine peptides at T = 298 K, $P_z = 38 \text{ atm}$, and $P_x = P_y = 1 \text{ atm}$.

2.5 Associated Content

In coarse-grained simulations of 18 polyvaline peptides, it was shown that order emerged from a disordered oligomer through the formation of antiparallel β -sheets. When a relatively large ordered nucleus of antiparallel β -sheets was formed, a few parallel sheets started to appear. The growing number of parallel β -sheets led eventually to a decrease in free-energy leading towards a minimum which was predominantly composed of parallel β -sheets [10]. A similar but more complex nucleation mechanism in which antiparallel -sheet formed easily in the beginning of a coarse-grained simulation followed by a transition to parallel β -sheet was also observed in simulations of $A\beta_{35-40}$ [11]. These studies suggest that antiparallel β conformations are more stable than parallel ones when a small number of peptides are involved in the sheets while for large sheets (involving 14 peptides in the case of polyvaline [10]) parallel β structures could be more stable [138].

In Figure 2.5, we compare PMF(s) for the interaction of two parallel and two antiparallel polyalanine peptides. The PMF for the antiparallel system was



Figure 2.6 Structure of two parallel polyalanine peptides at $\xi = 0.47$ nm.



Figure 2.7 Dependence of the number of hydrogen bonds on the distance ξ between centers-of-mass of peptides.

taken from Figure 2.6 of the manuscript and the umbrella sampling simulation for parallel peptides was performed as described in the method section of the manuscript. We find that antiparallel peptides are more stable than parallel ones by 0.75 kJ/mol/residue. It implies that they should occur more frequently in unconstrained simulationsconsistent with coarse-grained studies [10, 138]. Furthermore, an analysis of the structure at the global minimum ($\xi = 0.47$) reveals that parallel peptides do not form stable β -sheets with an hydrogen bond pattern extending continuously along the peptide—see Figure 2.6. Methyl groups of three (out of 10) sidechain-pairs point towards the backbone of the other peptide. We have made several attempts to create a stable parallel β -sheet. For example, we started with an ideal initial parallel β -sheet structure which after 30 ns became unstable producing structures like the one in Figure 2.6. In other simulations, we increased the pressure in the z-direction from 38 atm to 128 atm since parallel β -sheets are more extended that parallel one. However, all our attempts resulted in structures similar to Figure 2.6. Thus, we are confident that parallel β -sheet of polyalanine peptides are not stable in our set-up. Despite the lack a stability of parallel β -sheets, the total number of hydrogen bonds in the system does not change significantly as a function of the distance ξ between centers-of-mass between peptides—see Figure 2.7. In Figure 2.7, the formation of inter-peptide hydrogen bonds is preceded by the rupture of one peptide-water bond leading to the formation of one extra water-water hydrogen bond. Notice that in the global minimum, the average number of inter-peptide hydrogen bonds is 3 (out of 10) which is consistent with the structure shown in Figure 2.6.

CHAPTER 3

MOLECULAR INTERACTIONS ACCOUNTING FOR PROTEIN DENATURATION BY UREA

In Chapter 2, we use extended peptide models to study the mechanism of fibrillization of polyalanine peptides and compute free-energies required to form dimers, trimers and tetramers. In this chapter, we apply this extended model to investigate roles of urea on the interaction between extended non-polar peptides.

3.1 Introduction

One of the first scientific accounts of urea's effect on living systems dates back to 1902 when it was reported that dead frogs become translucent and fall into pieces in a saturated urea solution [151]. This molecule is now widely used as a denaturant in the lab whereby it perturbs native protein conformations to favor the unfolded state [78, 26]. However, urea's atomic mechanism has been a question of debate since its effect was quantify in the 60's [173]—although important insights have been obtained recently [76, 26, 81, 58, 80]. Urea may impede the formation of native structures either by forming strong bonds with the protein (direct mechanism) [28, 178, 83, 81, 109, 156, 123, 105, 130, 46] or by modifying the structure of water accounting for different solvation properties (indirect mechanism) [67, 15, 88, 44, 63]. Moreover, it is still not clear whether urea affects mainly the backbone or side chains of the protein. Answers to these open questions can be obtained using simplified computational frameworks that can quantify interaction strengths in the solvent and protein at the atomic level.

In recent years, several studies have emerged showing that, in solution, urea can replace water without changing significantly its structural properties since both molecules are highly polar [177, 173, 95, 155, 102, 83, 9, 90]. This has been a strong argument against urea's indirect mechanism which assumes that the formation of water structures is disrupted by this molecule [67, 129]. Urea's direct mechanism is, therefore, gaining increased acceptance in the scientific community. In particular, it is supported by explicit all-atom computer simulations [28, 178] wherein urea is shown to form strong non-bonded interactions with the protein providing the driving force for unfolding. Some studies indicate that these protein-urea interactions emerge mainly from hydrogen bonding with the backbone [18, 109, 57, 123, 9] whereas others point to dispersion interactions with non-polar groups of the proteins [159, 206, 102, 158, 80]. Separating contributions from backbone and side chain atoms is not an easy task even in computer simulations that have access to atomic detail [28, 81, 18].

Traditionally, model compounds have been used to estimate interaction strengths between different chemical groups in proteins [177, 156, 140, 170, 193, 179]. In particular, free-energies to transfer non-polar amino acids from pure water to urea solutions were shown to be negative implying that urea solutions can accommodate non-polar amino acids better than pure water [140]. As a result, hydrophobic interactions are weaker in urea solutions and this could account for protein denaturation. Small non-polar compounds (e.g., methane and alanine) as well as glycine were not found to be significantly affected by urea [140]. Computer simulations of model compounds point to the same conclusion [170, 193] but also highlight strong hydrogen bonding of urea with small charged solutes [57].

Different conclusions can be drawn from studies of simple model compounds because they do not account simultaneously for contributions from backbone and side chain atoms in the same proportion as in real proteins. Moreover, solvation of side chains are affected by the backbone as reported in a recent study [192, 77]. These limitations of model compounds can be overcome by studying homopeptides [81, 179, 102, 130]. Recently, Horinek and Netz computed free-energies τ to transfer homopeptides from pure water to aqueous solutions containing urea using all-atom molecular dynamics simulations and different force-fields [81]. Direct and indirect effects of urea were singled out by differentiating τ with respect to urea concentration. This provided important insights into effects of force-fields on urea's mechanism [81].

In this paper, we study effects of urea on non-restrained poly-glycine, polyalanine, and poly-leucine monomers as well as on the interaction of extended polyalanine and poly-leucine dimers. For the extended dimers, we compute the free-energy required to bring two peptides that are initially non-interacting all the way to a distance ξ , i.e., we compute potentials of mean force (PMF). Our choice of peptides and approach allows us to probe the effect of urea on two types of interactions that are essential to account for protein folding, i.e., hydrophobic interactions between side chains of alanine and leucine residues and interactions between backbone atoms. Notice that, in addition to non-polar residues, urea has also been shown to bind favorably to polar and charged residues [177, 178]. These types of residues are mainly located at the surface of proteins and, therefore, their energy of solvation is not expected to contribute significantly to protein unfolding.

Here, we show that poly-leucine monomers become less compact when urea is added to water whereas poly-glycine and poly-alanine monomers are only weakly affected by this co-solvent. Consistent with this result, changes in the PMF of extended poly-alanine and poly-leucine dimers when urea is added to water agree qualitatively with transfer free-energies of their respective side chains, i.e., only the PMF of poly-leucine dimers is destabilized (Figure 3.5). Also, we show that urea molecules are more densely populated around poly-leucine's side chains than in the bulk and this difference increases when peptides dissociate (Figure 3.7). This occurs because urea molecules can replace water in the first shell around peptides and they can occupy positions that are not accessible to water due to their planar distribution of partial charges [42, 100] (Figure 3.7). Thus, our results suggest that urea destabilize proteins through a direct mechanism whereby it binds favorably to side chains. Moreover, we observe a clear correlation between changes in non-bonded interactions (i.e., enthalpy) and changes in the PMF when urea is added to the system (Figure 3.8). These changes when peptides are brought close together emerge from contributions of protein-urea interactions (mostly Lennard-Jones) that favor the dissociated state and solvent-solvent interactions (electrostatic and Lennard-Jones) that favor the folded state (Figure 3.10).

3.2 Methodology

Microsecond-long molecular dynamics simulations are used to study unrestrained poly-glycine, poly-alanine and poly-leucine monomers in pure water and 3M urea solution. Two independent simulations, starting from different stretched conformations of the monomer, are performed for each combination of solvent and peptide. Systems are equilibrated for 60 ns, and they are analyzed along 940 ns trajectories. Simulations are performed in the NPT ensemble at 1 atm and 298 K.

Extended poly-alanine and poly-leucine dimers in this work are made of 10 residues which are immersed in pure water and aqueous urea solutions. Peptides are made "infinite" through the use of periodic boundary conditions by attaching the carbonyl-group of residue 1 to the amine-group of residue 10 in the z-direction. This method to describe extended peptides was first used by Horinek and Netz [81]. Advantages of this method are that it eliminates effects from chain ends and all amino acids are equivalent. To maintain peptides stretched at an average length of ~ 3.5 nm, pressures of -38 atm and -100 atm are applied along the z-direction for poly-alanine and poly-leucine, respectively. A pressure of 1 atm is applied along x and y directions accounting for an average density of 990 kg/m³. Thus, the anisotropic pressure used in this work ensures that peptides are stretched and their solvation resembles experimental conditions [133, 179].

To determine the free-energy landscape of extended dimers, we use an umbrella sampling protocol where the reaction coordinate is the x-y distance between centers-



Figure 3.1 x-y components of the distance vector $\vec{\xi}$. Trajectories of different umbrella-sampling windows are represented using different colors.

of-mass of peptides (ξ) in the range 0.4 nm to 2.0 nm. We simulate different windows in which peptides are restrained by a spring that has an equilibrium constant of 5,000 kJ/mol/nm². The equilibrium distances of springs in neighboring windows differ in steps of 0.05 nm. Each window is simulated for 150 ns. In Figure 3.1, we fix the position of one peptide at the origin and we show x and y components of the distance vector $\vec{\xi}$ for trajectories at the different windows. This figure shows that the twodimensional space defined by $\vec{\xi}$ is properly sampled in our simulations. Trajectories at the different windows are used to compute the "raw" PMF, i.e., PMF^{raw}, using the Weighted Histogram Analysis Method (WHAM) [84]. This quantities increases linearly with ξ . In this work we subtract this linear dependence on ξ and we use the average PMF, i.e., $\langle C \rangle$, computed in the ξ range 1.8–2.0 nm to define the reference PMF (i.e., its zero value):

$$PMF(\xi) = PMF^{raw}(\xi) + k_b T \log(\xi) - \langle C \rangle.$$
(3.1)

A detail account of the normalization procedure of the PMF is given in reference [43]. 38 windows were used in our umbrella sampling simulations of poly-leucine. Each one of these windows was sampled for 150 ns accounting for an accumulated simulation time of 5.7 μ s. The distribution of ξ in these windows is shown in Figure 3.2 for 7M urea solution. Significant overlap between neighboring distributions is observed which ensures repeated movement of the protein in the unbias simulations. In Figure 3.3, we show the PMF for the 7M urea solution computed using different time intervals. Clear convergence of the PMF is observed. Moreover, error-bars computed using the bootstrap method is shown in Figure 3.3. Table 3.1 shows details of the different systems studied.



Figure 3.2 Histogram of ξ for 38 windows in our 7M urea simulations.

Temperature (298 K) and pressure are controlled using the v-rescale thermostat $(\tau_{\rm T} = 1 \text{ ps})$ and the Parrinello–Rahman barostat ($\tau_{\rm P} = 1 \text{ ps}$), respectively. Simulations are carried out using GROMACS and the AMBER99SB-ildn-force-field [113]. We use the TIP3P model for water. Urea, $(NH_2)_2CO$, is a highly polarized molecule with a dipole moment 6-8 Debye (D) in liquid state [97] (the value for water is 2.95 D [75]). Therefore, urea can easily form hydrogen bonds with water molecules [128].


Figure 3.3 Convergence of the PMF for poly-leucine in 7M urea solution. PMF are computed using different time interval.

For urea we used the AMBER-force-field [147], as shown in Table 3.2. We use the leap-frog algorithm with a time-step of 2 fs to integrate the equation of motion and the neighbor list is updated every 10 steps. Electrostatics is treated using the Smooth Particle Mesh Ewald method with a grid spacing of 0.13 nm and a 1.3 nm real-space cutoff [20].

Table 3.1 System Details for Poly-leucine in Pure Water and Urea Solutions. C_{urea} Denotes the Concentration of Urea, N_{urea} and N_{water} Denote the Number of Urea and Water Molecules of the System

C_{urea}	N_{urea}	N_{water}	box size (\mathring{A})			
0M	0	2638	4.94 4.97 3.38			
2M	96	2417	4.98 4.98 3.44			
3.8M	201	2417	5.18 5.18 3.44			
7M	351	1912	5.14 5.14 3.40			

Atom	Partial Charge	Lennard-Jones parameter		
		$\sigma/arepsilon$		
С	0.88	3.34 / 0.360		
0	-0.613	2.96 / 0.879		
N	-0.924	3.25 / 0.711		
Н	0.395	1.07 / 0.066		

 Table 3.2
 Parameters of the AMBER-force-field for Urea

Stretched fragments of proteins are of fundamental importance in biology [145] as they form the building blocks of β -sheets. After α -helices, β -sheets are the most frequent secondary structures of globular proteins serving as template for 20–28% of all the residues [22]. In our setup, the two peptides in the simulation box run in opposite directions, i.e., the displacement vector between carbonyl- and aminegroups point in the positive z-direction for one peptide and in the negative direction for the other. As a result, the type of β -sheet structures adopted by peptides in our simulations is called antiparallel. β -sheet structures formed by peptides that run in the same direction are called parallel. In globular proteins, the later occur less frequently than antiparallel β -sheet structures and they will not be studied in this work. Both poly-alanine and poly-leucine have been found to form fibrils [59] which are stabilized by hydrogen bonds and side chain interactions from facing and surrounding residues. When a β -sheet is formed in our model, it is stabilized by hydrogen bonds and chain interactions from surrounding residues.

3.3 Results

In Figure 4.3, we study effects of urea on the conformation of unrestrained polyglycine, poly-alanine, and poly-leucine peptides by measuring the radius of gyration R_g of backbone atoms. Poly-glycine is a model of the protein backbone whereas



Figure 3.4 Distribution of the radius of gyration R_g of backbone atoms for polyglycine, poly-alanine, and poly-leucine in aqueous urea solutions at 298 K and 1 atm.

poly-alanine and poly-leucine represent peptides with small and large non-polar side chains, respectively. Figure 4.3a shows that conformations of poly-glycine are not significantly affected by urea: R_g distributions of poly-glycine in pure water (black line) and aqueous urea solution (red) are very similar to each other. Effects of urea are more pronounced for poly-alanine which becomes less compact when urea is added to water. This is shown in Figure 4.3b where the distribution of R_g shifts to larger values when urea is added to water. This shift in the distribution of R_g is even more significant for poly-leucine as shown in Figure 4.3c. This points to the importance of side chains in accounting for urea's destabilization of compact protein conformations. In agreement with Figure 4.3, a recent simulation of a fifteen residue poly-glycine peptide performed using the Kirkwood-Buff force field for urea reported that conformations of this peptide are not significantly affected by urea [80]. In this study, effects of urea on side chains of small proteins were found to be key in accounting for the unfolded state [80].

