Effect of Confinement on Protein Conformation in Presence of Osmolytes Urea and Trimethylamine N-Oxide: Replica Exchange Molecular Dynamics Simulation Study

A Thesis Submitted

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Gargi Borgohain

to the

Department of Chemistry

Indian Institute of Technology Guwahati, India

2016

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

DEDICATED TO
MY FAMILY MEMBERS

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Declaration

I hereby declare that the matter manifested in this thesis entitled "*Effect of Confinement on Protein Conformation in Presence of Osmolytes Urea and Trimethylamine N-Oxide: Replica Exchange Molecular Dynamics Simulation Study*" is the result of research carried out by me in the Department of Chemistry, Indian Institute of Technology Guwahati, India under the supervision of Dr. Sandip Paul.

In keeping with the general practice of reporting scientific observations, due acknowledgement has been made wherever the work described is based on the findings of other investigators.

Gargi Borgohain
IIT Guwahati

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Certificate

It is certified that the work contained in this thesis entitled, "*Effect of Confinement on Protein Conformation in Presence of Osmolytes Urea and Trimethylamine N-Oxide: Replica Exchange Molecular Dynamics Simulation Study*" has been carried out by Ms. Gargi Borgohain for the Degree of Doctor of Philosophy under my supervision and the same has not been submitted elsewhere for a degree.

Dr. Sandip Paul
Thesis Supervisor
Department of Chemistry
Indian Institute of Technology Guwahati
Guwahati-781039, India

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Preface

On accomplishment of my doctorate studies, I would like to express my sincere gratitude to everyone who directly or indirectly helped me during this wonderful journey of my research career. First and foremost, with a deepest sense of gratitude, I wish to express my sincere thanks to my supervisor Dr. Sandip Paul for giving me the opportunity to work in his research group. His valuable guidance, encouragement, inspiration and creative scientific ideas are helpful to raise my knowledge in research work. I am also thankful to him for giving me the freedom to pursue my own ideas and I find myself privileged to work under his kind guidance. His door was always open for the countless discussion sessions with him which I am going to treasure forever. My everlasting gratitude goes towards him.

I would like to acknowledge my sincere gratitude to the doctoral committee chair and members, Dr. Debasis Manna, Dr. Sumana Dutta and Dr. Padma Kumar Padmanabhan for periodically assessing my work and providing valuable suggestions for its improvement. My sincere thanks go to all other faculty members in the department for their kind help at various stages of my doctoral work. I am very grateful to the Ministry of Human Resource and Development (MHRD) and Department Of Science and Technology (DST), India for financial support and IIT Guwahati for all the facilities that were made available to me. I am also grateful to the Center for Development of Advanced Computing (CDAC) for the computational support.

Further, I would like to thank all my past and present group members for their support, helpful discussions and creating a wonderful environment in the lab. A big thanks to Dr. Rahul Sarma and Dr. Subrata Paul for their help, suggestions and sharing their many experiences during my work. I am also very grateful to all other group members Bhanita, Shubhadip, Krishna, Srijita, Saikat, Rabindra and Shubhajit. I extend my sincere thanks to Dr. Renjith Bhaskaran, Debashis, Harikrishna for their suggestions throughout my work and my friends Sabera, Nirmali, Sujit, Afsana, Hemanta, Bedika and the other research scholars of Chemistry department who have shared their thoughts and views with me and all the joyful moments we shared in this campus.

I take this opportunity to express my sincere thanks to all my teachers in school, college, and university days for helping and encouraging me in various aspects of my life and academics. All the learning from them will be an asset in every walk of my life.

Finally, my Ph. D. endeavor could not have been completed without the endless love, support, and blessings from my family members. I wish to express my sincere gratitude to my parents (Late Laxmi Prasad Borgohain and Mrs. Chandrama Borgohain), my brothers, parents-in-laws and at the end I express my acknowledgement to my better half Pradyumna Mazumdar. They are the source of inspiration for each and every step in my life and I dedicate this work to them.

Gargi Borgohain

2016

"Nature uses only the longest threads to weave her patterns, so that each small piece of her fabric reveals the organization of the entire tapestry."

— Richard P. Feynman

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Outline of the Thesis

Chapter 1: Introduction 1

Chapter 2: Model Dependency of TMAO's Counteracting Effect Against Action of Urea 17

Chapter 3: Stabilization of the Protein Trp Cage in Mixed Osmolyte Solutions of Urea and TMAO Without Confinement 41

Chapter 4: Effect of Non Polar Confinement on Protein Trp Cage Conformation in Aqueous Osmolyte Solutions 69

Chapter 5: Effect of Polar Confinement and Osmolyte Solutions on Folding/Unfolding Equilibrium of Protein Trp Cage 95

Chapter 6: Summary and Our View on Osmolyte and Confinement Induced Stabilization of Protein 119

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Chapter 1

Introduction

" The appeal of MD simulations is due to their ability to provide realistic descriptions of the actual molecular motions, and the ability to simulate the time evolution of reactive trajectories is very useful. Furthermore, MD simulations appear to provide a simple and effective way of averaging over protein configurations and obtaining the free energies that control biological functions."

– Arieh Warshel *Acc. Chem. Res.* **35**, 385 (2002)

■ PROTEIN DENATURATION AND COUNTERACTION

How and why a denatured polypeptide chain can spontaneously fold into a compact and highly ordered conformation has remained an important unresolved question in wide range of disciplines. In order to attain their respective biological activities, protein must have to fold into precise three-dimensional conformations. The folding of protein into its native state is the most fundamental example of biological self-assembly. Hence understanding this process will help to elucidate the process of evolutionary selection of properties of a biological system for functional advantage. Native state of a protein refers to the state of the highest thermodynamic stability under physiological conditions. If the folding process of a protein from its unfolded state undergoes all possible conformations, it will take an astronomical length of time to reach to the native state. However, literature [1, 2] revealed that folding process does not involve a series of mandatory steps between specific steps, but rather a stochastic search path is followed. Anfinsen and co-workers proposed that the compactness of the three dimensional structure of a protein depends intrinsically on the sequence of amino acid residues as well as on the multiple contributions from the cellular environment [3]. The uniqueness of energy landscape of a particular protein also pertains to amino acid sequence and this leads natural selection to evolve so that the protein gets native folded state rapidly and efficiently. Anfinsen's work shed light on understanding the process through which the sequence of polypeptide chain guides the whole protein to a particular three-dimensional conformation. To fully map the folding/unfolding process all conformational ensembles (from native state to denatured state) are required to characterize and the mechanism of conversion is also needed to explore. Experimental approaches can provide only limited information regarding these aspects. Molecular dynamics (MD) simulation is the most realistic technique available that allows time dependent monitoring of the detailed interactions of the protein with its surrounding chemical environments. In present days protein unfolding simulations are carried out equally as because they can be expected to reliably depict possible folding pathways of protein when viewed in reverse [4]. Theory of microscopic reversibility dictates that the mechanism of folding can be probed in both the directions of unfolding and refolding under same conditions [4]. Over recent years a huge number of research works have been reported in the approach of fully mapping the protein folding pathways. Anticipating the process of folding and unfolding of protein will contribute in understanding of biological processes also, such as protein degradation and translocation, aging, and human diseases. In view of this, unfolding of protein is relevant

to investigate since it plays crucial role in development of different amyloid diseases and many other cellular processes [5-7]. This leads us to focus on simulations of proteins in solutions to investigate the detailed processes of protein folding/unfolding equilibrium.

We consider two different widely studied proteins GB1 and Trp cage that are known to exhibit adequate folding behavior with simulation techniques. Small β-hairpin peptides have been used as model peptides for studying folding/unfolding mechanism of proteins [8, 9]. β-hairpin is the simplest example of an anti-parallel β-sheet found in protein structures that may be considered as a suitable system to study the fundamental issues in protein folding. A number of recent studies have also examined the osmolyte effects on the stability of β-hairpins [10-14]. Several computational studies are reported to explore the folding mechanism of β-hairpin and the effects of osmolytes on its stability. Because of the small size and fast kinetics, Trp cage is another extensively investigated protein in terms of computational and experimental works [15-20]. It has a well packed hydrophobic core and typical tertiary contacts that make it a highly stable paradigm in the field of protein research.

The challenges in the conformational sampling of complex biological systems prohibit the complete exploration of energy surfaces. In literature many significant efforts have been made till date for developing efficient simulation methods. Since from the birth of Replica Exchange Molecular Dynamics (REMD) in 1999 by Sugita and Okamoto [21], it has been a very popular approach to carry out biomolecular simulations. We have considered REMD simulation technique over the regular canonical molecular dynamics (MD) simulation because of the fact that for the latter technique it is usually difficult to obtain accurate canonical distributions both at low- and room temperatures [21]. Because simulations at such temperatures tend to get trapped in one of the huge number of local or minimum energy states. REMD overcomes the energy barriers by state exchange moves. One replica at a higher temperature exchanges its temperature with the immediate lower temperature replica and vice versa. This exchange facilitates barrier crossings of the energy landscape and helps to create the suited trajectories for configurational transitions. This in turn helps to trace out the trajectory of conformational transitions followed by the protein under consideration. REMD simulation technique is also used for enhanced conformational sampling which is always preferred for protein simulation studies. REMD ensures efficient sampling by crossing the energetic barriers. REMD method has numerous other advantages too. It produces information over a range of temperatures (so that we get the melting temperature curve of the protein), it works equally on explicit and implicit

solvents. In comparison to constant temperature MD (of protein simulation), it reduces the sampling time by a factor of 20 or so. The ability to run REMD on parallel processors has made it a popular tool for researchers, however, accurate measures of simulation data for large proteins remains to be exploited.

Protein conformation is sensitive to its surrounding environmental conditions such as temperature, pressure and assorted molecules comprising the volume around it [22]. Some small molecules that accumulate in high concentration inside the cell in response to osmotic stress are known as osmolytes. Nature exemplifies certain evidences of using protecting and denaturing osmolytes to restore cellular functions. For example, certain marine animals have adapted life at very high pressure and salinity by using osmolytes having denaturing effect such as urea [23] and protecting osmolyte such as betaine, trimethylamine N-oxide (TMAO) etc. to counteract the effect of denaturant [24]. Protecting osmolytes are often classified into two classes 'compatible' or 'counteracting' osmolytes [25, 26]. Compatible osmolytes increase protein stability but do not affect functional activity of the protein, whereas counteracting osmolytes cause changes in protein function. Some methylamines (e.g. trimethylamine-N-oxide (TMAO), betaine etc.) are considered as counteracting osmolytes. Some amino acids (e.g. proline and glycine) and polyols (e.g. trehalose, sucrose, and sorbitol) are considered compatible osmolytes. Urea is an extensively studied denaturing osmolyte, which has been postulated to exert its effect either directly [27] (by binding to the protein) or indirectly [28] (by altering the solvent structure which in turn offers denaturing effect on the protein). Some studies attributed the urea denaturation to both direct and indirect mechanisms [29, 30] while others emphasized the role of hydrogen bonding between urea and protein [31, 32]. However, the exact mechanism of urea-induced denaturation is not adequately established till date [33]. The osmolyte TMAO is found in the urea-rich cells of elasmobranchs and coelacanths to protect the protein against deleterious effect of urea. TMAO is considered as counteracting osmolyte which imparts its effect on both structural stability and functions of the protein. Studies based on TMAO-water interaction showed that the unfavorable interaction of TMAO with protein backbone is dominantly responsible for its stabilization effect [34]. Further, studies on the molecular interactions between TMAO and protein functional groups also give evidence of some favorable and unfavorable contributions towards the overall stabilizing effect of the osmolyte [34, 35]. Folding propensity of the protein was proposed to arise due to enhanced water structure in presence of TMAO which in turn affects the folding/unfolding equilibrium of the protein [36, 37]. Some studies are found to highlight the direct interaction between

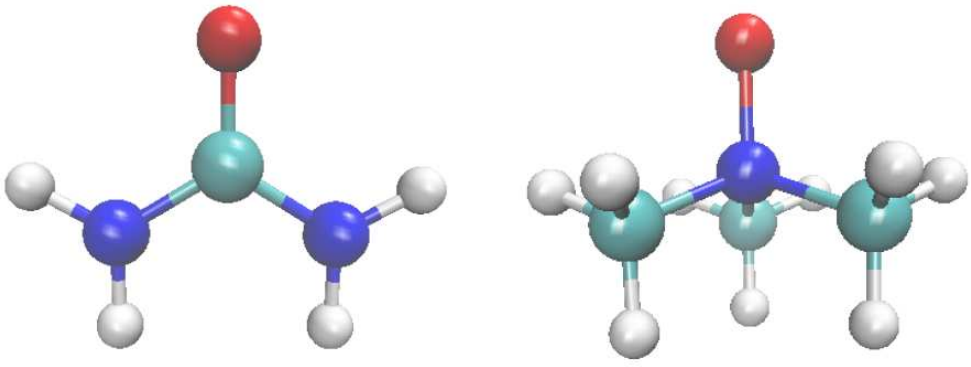
protein and TMAO molecules [36, 38, 39]. Thirumalai and co-workers [40] have proposed that TMAO acts as crowding agent that favors compactly folded native conformations of the protein through excluded volume effect. In other words TMAO increases stability of protein by acting as a nano-crowder. Crowding role of TMAO on native state stabilization is also supported by a recent study of F. Gai and co-workers [41]. Another study [42] proposed that TMAO produces a local environment surrounding the protein that increases the strength of backbone-backbone hydrogen bonds and reduces the conformational entropy of the entire system. Bolen and co-workers [43] suggested that the osmophobic interaction of the peptide backbone with TMAO molecules plays a vital role in protein stabilization. Pettit and co-workers [44] have explored the conformational preferences of decaalanine peptide in osmolyte solutions from the calculations of solvation free energy changes. They found that TMAO exerts its effect in ternary solutions from the most favorable van der Waals component. In a recent work Garcia and co-workers [33] reported the preferential exclusion of TMAO molecules from the surface of protein. Another study by Ganguly et al. [45] has shown that both urea and TMAO mutually exclude each other from the protein surface with simultaneous enhancement of urea-urea and TMAO-TMAO self-aggregation. They support indirect mechanism for TMAO stabilization which is established through strong TMAO-water interaction. In a recent study [46], Ganguly et al. emphasized on the delicate balance between TMAO-water and TMAO-urea interactions. TMAO-urea interaction results in depletion of urea from the protein surface. They have also reported that choice of different force field parameters of different atomic sites of TMAO have substantial impact on these interactions in mixed urea-TMAO solution. Simulation studies on TMAO-water mixture provide strong support of compactly bound water molecules around the polar head group of TMAO molecule [47]. Though a lot of research works has been devoted still study on this crucial osmolyte is captivating new area of surveillance everyday. In literature many experimental evidences also exist showing urea induced denaturation, TMAO induced protection and also counteraction of TMAO against urea action [42, 48]. The findings are substrate specific based on intermolecular interaction mechanism. These studies inspired us to investigate and characterize the protection and denaturation mechanisms of osmolytes urea and TMAO at molecular level using REMD simulation technique.

For possible quantitative characterization of different contributing factors of TMAO on protein stabilization different model systems are considered. In most of the studies, Kast model of TMAO [36, 49, 50] is used as osmoprotectant. Recently, Garcia and co-workers [33] have introduced another model of TMAO called Osmotic model based on osmotic pressure

measurement of TMAO solutions. Osmotic model incorporates more realistic interactions through parameterization of the strength of Lennard-Jones interactions of different atomic sites and scaling down the charges of TMAO atomic sites. Some other models of TMAO are also available viz. Netz model [51] and Shea model [52]. These are developed on the base of Kast model to reproduce thermodynamic properties of aqueous TMAO solution. But till date no such force field parameters for TMAO is available that adequately describes the balance of interactions in mixture of solution. So, we have restricted our study to Kast model based on our previous study. In each of the work we have compared the behavior of the protein in urea solutions as well as in mixture of urea and TMAO. In the first work, we have compared the protection behavior of Kast model and Osmotic model [53]. Throughout the study we have tried to shed lights on the protection of TMAO against urea denaturation to find out the best fitted mechanism of TMAO counteraction. From this study, we have found that Kast model of TMAO gives better protection against urea denaturation when used in conventional 2:1 ratio of urea:TMAO. Hence for the rest of the studies (from **Chapter 3** onwards) we have considered Kast model of TMAO.

Inside the cell, the fundamental process of protein folding is assisted by a class of conserved protein known as molecular chaperones. Molecular chaperone can be defined as any protein that interacts, stabilizes, or helps a non native protein to acquire its native conformation, but is not present in the final functional structure [54]. Chaperones initiate folding process of a newly synthesized protein by preventing its aggregation and misfolding [55]. The GroEL/GroES chaperonin system of E. coli has been studied most extensively [55-57]. Proteins are stabilized by many long-range interactions and they also have tendency to populate kinetically trapped intermediates [58]. Ring-shaped chaperonin GroEL binds with non-native protein within its central cavity and together with its co-factor GroES, that acts as a lid, accomplish folding [59, 60]. Both the folded protein and GroES are released from the GroEL unit after reaching the folded state. GroES binding causes substantial conformational change that significantly increases hydrophilicity of the cavity wall (i.e. more number of polar residues project towards inner wall of the cavity) [54]. This situation promotes folding by arrangement of ordered water molecules by the side of the surface and thereby generating a local environment where the protein buries its exposed hydrophobic regions effectively and thus accelerates folding [54, 61]. Brinker et al. [62] suggested that GroEL/GroES chaperonin system induces a confinement effect on the protein that smoothen up the energy landscape of the protein by preventing kinetically trapped intermediates. The effect of GroES is negligible. In literature many computational

and experimental studies have been reported to study protein folding/unfolding behavior under confinement in order to understand the functioning of the chaperone unit [63-67]. Recent computational studies considered the effects of confinement on protein folding and stability in the context of chaperons, pores, and other cellular environments [63, 66, 67]. Tian et al. [63] addressed the protein Trp cage under polar and nonpolar confinement and showed that protein folding stability was influenced by several factors, such as the nature of confined wall, sequence of the protein, and entropic stabilization. They established that the nonpolar confinement stabilized the folded state while the polar confinement destabilized the protein. Coarse-grained simulations of proteins showed that the effect of confinement is dependent on the size of the confining volume relative to the size of the protein [67-69]. The common observations from most of the studies are that confinement stabilizes the folded state of proteins, confined solvent also plays a crucial rule and the effects of confinement are reflected in the reduction of the conformational entropy associated with protein folding, which in turn leads to the stabilization of the folded states. This leads us to carry out REMD simulations [21] of Trp cage inside a non polar as well as polar confinement in pure water and in presence of binary osmolyte solution of urea and ternary osmolyte solutions of urea and TMAO. Effect of confinements on the osmolytes and their subsequent effects on the protein is investigated thoroughly.



Urea TMAO (Trimethylamine N-oxide)

CPK presentations of urea and TMAO molecules. Carbon, nitrogen, oxygen and hydrogen atoms are presented by cyan, blue, red and white color respectively.

How TMAO stabilizes proteins and counteracts the denaturing effect of urea remains to be an important issue in protein science. In the current study, an effort has been made to address this issue. First, we anticipate the results for binary and ternary solutions of urea and urea-TMAO respectively. Then the mechanisms lying behind denaturation and counteraction are investigated and illustrated. The simulation works and concluding remarks are presented in the subsequent chapters. The next section of the present chapter deals with the basic techniques of MD and REMD employed in our work. The details of the analyses of the applications of these techniques for specific systems are given in respective chapters. This is followed by a brief description of the works presented in the current thesis.

■ METHODOLOGY

In this thesis, we have used classical molecular dynamics (MD) as well as replica exchange molecular dynamics (REMD) simulation techniques. REMD is a widely used method to enhance conformational sampling of MD simulations. REMD simulation can improve sampling of conformational states when the states are separated by significant energy barriers. The fundamental tools and methods underlying MD and REMD are the same. Note that, throughout this thesis, specific AMBER12 force field parameter sets are used [70].

In MD simulation, the potential energy function (U) is described by all interactions between the atoms that are covalently bonded as well as non-bonded interactions between atoms and molecules in the condensed phase. The interactions between particles are governed by the so-called force field parameterization [71].

The potential energy function is written as a sum of bonded and non-bonded interaction terms

$$U = U_{bond} + U_{angle} + U_{dihedral} + U_{vdw} + U_{Coulomb} \tag{1.1}$$

The first three terms (U_{bond}, U_{angle}, $U_{dihedral}$) are the bonded terms, which describe the bond stretching, angle bending, and torsion rotation, and the last two terms are for the non-bonded potential. In bonded terms, the bond and angle contributions are described by harmonic potentials and all of the interactions between directly bonded atoms (1-2 interactions), angles (1-3 interactions, where two atoms bonded to a common atom), and torsion (interactions between pairs of 1-4 atoms) are defined as:

$$U_{bond} = \sum_{bonds} K_b(b_{ac} - b_{eq})^2 \tag{1.2}$$

$$U_{angle} = \sum_{angles} K_\theta(\theta_{ac} - \theta_{eq})^2 \tag{1.3}$$

$$U_{dihedral} = \sum_{dihedrals} \frac{V_n}{2}(1 + cos(n\phi - \delta)) \tag{1.4}$$

The letters b, θ, ϕ, and δ represent the bond length, bond angle, dihedral angle, and phase angle, respectively. The subscripts ac stands for actual and eq stands for equilibrium. The parameters K_b, K_θ, and V_n are the force constants for bond, bond angle, and dihedral angle respectively.

The non-bonded potentials are calculated using two terms, the first one is the Lennard-Jones term (U_{vdw}) [72] describing the van der Waals interaction [73], and the second one is the Coulomb term ($U_{coulomb}$) [74] that deals with the electrostatic interactions between particles having partial charges on them. The non-bonding interaction terms are defined as:

$$U_{vdw} = \sum_i \sum_{i<j} 4\epsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{1.5}$$

$$U_{coulomb} = \sum_i \sum_{i<j} [\frac{q_i q_j}{4\pi\epsilon_o r_{ij}}] \tag{1.6}$$

where the overall sum is over all the atom pairs i and j. Lennard-Jones parameters σ and ϵ are the diameter of atomic site and well depthenergy, respectively. r_{ij} is the inter-atomic distance. q_i and q_j are the partial charges on interaction sites i and j and ϵ_o is the electrical permittivity.

The aim of the MD simulation is to observe the evolution of atomic coordinates in time. We consider an N-particle system characterized by the following Hamiltonian

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + U(\mathbf{r}^N) \tag{1.7}$$

where m is the mass of each particle, p_i is the momentum of the i-th particle and $U(\mathbf{r}^N)$ is the total potential energy of the system which includes the all particle-particle interactions. The coordinates of the particles are denoted by $\mathbf{r}^N = \{\mathbf{r}_1, \cdots, \mathbf{r}_N\}$. The position and velocity of i-th particle is represented by $\mathbf{r}_i$ and $\mathbf{v}_i$, respectively. The method of molecular dynamics consists of solving the equation

$$a_i = \frac{\mathbf{F_i}}{m_i} \tag{1.8}$$

where $i = 1, 2, ...N$, m_i is the mass of i-th particle and $\mathbf{F}_i$ is the force acting on particle

i. This equation is obtained easily from the Lagrangian

$$L = \frac{1}{2}\sum_{i=1}^{N} m_i \mathbf{v_i}.\mathbf{v_i} - \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} u(r_{ij}) \tag{1.9}$$

where the potential U has been assumed to be the sum of pair potentials u_{ij}. The Lagrangian equation of motion is

$$\frac{d}{dt}(\frac{\partial L}{\partial \dot{q}_i}) - \frac{\partial L}{\partial \dot{q}_i} = 0 \tag{1.10}$$

It is clear from Eq. 1.10 that the dynamics of particles is described by 3N number of second order differential equations.