To shed light into urea denaturation mechanism, we study a simplified system, i.e., extended poly-alanine and poly-leucine dimers in aqueous urea solutions. These systems provide a framework to study effects of urea on atomic interactions responsible for the stability of native conformations, i.e., hydrophobic interactions



Figure 3.5 Potential of mean force (PMF) of poly-alanine (a) and poly-leucine (b) dimers in pure water (red) and aqueous urea solutions (blue). PMF values computed at ξ =2.0 nm are used as our reference, i.e., zero value. Characteristic configurations of dimers are also shown.

and hydrogen bonds. Figure 3.5 shows PMF of poly-alanine and poly-leucine dimers in pure water and in aqueous solutions containing urea. These PMF are characterized by a global minimum at $\xi \sim 0.5$ nm in which peptides assume anti-parallel β -sheet structures. For poly-leucine, PMF also exhibit a second minima at $\xi \sim 0.8$ nm. At this ξ value, we observe interdigitation of side chains with polar groups of the backbone pointing towards the solvent [133]. In addition to these minima, an energy barrier is observed between global and local minima for poly-leucine. Characteristic configurations at the barrier resemble the ones at $\xi \sim 0.8$ nm with side chains interdigitating more tightly.

In Figure 3.5a, the addition of urea does not account for significant changes in the PMF of poly-alanine chains. In contrast, urea weakens the interactions between poly-leucine chains as depicted in Figure 3.5b. Similar results for other concentrations of urea (2M, 3.8M, 5M, and 7M) are shown in Figure 3.6. A possible explanation



Figure 3.6 Potentials of mean force (PMF) of poly-alanine (a) and poly-leucine (b) dimers in pure water and different urea solutions.

for this difference in behavior is that urea affects mostly side chain (not backbone) interactions [25, 26] with side chains of leucine being more affected than side chains of alanine. This idea is supported by simulations of methane (CH₄) and neo-pentane (C₅H₁₂) dimers that are chemical compounds resembling side chains of alanine and leucine, respectively [170]. PMF describing the interaction of methane dimers do not change significantly in the presence of urea [170] whereas PMF of neo-pentane dimers become less negative with increasing urea concentration. If backbone interactions were the main energetic term affected by urea then PMF of both poly-leucine and poly-alanine in Figure 3.5 would have changed by a similar amount. This is not the case suggesting that the backbone is not the main part of the protein affected by urea. This is also consistent with transfer free-energies of glycine from pure water to urea solutions [140]. Glycine is often used as model for the backbone of proteins and its transfer free-energy is small (smaller than for alanine) and positive implying that urea has little or no destabilizing effect on the interaction between glycine residues. Another case in point regarding the PMF in Figure 3.5b is that the energy barrier separating global and local minima for poly-leucine increases significantly when urea is added to aqueous solution. A similar increase was observed in the PMF of non-polar model compounds [170]. This was related to enhanced cooperativity in folding/unfolding transitions of some proteins [114].

Additional support for the idea that urea weakens mainly side chain interactions and not backbone interactions is provided by studying the preferential interaction Γ . This quantity compares the distribution of urea and water as a function of the distances r from the closest peptide [25, 149]:

$$\Gamma(r) = \left\langle N_{\rm u}(r) - \frac{N_{\rm u}^{\rm bulk}}{N_{\rm w}^{\rm bulk}} N_{\rm w}(r) \right\rangle, \tag{3.2}$$

where $N_{\rm u}(r)$ and $N_{\rm w}(r)$ are the number of urea and water molecules, respectively, with minimal distance to peptide atoms between 0 and r. $N_{\rm u}^{\rm bulk}$ and $N_{\rm w}^{\rm bulk}$ are the number of urea and water molecules in the bulk. Distances r from the protein for which urea tends to accumulate will show positive Γ values whereas negative Γ values imply exclusion of urea from the protein. Figure 3.7 provides a comparison of Γ when dimers are at their preferred interaction distance, i.e., at $\xi = 0.47$ nm (full line), and when they are non-interacting, i.e., at $\xi = 2.0$ nm (dashed line). Γ measured with respect to all atoms (in black) of poly-alanine or poly-leucine increases when chains dissociate. These changes in Γ , i.e., $\Delta\Gamma(r) = \Gamma_{\rm monomer}(r) - \Gamma_{\rm dimer}(r)$ are shown in Figure 3.7c-d for poly-alanine and poly-leucine, respectively. For attractive co-solvents, $\Delta\Gamma(r)$ is proportional to the effect of the co-solvent on the free energy difference to unfold proteins [186]. In our simulations, $\Delta\Gamma(r)$ is significantly larger for poly-leucine compared to poly-alanine close to the peptide, i.e., for r < 0.6 nm.

To provide further insights, $\Gamma(r)$ is decomposed into contributions from backbone (Γ_{backbone} , in blue) and side chain ($\Gamma_{\text{sidechain}}$, in red) atoms. $\Gamma_{\text{backbone}}(r)$ is computed by counting the number of urea and water molecules (i.e., $N_{\text{u}}(r)$ and $N_{\text{w}}(r)$)



Figure 3.7 (a-b) Preferential interaction $\Gamma(r)$ around dimers ($\xi = 0.47$ nm) and monomeric peptides ($\xi = 2.0$ nm). Decomposition of the preferential interaction (black) into backbone (blue) and side-chain (red) contributions. (c-d) Changes in the preferential interaction $\Delta\Gamma = \Gamma_{\text{monomer}} - \Gamma_{\text{dimer}}$ as peptides dissociate. Spatial distribution functions of urea (panel e) and water (panel f) around poly-leucine dimers in 7M urea and pure water solutions, respectively. Isovalues for urea and water are 18 and 7.8.

that are closer to backbone than to side chains atoms with a distance between 0 and r from the peptide. A similar definition is used for $\Gamma_{\text{sidechain}}$. These definitions ensure additivity of $\Gamma(r)$, i.e., $\Gamma(r) = \Gamma_{\text{backbone}}(r) + \Gamma_{\text{sidechain}}(r)$. $\Gamma_{\text{sidechain}}$ and Γ_{backbone} are shown in Figure 3.7a-b. The corresponding change in these quantities, i.e., $\Delta\Gamma_{\text{backbone}}$ and $\Delta\Gamma_{\text{sidechain}}$, is shown in Figure 3.7c-d. For poly-alanine, both backbone (blue) and side chain atoms (red) contribute significantly to $\Delta\Gamma(r)$. In contrast, $\Delta\Gamma(r)$ for poly-leucine is mainly ascribed to side chains atoms (red). Notice that $\Delta\Gamma_{\text{backbone}}$ is positive and its magnitude is similar for both poly-alanine and poly-leucine chains. Although $\Delta\Gamma_{\text{backbone}}$ accounts for a significant contribution of the overall $\Delta\Gamma$ in poly-alanine, dimers of this peptides are not destabilized significantly–see Figure 3.5a.

The accumulation of polar urea molecules around the non-polar sidechain can be explained by the planar distribution of partial charges in urea that allows this molecule to surround planar non-polar environments without breaking hydrogen bonds with the solvent [42, 100]. These environments are usually unfavorable to water that would have at least one non-saturated hydrogen bond because of its tetrahedral distribution of partial charges. Figure 3.7e-f depicts spatial distribution functions of urea and water molecules around poly-leucine in aqueous urea and pure water solutions, respectively. To compute these functions, the simulation box was divided in bins of size $0.02 \times 0.2 \times 0.2 \text{ nm}^3$. Spatial distribution functions are given as the ratio between the density of solvent molecules in the bin and the density of the solvent in an ideal fluid. In addition to replacing water molecules around the peptide, urea is also found between side chains in the hydrogen bonded face of the β -sheet [196] (arrows in Figure 3.7e) whereas these positions are left unoccupied in pure water solutions (Figure 3.7f). Similarly, at other poly-leucine dimer distances (i.e., other ξ values), urea occupies spaces otherwise left empty by water.

In Figure 3.8a, we show changes in the PMF of poly-leucine when 7M urea is added to water, i.e., $\Delta PMF(\xi) = PMF^{urea}(\xi) - PMF^{water}(\xi)$. This quantity is the



Figure 3.8 (a) Changes in total energy (ΔE_{total}), entropic energy ($-T\Delta S$), and PMF (ΔPMF) when 7M urea is added to water as a function of the distance ξ between poly-leucine peptides. (b) Correlation between ΔPMF and changes in non-bonded energies when urea is added to water ($\Delta E_{non-bonded}$) as a function of ξ . Error bars are computed using block average of the different trajectories. (c) Projection of nonbonded energy into peptide-solvent (ΔE_{PS}), peptide-peptide (ΔE_{PP}), and solventsolvent (ΔE_{SS}) components.



Figure 3.9 Volumes of poly-leucine in pure water and 7M urea solution as a function of ξ . These volumes are computed with respect to their value at $\xi=2.0$ nm.

result of both enthalpic and entropic terms. Since the volume of the system does not change significantly as a function of ξ in Figure 3.9, the main term contributing to the enthalpy is the total energy of the system. In Figure 3.8a, we show how this total energy changes when 7M urea is added to water: $\Delta E_{total}(\xi) = E_{total}^{urea}(\xi) - E_{total}^{water}(\xi)$. The entropic term $(-T\Delta S)$ which is determined by subtracting ΔE_{total} from ΔPMF is also shown in Figure 3.8a. Notice that ΔPMF , ΔE_{total} , and $-T\Delta S$ are computed with respect to their values at $\xi = 2.0$ nm. Figure 3.8a shows that an increase in $\Delta E_{total}(\xi)$ is accompanied by an increase in $\Delta PMF(\xi)$. For example, both ΔPMF and ΔE peak at $\xi = 0.7$ nm and their local minima are located at $\xi = 0.5$ nm. The entropic term opposes these changes and it is negative, i.e., favorable to dimerization, for small ξ values. This is consistent with recent studies in which the entropic energy of urea was shown to favor the folded state of some macromolecules [158, 19, 154]. Thus, when urea is added to water changes in ΔPMF of poly-leucine are caused by changes in total energy of the system.

In Figure 3.8b, we show changes in the non-bonded component of the total energy, i.e., $\Delta E_{non-bonded}$, when urea is added to water. This component is the sum of van der Waals and electrostatic energies in the system and it is shown with respect to its value computed at $\xi = 2.0$ nm. As for the total energy, changes in $\Delta E_{non-bonded}$ and ΔPMF occur at the same ξ values and in the same direction. This correlation is, however, only qualitative since other energetic terms also contribute to ΔPMF , e.g., entropic energy in Figure 3.8a. In Figure 3.8c, we project $\Delta E_{non-bonded}$ into its peptide-peptide (ΔE_{PP}), peptide-solvent (ΔE_{PS}), and solvent-solvent (ΔE_{SS}) components. As peptides are brought close to each other, solvent molecules are released into the bulk breaking peptide-solvent bonds. These bonds are stronger in urea solution than in pure water accounting for the positive ΔE_{PS} in Figure 3.8c (red line). Congruently more solvent-solvent interactions are formed in this process. Solvent interactions involving urea molecules are stronger than in pure water accounting for the negative ΔE_{SS} in Figure 3.8c (blue line). In contrast, peptide-peptide interactions are not strongly affected by urea (black line) except at intermediate distances where side chains interdigitate. Changes in peptide-urea interactions have been extensively studied in the literature and related to the destabilizing effect of this co-solvent [25, 28, 26, 81]. However, as shown in Figure 3.8c, unfavorable changes in the peptide-solvent interactions occur simultaneously with a favorable change in solvent-solvent interactions. It is the result of these two interactions that account for the destabilizing effect of urea.

Additional insights into how ΔE_{PS} and ΔE_{SS} are affected by urea can be obtained by studying their van der Waals and electrostatic components. To that purpose, we write $\Delta E_{PS} = E_{PU}^{urea} + E_{PW}^{urea} - E_{PW}^{water}$. Now, the term $E_{PW}^{water} - E_{PW}^{urea}$ can be interpreted as the energy of water molecules that have been replaced by urea. We refer to this term as $\Delta E_{PW}^{replaced}$. Lennard-Jones (red lines) and electrostatic (blue) contributions of both ΔE_{PU}^{urea} and $\Delta E_{PW}^{replaced}$ are shown in Figure 3.10a-b. When



Figure 3.10 Lennard-Jones and electrostatic energies of (a-b) peptide-solvent and (c-d) solvent-solvent interactions. Peptide-solvent is decomposed into replaced peptide water interactions $\Delta PW^{\text{replaced}}$ (see text) and peptide-urea ΔPU^{urea} energies. Solvent-solvent is shown for pure water and aqueous urea solution. Quantities computed at $\xi=2.0$ nm are used as our reference.

the distance ξ of both poly-alanine and poly-leucine dimers is small, electrostatic contributions of these two terms are comparable in magnitude and they cancel out when accounting for ΔE_{PS} . This implies that the electrostatic interaction between water and the peptide is very similar to the electrostatic interaction between urea and the peptide [177]. This is consistent with results from another simulation [210]. In contrast, Lennard-Jones energies of these two terms are very different when peptides are close to each other with the magnitude of ΔE_{PU}^{urea} dominating over $\Delta E_{PW}^{replaced}$. Thus, peptide-solvent Lennard-Jones energies favors non-interacting configurations of the dimer. This energetic analysis is consistent with the molecular picture in Figure 4.10 where urea forms new Lennard-Jones bonds with non-polar side chains and it replaces some electrostatic bonds between peptide and water (e.g., hydrogen bonding).

In Figure 3.10c-d, we decompose solvent-solvent interactions (i.e., ΔE_{SS}) into their Lennard-Jones and electrostatic components for simulations performed in pure water and urea solutions. As peptides are brought close to each other, solvent molecules are released into the bulk and this accounts for a favorable decrease in solvent-solvent energy. Figure 3.10c-d shows that this reduction in energy is larger in urea solution and, therefore, the release of urea molecules into the bulk favors folding of the protein. This is consistent with experimental [173, 95, 155] and computational [14, 102, 177] studies in which urea was found to incorporate readily and tightly into water. In Figure 3.10c-d both Lennard-Jones and electrostatic energies contribute to urea insertion into the solvent.

From Figure 3.10 we can conclude that the increase in $\Delta E_{\rm PS}$ (see Figure 3.5b) when peptides are brought close together in urea solution compared to pure water emerges from loss of favorable Lennard-Jones interactions between peptide and urea. This occurs for both poly-alanine and poly-leucine dimers (see Figure 3.10a-b) although the effect in poly-alanine is not significant enough to produce a change in its PMF (see Figure 3.5a). In contrast, the reduction in $\Delta E_{\rm SS}$ (see Figure 3.5b) when peptides are brought close together in urea solution compared to pure water occurs because of changes in both Lennard-Jones and electrostatic energies—see Figure 3.10c-d.