It is also possible to write down the Hamiltonian (H) for the system and solve the the Hamiltonian equations of motion

$$\dot{\mathbf{q}}_k = \frac{\partial H}{\partial p_k} \tag{1.11}$$

$$\dot{\mathbf{p}}_k = -\frac{\partial H}{\partial q_k} \tag{1.12}$$

where $\mathbf{q}_k$ and $\mathbf{p}_k$ represent generalized coordinates and momenta. For a system with pairwise interaction potential, the Hamiltonian is

$$H = \frac{1}{2}\sum_{i=1}^{N} m_i \mathbf{v_i}.\mathbf{v_i} + \frac{1}{2}\sum_{i=1}^{N}\sum_{j\neq i}^{N} u(r_{ij}) \tag{1.13}$$

and Eqs. 1.11 and 1.12 yield

$$\frac{dr_i}{dt} = \frac{\mathbf{p_i}}{m_i} \tag{1.14}$$

$$-\dot{\mathbf{p}}_i = -\nabla \mathbf{u} = \mathbf{F_i} \tag{1.15}$$

where i=1,2,....N. There are now $6N$ first order differential equations to be solved.

The equation of motion is solved numerically to yield particle velocities and positions as a function of time. It is usually integrated by using finite difference approach. The Verlet algorithm is one of the most commonly used algorithm for this purpose [75]. The advantage of the use of Verlet algorithm is that its implementation is straightforward and storage requirements are modest. Although, it has the disadvantage of moderate precision during the calculation and velocity does not appear explicitly in the Verlet integration. As an improvement to the Verlet algorithm, the leap-frog algorithm [76] has been developed. But, it has a disadvantage that the positions and velocities are not synchronized. As an alternative of Verlet or the leapfrog algorithm, Velocity Verlet algorithm has been developed and the following relations are used to calculate new position and velocity at the same time:

$$r(t+dt) = r(t) + v(t)dt + \frac{1}{2}a(t)dt^2 \tag{1.16}$$

$$v(t+dt) = v(t) + \frac{1}{2}[a(t) + a(t+dt)]dt \tag{1.17}$$

To calculate the velocities at time t+dt, this method requires acceleration at time t and t+dt. In the present work, we have employed Velocity Verlet algorithm.

REMD algorithm of Sugita and Okamoto has become a widely-used tool for simulation of macromolecules that arises by applying the parallel tempering method to MD simulation. The REMD algorithm runs multiple isothermal MD simulations in parallel at a sequence of increasing temperatures $(T_0, T_1, ..., T_n)$ and intermittently attempts to swap simulations between temperatures i and j. The proposed swap accepted with probability is given by the Metropolis criterion below:

$$P_{acc} = min(1, exp\left[\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j}\right)(E_i - E_j)\right]) \tag{1.18}$$

where E_i and E_j and T_i and T_j are the energies and temperatures of the i-th and j-th replica, respectively. If the exchange is possible, the temperatures of the neighboring replicas will be exchanged, and the velocity will be scaled and reassigned according to the new temperature; otherwise, the two replicas will continue on their trajectories with the same temperature. In this way, replicas of the system which were trapped in local minima after the exchange have gained the kinetic energy to cross higher energy barriers.

Development of Metropolis criterion involves following steps: In standard replica exchange molecular dynamics, the simulated system consists of of M non interacting copies (or replicas) at M different temperatures. The position, momentum and temperature for each replica are denoted by $q^{[i]}, p^{[i]}$ and T_m, and here, i=1,....,M and m=1,...,M. The equilibrium probability for the generalized ensemble is given by:

$$W(p^{[i]}, q^{[i]}, T_m) = exp\left\{ -\sum_{i=1}^{M} \frac{1}{k_B T_m} H(p^{[i]}, q^{[i]}) \right\} \tag{1.19}$$

In the expression, the Hamiltonian $H(p^{[i]}, q^{[i]})$ is the sum of kinetic energy $K(p^{[i]})$ and potential energy $E(q^{[i]})$. At temperature T_m, we denote $p^{[i]}, q^{[i]}$ by $x_m^{[i]}$ and further we define $X = x_1^{[i(1)],...,x_M^{[i(M)]}}$ as one state of the generalized ensemble. Now we consider a pair of replicas that are exchanged. Suppose we exchange replicas i and j, which are at temperatures T_m and T_n, respectively. The exchange can be represented as:

$$X = ...; x_m^{[i]}; ...; x_n^{[j]}; ... \rightarrow X' = ...; x_m^{[j]}; ...; x_n^{[i]}; ... \tag{1.20}$$

To maintain detailed balance of the generalized system, microscopic reversibility has to be satisfied, thus giving

$$W(X)\rho(X \to X') = W(X')\rho(X' \to X) \tag{1.21}$$

where $\rho(X \to X')$ is the exchange probability between two states X and X'. An important step in the derivation of the exchange criterion is the substitution of the Boltzmann factor for the weight of each conformation into Eq. 1.21, yielding Eq. 1.22 (given below). Note that, it is not strictly correct until equilibrium has been reached, (this is the point at which the structures are actually considered for exchange with this probability).

$$\begin{aligned} exp\left\{ -\frac{1}{k_B T_m} H(p^{[i]}, q^{[i]}) - \frac{1}{k_B T_m} H(p^{[j]}, q^{[j]}) \right\} . \rho(X \to X') = \\ exp\left\{ -\frac{1}{k_B T_m} H(p^{[j]}, q^{[j]}) - \frac{1}{k_B T_m} H(p^{[i]}, q^{[i]}) \right\} . \rho(X' \to X) \end{aligned} \tag{1.22}$$

By rearranging Eq. 1.22, we obtain the Metropolis exchange probability (see Eq. 1.18). It must be reiterated that Eq. 1.22 is valid only for equilibrated ensembles that follow Boltzmann distribution. This assumption is followed till the end of the simulation. Use of this exchange probability results into adoption of the correct ensemble by each replica.

In the standard REMD, several replicas at different temperatures are simulated simultaneously and independently for a chosen number of MD steps. Then exchange is attempted between a pair of replicas (the probability of success is calculated using Eq. 1.18). If the exchange is accepted, the temperatures of the replicas will be swapped, and the velocities will be scaled accordingly. Otherwise, if the exchange is rejected, both of the replicas will continue with their current trajectories with the same thermostat temperature.

■ PRESENT WORK

TMAO, a widely studied osmolyte, counteracts the denaturing effect of urea, typically in 8:4 molar ratio of urea:TMAO, and thereby stabilizes the native conformation of the protein to maintain cellular function. The mechanism by which TMAO provides counteraction against the action of urea and stabilizes protein has been investigated. We focus on the effects of urea and TMAO on the folding/unfolding equilibrium of protein and put an effort to uncover the molecular basis of denaturation and counteraction mechanisms. In **Chapter 1**, we present a detailed literature review of the mechanism of action of the osmolytes, their effects on the protein. The reason of choosing REMD (Replica Exchange Molecular

Dynamics) simulation technique over the classical MD technique is also explained. **Chapters 2** deals with the Smith model of urea and two different models of TMAO viz. Kast model and Osmotic model. Osmotic model was developed based on Kast model that yield a higher osmotic pressure of TMAO-water solution. It enhances TMAO-water interactions and reduces TMAO-TMAO interactions when compared to that of the Kast model. We have performed a comparative study utilizing these two models of TMAO to examine the extent of protein protecting ability of the two models. We have tried to explore the underlying mechanism by which urea and TMAO exert their effects on the protein β-hairpin. A close examination of the occupancy of different intra-protein hydrogen bonds shows that TMAO molecules protect the terminal hydrogen bonds of the protein effectively. Urea causes denaturation of the protein through breaking of the terminal intra-protein hydrogen bonds. Kast model of TMAO, in conventional 2:1 ratio of urea:TMAO provides counteraction against denaturing effect of urea. Osmotic model provides maximum counteraction when used in 1:2 ratio, but fails to provide sufficient conformational stability to the protein in conventional 2:1 ratio of urea and TMAO. These findings lead us to carry out further studies with Kast model of TMAO. Errors in the results obtained from molecular dynamics simulation arise from insufficient sampling. Assessing the quality of the sampling shows how accurately a given quantity is computed. Without adequate sampling, the predictions of the results remain incompletete. Conclusions can not be drawn from an undersampled calculation. We will address the statistical uncertainty for individual observables wherever needed in the following chapters. Statistical analyses are of utmost importance in establishing the reliability of simulation data for any given study.

In **Chapter 3**, we have emphasized on the counteraction behavior of TMAO against action of urea without using any confinement. Using REMD simulation technique we have studied urea-induced denaturation and TMAO-induced counteraction of the protein Trp cage. We have particularly emphasized on three factors that persuade stability on the protein Trp cage: (i) salt bridge between residues Asp9 and Arg16, (ii) the buried hydrophobic core surrounding the indole plane of residue Trp6 and (iii) orientational preference of different aromatic planes constituting the protein with respect to the indole plane of residue Trp6. Existence of the salt bridge and hydrophobic core of the protein are investigated thoroughly. Angles between vectors normal to the planes are considered for the comparison of orientational preference among different pair of planes. We have analyzed RMSF and secondary structure of the protein that reflect the behavior of the protein under the effect of urea and mixture of urea-TMAO solution. Analyses of melting curves and free

energy landscapes indicate the difference of folding behaviors of the protein in presence of different chemical environments. We have calculated average number of hydrogen bonds among different components. Large number of protein-urea hydrogen bonds indicate direct interaction of urea with the protein that result in unfolding of the protein. Small number of protein-TMAO hydrogen bonds provide insufficient support for direct interaction mechanism of TMAO with the protein. Since we have not observed any evidence of direct interaction, the study provokes indirect mechanism of protein stabilization by TMAO molecules.

In **Chapter 4**, we have examined the effects of non polar confinement on the folding/unfolding equilibrium of the protein Trp cage by means of REMD simulation technique. We have used a near spherical non polar fullerene ball consisting of 2940 carbon atoms in order to mimic the role of chaperonin unit in protein folding. Free energy landscapes (FELs) are calculated as a function of temperature and RMSD (root mean square deviation) that depict conformational space visited by the protein. Analyses of radius of gyration and RMSF (root mean square fluctuation) indicate that the protein experiences decreased conformational space due to limited volume provided by the confinement. Protein-urea interaction energy is also analyzed by splitting it into electrostatic and van der Waals components. Analysis of density profiles of water for pure water system reveals that there is dewetting of the hydrophobic surface of the fullerene ball. In binary urea solution all the residues of the protein show high probability to remain near the surface of the fullerene ball. Under the effect of confinement urea molecules get more access to solvate the protein that cause denaturation. Average number of intra- and inter- molecular hydrogen bonds are also calculated that shed light on denaturation and counteraction mechanisms of urea and TMAO respectively.

Similar to the work in **Chapter 4**, in **Chapter 5** we have studied the effects of polar confinement on the folding/unfolding equilibrium of the protein Trp cage using replica exchange molecular dynamics simulation. We have used a spherical polar confinement consisting of 2940 carbon atoms with randomly distributed charges on the different carbon atoms of the fullerene like ball. Like the preceding chapters here also we have explored FELs as a function of temperature and RMSD. Comparison of FELs and other properties like radius of gyration, RMSF, DSSP indicate that in polar confinement the folding propensity of the protein becomes lower than that in non polar confinement. Urea induced denaturation is found to be more effective in presence of polar confinement. Equivalent folding propensity of the protein in pure water and mixed ternary solution indicates that

TMAO counteracts urea's action effectively in presence of polar confinement. Analysis of density profiles of water ensures an enriched layer of water that wets the polar surface of the confinement. The protein residues in all the solutions show small tendency to occupy peripheral region of the confinement. The roles of salt bridge and the hydrophobic core of the protein pertaining stabilization are also analyzed. Calculation of intra- and intermolecular hydrogen bonds are also employed to analyze the mechanisms of urea and TMAO inside the polar confinement.

In **Chapter 6**, we have discussed the results of each of the work and tried to bring our view on the counteraction mechanism of TMAO and the effect of confinement on the protein folding/unfolding equilibrium. A comparative view of the effect of polar/non polar confinement on protein conformation is also elucidated.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Chapter 2

Model Dependency of TMAO's Counteracting Effect Against Action of Urea

"Proteins can do Almost Everything....Before we can hope to understand how genes work, how muscles contract, how nerves conduct electricity, how embryos develop, or how our bodies function, we must attain a deep understanding of proteins..."

– B. Alberts et al., Molecular Biology of the Cell, Fifth Edition, 2008, Garland Science, New York, USA, p. 125

Folding and unfolding of proteins are of vital importance for the regulation of their biological activities. Native state of a protein refers to the state of the highest thermodynamic stability under physiological conditions. The compactness of the three dimensional structure of a protein depends intrinsically on the sequence of amino acid residues as well as on the multiple contributions from the cellular environment. Protein is sensitive to the surrounding environmental conditions such as temperature, pressure and assorted molecules comprising the volume around it [22]. Urea is an extensively studied denaturing osmolyte, which has been postulated to exert its effect either directly [27] (by binding to the protein) or indirectly [28] (by altering the solvent structure which in turn offers denaturing effect on the protein). Most of the recent studies have established direct interaction of urea with protein. The counteracting osmolyte TMAO is found in the urea-rich cells to protect the protein against deleterious effect of urea. Studies based on TMAO-water interaction showed that the unfavorable interaction of TMAO with protein backbone is dominantly responsible for its stabilization effect [34]. Though lot of research works have been devoted still study on this crucial osmolyte is unraveling new facets everyday.

In literature many experimental evidences exist showing urea induced denaturation, TMAO induced protection and also counteraction of TMAO against urea action [42, 48]. The findings are substrate specific based on intermolecular interaction mechanism. These studies inspired us to investigate and characterize the protection and denaturation mechanisms of osmolytes urea and TMAO at molecular level using simulation technique. To bring out the detailed mechanism of protein-solvent interactions and their roles in overall conformational behavior of the protein are the core objective of this study.

Comparison of counteracting effect of Kast model versus Osmotic model

Overview: We have performed molecular dynamics simulation of GB1 peptide (a 16-residue β-hairpin) in different osmotic environment. Urea is used for denaturation of the peptide and TMAO is used as the counteracting osmolyte to offset the denaturing effect of urea. Direct mechanism of urea-induced denaturation is evident from interaction energy calculations. Electrostatic interactions between urea and protein plays a major role in denaturation process. To emphasize on protein protecting action of TMAO against urea, two different models of TMAO are used viz. Kast model and Osmotic model. The study reveals that Osmotic model of TMAO gives the best protection to counteract urea's action when it is used in ratio 1:2 of urea:TMAO. The cause of better protection with Osmotic model (in 1:2 ratio) arises due to the fact that presence of TMAO makes urea-protein electrostatic energy more unfavorable. Preferential solvation of TMAO molecules by urea (and water) molecules is also observed. This preferential solvation of TMAO molecules causes a depletion in the number of urea molecules in the vicinity of the protein. Intra-protein hydrogen bonds between different residues of protein further reveals the breaking of backbone hydrogen bonds of residues 2 and 15 in presence of urea and the same hydrogen bond is preserved in presence of TMAO. Free energy landscapes show that the narrowest distribution is obtained for osmotic TMAO model when used in 1:2 ratio of urea and TMAO.

■ INTRODUCTION

In some marine organisms urea is accumulated in their cells as a result of osmotic adaptation [31, 77, 78]. Urea being a denaturant, accumulation of another stabilizing osmolyte is commenced to offset urea's action. Trimethylamine-N-oxide (TMAO) is one such kind of protecting osmolyte found in urea-rich cells of marine animals. It tends to stabilize the folded state of protein and thus provide favorable conditions for cellular functions.

Urea is required in high concentration to cause protein denaturation. Two mechanisms are often proposed to explain urea-induced denaturation, viz. direct and indirect mechanism [22, 27-30, 77-79]. Direct mechanism assumes that preferential interaction between protein and urea play a significant role, whereas indirect mechanism assumes that urea alters the water structure that causes an enhancement of hydration of the protein and perturbs the protein conformation (discussed in Chapter 1). Some studies support working of both of the mechanisms [30]. Most of the recent studies emphasize on the protein-urea direct interaction which makes dominant contribution in urea-induced protein denaturation. Though a great deal of effort has been made in order to understand the underlying mechanism by which TMAO counteracts the deleterious effect of urea on protein molecules but a definitive answer towards this is yet to be obtained [4, 33, 36, 40, 42, 49, 80-82]. Different hypotheses have been proposed for the counteracting ability of TMAO on urea denaturation: (i) preferential exclusion of TMAO molecules from the protein surface [33, 80], (ii) TMAO's ability to act as a molecular crowder which states that TMAO molecules help in strengthening protein backbone-backbone hydrogen bonds by creating a local protein environment [40, 42], (iii) Preferential solvation of TMAO molecules by urea and water [49] and (iv) TMAO-induced strengthening of water-water and water-urea interactions [4]. In this context it is worth to mention that in a recent study J. E. Shea and co-workers [82] showed that peptide hydration is greatly affected in presence of osmolytes urea and TMAO. For urea, once it is bound to the protein, it repels water molecules and favors urea-urea and urea-protein interactions. Whereas, TMAO molecule does not favor interacting with the like molecules rather, it helps to redistribute the water molecules (from hydrophilic residues of the protein to hydrophobic residues) which causes the formation of stable compact conformation of the protein. Thus, it can be said that the mechanism of protection and deprotection (of protein) is still a matter of dispute, and a generic mechanism is still in demand.

Small β-hairpin peptides have been used as model peptides for studying folding/unfolding mechanism of proteins [8, 9]. β-hairpin is the simplest example of an anti-parallel β-sheet found in protein structures that may be considered as a suitable system to study the fundamental issues in protein folding. They may also provide a clue into the problem of conformational diseases such as prion diseases, Alzheimer's diseases, etc. Recent studies are reported that have examined osmolyte effects on the stability of β-hairpins [10-14]. Several computational studies are reported to explore the folding mechanism of β-hairpin and the effects of osmolytes on its stability [83].

For possible quantitative characterization of different contributing factors of TMAO on protein stabilization different model systems are studied. In most of the studies, Kast model of TMAO [36, 49, 50] is used as osmoprotectant. Recently, Garcia and co-workers [33] have introduced another model of TMAO called Osmotic model based on osmotic pressure measurement of TMAO solutions. Osmotic model incorporates more realistic interactions through parameterization of the strength of Lennard-Jones interactions and scaling down the charge of TMAO molecule.

In this work we have reported the MD simulation of GB1 hairpin in water and compared its behavior in urea solutions as well as in mixture of urea and TMAO with two different models of TMAO. Model dependent study of the counteracting effect of this protecting osmolyte has not been explored till date. Several computational studies have addressed the folding mechanism of GB1 hairpin because of its strong ability to retain its secondary structure element in solution [84]. Throughout this study we have tried to shed light on the protection of TMAO against urea denaturation of β-hairpin and use of different model of TMAO to find out the best fitted mechanism of TMAO counteraction. In the next section, we have described the models and simulation method. Thereafter we have discussed the results. Finally, conclusions are summarized in the last section.

■ MODELS AND SIMULATION METHOD

In this study we have carried out classical molecular dynamics (MD) simulations of counteraction of TMAO on urea conferred denaturation of a 16-residue β-hairpin, GEWTY-DDATKTFTVTE, corresponding to the C-terminal of the immunoglobulin binding domain of streptococcal protein G (PDB code: 1GB1). Three sodium ions were added to neutralize the charged side chains of the protein. For protein AMBERff99 [85] force field parameterization was adopted. SPC/E [86] model of water was used for simulation of each of the system as described in Table 2-1. Smith model of urea [87], Kast [88] and Osmotic model

[33] of TMAO were used. Each of the simulations was performed for 50 ns using AMBER12 [70] simulation package.

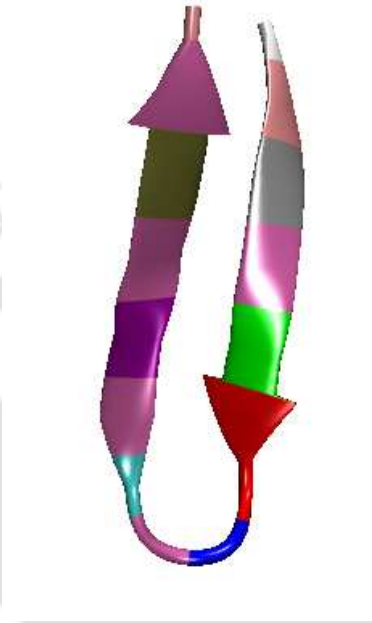


Figure 2-1. *NMR structure of the protein β-hairpin with residue-wise color identity.*

With Packmol [89] package the initial configurations of all the systems were constructed, and then they were subjected to 10000 steps of minimization (steepest descent followed by conjugate gradient method). The resultant systems were then heated gradually from 0 K to 300 K followed by an equilibration run. Then the simulations were carried out up to 50 ns in NPT ensembles (at 300K temperature and 1 atmospheric pressure). A 10.0 Å cutoff was used for all non-bonded interactions. SHAKE [90] algorithm was implemented in order to put constrain in all bonds involving hydrogen. Temperature was controlled by Langevin dynamics [91] with a collision frequency of 1 ps^{-1} and pressure was controlled by using Berendsen barostat [92] with pressure coupling constant 2.0 ps. Electrostatic interactions were evaluated fully by particle-mesh Ewald (PME) method and periodic boundary conditions [93] were applied in all three directions. To carry out post-simulation analyses AMBER ptraj module as well as VMD [94] (Visual Molecular Dynamics package) were used on the trajectories obtained from constant temperature molecular dynamics simulation.

Table 2-1. Overview of Systems[a]

System	N_{water}	N_{urea}	N_{TMAO}	$Density(\rho)$ (g/cc)	$BoxLength$ (Å)
Protein-Water (PW)	3000	0	0	1.0102	45.190
Protein-Water-Urea (PWU, 5M)	2293	283	0	1.0915	45.080
Protein-Water-Urea-TMAO (PWUT, 2:1 i.e. 5M:2.5M, Kast)	1727	283	141	1.1178	44.779
Protein-Water-Urea-TMAO (PWUT, 2:1 i.e. 5M:2.5M, Osmotic)	1727	283	141	1.1326	44.619
Protein-Water-Urea-TMAO (PWUT, 1:2 i.e. 2.5M:5M, Osmotic)	1627	141	283	1.5192	43.839

[a] N_{water}, N_{urea} and N_{TMAO} represent the number of water, urea and TMAO molecules respectively. ρ represents density of the system.

■ RESULTS AND DISCUSSION

Root Mean Square Deviation (RMSD)

The calculations of the atom positional root-mean-square deviation (RMSD) of the C_α-atoms of the protein provide information about the conformational changes of the protein in different chemical environments. For this, following earlier work [95], we calculated windowed RMSD for different systems as a function of time. It is evident from the literature [96] that RMSD value up to 1.5 Å can be considered as a criterion for the folded state of the β-hairpin. From the plot of RMSD vs. time for all the systems studied (Figure 2-2), it is observed that binary 5 M urea solution perturbs the protein conformation as RMSD value exceeds 3.0 Å.

For the systems containing TMAO along with urea, the counteraction behavior of TMAO is apparent from the much lower RMSD values distributed around 1.5 Å. A comparison of the three RMSD values among three different mixed osmolyte systems shows that the most stable conformation of the protein is achieved when Osmotic model of TMAO is used in the osmolyte mixture with 1:2 ratio of urea:TMAO. Kast model of TMAO is, however, able to counteract the protein against urea denaturation in 2:1 ratio. Noticeable fluctuations in RMSD values for the Osmotic model containing system is observed when

the mixture is used in 2:1 ratio. Thus Kast model of TMAO in 2:1 ratio and Osmotic model in the reverse ratio i.e., 1:2 provide the most effective counteraction of TMAO against the action of urea.