3.4 Conclusions

In this work we study effects of urea on poly-peptides. We show that unrestrained poly-leucine peptides become less compact when urea is added to water whereas poly-glycine and poly-alanine peptides are only weakly affected by this co-solvent. To rationalize these results, we study interactions between stretched poly-alanine and poly-leucine dimers. In particular, we compute potential of mean force of these dimers in pure water and aqueous solutions containing urea. Urea is found to significantly destabilize poly-leucine dimers whereas it has little effect on the stability of polyalanine dimers (Figure 3.5). This is consistent with our simulations of unrestrained poly-peptides and it leads to the conclusion that urea has a main effect on interactions involving side chains since poly-alanine and poly-leucine are different only in their side chains. This is further confirmed by studying the distribution of urea molecules around peptides using the preferential interaction defined in Equation 4.10. We found that the number of urea molecules close to side chains is larger than the expected number for the same volume in bulk solvent (Figure 3.7a-b). Moreover, the difference between the number of urea molecules around peptides and the expected number increases when peptides dissociate (Figure 3.7c-d). This is significantly larger for polyleucine than for poly-alanine. Therefore, our results suggest that urea destabilizes proteins through a direct mechanism whereby it binds favorably to large non-polar side chains. An analysis of the energy of the system shows that urea forms mainly Lennard-Jones interactions with side chains while maintaining electrostatic bonds with the solvent (Figure 3.10). This is possible due to the planar distribution of partial charges in urea that allows it to occupy positions between non-polar side chains which are not accessible to water (Figure 3.7e-f).

In summary, we propose a simplified framework to study effects of urea on hydrophobic interactions and hydrogen bonds in proteins. This framework can be used to provide insights into effects of other cosolvents [167]. Results for urea agree with other studies in which this molecule is found to destabilize proteins through dispersion interactions with non-polar groups of the proteins [159, 206, 102, 158, 80]. The simplified nature of the framework used here allows us to provide a comprehensive description of the effects of urea on the interactions leading to the "association" of peptides. Despite the insights brought by the current study, limitations should also be noted. In particular, our framework does not account for the reduction in the conformational entropy of proteins associated with folding since peptides in our simulations are stretched. Conformational entropies account for the main energetic term $(-T\Delta S)$ opposing dimerization and simulations without restrains are required to account for this "dissociation energy".

CHAPTER 4

EFFECTS OF TMAO ON THE CONFORMATION OF PEPTIDES AND ITS IMPLICATIONS FOR SMALL PROTEINS

In Chapter 3, urea was shown to denature proteins through a direct mechanism, by interacting favorably with peptide hydrophobic side chains. Conversely to denaturants, the presence of protecting osmolytes in solution favors the native state of proteins. TMAO is an example of protecting osmolyte, whereas its molecular mechanism is much less clear. In this chapter, we provide insights into the stabilizing mechanisms of TMAO on protein structures by performing microsecond long all-atom molecular dynamics simulations of peptides and replica exchange molecular dynamics simulation of the Trp-cage miniprotein.

4.1 Introduction

Biochemical reactions in living systems take place in aqueous environments containing small organic molecules [202, 203]. These molecules affect osmosis in cells and they are known as protecting or denaturing osmolytes depending on how they modulate the stability of proteins [203, 201, 168]. Denaturing osmolytes, e.g., urea and guanidine, are widely used in protein folding studies to destabilize the native state [176]. Effects of these molecules emerge from favorable protein-osmolyte interactions which enhances the preference for protein conformations with greater solvent exposed surface area, i.e., the unfolded state [90, 80, 46, 180, 208, 114]. In contrast, protecting osmolytes, e.g., trimethylamine N-oxide (TMAO) and proline, favor the native state of proteins [3]. These osmolytes counteract effects of water stresses enabling organisms to cope with extreme conditions [202]. For example, deep-sea animals counteract effects of hydrostatic pressure by increasing the concentration of TMAO in their muscles [72, 204]. Despite intensive studies on the effects of TMAO in proteins, its molecular mechanisms remain a question of debate [153].

The stabilizing effect of TMAO on native structures is often explained by its exclusion from the vicinity of the protein surface [41, 110, 171]. Exclusion is possible because TMAO is more strongly attracted to water [86, 16, 211, 132, 162, 99, 167] than to the protein surface. Moreover, exclusion may be more pronounced close to the main chain of the protein, i.e., the backbone [167, 21, 7, 8, 6] which explains the adoption of more compact conformations by poly-glycine (which is commonly used as a model of the protein backone) in aqueous TMAO solution compared to pure water [82]. Despite these results, TMAO exclusion from the protein surface as a mechanism to protect the native state has been challenged by recent studies. In particular, computer simulations and experiments have shown that TMAO stabilizes compact conformations of some non-polar polymers through direct interactions [157, 131. Recently, TMAO's effect on proteins was reported to emerge from favorable interactions of this osmolyte with the heterogeneous protein surface that emerges upon folding [108]. Other studies are suggesting that protecting effects of TMAO emerge because this molecule acts as a crowding agent reducing the conformational entropy of the unfolded state [37, 120] and/or by weakening the strength of hydrogen bonds between the protein and water molecules [120].

In this chapter, we highlight effects of TMAO on the molecular forces stabilizing native protein structures by studying small peptides and the Trp-cage miniprotein. Consistent with other studies [82, 37], we find that TMAO favors compact conformations of a peptide model of the protein backbone, i.e., poly-glycine. However, the addition of even the smallest non-polar side chain (i.e., $-CH_3$ group of alanine) to the backbone counteract this effect while larger non-polar side chains account for peptide swelling. This suggests that TMAO can destabilize the hydrophobic core of proteins. We also study conformations of non-polar peptides with charged terminal residues. We find that these peptides become more compact when TMAO is added to water due to stronger interactions between charged residues. In light of these results, we hypothesize that competition of TMAO's effect on hydrophobic and charged interactions accounts for its net stabilizing role in proteins. Extensive replica exchange molecular dynamics (REMD) simulations of the small Trp-Cage protein in pure water and TMAO solution are performed to test this hypothesis. Accordingly, we find that residues that form the hydrophobic core of Trp-cage sample more extended conformations while distances between its charged residues decrease when TMAO is added to water.

4.2 Methodology

All the simulations are performed using Gromacs 4.5 [197] with the AMBER99SBildn-force-field to describe peptides [113], TIP3P water, and Kast model to mimic the behavior of TMAO molecules. The leap-frog algorithm with a time-step of 2 fs is used to integrate the equations of motion and the neighbor list is updated every 10 steps. A Lennard-Jones cutoff of 1.3 nm was used. Electrostatics is treated using the Smooth Particle Mesh Ewald method with a grid spacing of 0.13 nm and a 1.3 nm real-space cutoff. The initial 60 ns of each simulation is ignored and the remaining is used for analysis. Temperature is controlled using the v-rescale thermostat ($\tau_T=1$ ps) and pressure using the Parrinello–Rahman barostat ($\tau_P=1$ ps). Details of the simulations performed in the manuscript for poly-peptides and the Trp-cage miniprotein are described in Table 4.1.

Equilibrium conformations of peptides in this work are determined using NPT molecular dynamics simulations (1 atm and 298 K) of at least two independent simulations. Peptides are terminated with COOH and NH_2 groups, and they are initially in an extended conformation. To study equilibrium configurations of the Trp-cage miniprotein, we use REMD simulations. Trp-cage has one positive net

Dontido	Timo	Number of water	Number of TMAO	TMAO	
1 epude	TIME	molecules	molecules	Concentration	
	$1 \times 800 \text{ ns}$	2149	2149 0		
G10	$1 \times 800 \text{ ns}$	1554	103	3M	
	$1 \times 800 \text{ ns}$	905	228	7M	
	$1 \times 1000 \text{ ns}$	4100	0	0M	
A10	$2 \times 1000 \text{ ns}$	3008	302	3M	
	$2 \times 1000 \text{ ns}$	2341	333	$5\mathrm{M}$	
	$1 \times 1800 \text{ ns}$	2126	0	0M	
V10	$3 \times 400 \text{ ns}$	3014	199	3M	
	$2 \times 900 \text{ ns}$	1768	457	$7\mathrm{M}$	
	$1 \times 1100 \text{ ns}$	2069	0	0M	
L10	$2 \times 1000 \text{ ns}$	3008	199	3M	
	$2 \times 1000 \text{ ns}$	1762	454	$7\mathrm{M}$	
	$2 \times 1000 \text{ ns}$	2135	0	0M	
A16-22	$2 \times 1000 \text{ ns}$	1542	104	3M	
	$2 \times 1000 \text{ ns}$	894	230	7M	
	$3 \times 1200 \text{ ns}$	2137	0	0M	
KV5E	$3 \times 1200 \text{ ns}$	1545	104	3M	
	$3 \times 1200 \text{ ns}$	898	230	7M	
	$3 \times 1200 \text{ ns}$	2133	0	0M	
KL5E	$3 \times 1200 \text{ ns}$	1540	104	3M	
	$3 \times 1200 \text{ ns}$	893	230	7M	
T	$32 \times 450 \text{ ns}$	2998	0	0M	
Trp-cage	$32 \times 250 \text{ ns}$	1528	221	5M	

 $\label{eq:table 4.1} \begin{array}{c} \textbf{Table 4.1} \\ \textbf{Details of the Simulations Performed for Poly-peptides and the Trp-cage} \\ \textbf{Miniprotein} \end{array}$

charge which is neutralized by adding one chloride ion to the solvent. Simulations are performed in pure water and 5 M TMAO solution at a constant pressure of 1 atm. We use 32 temperatures 298.00, 300.50, 303.52, 306.56, 309.63, 312.72, 315.84, 318.98, 322.14, 325.33, 328.55, 331.79, 335.06, 338.36, 341.67, 345.03, 348.40, 351.80, 355.22, 358.68, 362.16, 365.66, 369.20, 372.73, 376.32, 379.94, 383.58, 387.26, 390.97,394.70, 398.47, and 402.26 which provide an exchange rate of approximately 0.2 [40]. Exchanges between neighboring replicas are attempted at every 750 steps. Simulations are extended over 300 ns per replica and the last 200 ns are used for analysis. Thermodynamics of Trp-Cage are shown in Figure 4.1, Figure 4.2, Table 4.2 and Table 4.3. To show that results from simulations are force-field independent, we also use Osmotic [27] and Netz [167] models for TMAO as well as SPCE water. The parameters of Kast, Netz and Garcia force-fields of TMAO are summarized in Table 4.4 - 4.7. We use high concentrations of TMAO molecules to allow for significant changes in protein conformations within the time-scale of simulations. At these high concentrations, Kast model tends to underestimate the effective repulsion between TMAO molecules [27].

	Calculated	Experimental [4]		
$\Delta G(\mathrm{kJ/mol})$	5.88	3.2		
$\Delta H({\rm kJ/mol})$	-77	50		
$C_P(kJ/mol/K)$	2.34	0.30		
$T_c(\mathbf{K})$	360	317		

 Table 4.2
 Thermodynamic Parameters for Trp-Cage at 298K in Pure Water

To determine the free-energy landscape describing charge-charge interaction, we use an umbrella sampling protocol where the reaction coordinate is the distance between NA^+ and CL^- ions in the range 0.2 to 1.0 nm. We simulate different windows

	Calculated
$\Delta G({ m kJ/mol})$	7.45
$\Delta H(\rm kJ/mol)(\rm kJ/mol)$	-62.06
$C_P(kJ/mol/K)$	2.00
$T_c(\mathbf{K})$	356

Table 4.3Thermodynamic Parameters for Trp-Cage at 298K in 5M TMAO Solution



Figure 4.1 Temperature dependence of RMSD distributions of Trp-cage miniprotein in pure water



Figure 4.2 Fraction of folded replicas as a function of temperature for the Trp-cage miniprotein in pure water. The number of folded replicas were computed by integrating the RMSD distribution up to the cut-off value. Different curves are computed when different portions of the simulation are ignored: 120 ns (black), 180 ns (red), 240 ns (green), 300 ns (blue), 360 ns (orange). The total simulation time is 480 ns. This figure shows convergence of the simulation. The coexistence temperature T_e 360 K is defined when the fraction of folded states is 0.5.

	Atoms type	Partial charge	Well depth	Contact distance	
	Atoms type	q [e]	$\varepsilon \; [{\rm kJ/mol}]$	$\sigma = 2r \text{ [nm]}$	
	О	-0.65	0.6385	0.3266	
Veet	Ν	0.44	0.8368	0.2926	
Kast	С	-0.26	0.2828	0.3041	
	Н	0.11	0.0774	0.1775	
	О	-0.91	0.6385	0.3266	
Nata	Ν	0.70	0.8368	0.2926	
Netz	С	-0.26	0.2828	0.3600	
	Н	0.11	0.0774	0.2101	
	О	-0.78	0.6385	0.3266	
	Ν	0.528	0.8368	0.2926	
Garcia	С	-0.312	0.2828	0.3041	
	Н	0.132	0.0774	0.1775	

 Table 4.4
 Partial Charges and Lennard-Jones Parameters

Table 4.5Bond Parameters

Bonds	Well depth ε [kJ/(mol nm2)]	r^0 [nm]
O-N	143335.472	0.1407
N-C	107181.528	0.1506
C-H	247257.664	0.1082
O-C*	30133.168	0.233
N-H*	65814.32	0.2101
H-H*	5493.592	0.1768
C-C*	40597.352	0.2414

Table 4.6 Angle Parameters

Angles	Force Constant $k [kJ/mol]$	$a_0 [\mathrm{deg}]$
O-N-C	254.97296	109.99
N-C-H	208.94896	108.07
C-N-C	576.1368	108.16
Н-С-Н	229.57608	108.25

 Table 4.7
 Dihedral Angle Parameters

Dihedrals	Well depth ε [kJ/mol]	Multiplicity n	Phase $t_0[deg]$
O-N-C-H	1.12968	3	0
C-N-C-H	1.12968	3	0

in which ions are restrained by a spring that has an equilibrium constant of 4000 kJ/mol/cm². The equilibrium distances of springs in neighboring windows differ in steps of 0.05 nm. Each window is simulated for 150 ns and the Potential of Mean Force (PMF) is computed using the Weighted Histogram Analysis Method (WHAM). These simulations were performed using TIP3P.

4.3 Results

Figure 4.3a shows extended conformations of deca-homopeptides made of glycine (G_{10}) , alanine (A_{10}) , valine (V_{10}) and leucine (L_{10}) . The side chain of glycine is made of one hydrogen atom and, therefore, poly-glycine is commonly used as a model of the protein backbone. The side chain of alanine is made of a small non-polar group $(-CH_3)$ whereas value and leucine are decorated with large non-polar groups $(-C_3H_7 \text{ and } -C_4H_9)$. Thus, L_{10} and V_{10} are more hydrophobic in nature than A_{10} . Distributions of the radius of gyration R_g of backbone atoms of these peptides are



Figure 4.3 Conformations of Homopeptides in Different Aqueous Solutions. (a) Schematics of homopeptides using a cartoon-like representation for the backbone and a van der Waals representation for side chains highlighting the weaker hydrophobic nature of poly-alanine compared to poly-valine and poly-leucine. TMAO (C_3H_9NO) and water (H_2O) are also shown. Distributions of the radius of gyration R_g of backbone atoms for (b) poly-glycine, (c) poly-alanine, (d) poly-valine, and (e) poly-leucine in pure water as well as 3 M and 7 M TMAO solutions at 298 K and 1 atm.

studied in Figure 4.3b-e in pure water (black) as well as 3 M (red) and 7 M (blue) TMAO solutions.