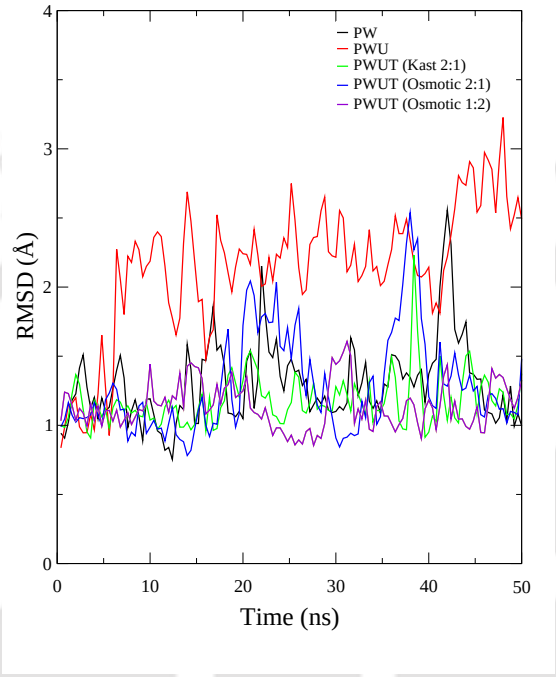


Figure 2-2. *Root mean square deviation (RMSD) of C_α-atoms of the protein as a function of simulation time. PW (black), PWU (red), PWUT (Kast 2:1) (green), PWUT (Osmotic 2:1) (blue) and PWUT (Osmotic 1:2) (magenta).*

Solvent Accessible Surface Area (SASA)

The solvent accessible surface area (SASA) measures the surface area of the solute protein over which contact between the solute and solvent molecules takes place. Using VMD [97] we have calculated SASA by considering a probe radius of 1.4 Å . In Table 2-2 we include the SASA values for all the systems. In the same table we also present the standard errors of SASA values for all the systems. The standard errors were estimated by dividing total simulation runs in to five blocks, each of 10 ns. Then, the SASA values were obtained for each block, and standard errors were calculated assuming that the block values provide independent estimates of the SASA values. Note that, the similar procedure is applied for the estimation of standard standard errors of other properties mentioned

in this article. Nevertheless, these values indicate that the maximum value of SASA is obtained for binary urea system while its smallest value is obtained for the system without any osmolyte. This observation suggests that urea delivers denaturation to the protein such that the unwinding of the two anti-parallel sheets takes place resulting into increase of surface area of the protein to be accessed by solvent molecules, whereas the smallest value of SASA indicates that in absence of any osmolyte i.e. in aqueous environment the protein continues to maintain its native conformation. For the mixed osmolyte systems, the value of SASA is found to be almost identical irrespective of the type of TMAO model used. Nevertheless, these values are less than that of binary urea system indicating more compact or folded form of the protein than binary urea system.

Table 2-2. SASA for respective systems[a]

System	$SASA$(Å^2)
PW	1715.08 (±16.40)
PWU	1844.69 (±8.44)
PWUT (Kast 2:1)	1782.04 (±3.84)
PWUT (Osmotic 2:1)	1796.79 (±5.88)
PWUT (Osmotic 1:2)	1786.34 (±10.34)

[a] Solvent accessible surface area (SASA) for respective systems. The numbers inside parentheses are the standard errors that were estimated by dividing the total simulation runs in to five blocks.

Free Energy Landscapes (FELs)

Two suitable reaction coordinates that describe the folding dynamics of the protein (viz. RMSD and intra-protein hydrogen bond numbers in this study) are chosen to construct free energy landscape of the β-hairpin in different osmolyte solutions as well as in pure water. The free energy landscapes (FEL) are constructed according to the following equation:

$$\Delta G(V) = -k_B T[lnP(V) - lnP_{max}] \tag{2.1}$$

where P(V) is the probability distribution of the coordinate (V) which is calculated from the entire trajectory and is subtracted from the maximum of the distribution P_{max} so that

ΔG=0 for the free energy minimum. In this study, ΔG vs C_α-RMSD is plotted for the respective intra-protein hydrogen bond numbers (see Figure 2-4).

The energy landscapes indicate different intermediates throughout the trajectory. In order to verify convergence of the simulation run we have constructed free energy landscapes for different time frame (see Figure 2-5) and compared with the energy landscape of the entire run. First 5 ns from each of the simulations is discarded and rest 45 ns is split into three windows (i.e. 15 ns each) and FEL is calculated for each 15 ns window. Convergence of sampling is indicated from the resemblance of FEL for last 15 ns with the overall FEL. Representative conformations of the protein at the end of simulation is shown in Figure 2-3. Figure 2-4 (a) shows that in pure water the conformation of the hairpin remains almost intact throughout the entire simulation path. In presence of 5 M urea (Figure 2-4 (b)) the landscape shows two distinctive minima. The minimum below 1.5 Å of RMSD indicates folded form of the protein and above 1.5 Å indicates unfolded conformations. This figure further reveals that urea acts on the protein and decreases the number of intra-protein hydrogen bonds so that the native conformation is not retained, in fact the protein attains denatured conformation with much higher RMSD values.

In mixed osmolyte systems, (Figure 2-4 (c)-(e)) the maximum stable behavior of the hairpin is observed in Osmotic model of TMAO system when reverse ratio of osmolytes is maintained. Mixed osmolyte system with Kast model of TMAO also shows significant protection of the protein. But Osmotic model in regular 2:1 ratio does not protect the protein properly which is indicated by much rugged energy surface with multiple number of minima. The narrowest distribution of FEL in reverse Osmotic system further suggests that the maximum efficiency of the Osmotic model to protect the protein when it is used in 1:2 ratio. The cause lying behind better protection of Osmotic model can be interpreted as due to the removal of more number of urea molecules from the protein surface as well as the preferential solvation of TMAO molecules by urea (and water) molecules (addressed below).

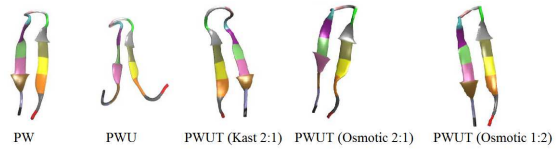


Figure 2-3. *Snapshots of the protein in different systems at the end of simulation.*

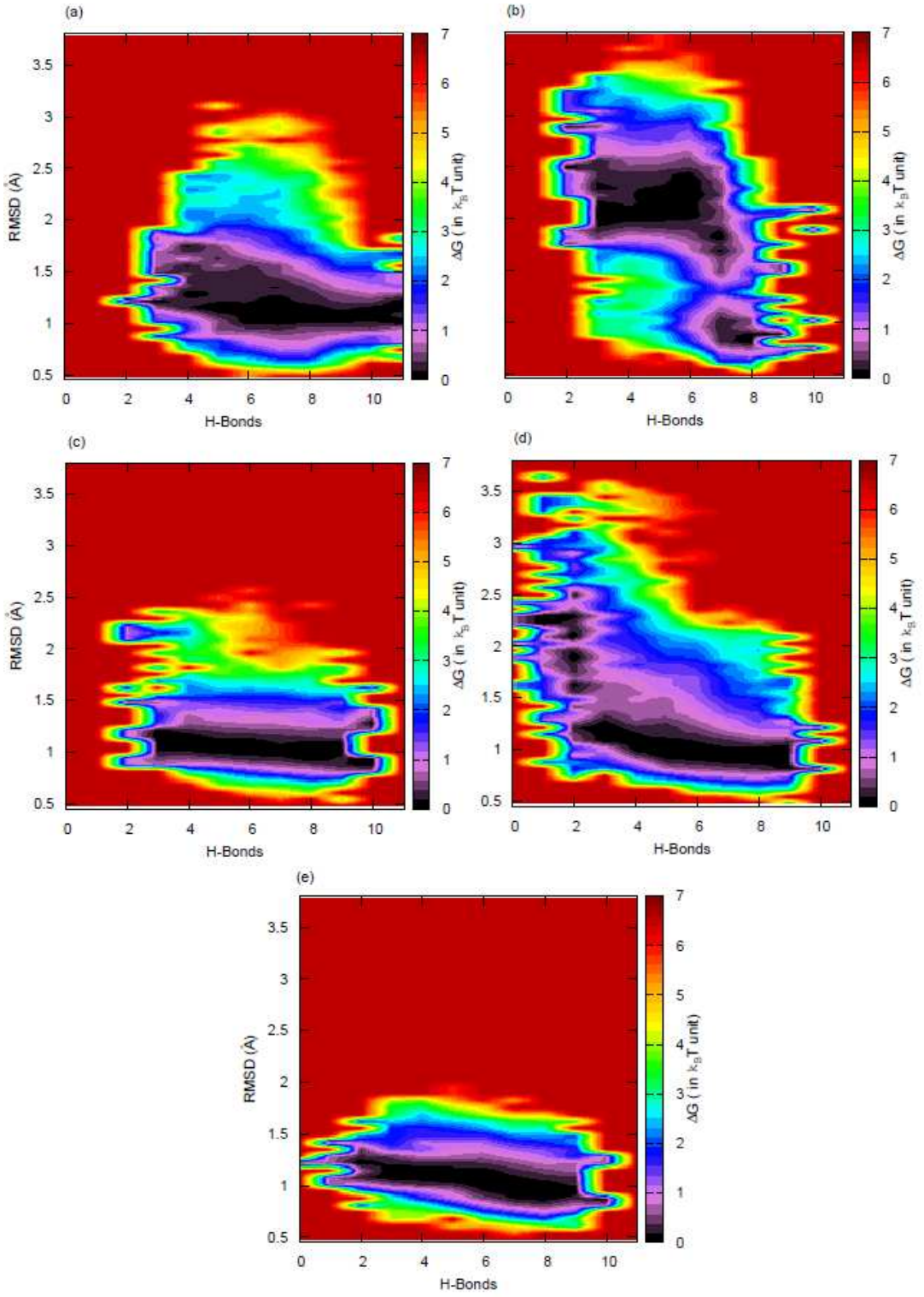


Figure 2-4. *Free energy landscapes (FEL) for systems (a) PW, (b) PWU, (c) PWUT (Kast 2:1), (d) PWUT (Osmotic 2:1) and (e) PWUT (Osmotic 1:2).*