Figure 4.3b shows that compact and extended conformations of G_{10} become more and less populated, respectively, when TMAO is added to water. Previous studies have suggested that this effect of TMAO is the dominant stabilizing effect of this osmolyte on native protein structures [21]. In Figure 4.3c-e, we probe effects of TMAO on non-polar peptides. Distributions of R_g for the A_{10} peptide are not strongly affected by TMAO—see Figure 4.3c. This is consistent with computational studies showing that the interaction between non-polar compounds that are comparable in size to alanine's side chain $(-CH_3)$, i.e., methane molecules (CH_4) , is only weakly affected by TMAO [4]. In contrast, V_{10} and L_{10} become more extended as the concentration of TMAO increases—see Figure 4.3d-e. In Figure 4.3e, we also show distributions of R_g for L_{10} in aqueous solutions containing Osmotic and Netz TMAO molecules (in green). Independent of TMAO force-field, Figure 4.3e shows that the population of extended conformations of L_{10} increases when TMAO is added to water. Swelling of L_{10} is also observed with a different water model, i.e., SPCE water, using Netz TMAO molecules—see Figure 4.4b. These results show that TMAO's collapsing effect on the backbone can be overcompensated by its effect on hydrophobic residues which causes peptides to swell. This is consistent with the hydrophobic effect becoming weaker in aqueous TMAO solutions as reported for non-polar compounds that are comparable in size to leucine's side chain $(-C_4H_9)$, e.g., neo-pentane (C_5H_{10}) [144, 121, 122, 68, 157]. Moreover, since the hydrophobic core is key to protein folding this result suggests that, in addition to the backbone, other elements of the protein may also play an important role in TMAO's stabilizing effect.

Insights into another stabilizing effect of TMAO can be obtained by studying the $A\beta_{16-22}$ peptide (<u>KLVFFAE</u>) which is made of five non-polar residues (bold letters) flanked by opposite charged residues (underline letters). Based on results



Figure 4.4 Distributions of the radius of gyration Rg of backbone atoms for (a) $A\beta_{16-22}$ and (b) poly-leucine in pure water and 7M TMAO solutions at 298 K and 1 atm. These simulations are performed using SPCE water model and Netz TMAO model.



Figure 4.5 Effects of TMAO on the conformation of non-polar peptides flanked by opposite charged residues. Distribution of the radius of gyration R_g of backbone atoms for (a) $A\beta_{16-22}$, (b) KV₅E and (c) KL₅E in different aqueous solutions at 298 K and 1 atm.

from Figure 4.3, the addition of TMAO to water is expected to cause the $A\beta_{16-22}$ peptide to swell due to its highly non-polar nature. However, all-atom molecular dynamics simulations using the CHARMM22 force field with the CMAP modification have reported that TMAO favors compact conformations of the $A\beta_{16-22}$ peptide possible due to a coil-helix transition [37]. In our simulations using the AMBER99SBildn-force-field, the peptide also becomes more compact when TMAO is added to water (see Figure 4.5a) but without favoring any type of secondary structures-see Table 4.8 and Figure 4.6. An analysis of the structures of $A\beta_{16-22}$ shows that TMAO favors conformations where charged residues are close to each other-see Figure 4.7 Therefore, we speculate that charged residues at the end of ${\rm A}\beta_{16-22}$ are responsible for the collapsing effect of this peptide. To verify this idea, we flank short poly-valine and poly-leucine peptides (which adopt more extended conformation in the presence of TMAO-see Figure 4.3d-e) with opposite charged residues: K-L₅-E and K-V₅-E. Results from simulations of these peptides are shown in Figure 4.5b-c. These simulations show that the presence of charged residues makes TMAO to favor compact peptide conformations. A similar result is also observed using SPCE water and Netz model for $A\beta_{16-22}$ -see Figure 4.4 and Table 4.9. It suggests that TMAO enhances the magnitude of charge-charge interactions. Accordingly, in Figure 4.8 we compute the potential of mean force (PMF) for the interaction between Na^+ and $Cl^$ in pure water and 7M TMAO solution. The PMF to form a contact between these ions increases significantly ($\sim 2 \text{ kJ/mol}$) when TMAO is added to water. Thus, effects of TMAO on charge-charge interactions contribute to overcome TMAO's swelling effect on non-polar peptide segments and they may play an important role in stabilizing compact protein structures.

To show that the results obtained for peptides in Figures. 4.3 and 4.5 also apply to proteins that fold into a native state, we show in Figure 4.9 results from REMD simulations of the Trp-cage miniprotein. The key charge-charge interaction

Table 4.8 Average fraction of Helix (α helix + 3-helix), β -sheet, and Rest (coil + bend + turn) in the Simulations of $A\beta_{16-22}$, KV_5E , and KL_5E Peptides in Pure Water (0M) and in Aqueous TMAO Solutions (3M and 7M). Secondary Structures were Computed Using DSSP Definitions

	Helix		β -Bridge			Rest			
	0M	3M	7M	0M	3M	7M	0M	3M	7M
$A\beta_{16-22}$	0.01	0.01	0.00	0.00	0.00	0.00	0.99	0.99	1.00
KV_5E	0.00	0.00	0.00	0.00	0.00	0.00	1.00	1.00	1.00
KL_5E	0.01	0.01	0.02	0.00	0.01	0.05	0.99	0.98	0.94



Figure 4.6 Secondary structures of sample trajectories of $A\beta_{16-22}$, KV_5E , and KL_5E peptides in pure water (0M) and in aqueous TMAO solutions (3M and 7M). Secondary structures were computed using DSSP definitions.



Figure 4.7 Probability distribution of the distance between charged hydrogen atoms of Lysine residue (K) and charged oxygen atoms of glutamic acid (E) for $A\beta_{16-22}$, KV_5E , and KL_5E in pure water (0M) and in TMAO solutions (3M and 7M). Characteristic conformations of flanked peptides at the peak with shortest charge to charge distance are also showing. Lysine and glutamic acid are shown in blue and red, respectively.



Figure 4.8 Free-energy landscape between $\rm Na^+$ and $\rm Cl^-$ in pure water and TMAO solutions .



Figure 4.9 Effects of TMAO on the Trp-cage miniprotein. (a) Amino acid sequence of Trp-cage highlighting residues that form the hydrophobic core and a key charge-charge interaction of this miniprotein. The native structure of Trp-cage with side chains of charged and the hydrophobic residues are depicted on the right hand side. (b-c) R_g distributions of charged and hydrophobic residues are depicted for simulations performed in pure water and 5 M TMAO solution. (d-e) Free-energy landscape of Trp-cage divided by the thermal energy as a function of hydrophobic and charge R_g in pure water and 5 M TMAO solution.

of Trp-cage flanks the loop region of this protein while the hydrophobic core holds one side of the α -helix bonded to the straight segment in the native state [136]—see Fig.4.9a. Experimental and computational studies have highlighted the importance of these residues in accounting for the stability and folding of Trp-cage [85, 209]. Results from our simulations of peptides (see Figures. 4.3 and 4.5) predict that adding TMAO to water destabilizes the hydrophobic core of Trp-cage while it increases the stability of the charged residues. To verify this prediction, distributions of the radius of gyrations of hydrophobic $\mathbf{R}_g^{\text{hydro}}$ and charged $\mathbf{R}_{\text{g}}^{\text{charge}}$ residues are shown in Figure 4.9b-c for simulations performed in pure water and 5 M TMAO solution. These distributions are shown for the unfolded state at the coexistent temperature $T_e = 350$ K where native and unfolded states are equally populated. Figures 4.9b-c show that when TMAO is added to water charged and hydrophobic residues sample more compact and extended conformations, respectively. This is consistent with results from Figures 4.3 and 4.5. To provide further detail, we show the free-energy of Trp-cage divided by the thermal energy as function of $R_g^{\rm hydro}$ and $R_g^{\rm charge}$ in pure water (Figure 4.9d) and TMAO solution (Figure 4.9e) at T_e . They show that in pure water, the unfolded state of Trp-cage samples mostly conformations in which charged residues are far apart, i.e., states U3, and U4 in Figure 4.9d. In contrast, the unfolded state in TMAO solution is mostly characterized by hydrophobic residues that are far apart from each other whereas charged residues remain as compact as the native state-see states U5, U6 and U7 in Figure 4.9e. Atomic insights into the structures adopted by Trp-cage are provide in Figure 4.11. These structures correspond to the most popular cluster at different R_g^{hydro} and R_g^{charge} values. Arrows linking the different structures are drawn based on the proximity of these states in the free-energy landscape in Figure 4.9d-e. In pure water, the largest distances between key charged residues is achieved when the segment of the protein linking the two charges form a beta-strand (see state U4 in Figure 4.11). This conformation can be reached from the native state

by first breaking the key charge-charge interaction (U1) followed by swelling of the hydrophobic core (U3) and the formation of the beta-sheet (U4). Alternatively, state U4 can be reached by first breaking the hydrophobic core (U2) followed by breaking the bond between charged residues (U3) and the formation of the beta-sheet (U4). The existence of these two folding/unfolding scenarios of Trp-cage is consistent with other computational studies of this protein [7]. Trp-cage configurations with the least compact hydrophobic core are found in TMAO solutions (state U6 in Figure 4.11). This state is reached from the native state by first breaking the alpha-helix (U5) followed by swelling of the hydrophobic core (U6). Notice that extending the two key charged residues from U6 leads to the formation of a beta-sheet (U7). Thus, Figure 4.11 shows that the unfolded state of Trp-cage in pure water and TMAO solutions samples different sets of conformations.

To provide insight into how TMAO destabilizes hydrophobic interactions while enhancing the magnitude of charge-charge interactions, we compute the preferential interaction defined as [27, 25]

$$\Gamma(r) = \left\langle N_{\text{TMAO}}(r) - \left(\frac{N_{\text{TMAO}}^{\text{bulk}}}{N_{\text{water}}^{\text{bulk}}}\right) N_{\text{water}}(r) \right\rangle, \tag{4.1}$$

where $N_{\text{TMAO}}(r)$ and $N_{\text{water}}(r)$ are the number of TMAO and water molecules with minimal distance to peptide atoms between 0 and r. $N_{\text{TMAO}}^{\text{bulk}}$ and $N_{\text{water}}^{\text{bulk}}$ are numbers of TMAO and water molecules in the bulk. $\Gamma(r)$ compares the number of TMAO molecules within a distances r from the protein with the expected number of TMAO in a similar water region in the bulk. It has been proposed that osmolytes which accumulate in the vicinity of proteins and, therefore, are characterized by positive $\Gamma(r)$ values, interact favorably with the protein surface favoring the unfolded state. In contrast, osmolytes that are repelled from the protein surface (i.e., $\Gamma(r) < 0$) favor the folded state of proteins. Insights into how osmolytes are partitioned close to the air-water interface can be obtained from measurements of surface tension which



Figure 4.10 Preferential interaction $\Gamma(r)$ of poly-glycine, poly-leucine, $A\beta_{16-22}$ and Trp-cage. (b-d) $\Gamma(r)$ is decomposed into contributions from backbone (red) and side chain (blue) atoms. (c-d) $\Gamma(r)$ of side chains is further analyzed in terms of contributions from polar (doted blue), and non-polar (dashed blue) side chains. Only key residues comprising charge-charge interactions and the hydrophobic core (see Figure 3a) are used to compute contributions of polar and non-polar side chains of Trp-cage.



Figure 4.11 Schematics of the most populated conformations and their allowed transitions. Side chains of the key charged residues (D9 and R16) are shown and the backbone is colored based on secondary structures: beta-sheet in yellow, alpha-helix in purple, coil in gray, and loop in green.

are often, but not always, consistent with their effects on proteins. Upon addition of TMAO to water, the air-water surface tension decreases suggesting that this osmolyte accumulates at this interface [64]. In contrast, positive transfer free-energies of proteins from pure water to TMAO solutions imply that this osmolyte is repelled from the vicinity of the protein [21].

In Figure 4.10a, we show that for poly-glycine $\Gamma(r)$ is negative implying that TMAO is excluded from the surface of this peptide. For the other peptides, we provide insights into how backbone and side chain atoms contribute to $\Gamma(r)$ by assigning solvent molecules to a particular group (backbone or side chain) if it is closest to that group. This assignment is independent of r. Solvent molecules assigned to a group are used to compute $\Gamma(r)$ of that particular group. Consistent with our results for poly-glycine, $\Gamma(r)$ computed for solvent molecules associated to backbone atoms of poly-leucine, $A\beta_{16-22}$, and Trp-cage (red lines in Fig. 4.10b-d) are negative confirming that TMAO is excluded from the backbone of proteins. Similarly, TMAO is excluded from the proximity of charged residues of $A\beta_{16-22}$ and the Trp-cage protein as $\Gamma(r)$ for these groups (dotted blue lines in Figure 4.10c,d) are negative. In contrast, $\Gamma(r)$ computed for solvent molecules associated with non-polar side chain atoms of poly-leucine (blue line in panel b) as well as of $A\beta_{16-22}$ and Trp-cage (dashed blue lines in panels c,d) are positive for distances greater than 0.4 nm implying that TMAO is attracted to non-polar groups of proteins. Notice that the exclusion of TMAO from backbone and charged amino acids is consistent with peptides adopting more compact structures in Figures 4.3b and 4.5. Also, attraction of TMAO to non-polar residues is consistent with non-polar peptides adopting more extended conformation in Figure 4.3d-e.

4.4 Conclusions

In summary, we find that TMAO accounts for swelling of non-polar peptides suggesting that it can destabilize the hydrophobic core of proteins. Accordingly, our simulations of the Trp-cage protein reveal that its non-polar residues adopt more extended conformations in TMAO solutions. This result may provide rationalization for the R_g of the Snase protein which was found to be larger in TMAO solutions (17.3 \pm 1.5 Å) than in water (15.6 \pm 0.2 Å) measured using SAXS [103]. Moreover, we find that effects of TMAO on backbone and charged residues are found to stabilize compact peptide structures. Traditionally, the former has been related to TMAO's main stabilizing mechanisms [82, 21]. However, the observed swelling of non-polar peptides in our simulations provide evidence that, at least for these amino acid sequences, TMAO's effects on the backbone are not enough to counteract its effects on non-polar residues. For non-polar peptides flanked by charged residues as well as the Trp-cage miniprotein, we find that charged residues contribute significantly to counteract effects of TMAO on non-polar residues. Evidence that charged residues
contribute to the stability of globular proteins is provided by thermophilic proteins [29, 152]. While hydrophobic interactions are the main interaction accounting for the increases stability of thermophilic proteins, 68 % of these proteins showed an increased number of salt bridges when compared to their mesophilic homologs [74]. These additional salt bridges contribute to enable thermophilic proteins to function at higher temperatures [184]. Despite the novel insights brought up by our simulations, this work does not exclude the existence of other stabilization mechanisms of TMAO, e.g., the recently proposed surfactant mechanism which may explains TMAO's effect on elastin that is made of alternating glycine and non-polar residues with no charged amino acids [108].

Poly-glycine	Charmm-27	Charmm-22	Amber99SB-ildn	Amber99SB-ildn
	Kast	Kast	Kast	Netz
	TIP3P [82]	TIP3P [37]	TIP3P [Fig.1]	SPCE [Fig.S6]
$A\beta_{16-22}$		Charmm-22	Amber99SB-ildn	Amber99SB-ildn
		Kast	Kast	Netz
		TIP3P [37]	TIP3P [Fig.1]	SPCE [Fig.S7]
Poly-leucine			Amber99SB-ildn	Amber99SB-ildn
			Kast, Netz, Osmotic	Netz
			TIP3P [Fig.1]	SPCE [Fig.S7]

 Table 4.9 Details of the Simulation Force Fields

CHAPTER 5

EFFECTS OF TMAO ON CHARGED INTERACTIONS

In Chapter 4, TMAO was shown to strengthen charge-charge interactions which favored compact conformations of peptides flanked with oppositely charged termini and increased the stability of the native state of Trp-cage. In this chapter, we provide an in-depth understanding of the effect of TMAO on charge-charge interaction between Na⁺ and Cl⁻.