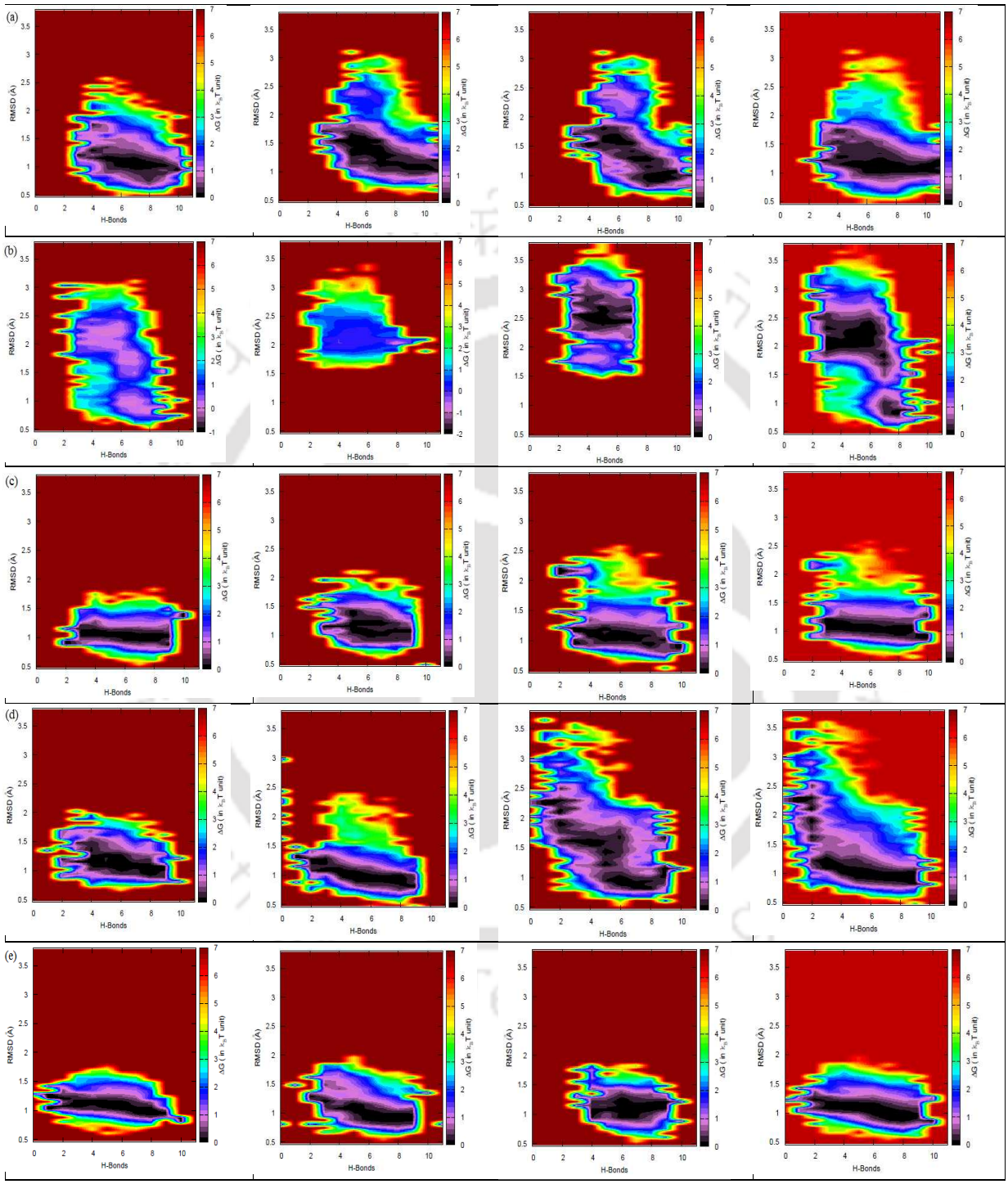


Figure 2-5. *The windowed FELs (15 ns each) are plotted from left to right for (a) PW, (b) PWU, (c) PWUT (Kast 2:1), (d) PWUT (Osmotic 2:1) and (e) PWUT (Osmotic 1:2). Last FEL is for the whole simulation run.*

Dihedral Angle Principal Component Analysis (dPCA)

For principal component analysis (PCA) of a system containing M atoms the correlated internal motions are represented by the covariance matrix:

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \tag{2.2}$$

where $x_i, ...x_{3M}$ are the mass-weighted Cartesian coordinates of the atoms and $\langle ... \rangle$ denotes the average over all sampled conformations. The eigenvectors and eigenvalues of the covariance matrix contain information about the collective motions of the atoms and their amplitudes.

To obtain dPCA [98] we have performed PCA on sin- and cos-transformed dihedral angles. Five different free energy surfaces are constructed for five different systems by using 'Carma: a molecular dynamics analysis program' [99]. Figure 2-6 shows the free energy change as a function of first two dihedral angle principal components. dPCA can clearly resolve the minima of the free energy surface. The different landscapes efficiently show the protein conformational changes under different environments.

In Figure 2-6 (a), the two minima belong to two folded conformations with one major dominant conformation which are separated from each other by an energy barrier. In binary urea system, the occurrence of folded conformations is found to be small. Note that the low energy barriers depict easy transitions between closely related conformations. But in binary urea system all the minima are separated. Highly populated folded conformations are observed for Kast and Osmotic model in reverse ratio (as shown in Figure 2-6 (c) and (e)). On the other hand, Osmotic model in 2:1 urea and TMAO ratio shows a very narrow distribution (Figure 2-6 (d)) indicating low occurrence of folded state.

Thus, it becomes clearly visible that in pure water the protein exhibits one state folding with a small occurrence of another folding state which are closely spaced energetically. In 5 M binary urea solutions, urea molecules establish direct interaction (as is evident from energy decomposition analysis addressed below) with the protein delivering denaturation. It results into small occurrence of folded conformations. Because of huge changes in internal structure of the protein, the minima are separated by large energy barriers. High occurrence of folded conformations are found when Osmotic model of TMAO is used in 1:2 ratio. But when the Osmotic model is used in 2:1 ratio, folded conformation of the protein is found with a relatively smaller occurrence due to insufficient counteraction provided by the TMAO molecules. When Kast model of TMAO is used much better counteraction against action of urea is achieved in comparison to Osmotic model of TMAO used in reverse

ratio.

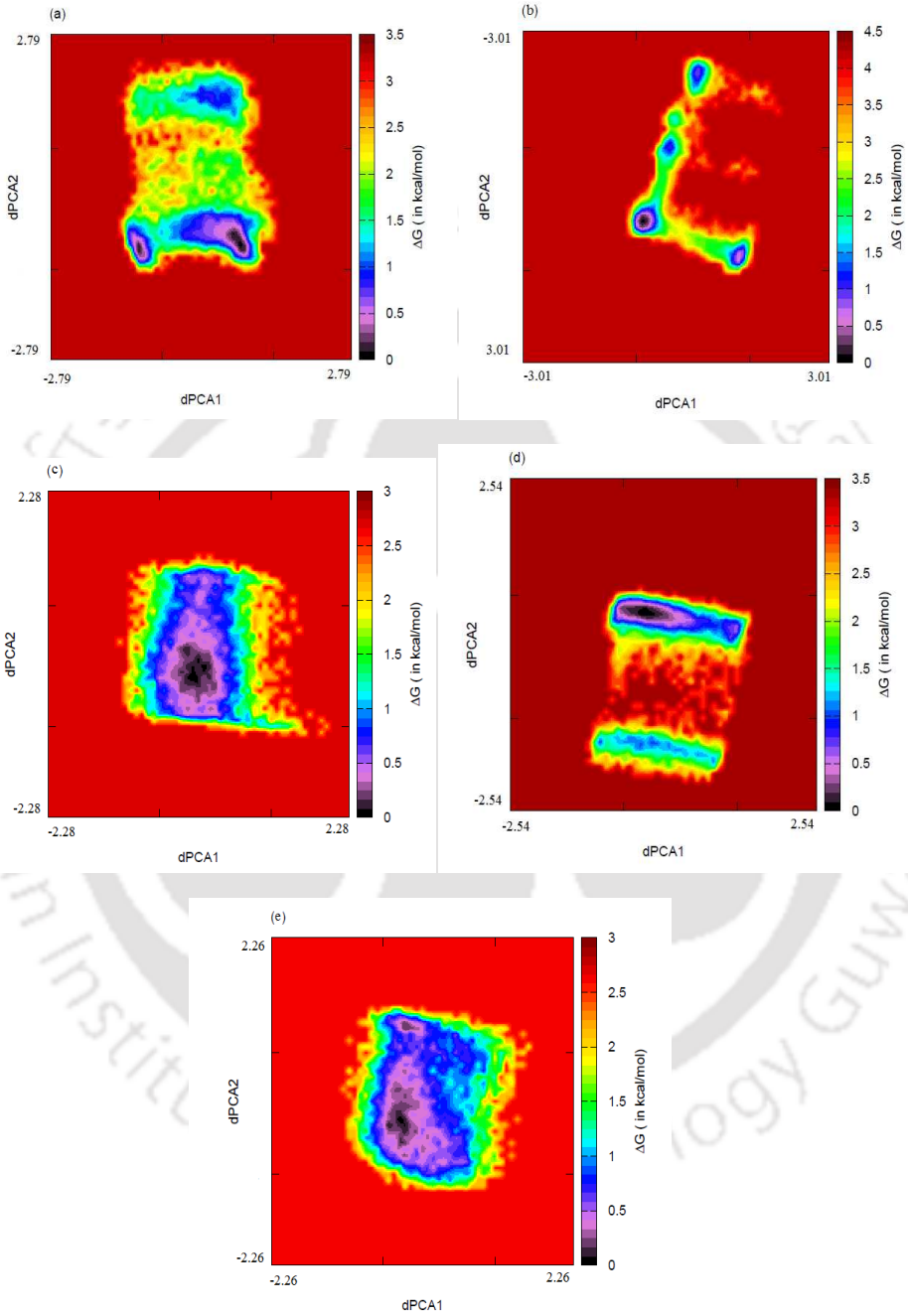


Figure 2-6. *Dihedral angle principal component analysis (DPCA) for systems (a) PW, (b) PWU, (c) PWUT (Kast 2:1), (d) PWUT (Osmotic 2:1) and (e) PWUT (Osmotic 1:2).*

Effect of TMAO on Urea-Protein Interactions

In order to examine urea's direct interaction pattern with the protein we have decomposed the urea-protein interaction energy into electrostatic and van der Waals (vdW) energy components and the same is shown in Figure 2-7. Note that in protein-urea interaction energy calculations we consider the interaction between the protein and all urea molecules present in a given system. The electrostatic energy refers to the direct interaction of urea with the protein through hydrogen bonds or other electrostatic interactions and van der Waals energy refers to the favorable interaction of the protein and urea through dispersion interactions [100]. Literature reveals that [101, 79] dominant interaction amongst the two is system dependent. Depending upon the other osmolyte present and nature of the protein, the degree of interaction energy dependency changes and it is not necessarily mutually exclusive. Figure 2-7 indicates that electrostatic energy plays a dominant role in protein-urea interaction. Both electrostatic and van der Waals energies between protein and urea (and hence total interaction energies) become more unfavorable when TMAO is introduced. But the effect of TMAO on the electrostatic energy is noticeable more than the van der Waals energy. This can be explained on the basis of the fact that urea molecules are being replaced from the solvation shell of the protein when TMAO is introduced. This prompted us to calculate radial distribution function (rdf) involving urea carbon and peptide heavy atoms and the same is shown in Figure 2-8 (a). In Figure 2-8 (b), we show rdf involving nitrogen atom of TMAO and peptide heavy atoms. A decreased peak height of the first peak of urea carbon-peptide heavy atoms rdf as well as the presence of a small peak at around 4.4 Å in TMAO nitrogen-peptide heavy atoms rdf for mixed osmolyte systems indicate that some of the urea molecules may be replaced (by TMAO molecules) from the protein surface. The calculations of average number of heavy atoms of urea and TMAO around the protein (as shown in Table 2-3) confirms this fact (i.e., a depletion in the number of heavy atoms of urea molecules around the peptide heavy atoms for mixed osmolyte system). Behavior of Kast and Osmotic model of TMAO are found to be similar when the ratio of urea:TMAO is the same. On the other hand for the reverse ratio, TMAO being more in number (a much lower urea number density), a much lowering in the number of urea molecules is observed in the solvation shell of protein and hence protein-urea interaction energy becomes more unfavorable.