5.1 Methodology

To investigate the effects of TMAO on ionic interaction, we perform molecular dynamics simulations in the NPT ensemble at 1 atm and four temperatures (273K, 298K, 338K and 368K). The open-source Gromacs suite version 4.6.5 is used to perform these simulations [197]. The temperature is controlled using the v-rescale thermostat ($\tau_T = 1$ ps) and pressure is fixed using the Parrinello-Rahman barostat ($\tau_P = 1$ ps). A cut-off of 1.3 nm is used to account for short-range non-bonded interactions. Long-range electrostatics were calculated using the Particle Mesh Ewald (PME) algorithm with a grid spacing of 0.13 nm and a 1.3 nm real-space cutoff. Simulations are conducted using the TIP3P water model, Kast and Netz [167] models for TMAO and the AMBER model for Na⁺ and Cl⁻.

To compute the free-energy landscape describing Na^+-Cl^- interaction, we use an umbrella sampling protocol. For the reaction coordinate of these interactions, we use the distance ξ between between ions in the range of 0.2 nm to 1.1 nm. In the different windows of the umbrella sampling protocol, ions are restrained to their equilibrium distance by a spring. The equilibrium distances of springs in neighboring windows differ in steps of 0.05 nm and their spring constant is 4,000 kJ mol⁻¹ nm⁻². Each window is simulated for 150 ns and the potential of mean force (PMF) is computed using the Weighted Histogram Analysis Method (WHAM) [84]. Notice that the PMF increases with $-k_bT\log(\xi^2)$ due to the three-dimensional nature of ξ , where $-k_bT$ is the thermal energy. We subtract this dependence of the PMF on ξ and the PMF at the last umbrella sampling window ($\xi = 1.75$ nm for the neopentane pair or $\xi = 1.1$ for Na⁺-Cl⁻) is shifted to zero.

The enthalpy and entropic energy as a function of ξ are computed from the temperature dependence of the PMF. For each ξ distance, the temperature dependence of the PMF is fitted to the thermodynamic relation [51]:

$$PMF(\xi, T) = \Delta H_{o}(\xi) - T\Delta S_{o}(\xi,) + \Delta C_{op}(\xi) \left[(T - T_{o}) - T \log \left(\frac{T}{T_{o}} \right) \right], \quad (5.1)$$

where $\Delta S_o(\xi)$, $\Delta H_o(\xi)$ and $\Delta C_{op}(\xi)$, correspond to changes in entropy, enthalpy, and heat capacity at the reference temperature T_o , respectively.

5.2 Results

Potential of Mean Force In Figure 5.1a, we show PMF describing the interaction between Na⁺ and Cl⁻ ions in pure water (black) and aqueous TMAO solutions (red and blue). These PMF are characterized by the presence of a contact minimum (cm), a desolvation barrier (*db*), a first-solvent-separated minimum (1st ssm), and a second-solvent-separated minimum (2nd ssm). Notice that several force-fields, including the AMBER force-field used in this work, have been shown to reproduce the experimental transfer free-energy of sodium (-375 kJ/mol) and cloride (-347 kJ/mol) to pure water, which combined account for a free-energy change of 722 kJ/mol [143]. However, the distance-dependent interaction for Na⁺ and Cl⁻ ions, i.e., position and magnitude of the PMF at different states, has been shown to depend on the force-field used in the simulation [191]. While there are no experimental estimates of the PMF, most classical force-fields overestimate the magnitude of the PMF at cm when compared to *ab initio* simulations. This overestimated cm in classical



Figure 5.1 (a) Potential of mean force (PMF) for the interaction between Na⁺ and Cl⁻ ions in pure water (black) and aqueous TMAO solutions (red and blue). PMF values computed at ξ =1.1 nm are used as our reference, i.e., zero value. (b) Characteristic configurations of sodium (cyan) and choride (yellow) at cm, db, 1st ssm, and 2nd ssm. Water molecules in the first solvation shell are shown in a licorice representation. Dotted lines correspond to hydrogen bonds. (c-d) RDF of ion-O_{water} (red) and ion-H_{water} (blue) in pure water simulations when ions are fart apart from each other, i.e., $\xi = 1.1$ nm.

molecular dynamics simulations, which is enhanced in AMBER when compared to other force-fields, may account for abnormal ion clustering at high concentrations [191]. Arrows in Fig. 5.1a show increased stability of cm states when Kast or Netz TMAO molecules are added to water. db becomes smaller in simulations performed with Kast TMAO whereas Netz model has little effect on the stability of this state. Surprisingly, we find that Kast model stabilizes the 1st ssm while Netz model destabilizes this state.

In Figure 5.1b, we show characteristic configurations of the different states of PMF. One and two layers of water molecules separate Na⁺ and Cl⁻ ions in the 1st and 2nd ssm, respectively. These ions are in contact, i.e., at close proximity, at cm whereas cavities in the space between them account for a large positive PMF at *db*. In Figure 5.1c-d, we show RDF for ion-O_{water} and ion-H_{water} in simulations performed in 7 M Kast TMAO solution when ions are fart apart from each other, i.e., $\xi = 1.0$ nm. From these RDF, O_{water} peaks closer to Na⁺ than H_{water} whereas the opposite order in first peaks is observed for Cl–. This reflects the orientation of water molecules in the solvent first-shell which is such that their hydrogen atoms, which have positive partial charges, point towards and away from Cl⁻ and Na⁺ ions, respectively. These water configurations minimize the electrostatic energy between ions and the dipole moment of water. Notice that water structures around neopentane optimize the network of hydrogen bond in the solvation shell.

Enthalpy and Entropy To provide insights into effects of TMAO on the distance-dependent interaction of Na⁺ and Cl⁻ ions, we decompose the PMF into enthalpy (Δ H) and entropic energy ($-T\Delta$ S) in Figure 5.2. In pure water (panel a), the interaction between Na⁺ and Cl⁻ ions is favored by entropic energy (blue) and it is opposed by enthalpy (red) [207]. This reflects the reduced mobility of water in the first-solvation shell around ions due to strong electrostatic interactions between



Figure 5.2 Decomposition of PMFs (black) of sodium chloride ion pair into enthalpies (red) and entropies (blue) contribution in pure water (a) and aqueous TMAO solutions ((b),(c)) at 298 K as a function of seperation distance between the two ions. Quantities computed at $\xi/\xi_o=1.0$ are used as our reference, i.e., zero value.

ions and the dipole moment of water molecules. At short ξ distances, i.e., when Na⁺ and Cl⁻ are close to each other, some water in the first-solvation shell are released into the bulk where they become more disordered accounting for a favorable change in the entropic energy—see Figure 5.2a. Concurrently, the enthalpy of these released water molecules increases as they stop interacting with the ions.

When TMAO is added to water, transferring shell water to the bulk reduces both the enthalpic penalty and the entropic gain associated with the formation of cm. This is represented by arrow in Figure 5.2b-c. Notice that these effects are more significant in the Netz model for which TMAO has a larger dipole moment associated with its nitrogen and oxygen atoms as well as a larger hydrophobicity related to CH_3 groups. Also, TMAO reduces the enthalpic barrier and it increases the entropic energy associated with the formation of cavities at db.

Solvation Shell around Na⁺ and Cl⁻ ions In Figure 5.3a-b, we show RDF for the interaction between ions and O_{TMAO} (red line), N_{TMAO} (blue), or C_{TMAO} (black). RDF describing the interactions between ions and O_{water} are also shown in green. These function are computed in simulations performed at 7 M Kast TMAO solution when Na⁺ and Cl⁻ ions are far apart from each other, i.e, at $\xi = 1.1$ nm. In Figure 5.3a, first peaks in the RDF of Na⁺ and oxygen atoms of both water and TMAO occur at the same short distance of 0.24 nm. This attraction between Na⁺ and oxygen atoms of these molecules accounts for a solvation shell containing 1.4 TMAO and 4.5 water molecules, respectively. In 3 M Kast TMAO solution (RDF not shown), the number of TMAO and water molecules in the solvation shell is 0.3 and 5.5, respectively. Notice that in pure water (RDF not shown), the number of solvent molecules around Na⁺ is 5.9. These numbers were obtained by integrating first peaks in RDF of Na⁺–O_{TMAO}, and Na⁺–O_{water}. The lack of C_{TMAO} atoms in the first solvation shell of Na⁺ (see Figure 5.3a) suggests that methyl groups of TMAO are



Figure 5.3 (a-b) RDF of ions and different atoms of TMAO (oxygen in red, nitrogen in blue and carbon in black) or water (green) for simulations performed in 7 M Kast TMAO solution when ions are far apart, i.e., $\xi = 1.1$ nm. Inset panels show the distribution of θ in the solvation shell. (c) Spatial distribution functions of TMAO (pink color) and water (silver color) around ions in 7 M Kast TMAO solution at $\xi = 0.375$ nm (left) and at $\xi = 1.1$ nm (right). Isovalues for TMAO and water are 7.5 and 10.

solvated away from this ion. Accordingly, the distribution of θ (defined in Figure 5.3a for TMAO molecules that are in the solvation shell) is biased towards large angles–see inset of Figure 5.3a. This characterizes TMAO orientations in which O_{TMAO} point towards Na⁺ and C_{TMAO} atoms points away from this ion.

Figure 5.3b shows that RDF of Cl⁻ and different atoms of TMAO are not characterized by a strong first peak. Moreover, only a reduced number of C_{TMAO} atoms are located within the first solvation shell around the Cl⁻ ion. This suggests exclusion of TMAO's polar group from the vicinity of Cl⁻. Accordingly, the distribution of θ in the inset of Figure 5.3b is biased towards short angles corresponding to oxygen atoms of TMAO pointing away from Cl⁻.

In Figure 5.3c, we show spatial distributions of TMAO (in pink) and water (silver) around Na⁺ and Cl⁻ when these ions are far apart ($\xi = 1.1$ nm) and at cm ($\xi = 0.375$ nm). In agreement with RDF in Figure 5.3a-b, these spatial distribution functions provide evidence that the solvation shell of Na⁺ contains both water and TMAO molecules whereas the solvation shell around Cl⁻ contains mainly water molecules.

5.3 Conclusions

To provide insights into how TMAO affects charged interactions in aqueous solution, we study the association of Na⁺ and Cl⁻ ions. Kast and Netz models are used to mimic TMAO. Both TMAO models are found to stabilize the interaction between Na⁺ and Cl⁻ as they reduce the enthalpic penalty of bringing these ions together. We find that TMAO is attracted to Na⁺ becoming part of its solvation shell whereas it is excluded from the vicinity of Cl⁻.

CHAPTER 6

SUMMARY AND FUTURE WORK

6.1 Major Results of this Dissertation

In this dissertation, we addressed effects of cosolvents on protein stability using all-atom molecular dynamics simulations. This is of fundamental importance to understand living systems as most biological reaction occur in solutions containing organic molecules. Moreover, understanding how cosolvents affect protein stability may lead to new principles for drug design. In particular, we have studied the effect of denaturant urea and protecting osmolyte TMAO. This dissertation has has provided physical insights and it has made advances in the methodology for studying these phenomena.

In Chapter 2 of this dissertation, we have proposed a simplified framework in which extend model peptides were used to study contributions of backbone and side chain atoms to fibrillation. This framework was used in Chapter 3 to study effects of urea in this process. The simplified nature of our methodology has enabled us to provide a comprehensive quantitative description of the effects of urea on the interactions leading to the "association" of peptide. In particular, potential of mean force of extended peptides in pure water and aqueous solutions containing urea. Urea is found to significantly destabilize dimers with large hydrophobic sidechain. This is consistent with our simulations of unrestrained poly-peptides, and it is further confirmed by studying the distribution of urea molecules around peptides using the preferential interaction. Moreover, an analysis of the energy of the system shows that urea forms mainly Lennard-Jones interactions with side chains while maintaining electrostatic bonds with the solvent. Our study on urea provides a clear picture on the balance of driving forces that govern the process at thermodynamic and microscopic level.

In Chapter 4, we studied the molecular mechanism of stabilization of proteins by TMAO. Despite intensive and recent studies, the molecular mechanisms of TMAO (and protecting osmolytes in general) remain a question of debate. Most studies have focused on the effect of this osmolyte on the protein backbone. Here, we show that, through microsecond long all-atom molecular dynamics simulations of peptides and REMD of the Trp-cage miniprotein, effects of TMAO on the backbone are not dominant. In particular, TMAOs effect on the backbone is overcompensated by its destabilizing effect on the hydrophobic core: non-polar peptides and residues forming the hydrophobic core of the Trp-cage protein adopt more extended conformations in solutions containing TMAO. We find that a main interaction that can stabilize folded proteins are charge-charge interactions. In light of these results, we propose that competing effects of TMAO on hydrophobic and charge-charge interactions account for its net stabilizing role on proteins.

In Chapter 4, it is suggested that effects of TMAO on charge-charge interactions can provide an important stabilization effect on native protein structures. However, the stabilizing mechanism of TMAO on charged interaction has not been documented yet. In Chatper 5, we provide in-depth insights into the stabilizing effects of TMAO on the distance-dependent interaction of Na⁺ and Cl⁻ ions. We find that this charge-charge interaction become stronger when TMAO is added to water because this osmolyte decreases the enthalpic penalty of bringing Na⁺ and Cl⁻ close together. At the molecular level, we find that TMAO is attracted to Na⁺ becoming part of its solvation shell whereas it is excluded from the vicinity of Cl⁻. These results for charged interactions are more pronounced in simulation performed using Netz model which has larger dipole moment when compared to the Kast model.

6.2 Foreseeing Future Work

Recent studies indicate that TMAO not only benefits living organisms by regulation osmotic pressure, but it may also have a deleterious effects on human health [195]. In particular, consumption of choline-rich foods such as red meat, egg and processed foods, promote TMAO concentration in the blood [199, 188]. The has been found to correlate with an increased the risk of cardiovascular disease (CVD), which may lead to stroke or heart attack [127], leading cause of death in the U.S. and worldwide. The critical questions is what is the role of TMAO in CVD [198] and how TMAO promote the probability of CVD.

Remarkably, MD simulations have been applied to explore important determinants in processes leading to CVD, and these determinants include lipid transportations [106, 34, 101], lipid interactions [183] and atrial fibrillation [91, 39]. Therefore, we are encouraged to investigate the effect of TMAO in CVD, e.g., role of TMAO on Cholesteryl Ester Transfer Protein by MD simulation.

Role of TMAO on Cholesteryl Ester Transfer Protein. Cholesteryl Ester Transfer Protein (CETP in Figure 6.1) plays a critical role in lipid metabolism by facilitating the net transfer of lipids from high-density lipoprotein (HDL) to low density lipoprotein (LDL). The activity of CETP directly lowers cholesterol levels of HDL (good cholesterol) and enhances LDL (bad cholesterol). LDL cholesterol is considered bad cholesterol since high levels of LDLs lead to plaque buildup and atherosclerosis. HDL cholesterol is considered good cholesterol since HDLs absorb cholesterol and carry it to the liver, where it is then flushed from the body. This helps to prevent plaque buildup in the arteries. CETP deficiency and its inhibition in humans and rabbits have shown to reduce the probability of CVD. Moreover, this reduced risk account for increased levels of HDL and decreased levels of LDL through CETP inhibition [73, 89]. Therefore, to prevent CVD, discovery of small molecule inhibitors became an active stratege to hinder CETP's lipid transfer activity [115, 96]. In particular, MD simulation presenting atomic-level details have been accelerating the drug-discovery processes targeting CETP for treatment of CVD [106, 34, 101].



Figure 6.1 Structure of CETP showing N and C terminal regions. Two dioleoylphosphocholine (DOPC) plugging the tunnel openings are colored in orange, and cholesterol ester (CE) located inside the tunnel are colored in cyan. The functional loops from $\Omega 1$ to $\Omega 6$ and beta-bridge are marked in the figure.

We have studied effects of TMAO on small proteins. Our results show that TMAO strongly affects hydrophobic and ionic interactions. Since CETP is made of many hydrophobic amino acids and several key salt-bridges, this suggests that TMAO might have a strong effect on CETP. A possible mechanism by which TMAO may affect regulation of cholesterol is by changing the conformation/function of CETP. The aim of this project is to study effects of TMAO on the conformation and function of CETP at the atomic level. This will be accomplished using all-atom molecular dynamics simulations to observe the effect of TMAO on CETP.

BIBLIOGRAPHY

- M. P. Allen et al. Introduction to molecular dynamics simulation. Computational soft matter: from synthetic polymers to proteins, 23:1–28, 2004.
- [2] C. B. Anfinsen. Principles that govern the folding of protein chains. Science, 181:223, 1973.
- [3] T. Arakawa and S. Timasheff. The stabilization of proteins by osmolytes. *Biophysical Journal*, 47(3):411, 1985.
- [4] M. V. Athawale, J. S. Dordick, and S. Garde. Osmolyte trimethylamine-n-oxide does not affect the strength of hydrophobic interactions: origin of osmolyte compatibility. *Biophysical journal*, 89(2):858, 2005.
- [5] M. V. Athawale, S. Sarupria, and S. Garde. Enthalpy-entropy contributions to salt and osmolyte effects on molecular-scale hydrophobic hydration and interactions. *The Journal of Physical Chemistry B*, 112(18):5661–5670, 2008.
- [6] M. Auton and D. W. Bolen. Chapter twenty-three-application of the transfer model to understand how naturally occurring osmolytes affect protein stability. *Methods* in Enzymology, 428:397–418, 2007.
- [7] M. Auton, D. W. Bolen, and J. Rösgen. Structural thermodynamics of protein preferential solvation: osmolyte solvation of proteins, aminoacids, and peptides. *Proteins: Structure, Function, and Bioinformatics*, 73(4):802–813, 2008.
- [8] M. Auton, A. C. M. Ferreon, and D. W. Bolen. Metrics that differentiate the origins of osmolyte effects on protein stability: a test of the surface tension proposal. *Journal of Molecular Biology*, 361(5):983–992, 2006.
- [9] M. Auton, L. M. F. Holthauzen, and D. W. Bolen. Anatomy of energetic changes accompanying urea-induced protein denaturation. *Proceedings of the National Academy of Sciences*, 104(39):15317–15322, 2007.
- [10] F. Baftizadeh, X. Biarnes, F. Pietrucci, F. Affinito, and A. Laio. Multidimensional view of amyloid fibril nucleation in atomistic detail. *Journal of the American Chemical Society*, 134:3886, 2012.
- [11] F. Baftizadeh, F. Pietrucci, X. Biarnés, and A. Laio. Nucleation process of a fibril precursor in the c-terminal segment of amyloid-β. *Physical review letters*, 110(16):168103, 2013.
- [12] R. L. Baldwin. In search of the energetic role of peptide hydrogen bonds. Journal of Biological Chemistry, 278:17581, 2003.

- [13] R. L. Baldwin. Energetics of protein folding. Journal of Molecular Biology, 371(2):283, 2007.
- [14] D. Bandyopadhyay, S. Mohan, S. K. Ghosh, and N. Choudhury. Molecular dynamics simulation of aqueous urea solution: is urea a structure breaker. *The Journal* of Physical Chemistry B, 118(40):11757–11768, 2014.
- [15] B. J. Bennion and V. Daggett. The molecular basis for the chemical denaturation of proteins by urea. Proceedings of the National Academy of Sciences, 100(9):5142–5147, 2003.
- [16] B. J. Bennion and V. Daggett. Counteraction of urea-induced protein denaturation by trimethylamine n-oxide: a chemical chaperone at atomic resolution. *Proceedings of the National Academy of Sciences of the United States of America*, 101(17):6433-6438, 2004.
- [17] J. P. Bernacki and M. M. Regina. Length-dependent aggregation of uninterrupted polyalanine peptides. *Biochemistry*, 50:9200, 2011.
- [18] A. Berteotti, A. Barducci, and M. Parrinello. Effect of urea on the β -hairpin conformational ensemble and protein denaturation mechanism. Journal of the American Chemical Society, 133:17200, 2011.
- [19] A. K. Bhuyan. Protein stabilization by urea and guanidine hydrochloride. Biochemistry, 41(45):13386–13394, 2002.
- [20] P. Bjelkmar, P. Larsson, M. A. Cuendet, B. Hess, and E. Lindahl. Implementation of the charmm force field in gromacs: Analysis of protein stability effects from correction maps, virtual interaction sites, and water models. *Journal of Chemical Theory and Computation*, 6:459, 2010.
- [21] D. W. Bolen and G. D. Rose. Structure and energetics of the hydrogen-bonded backbone in protein folding. *Annual Review of Biochemistry*, 77:339–362, 2008.
- [22] C. I. Branden et al. Introduction to protein structure. Garland Science, 1999.
- [23] J. F. Brandts. The thermodynamics of protein denaturation. i. the denaturation of chymotrypsinogen. Journal of the American Chemical Society, 86:4291, 1964.
- [24] J. F. Brandts. The thermodynamics of protein denaturation. ii. a model of reversible denaturation and interpretations regarding the stability of chymotrypsinogen. *Journal of the American Chemical Society*, 86:4302, 1964.
- [25] D. R. Canchi and A. E. García. Backbone and side-chain contributions in protein denaturation by urea. *Biophysical Journal*, 100:1526–1533, 2011.
- [26] D. R. Canchi and A. E. García. Cosolvent effects on protein stability. *Annual review* of physical chemistry, 64:273, 2013.

- [27] D. R. Canchi, P. Jayasimha, D. C. Rau, G. I. Makhatadze, and A. E. Garcia. Molecular mechanism for the preferential exclusion of tmao from protein surfaces. *The Journal of Physical Chemistry B*, 116(40):12095–12104, 2012.
- [28] D. R. Canchi, D. Paschek, and A. E. Garcia. Equilibrium study of protein denaturation by urea. Journal of the American Chemical Society, 132(7):2338–2344, 2010.
- [29] H. Chan, Z. Zhang, S. Wallin, and Z. Liu. Cooperativity, local-nonlocal coupling, and nonnative interactions: Principles of protein folding from coarse-grained models. Annual Review of Physical Chemistry, 62:301, 2011.
- [30] Y. Chebaro, P. Jiang, T. Zang, Y. Mu, P. H. Nguyen, N. Mousseau, and P. Derreumaux. Structures of aβ17-42 trimers in isolation and with five small-molecule drugs using a hierarchical computational procedure. *Journal* of Physical Chemistry B, 116:84, 2012.
- [31] W. Chen, N. Mousseau, and P. Derreumaux. The conformations of the amyloid- β (21-30) fragment can be described by three families in solution. *The Journal of Chemical Physics*, 125:084911, 2006.
- [32] P.-N. Cheng, J. D. Pham, and J. S. Nowick. The supramolecular chemistry of β -sheets. Journal of the American Chemical Society, 135(15):5477, 2013.
- [33] M. Cheon, I. Chang, S. Mohanty, L. M. Luheshi, C. M. Dobson, M. Vendruscolo, and G. Favrin. Structural reorganisation and potential toxicity of oligomeric species formed during the assembly of amyloid fibrils. *PLoS Computational Biology*, 3(9):e173, 2007.
- [34] V. R. Chirasani, R. Sankar, and S. Senapati. Mechanism of inhibition of cholesteryl ester transfer protein by small molecule inhibitors. J. Phys. Chem. B, 120(33):8254–8263, 2016.
- [35] F. Chiti, M. Calamai, N. Taddei, M. Stefani, G. Ramponi, and C. M. Dobson. Studies of the aggregation of mutant proteins in vitro provide insights into the genetics of amyloid diseases. *Proceedings of the National Academy of Sciences*, 99(suppl 4):16419–16426, 2002.
- [36] F. Chiti and C. M. Dobson. Protein misfolding, functional amyloid, and human disease. Annual Review of Biochemistry, 75:333, 2006.
- [37] S. S. Cho, G. Reddy, J. E. Straub, and D. Thirumalai. Entropic stabilization of proteins by tmao. *The Journal of Physical Chemistry B*, 115(45):13401, 2011.
- [38] S. Collins, A. Douglass, R. Vale, and J. Weissman. Mechanism of prion propagation: Amyloid growth occurs by monomer addition. *PLOS Biology*, 2:e321, 2004.

- [39] B. A. Colson, A. R. Thompson, L. M. Espinoza-Fonseca, and D. D. Thomas. Sitedirected spectroscopy of cardiac myosin-binding protein c reveals effects of phosphorylation on protein structural dynamics. *Proceedings of the National Academy of Sciences*, 113(12):3233–3238, 2016.
- [40] B. Cooke and S. C. Schmidler. Preserving the boltzmann ensemble in replica-exchange molecular dynamics. *The Journal of Chemical Physics*, 129(16):164112, 2008.
- [41] E. Courtenay, M. Capp, C. Anderson, and M. Record. Vapor pressure osmometry studies of osmolyte-protein interactions: implications for the action of osmoprotectants in vivo and for the interpretation of osmotic stress experiments in vitro. *Biochemistry*, 39(15):4455–4471, 2000.
- [42] D. Cui, S.-C. Ou, and S. Patel. Protein denaturants at aqueous-hydrophobic interfaces: Self-consistent correlation between induced interfacial fluctuations and denaturant stability at the interface. The Journal of Physical Chemistry B, 119(1):164–178, 2014.
- [43] C. Czaplewski, S. Kalinowski, A. Liwo, and H. A. Scheraga. Comparison of two approaches to potential of mean force calculations of hydrophobic association: particle insertion and weighted histogram analysis methods. *Molecular Physics*, 103(21-23):3153–3167, 2005.
- [44] A. Das and C. Mukhopadhyay. Urea-mediated protein denaturation: a consensus view. The Journal of Physical Chemistry B, 113(38):12816–12824, 2009.
- [45] M. L. de la Paz, G. M. de Mori, L. Serrano, and G. Colombo. Sequence dependence of amyloid fibril formation: Insights from molecular dynamics simulations. *Journal of Molecular Biology*, 349(3):583, 2005.
- [46] G. A. de Oliveira and J. L. Silva. A hypothesis to reconcile the physical and chemical unfolding of proteins. *Proceedings of the National Academy of Sciences*, 112(21):E2775–E2784, 2015.
- [47] S. Deechongkit, P. E. Dawson, and J. W. Kelly. Toward assessing the positiondependent contributions of backbone hydrogen bonding to β-sheet folding thermodynamics employing amide-to-ester perturbations. Journal of the American Chemical Society, 126(51):16762, 2004.
- [48] S. Deechongkit, H. Nguyen, E. T. Powers, P. E. Dawson, M. Gruebele, and J. W. Kelly. Context-dependent contributions of backbone hydrogen bonding to β-sheet folding energetics. *Nature*, 430(6995):101, July 2004.
- [49] C. L. Dias. Unifying microscopic mechanism for pressure and cold denaturations of proteins. *Physical Review Letters*, 109:048104–048109, 2012.
- [50] C. L. Dias, T. Ala-Nissila, M. Karttunen, I. Vattulainen, and M. Grant. Microscopic mechanism for cold denaturation. *Physical Review Letters*, 100:118101–118105, 2008.

- [51] C. L. Dias, T. Ala-Nissila, J. Wong-ekkabut, I. Vattulainen, M. Grant, and M. Karttunen. The hydrophobic effect and its role in cold denaturation. *Cryobiology*, 60:91–99, 2010.
- [52] C. L. Dias and H. S. Chan. Pressure-dependence properties of elementary hydrophobic interactions: Ramifications for activation properties of protein folding. *Journal* of Physical Chemistry B, 118:7488–7509, 2014.
- [53] K. Dill. Dominant forces in protein folding. *Biochemistry*, 29:7133, 1990.
- [54] F. Ding, J. J. LaRocque, and N. V. Dokholyan. Direct observation of protein folding, aggregation, and a prion-like conformational conversion. *Journal of Biological Chemistry*, 280:40235, 2005.
- [55] C. M. Dobson. Protein folding and misfolding. Nature, 426:884, 2003.
- [56] E. D. Eanes and G. G. Glenner. X-ray diffraction studies on amyloid filaments. J. Histochem. Cytochem., 16(11):673, Nov. 1968.
- [57] B. B. Edward P. O'Brien, Ruxandra I. Dima and D. Thirumalai. Interactions between hydrophobic and ionic solutes in aqueous guanidinium chloride and urea solutions:? lessons for protein denaturation mechanism. *Journal of the American Chemical Society*, 2007.
- [58] H. G. England JL. Role of solvation effects in protein denaturation: from thermodynamics to single molecules and back. *Annu Rev Phys Chem*, 2011.
- [59] M. Fändrich and M. Dobson, Christopher. The behaviour of polyamino acids reveals an inverse side chain effect in amyloid structure formation. *EMBO J.*, 21:5682, 2002.
- [60] M. Fändrich, M. A. Fletcher, and C. M. Dobson. Amyloid fibrils from muscle myoglobin. *Nature*, 410(6825):165, Mar. 2001.
- [61] A. R. Fersht. The hydrogen bond in molecular recognition. *Trends Biochem. Sci*, 12(0):301, 1987.
- [62] A. R. Fersht, J.-P. Shi, J. Knill-Jones, D. M. Lowe, A. J. Wilkinson, D. M. Blow, P. Brick, P. Carter, M. M. Y. Waye, and G. Winter. Hydrogen bonding and biological specificity analysed by protein engineering. *Nature*, 314:235, 1985.
- [63] E. Finer, F. Franks, and M. Tait. Nuclear magnetic resonance studies of aqueous urea solutions. *Journal of the American Chemical Society*, 94(13):4424–4429, 1972.
- [64] A. Fiore, V. Venkateshwaran, and S. Garde. Trimethylamine n-oxide (tmao) and tert-butyl alcohol (tba) at hydrophobic interfaces: Insights from molecular dynamics simulations. *Langmuir*, 29(25):8017–8024, 2013.