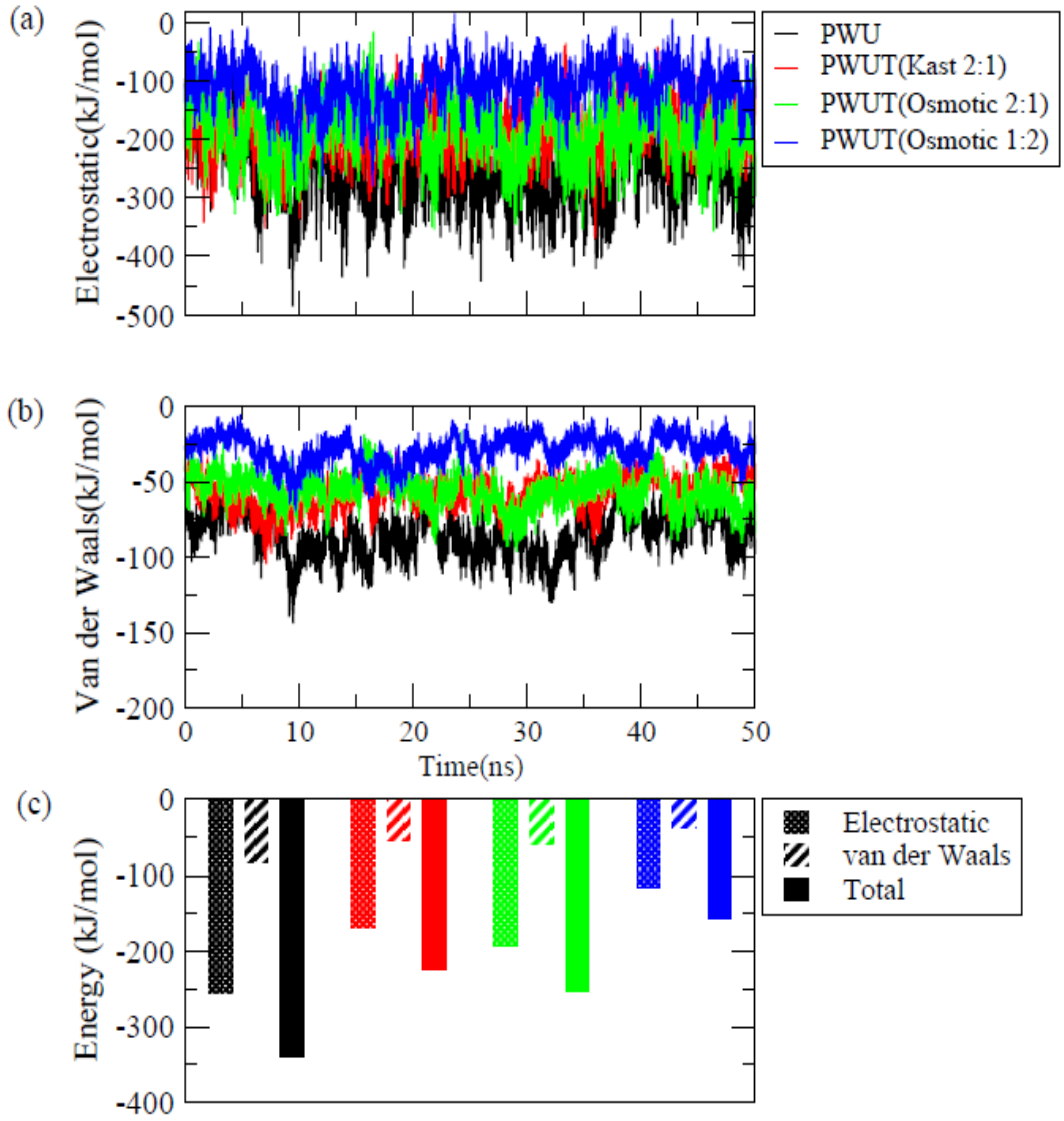


Figure 2-7. *Intermolecular (a) electrostatic, and (b) van der Waals interaction energies between protein and urea molecules as a function of simulation time. Part c represents electrostatic, van der Waals and total protein-urea interaction energies averaged over simulation time.*

In this context it is worth to mention here that Kokubo and coworkers [44] reported that urea forms better vdW contacts with the protein than TMAO. Favorable vdW interaction of urea with denatured states of protein helps it to stabilize the extended denatured state than the compact helix state. For binary TMAO solution vdW contribution is more (for both native and denatured conformations) than water. This favorability of vdW component cannot overcome the unfavorability of electrostatic component which makes it

a higher free energy system with low protein solubility. This effect becomes highly significant for denatured protein with more exposed area. Thus, a binary TMAO solution with a compact protein is more stable than the one containing a denatured protein. As a result, TMAO solution stabilizes the native state of a protein. When mixture of urea and TMAO is considered, it provides the most favorable vdW component. As like the binary TMAO solutions, similar explanations are provided for ternary mixed osmolyte solutions. Counteraction of TMAO does not stem from electrostatic component (since electrostatic component is more unfavorable for ternary osmolyte solution than binary TMAO solution).

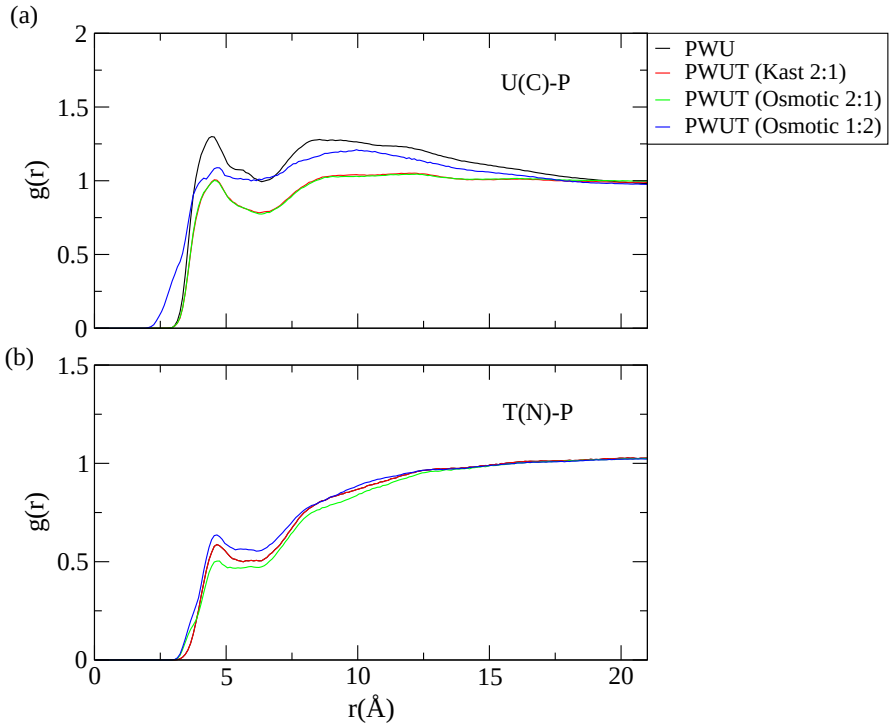


Figure 2-8. *Radial distribution functions involving (a) urea carbon and (b) TMAO nitrogen and the peptide heavy atoms for the systems PWU (black) , PWUT (Kast 2:1) (red), PWUT (Osmotic 2:1) (green) and PWUT (Osmotic 1:2) (blue) respectively.*

Protein Solvation and Hydrogen Bond Properties

According to previous study [102], when a heavy atom of solution species fell within a distance of 3.5 Å from any heavy atom of the protein, it was considered to be inside solvation shell of the protein. The calculations of average number of solvent molecules within 3.5 Å of the protein heavy atoms also reveal similar observations (see Table 2-3).

It is apparent that in mixed osmolyte systems urea molecules are more likely to occupy nearby positions of the protein than TMAO. But what is remarkable is that the presence of TMAO molecules decreases the number of urea molecules at the protein surface and this effect is more profound for Kast model and Osmotic model (in 1:2 ratio) of TMAO. This observation can act as a corroborative evidence of what we observed for protein-urea hydrogen bonds (discussed below).

Table 2-3. Average number of heavy atoms of water, urea and TMAO within first solvation shell[a]

System	$N_{W(O)}$	$N_{U(C)}$	$N_{U(N)}$	$N_{U(O)}$	$N_{T(C)}$	$N_{T(N)}$	$N_{T(O)}$
PW	50.33 (±1.04)	−−	−−	−−	−−	−−	−−
PWU	32.77 (±1.79)	2.44 (±0.16)	14.08 (±0.77)	6.13 (±0.48)	−−	−−	−−
PWUT (Kast 2:1)	28.78 (±1.34)	1.81 (±0.10)	12.67 (±0.51)	4.75 (±0.27)	4.06 (±0.13)	0.11 (±0.01)	0.79 (±0.05)
PWUT (Osmotic 2:1)	29.13 (±1.32)	1.94 (±0.15)	13.96 (±0.69)	5.24 (±0.34)	4.43 (±0.14)	0.39 (±0.01)	1.23 (±0.04)
PWUT (Osmotic 1:2)	23.93 (±0.75)	0.91 (±0.04)	6.64 (±0.37)	2.58 (±0.20)	10.44 (±0.26)	0.63 (±0.03)	2.09 (±0.17)

[a] $N_{W(O)}$, $N_{U(C)}$, $N_{U(N)}$, $N_{U(O)}$, $N_{T(C)}$, $N_{T(N)}$, $N_{T(O)}$ are the average number of heavy atoms of water, urea and TMAO within 3.5 Å of any heavy atom the protein residue. W, U, and T refer to water, urea, and TMAO molecules respectively. The numbers inside parentheses are the standard errors that were estimated by dividing the total simulation runs in to five blocks.

Following earlier work [103], a maximum distance of 3.5 Å and a minimum D-H-A angle of 120° are used as geometric criteria to estimate the number of different types of hydrogen bonds (where D and A represent donor and acceptor atoms respectively). Note that, we have used the same hydrogen bond criteria for defining intra- as well as intermolecular hydrogen bonds. The average number of hydrogen bonds formed between protein and different species as well as intra-protein hydrogen bonds are shown in Table 2-4. In case of pure water system, i.e., in absence of any osmolyte the number of intra-protein hydrogen bond is maximum. This is in consistent with the idea of spontaneous

folding behavior of proteins when they are surrounded by aqueous environment [103]. It can be seen that the minimum number of intra-protein hydrogen bond is formed in case of binary urea system. It indicates unfolding of the protein under the influence of urea. Urea molecules form considerable number of hydrogen bonds with the protein. This allows urea molecules to deliver their effect on the protein which in turn delivers denaturation. The total number of hydrogen bonds formed by the protein with the solution species (n_{HB}^{total} in Table 2-4) indicates that in presence of Kast model of TMAO, the protein is less likely to form hydrogen bond with urea molecules than in presence of Osmotic model of TMAO when used in 2:1 ratio. This table further reveals that the minimum number of protein-urea hydrogen bond is formed for Osmotic TMAO model (in reverse ratio). Furthermore, the addition of TMAO decreases urea-water hydrogen bonds. This is because of the fact of formation of considerable amount of TMAO-water hydrogen bonds. What is interesting is that in mixed osmolyte systems TMAO molecules are also solvated by urea molecules significantly (as revealed by urea-TMAO hydrogen bond numbers). This in turn makes urea molecules somewhat unavailable for the protein to deliver denaturing effect. We note that, Kuffel et al. [104] investigated the effects of osmolytes urea, TMAO and TMU (tetramethylurea) on the structure of surrounding water. They concluded that these osmolytes fit reasonably well into the structure of surrounding water and a slight enhancement in the stability of vicinal water-water hydrogen bonds is also observed. These findings agree well with the previous findings [49]. Here we also note that, by using a combination of thermodynamic experiments and Kirk-Buff theory approach Rosgen and co-workers [105] suggested preferential solvation of TMAO molecules by water and urea. They proposed that TMAO's strong hydrogen-bonding ability with water and urea molecules as well the excluded volume effect account for the stabilization of protein native conformation in presence of TMAO. According to their proposed hypothesis, in binary TMAO solution (in absence of urea) protein is preferentially hydrated with at least one water layer between the protein and TMAO. TMAO molecule interacts with the protein through an intercalated water molecule. When urea is added, TMAO can interact with the protein through either an intercalated urea, or an intercalated water molecule. Since some of the water molecules in the intermediate layer are replaced by urea, and TMAO molecules interact with these urea molecules this effectively increases the spacing between the TMAO and protein molecule. As a result of this TMAO molecules are more excluded from the protein surface in ternary mixed osmolyte solution when compared to binary TMAO solution (without urea). This makes the TMAO molecules a powerful protein stabilizer in presence of urea.

Table 2-4. Average number of hydrogen bonds[a]

System	n_{HB}^{W}	n_{HB}^{U}	n_{HB}^{T}	n_{HB}^{Total}	n_{HB}^{P-P}	n_{HB}^{U-W}	n_{HB}^{T-W}	n_{HB}^{U-T}
PW	43.78	−−	−−	43.78	7.35	−−	−−	−−
PWU	31.06	14.51	−−	45.57	5.71	4.56	−−	−−
PWUT (Kast 2:1)	26.69	13.28	0.62	40.59	6.56	3.69	2.29	0.26
PWUT (Osmotic 2:1)	27.83	14.25	1.62	43.70	6.07	3.74	3.01	0.21
PWUT (Osmotic 1:2)	22.63	7.63	2.24	32.50	6.35	3.53	2.95	0.53

[a] n_{HB}^{W}, n_{HB}^{U} and n_{HB}^{T} are the average number of hydrogen bonds formed by the protein with water, urea and TMAO respectively. n_{HB}^{Total} is the total number of hydrogen bond formed by the protein with the solution species and n_{HB}^{P-P} is the average number of intra-protein hydrogen bond. n_{HB}^{U-W} and n_{HB}^{T-W} are the average number of hydrogen bond formed per urea molecule and per TMAO molecule. n_{HB}^{U-T} represents number of hydrogen bond formed between urea and TMAO per urea molecule.