- [65] A. W. Fitzpatrick, T. P. J. Knowles, C. A. Waudby, M. Vendruscolo, and C. M. Dobson. Inversion of the balance between hydrophobic and hydrogen bonding interactions in protein folding and aggregation. *PLoS Computational Biology*, 7:e10002169, 2011.
- [66] H. S. Frank and M. W. Evans. 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. *Journal of Chemical Physics*, 13:507, 1945.
- [67] H. S. Frank and F. Franks. Structural approach to the solvent power of water for hydrocarbons; urea as a structure breaker. *The Journal of Chemical Physics*, 48(10):4746–4757, 1968.
- [68] P. Ganguly, N. F. van der Vegt, and J.-E. Shea. Hydrophobic association in mixed urea-tmao solutions. *The Journal of Physical Chemistry Letters*, 2016.
- [69] J. Gao, D. A. Bosco, E. T. Powers, and J. W. Kelly. Localized thermodynamic coupling between hydrogen bonding and microenvironment polarity substantially stabilizes proteins. *Nature Structural & Molecular Biology*, 16(7):684, 2009.
- [70] E. Gazit. A possible role for π -stacking in the self-assembly of amyloid fibrils. *FASEB* J., 16(1):77, 2002.
- [71] A. Geddes, K. Parker, E. Atkins, and E. Beighton. Cross- β conformation in proteins. Journal of Molecular Biology, 32:343, 1968.
- [72] M. B. Gillett, J. R. Suko, F. O. Santoso, and P. H. Yancey. Elevated levels of trimethylamine oxide in muscles of deep-sea gadiform teleosts: A high-pressure adaptation. *Journal of Experimental Zoology*, 279(4):386–391, 1997.
- [73] T. Gordon, W. P. Castelli, M. C. Hjortland, W. B. Kannel, and T. R. Dawber. High density lipoprotein as a protective factor against coronary heart disease: the framingham study. *The American journal of medicine*, 62(5):707–714, 1977.
- [74] M. M. Gromiha, M. C. Pathak, K. Saraboji, E. A. Ortlund, and E. A. Gaucher. Hydrophobic environment is a key factor for the stability of thermophilic proteins. *Proteins: Structure, Function, and Bioinformatics*, 81(4):715–721, 2013.
- [75] A. V. Gubskaya and P. G. Kusalik. The total molecular dipole moment for liquid water. *The Journal of Chemical Physics*, 117(11):5290–5302, 2002.
- [76] E. J. Guinn, L. M. Pegram, M. W. Capp, M. N. Pollock, and M. T. Record. Quantifying why urea is a protein denaturant, whereas glycine betaine is a protein stabilizer. *Proceedings of the National Academy of Sciences*, 108(41):16932–16937, 2011.

- [77] T. Hajari and N. F. van der Vegt. Peptide backbone effect on hydration free energies of amino acid side chains. *The Journal of Physical Chemistry B*, 118(46):13162– 13168, 2014.
- [78] F. U. Hartl and M. Hayer-Hartl. Molecular chaperones in the cytosol: from nascent chain to folded protein. *Science*, 295(5561):1852–1858, 2002.
- [79] R. D. Hills Jr. and C. L. Brooks III. Hydrophobic cooperativity as a mechanism for amyloid nucleation. *Journal of Molecular Biology*, 368(3):894, 2007.
- [80] A. S. Holehouse, K. Garai, N. Lyle, A. Vitalis, and R. V. Pappu. Quantitative assessments of the distinct contributions of polypeptide backbone amides versus side chain groups to chain expansion via chemical denaturation. *Journal* of the American Chemical Society, 137(8):2984–2995, 2015.
- [81] D. Horinek and R. R. Netz. Can simulations quantitatively predict peptide transfer free energies to urea solutions? thermodynamic concepts and force field limitations. Journal of Physical Chemistry A, 115(23):6125–6136, 2011.
- [82] C. Y. Hu, G. C. Lynch, H. Kokubo, and B. M. Pettitt. Trimethylamine n-oxide influence on the backbone of proteins: An oligoglycine model. *Proteins: Structure, Function, and Bioinformatics*, 78(3):695–704, 2010.
- [83] L. Hua, R. Zhou, D. Thirumalai, and B. J. Berne. Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proceedings of the National Academy of Sciences of the United States of America*, 105:16928, 2008.
- [84] J. S. Hub, B. L. de Groot, and D. van der Spoel. g_wham—a free weighted histogram analysis implementation including robust error and autocorrelation estimates. *Journal of Chemical Theory and Computation*, 6(12):3713, 2010.
- [85] P. Hudáky, P. Stráner, V. Farkas, G. Váradi, G. Tóth, and A. Perczel. Cooperation between a salt bridge and the hydrophobic core triggers fold stabilization in a trp-cage miniprotein. *Biochemistry*, 47(3):1007–1016, 2008.
- [86] J. Hunger, K.-J. Tielrooij, R. Buchner, M. Bonn, and H. J. Bakker. Complex formation in aqueous trimethylamine-n-oxide (tmao) solutions. *The Journal* of Physical Chemistry B, 116(16):4783–4795, 2012.
- [87] E. G. Hutchinson, R. B. Sessions, J. M. Thornton, and D. N. Woolfson. Determinants of strand register in antiparallel β-sheets of proteins. *Protein Science*, 7:2287, 1998.
- [88] A. Idrissi, F. Sokolić, and A. Perera. A molecular dynamics study of the urea/water mixture. The Journal of Chemical Physics, 112(21):9479–9488, 2000.

- [89] A. Inazu, M. L. Brown, C. B. Hesler, L. B. Agellon, J. Koizumi, K. Takata, Y. Maruhama, H. Mabuchi, and A. R. Tall. Increased high-density lipoprotein levels caused by a common cholesteryl-ester transfer protein gene mutation. *New England Journal of Medicine*, 323(18):1234–1238, 1990.
- [90] G. S. Jas, E. C. Rentchler, A. M. Słowicka, J. R. Hermansen, C. K. Johnson, C. R. Middaugh, and K. Kuczera. Reorientation motion and preferential interactions of a peptide in denaturants and osmolyte. *The Journal of Physical Chemistry B*, 120(12):3089–3099, 2016.
- [91] J. J. Jayasundar, J. Xing, J. M. Robinson, H. C. Cheung, and W.-J. Dong. Molecular dynamics simulations of the cardiac troponin complex performed with fret distances as restraints. *PloS one*, 9(2):e87135, 2014.
- [92] C. G. Ji, X. Xiao, and J. Z. H. Zhang. Studying the effect of site-specific hydrophobicity and polarization on hydrogen bond energy of protein using a polarizable method. *Journal of Chemical Theory and Computation*, 8(6):2157–2164, 2012.
- [93] C. G. Ji and J. Z. H. Zhang. Quantifying the stabilizing energy of the intraprotein hydrogen bond due to local mutation. *Journal of Physical Chemistry B*, 115(42):12230, 2011.
- [94] W. L. Jorgensen. Interactions between amides in solution and the thermodynamics of weak binding. Journal of the American Chemical Society, 111(10):3770, 1989.
- [95] A. T. Joseph D. Batchelor, Alina Olteanu and G. J. Pielak. Impact of protein denaturants and stabilizers on water structure. JACS, 126:1958, 2004.
- [96] T. Joy and R. A. Hegele. Is raising hdl a futile strategy for atheroprotection? Nature reviews Drug discovery, 7(2):143–155, 2008.
- [97] B. Kallies. Coupling of solvent and solute dynamicsmolecular dynamics simulations of aqueous urea solutions with different intramolecular potentials. *Physical Chemistry Chemical Physics*, 4(1):86–95, 2002.
- [98] W. Kauzmann. Some factors in the interpretation of protein denaturation. Adv Protein Chem., 14:1, 1959.
- [99] Y. Koga, P. Westh, K. Nishikawa, and S. Subramanian. Is a methyl group always hydrophobic? hydrophilicity of trimethylamine-n-oxide, tetramethyl urea and tetramethylammonium ion. *The Journal of Physical Chemistry B*, 115(12):2995–3002, 2011.
- [100] T. Koishi, K. Yasuoka, S. Y. Willow, S. Fujikawa, and X. C. Zeng. Molecular insight into different denaturing efficiency of urea, guanidinium, and methanol: a comparative simulation study. *Journal of Chemical Theory and Computation*, 9(6):2540–2551, 2013.

- [101] A. Koivuniemi, T. Vuorela, P. T. Kovanen, I. Vattulainen, and M. T. Hyvönen. Lipid exchange mechanism of the cholesteryl ester transfer protein clarified by atomistic and coarse-grained simulations. *PLoS computational biology*, 8(1):e1002299, 2012.
- [102] H. Kokubo and B. Pettitt. Preferential solvation in urea solutions at different concentrations: Properties from simulation studies. *Journal of Physical Chemistry B*, 111:5233, 2007.
- [103] C. Krywka, C. Sternemann, M. Paulus, M. Tolan, C. Royer, and R. Winter. Effect of osmolytes on pressure-induced unfolding of proteins: A high-pressure saxs study. *ChemPhysChem*, 9(18):2809, 2008.
- [104] K. J. Laidler and M. C. King. The development of transition-state theory. J. phys. Chem, 87(15):2657–2664, 1983.
- [105] S. Lee, Y. L. Shek, and T. V. Chalikian. Urea interactions with protein groups: A volumetric study. *Biopolymers*, 93(10):866–879, 2010.
- [106] D. Lei, M. Rames, X. Zhang, L. Zhang, S. Zhang, and G. Ren. Insights into the tunnel mechanism of cholesteryl ester transfer protein through all-atom molecular dynamics simulations. *Journal of Biological Chemistry*, 291(27):14034–14044, 2016.
- [107] H. Li, C. Tang, and N. S. Wingreen. Nature of driving force for protein folding: A result from analyzing the statistical potential. *Physical Review Letters*, 79:765, 1997.
- [108] Y.-T. Liao, A. C. Manson, M. R. DeLyser, W. G. Noid, and P. S. Cremer. Trimethylamine n-oxide stabilizes proteins via a distinct mechanism compared with betaine and glycine. *Proceedings of the National Academy of Sciences*, 114(10):2479–2484, 2017.
- [109] W. K. Lim, J. Rösgen, and S. W. Englander. Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proceedings of the National Academy of Sciences*, 106(8):2595–2600, 2009.
- [110] T.-Y. Lin and S. N. Timasheff. Why do some organisms use a urea-methylamine mixture as osmolyte? thermodynamic compensation of urea and trimethylamine n-oxide interactions with protein. *Biochemistry*, 33(42):12695–12701, 1994.
- [111] E. R. Lindahl. Molecular dynamics simulations. *Molecular modeling of proteins*, pages 3–23, 2008.
- [112] K. Lindorff-Larsen, S. Piana, R. O. Dror, and D. E. Shaw. How fast-folding proteins fold. *Science*, 334:6055, 2011.

- [113] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw. Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins: Structure, Function, and Bioinformatics*, 78(8):1950– 1958, 2010.
- [114] A. Linhananta, S. Hadizadeh, and S. S. Plotkin. An effective solvent theory connecting the underlying mechanisms of osmolytes and denaturants for protein stability. *Biophysical journal*, 100(2):459–468, 2011.
- [115] P. Linsel-Nitschke and A. R. Tall. Hdl as a target in the treatment of atherosclerotic cardiovascular disease. *Nature reviews Drug discovery*, 4(3):193–205, 2005.
- [116] H. Lodish, A. Berk, S. L. Zipursky, P. Matsudaira, D. Baltimore, and J. Darnell. Molecular cell biology 4th edition. National Center for Biotechnology InformationÕs Bookshelf, 2000.
- [117] Y. Lu, P. Derreumaux, Z. Guo, N. Mousseau, and G. Wei. Thermodynamics and dynamics of amyloid peptide oligomerization are sequence dependent. *Proteins: Structure, Function, and Bioinformatics*, 75(4):954, 2009.
- [118] B. Ma and R. Nussinov. Molecular dynamics simulations of a β -hairpin fragment of protein g: Balance between side-chain and backbone forces. Journal of Molecular Biology, 296:1091, 2000.
- [119] B. Ma and R. Nussinov. Molecular dynamics simulations of alanine rich β -sheet oligomers: Insight into amyloid formation. *Protein Sci.*, 11:2335, 2002.
- [120] J. Ma, I. M. Pazos, and F. Gai. Microscopic insights into the protein-stabilizing effect of trimethylamine n-oxide (tmao). Proceedings of the National Academy of Sciences of the United States of America, 111:8476, 2014.
- [121] R. D. Macdonald and M. Khajehpour. Effects of the osmolyte tmao (trimethylaminen-oxide) on aqueous hydrophobic contact-pair interactions. *Biophysical chemistry*, 184:101–107, 2013.
- [122] R. D. Macdonald and M. Khajehpour. Effects of the protein denaturant guanidinium chloride on aqueous hydrophobic contact-pair interactions. *Biophysical chemistry*, 196:25–32, 2015.
- [123] G. I. Makhatadze and P. L. Privalov. Protein interactions with urea and guanidinium chloride: a calorimetric study. *Journal of Molecular Biology*, 226(2):491–505, 1992.
- [124] D. Markthaler, J. Zeman, J. Baz, J. Smiatek, and N. Hansen. Validation of trimethylamine-n-oxide (tmao) force fields based on thermophysical properties of aqueous tmao solutions. *The Journal of Physical Chemistry B*, 2017.
- [125] J. Maupetit, P. Derreumaux, and P. Tufféry. Pep-fold: an online resource for de novo peptide structure prediction. Nucl. Acids Res., 37:W498, 2009.

- [126] J. Maupetit, P. Derreumaux, and P. Tufféry. A fast method for large-scale de novo peptide and miniprotein structure prediction. *Journal of Computational Chemistry*, 31:726, 2010.
- [127] M. F. McCarty and J. J. DiNicolantonio. Trimethylamine-n-oxide and heart failure. Journal of the American College of Cardiology, 66(1):94–96, 2015.
- [128] R. McNulty, J. P. Ulmschneider, H. Luecke, and M. B. Ulmschneider. Mechanisms of molecular transport through the urea channel of helicobacter pylori. *Nature* communications, 4, 2013.
- [129] S. Moelbert, B. Normand, and P. De Los Rios. Kosmotropes and chaotropes: modelling preferential exclusion, binding and aggregate stability. *Biophysical chemistry*, 112(1):45–57, 2004.
- [130] B. Moeser and D. Horinek. Unified description of urea denaturation: backbone and side chains contribute equally in the transfer model. *The Journal of Physical Chemistry B*, 118(1):107–114, 2013.
- [131] J. Mondal, G. Stirnemann, and B. Berne. When does trimethylamine n-oxide fold a polymer chain and urea unfold it? The Journal of Physical Chemistry B, 117(29):8723–8732, 2013.
- [132] K. L. Munroe, D. H. Magers, and N. I. Hammer. Raman spectroscopic signatures of noncovalent interactions between trimethylamine n-oxide (tmao) and water. *The Journal of Physical Chemistry B*, 115(23):7699–7707, 2011.
- [133] C. Narayanan and C. L. Dias. Hydrophobic interactions and hydrogen bonds in β -sheet formation. Journal of Chemical Physics, 139:115103, 2013.
- [134] C. Narayanan and C. L. Dias. Exploring the free energy landscape of a model β -hairpin peptide and its isoform. *Proteins: Structure, Function, and Bioinformatics*, 82(10):2394–2402, 2014.
- [135] J. Nasica-Labouze, M. Meli, P. Derreumaux, G. Colombo, and N. Mousseau. A multiscale approach to characterize the early aggregation steps of the amyloidforming peptide gnnqqny from the yeast prion sup-35. *PLoS Computational Biology*, 7:e1002051, 2011.
- [136] J. W. Neidigh, R. M. Fesinmeyer, and N. H. Andersen. Designing a 20-residue protein. Nature Structural & Molecular Biology, 9(6):425–430, 2002.
- [137] R. Nelson, M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe, and D. Eisenberg. Structure of the cross-β spine of amyloid-like fibrils. *Nature*, 435(7043):773, June 2005.
- [138] P. Nguyen and P. Derreumaux. Understanding amyloid fibril nucleation and $a\beta$ oligomer/drug interactions from computer simulations. Accounts of chemical research, 47(2):603–611, 2013.

- [139] P. Nguyen and P. Derreumaux. Understanding amyloid fibril nucleation and $a\beta$ oligomer/drug interactions from computer simulations. Accounts of Chemical Research, 47:603, 2014.
- [140] Y. Nozaki and C. Tanford. The solubility of amino acids and related compounds in aqueous urea solutions. Journal of Biological Chemistry, 238(12):4074–4081, 1963.
- [141] D. E. Otzen, O. Kristensen, and M. Oliveberg. Designed protein tetramer zipped together with a hydrophobic alzheimer homology: A structural clue to amyloid assembly. *Proceedings of the National Academy of Sciences of the United States* of America, 97:9907, 2000.
- [142] J. D. P. Sulc and A. A. Louis. Introduction to molecular simulation. MIT Press, 2015.
- [143] M. Patra and M. Karttunen. Systematic comparison of force fields for microscopic simulations of nacl in aqueous solutions: diffusion, free energy of hydration, and structural properties. *Journal of computational chemistry*, 25(5):678–689, 2004.
- [144] S. Paul and G. Patey. The influence of urea and trimethylamine-n-oxide on hydrophobic interactions. The Journal of Physical Chemistry B, 111(28):7932– 7933, 2007.
- [145] L. Pauling and R. B. Corey. Configurations of polypeptide chains with favored orientations around single bonds: Two new pleated sheets. Proceedings of the National Academy of Sciences of the United States of America, 37(11):729–740, Nov. 1951.
- [146] L. Pauling and R. B. Corey. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy* of Sciences of the United States of America, 37:205, 1951.
- [147] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. DeBolt, D. Ferguson, G. Seibel, and P. Kollman. Amber, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications*, 91(1):1–41, 1995.
- [148] A. Perczel, P. Hudáky, and V. K. Pálfi. Dead-end street of protein folding: Thermodynamic rationale of amyloid fibril formation. Journal of the American Chemical Society, 129(48):14959, 2007.
- [149] V. Pierce, M. Kang, M. Aburi, S. Weerasinghe, and P. E. Smith. Recent applications of kirkwood-buff theory to biological systems. *Cell biochemistry* and biophysics, 50(1):1–22, 2008.
- [150] P. L. Privalov. Cold denaturation of protein. Critical Reviews in Biochemistry and Molecular Biology, 25(4):281, Jan. 1990.

- [151] W. Ramsden. On some new properties of urea (with demonstrations). The British Medical Journal, pages 783–784, 1902.
- [152] A. Razvi and J. M. Scholtz. Lessons in stability from thermophilic proteins. Protein Science, 15(7):1569–1578, 2006.
- [153] D. R.Canchi and A. E. Garcia. Cosolvent effects on protein stability. *Annual Review* of Physical Chemistry, 2013.
- [154] R. L. Remmele Jr, J. Zhang-van Enk, D. Phan, and L. Yu. Stabilization by urea during thermal unfolding-mediated aggregation of recombinant human interleukin-1 receptor (type ii): does solvation entropy play a role? *The Journal of Physical Chemistry B*, 116(24):7240–7251, 2012.
- [155] Y. L. A. Rezus and H. J. Bakker. Effect of urea on the structural dynamics of water. Proceedings of the National Academy of Sciences of the United States of America, 103:18417, 2006.
- [156] D. R. Robinson and W. P. Jencks. The effect of compounds of the urea-guanidinium class on the activity coefficient of acetyltetraglycine ethyl ester and related compounds1. Journal of the American Chemical Society, 87(11):2462–2470, 1965.
- [157] F. Rodriguez-Ropero, P. Rotzscher, and N. F. van der Vegt. Comparison of different tmao force fields and their impact on the folding equilibrium of a hydrophobic polymer. *The Journal of Physical Chemistry B*, 120(34):8757–8767, 2016.
- [158] F. Rodriguez-Ropero and N. F. van der Vegt. Direct osmolyte-macromolecule interactions confer entropic stability to folded states. The Journal of Physical Chemistry B, 118(26):7327-7334, 2014.
- [159] R. Z. Ronen Zangi and B. J. Berne. Urea's action on hydrophobic interactions. Journal of the American Chemical Society, 2009.
- [160] G. D. Rose, P. J. Fleming, J. R. Banavar, and A. Maritan. A backbone-based theory of protein folding. *Proceedings of the National Academy of Sciences of the* United States of America, 103(45):16623, 2006.
- [161] G. D. Rose and R. Wolfenden. Hydrogen bonding, hydrophobicity, packing, and protein folding. Annual Review of Biophysics and Biomolecular Structure, 22:381, 1993.
- [162] L. B. Sagle, K. Cimatu, V. A. Litosh, Y. Liu, S. C. Flores, X. Chen, B. Yu, and P. S. Cremer. Methyl groups of trimethylamine n-oxide orient away from hydrophobic interfaces. *Journal of the American Chemical Society*, 133(46):18707–18712, 2011.
- [163] S. Santini, G. Wei, N. Mousseau, and P. Derreumaux. Pathway complexity of alzheimer's β -amyloid a β_{16-22} peptide assembly. *Structure*, 12(7):1245, 2004.

- [164] M. R. Sawaya, S. Sambashivan, R. Nelson, M. I. Ivanova, S. A. Sievers, M. I. Apostol, M. J. Thompson, M. Balbirnie, J. J. W. Wiltzius, M. A. McFarlane, Heather T., C. Riekel, and D. Eisenberg. Atomic structures of amyloid cross-beta spines reveal varied steric zippers. *Nature*, 447:453, 2007.
- [165] C. Scharnagl, O. Pester, P. Hornburg, D. Hornburg, A. Götz, and D. Langosch. Side-chain to main-chain hydrogen bonding controls the intrinsic backbone dynamics of the amyloid precursor protein transmembrane helix. *Biophysical Journal*, 106:1318, 2014.
- [166] J. Schmit, K. Ghosh, and K. A. Dill. What drives amyloid molecules to assemble into oligomers and fibrils? *Biophysical Journal*, 100:450, 2011.
- [167] E. Schneck, D. Horinek, and R. R. Netz. Insight into the molecular mechanisms of protein stabilizing osmolytes from global force-field variations. *The Journal of Physical Chemistry B*, 117(28):8310–8321, 2013.
- [168] M. A. Schroer, Y. Zhai, D. C. F. Wieland, C. J. Sahle, J. Nase, M. Paulus, M. Tolan, and R. Winter. Exploring the piezophilic behavior of natural cosolvent mixtures. Angewandte Chemie, 123(48):11615, 2011.
- [169] D. E. Shaw, P. Maragakis, K. Lindorff-Larsen, S. Piana, R. O. Dror, M. P. Eastwood, J. A. Bank, J. M. Jumper, J. K. Salmon, Y. Shan, and W. Wriggers. Atomic-level characterization of the structural dynamics of proteins. *Science*, 330:60002, 2010.
- [170] S. Shimizu and H. S. Chan. Origins of protein denatured state compactness and hydrophobic clustering in aqueous urea: Inferences from nonpolar potentials of mean force. *Proteins: Structure, Function, and Bioinformatics*, 49:560, 2002.
- [171] N. Smolin, V. Voloshin, A. Anikeenko, A. Geiger, R. Winter, and N. N. Medvedev. Tmao and urea in the hydration shell of the protein snase. *Physical Chemistry Chemical Physics*, (accepted for publication), 2017.
- [172] W. Song, G. Wei, N. Mousseau, and P. Derreumaux. Self-assembly of the β 2-microglobulin nhvtlsq peptide using a coarse-grained protein model reveals a β -barrel species. Journal of Physical Chemistry B, 112:4410, 2008.
- [173] A. Soper, E. Castner, and A. Luzar. Impact of urea on water structure: a clue to its properties as a denaturant? *Biophysical chemistry*, 105(2):649–666, 2003.
- [174] F. Sterpone, P. H. Nguyen, M. Kalimeri, and P. Derreumaux. Importance of the ion-pair interactions in the opep coarse-grained force field: Parametrization and validation. *Journal of Chemical Theory and Computation*, 9:4574, 2013.
- [175] J. E. Straub and D. Thirumalai. Principles governing oligomer formation in amyloidogenic peptides. Current Opinion in Structural Biology, 20(2):187, 2010.

- [176] T. O. Street, D. W. Bolen, and G. D. Rose. A molecular mechanism for osmolyteinduced protein stability. *Proceedings of the National Academy of Sciences*, 103(38):13997–14002, 2006.
- [177] M. C. Stumpe and H. Grubmuller. Aqueous urea solutions: Structure, energetics, and urea aggregation. *Journal of Physical Chemistry B*, 111:6220, 2007.
- [178] M. C. Stumpe, H. Grubmüller, and V. S. Pande. Polar or apolar the role of polarity for urea-induced protein denaturation. *PLoS Comput Biol*, 4(11):e1000221, 2008.
- [179] Z. Su and C. L. Dias. Driving beta-strands into fibrils. The Journal of Physical Chemistry B, 118(37):10830–10836, 2014.
- [180] Z. Su and C. L. Dias. Molecular interactions accounting for protein denaturation by urea. Journal of Molecular Liquids, 228:168, 2017.
- [181] Z. Su, F. Mahmoudinobar, and C. L. Dias. Effects of trimethylamine-n-oxide on the conformation of peptides and its implications for proteins. *Physical Review Letters*, 119(10):108102, 2017.
- [182] Y. Sugita and Y. Okamoto. Replica-exchange molecular dynamics method for protein folding. *Chemical physics letters*, 314(1):141–151, 1999.
- [183] S. Sukenik, S. Dunsky, A. Barnoy, I. Shumilin, and D. Harries. Tmao mediates effective attraction between lipid membranes by partitioning unevenly between bulk and lipid domains. *Physical Chemistry Chemical Physics*, 19(44):29862– 29871, 2017.
- [184] A. Szilágyi and P. Závodszky. Structural differences between mesophilic, moderately thermophilic and extremely thermophilic protein subunits: results of a comprehensive survey. *Structure*, 8(5):493–504, 2000.
- [185] T. Takeda and D. K. Klimov. Computational backbone mutagenesis of $a\beta$ peptides: Probing the role of backbone hydrogen bonds in aggregation. Journal of Physical Chemistry B, 114(14):4755, 2010.
- [186] C. Tanford. Extension of the theory of linked functions to incorporate the effects of protein hydration. Journal of Molecular Biology, 39(3):539–544, 1969.
- [187] C. Tanford and J. Reynolds. *Nature's robots: a history of proteins*. OUP Oxford, 2003.
- [188] W. W. Tang and S. L. Hazen. The contributory role of gut microbiota in cardiovascular disease. The Journal of clinical investigation, 124(10):4204, 2014.
- [189] B. Tarus, J. E. Straub, and D. Thirumalai. Dynamics of asp23-lys28 salt-bridge formation in $a\beta_{10-35}$ monomers. Journal of the American Chemical Society, 128(50):16159, 2006.

- [190] D. Thirumalai, G. Reddy, and J. E. Straub. Role of water in protein aggregation and amyloid polymorphism. *Accounts of Chemical Research*, 45(1):83, 2012.
- [191] J. Timko, D. Bucher, and S. Kuyucak. Dissociation of nacl in water from ab initio molecular dynamics simulations. The Journal of chemical physics, 132(11):114510, 2010.
- [192] D. S. Tomar, V. Weber, B. M. Pettitt, and D. Asthagiri. Conditional solvation thermodynamics of isoleucine in model peptides and the limitations of the group-transfer model. *The Journal of Physical Chemistry B*, 118(15):4080– 4087, 2014.
- [193] D. Trzesniak, N. F. van der Vegt, and W. F. van Gunsteren. Computer simulation studies on the solvation of aliphatic hydrocarbons in 6.9 m aqueous urea solution. *Physical Chemistry Chemical Physics*, 6(4):697–702, 2004.
- [194] R. Tycko. Insights into the amyloid folding problem from solid-state nmr. Biochemistry, 42:3151, 2003.
- [195] M. Ufnal, A. Zadlo, and R. Ostaszewski. Tmao: A small molecule of great expectations. *Nutrition*, 31(11):1317–1323, 2015.
- [196] T. Urbic and C. Dias. Hydration of non-polar anti-parallel beta-sheets. J. Chem. Phys., 140:165101, 2014.
- [197] D. van der Spoel et al. *GROMACS USER MANUAL Version 4.0.* The GROMACS development team, 2005.
- [198] M. T. Velasquez, A. Ramezani, A. Manal, and D. S. Raj. Trimethylamine n-oxide: the good, the bad and the unknown. *Toxins*, 8(11):326, 2016.
- [199] Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. DuGar, A. E. Feldstein, E. B. Britt, X. Fu, Y.-M. Chung, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*, 472(7341):57– 63, 2011.
- [200] G. Wei, N. Mousseau, and P. Derreumaux. Sampling the self-assembly pathways of kffe hexamers. *Biophysical Journal*, 87:3648, 2004.
- [201] P. Yancey. Proteins and counteracting osmolytes. *Biologist*, 50(3):126–131, 2003.
- [202] P. H. Yancey. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *Journal of Experimental Biology*, 208(15):2819–2830, 2005.
- [203] P. H. Yancey, M. E. Clark, S. C. Hand, R. D. Bowlus, and G. N. Somero. Living with water stress: evolution of osmolyte systems. *Science*, 217(4566):1214–1222, 1982.

- [204] P. H. Yancey, M. E. Gerringer, J. C. Drazen, A. A. Rowden, and A. Jamieson. Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *Proceedings of the National Academy of Sciences*, 111(12):4461– 4465, 2014.
- [205] A.-S. Yang and B. Honig. Free energy determinants of secondary structure formation: Ii. antiparallel β-sheets. Journal of Molecular Biology, 252(3):366, 1995.
- [206] Z. Yang, P. Xiu, B. Shi, L. Hua, and R. Zhou. Coherent microscopic picture for ureainduced denaturation of proteins. *Journal of Physical Chemistry B*, 116:8856, 2012.
- [207] K. Yui, M. Sakuma, and T. Funazukuri. Molecular dynamics simulation on ion-pair association of nacl from ambient to supercritical water. *Fluid Phase Equilibria*, 297(2):227–235, 2010.
- [208] R. Zangi, R. Zhou, and B. Berne. Ureas action on hydrophobic interactions. Journal of the American Chemical Society, 131(4):1535–1541, 2009.
- [209] R. Zhou. Trp-cage: folding free energy landscape in explicit water. Proceedings of the National Academy of Sciences, 100(23):13280–13285, 2003.
- [210] R. Zhou, J. Li, L. Hua, Z. Yang, and B. Berne. Comment on urea-mediated protein denaturation: A consensus view. The Journal of Physical Chemistry B, 115(5):1323–1326, 2011.
- [211] Q. Zou, B. J. Bennion, V. Daggett, and K. P. Murphy. The molecular mechanism of stabilization of proteins by tmao and its ability to counteract the effects of urea. Journal of the American Chemical Society, 124(7):1192–1202, 2002.