A plot of occupancy of different intra-protein hydrogen bonds against time (Figure 2-9) shows that TMAO tends to protect the terminal hydrogen bonds (viz. in between residues 2 and 15) joining the two β-sheets more effectively throughout the entire simulation. But in binary urea system (Figure 2-9 (b)) shows no existence of this terminal hydrogen bond. This leads us to conclude that in presence of urea, denaturation of the protein takes place through breaking of terminal intra-protein hydrogen bonds. A comparison of TMAO containing systems depicts that Osmotic model containing system in 1:2 ratio (Figure 2-9 (e)) provides the maximum occupancy for terminal hydrogen bonds. This results in elongated sheet conformation of the protein and the conformation attains more stability. Although other intra-protein hydrogen bonds in between residues 13 and 4, 11 and 6, 10 and 6 (as shown in Figure 2-9 (a) and (b)) have high occupancy for urea and pure water containing systems than TMAO containing systems, it can be concluded that terminal intra-protein hydrogen bonds play a major role in stabilizing overall conformation of the protein.

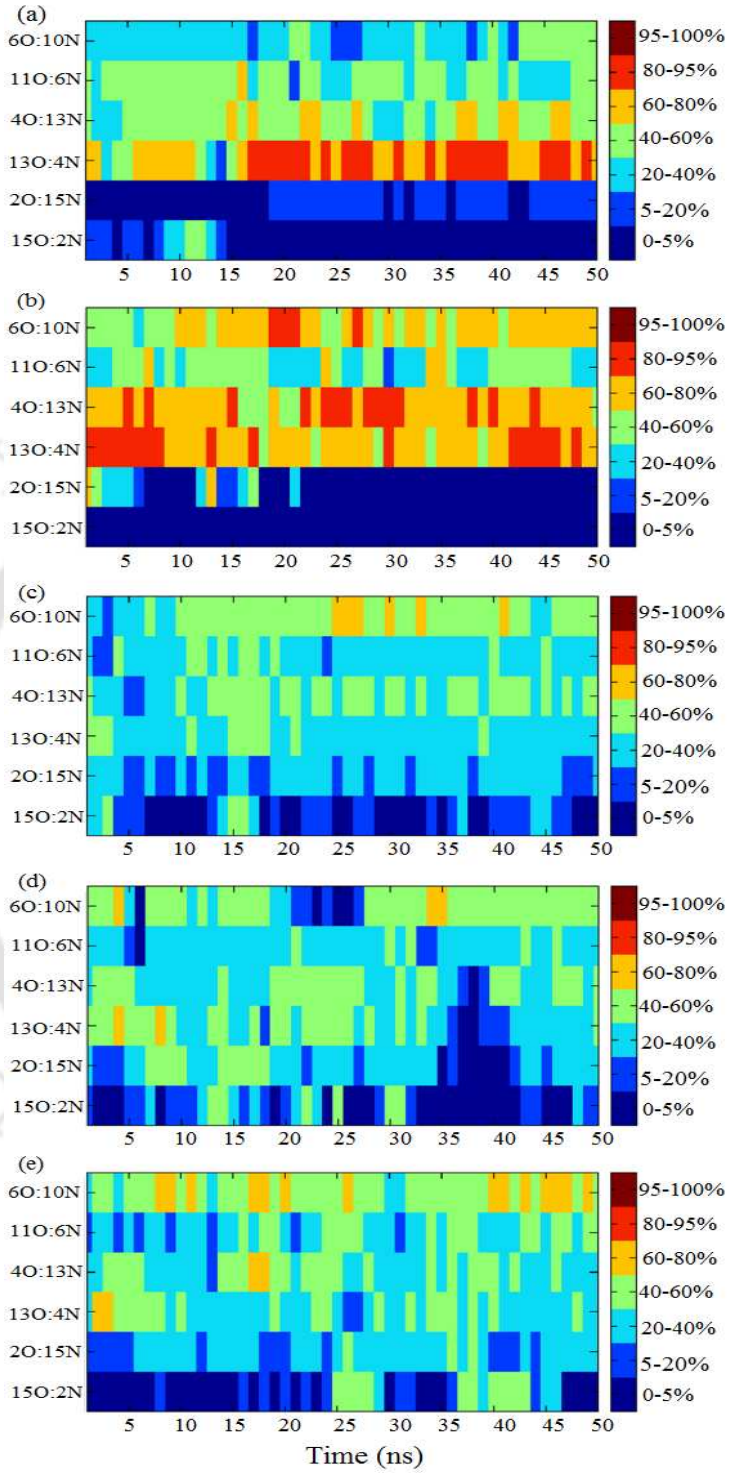


Figure 2-9. *Occupancy of different intra-protein hydrogen bonds formed between the anti-parallel sheets of the β-hairpin for all the systems studied viz. (a) PW, (b) PWU, (c) PWUT (Kast 2:1), (d) PWUT (Osmotic 2:1) and (e) PWUT (Osmotic 1:2).*

As a further investigation we calculated average number of hydrogen bond formed by residues 2 and 15 with urea. For binary urea system the value is 1.031 and for mixed osmolyte systems viz., in Kast model it is 0.692, for Osmotic model (in 2:1 ratio) is 0.451 and for reverse ratio it is 0.398. For the binary urea system, both side chain and backbone of residue 2 and 15 forms hydrogen bond with urea with almost equal weights (0.541 and 0.490), whereas for mixed osmolyte systems side chain contribution is more than backbone (0.508 and 0.040, 0.230 and 0.221, 0.227 and 0.171 respectively for above three TMAO containing systems). Thus it is the backbone intra-protein hydrogen bond (of residue 2 and 15) which is more affected in binary urea system and in presence of TMAO this hydrogen bond is being protected. These results lead us to conclude that involvement of backbones of residue 2 and 15 to form hydrogen bond with urea plays a vital role in determining overall conformation of the protein.

Above discussions lead us to infer that urea delivers its action through direct interaction with protein through the breaking of terminal hydrogen bonds that link the two sheets in parallel fashion while TMAO invokes an environment surrounding the protein which is deficient in water and urea molecules and keeps the protein in its folded form.

■ SUMMARY AND CONCLUSIONS

In this work, by employing classical molecular dynamics simulations, we have examined the extent of protecting ability of two different models of TMAO to counteract urea's action. We have also explored the underlying mechanism by which TMAO can protect the deleterious effect of urea on protein conformation. The folding behavior of the β-hairpin is found to be dependent on the nature of surrounding osmolyte used. Urea causes denaturation of the initial conformation of the protein through breaking of intra-protein hydrogen bonds joining its terminal residues. The turn region joining the two β-sheets are also effected in presence of urea. Electrostatic interaction between urea and protein is found to play an important role suggesting direct interaction mechanism is operative. Both electrostatic and vdW interactions become more unfavorable when TMAO is introduced. Effect being more noticeable on electrostatic than vdW component suggests that urea molecules are replaced from the solvation shell of the by TMAO molecules. This is evident from the calculations of average number of heavy atoms of urea and TMAO in the solvation shell of the protein. Mixed osmolyte systems are studied in order to explore the counteracting ability of different models of TMAO. Kast model of TMAO is found to be capable of offsetting the denaturing effect of urea. Osmotic model of TMAO (in 2:1 ratio)

is also found to exhibit a close behavior with that of Kast model except for the fact that Osmotic model samples some number of denatured conformations of the protein. However, when Osmotic model is used in 1:2 ratio, it provides maximum counteraction against action of urea. Highly populated terminal intra-protein hydrogen bonds are formed resulting into a highly compact conformation of the protein. As is evident from dPCA analysis that the maximum number of stable conformations are sampled when Osmotic model is used in 1:2 ratio. It is achieved by retaining the H-bonds between the two strands of the protein, which facilitates the protein to remain in folded state. Calculation of average number of intra-protein hydrogen bonds between terminal residues 2 and 15 indicates that it is mostly affected in binary urea solution and the same remains protected in presence of TMAO.

From the findings of the work carried out in this chapter lead us to conclude that Kast model of TMAO provides better protection to the protein when used in conventional 2:1 ratio of urea:TMAO. This finding inspired us to carry out rest of the studies with Kast model of TMAO.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Chapter 3

Stabilization of the Protein Trp Cage in Mixed Osmolyte Solutions of Urea and TMAO Without Confinement

"Random walks allow the simulation to pass any energy barrier and to sample a much wider phase space than by conventional methods. Monitoring the energy in a single simulation run, one can obtain not only the global minimum-energy state but also any thermodynamic quantities as a function of temperature for a wide temperature range."

– Y. Sugita and Y. Okamoto *Chem. Phys. Lett.* **314**, 141 (1999)

REMD simulation is one of the most popular computational tools to explore the free energy landscape of a protein due to efficient sampling. In REMD efficient sampling arises by escaping a local minimum (where a trajectory is trapped), by exchange with a higher temperature configuration. Hence REMD simulations can provide an atomic-level picture of the protein's conformational transition pathway. In this study, we present folding/unfolding equilibrium of the protein Trp cage in presence of denaturing osmolyte urea and mixture of denaturing and counteracting osmolytes urea and TMAO by REMD simulation technique. We have used REMD method to elucidate different structural properties of the protein at different temperatures. The study shows how the energy landscapes are affected depending on the surrounding osmolytes. Proteins exhibit marginal stability and literature shows that urea shifts the equilibrium toward unfolded state and TMAO shifts the equilibrium toward folded state [27-47]. Different molecular mechanisms are proposed to explain their actions, however, a definitive answer is yet to be obtained. In this work we have put molecular level insights on the denaturing action of urea and counteracting action of TMAO on the protein Trp cage.

To be more specific on the denaturing and counteracting mechanisms of urea and TMAO on the protein Trp cage we have emphasized on the salt bridge, hydrophobic core and orientational preferences of different planes of amino acid side chains. Role of salt bridge on determining stability of the protein has been a matter of dispute. It is not always a necessary factor for protein stabilization. Trp cage possesses a well packed hydrophobic core. This invokes us to examine the existence of the core and role of TMAO in the protection of the hydrophobic core. Interactions between side chains of the protein residues are regarded as a decisive factor for protein stabilization. Thus we have studied preferential orientations of different planes and how they are affected with the change of chemical environment surrounding the protein. We have carried out analysis of the hydrogen bond and effect of TMAO on protein-urea interaction energy in order to highlight the stabilizing effect of TMAO to compensate the deleterious effect of urea.

Use of REMD to explore the effects of denaturing and mixed osmolyte solutions on the protein Trp cage

Overview: Using replica exchange molecular dynamics simulation technique we have studied urea-induced denaturation and its counteraction by TMAO of the protein Trp cage. We have found that in binary solution urea molecules are very much prone to stay nearby the salt bridge. In presence of TMAO, salt bridge is found to be protected from urea molecules. TMAO also prevents hydration of the hydrophobic core of the protein and thus provides counteraction. Emphasizing on the orientational preferences of different aromatic planes reveals that near parallel orientation of Trp6-Trp18 and near perpendicular orientations of Trp6-Trp17 and Trp6-Trp19 have significant contributions towards stabilization. Trp6-Tyr3 pair is more likely to maintain a tilt angle of 25°. These orientational preferences of the planes that are lost in binary urea solution regained significantly in mixed urea/TMAO solutions. We have also found a highly favorable interaction energy between protein and urea contributed from both electrostatic and van der Waals components. Addition of TMAO makes these interactions more unfavorable. Protein solvation behavior and hydrogen bond calculations show that TMAO molecules replace some urea molecules from the surface of the protein. As a result, a reduction in the protein-urea hydrogen bonds is observed when compared to that of binary urea system. Since no evidence of direct interaction is observed, our findings provoke indirect mechanism of protein stabilization by TMAO molecules.

■ INTRODUCTION

Intrinsic complexities underlying in the protein folding process makes it a never ending field for the researchers. Residue sequence, physiological condition of the surrounding environments and many other factors play crucial role in determining the folding landscapes of a protein [23, 106]. Significant perturbations in those conditions can cause loss of the native structure of the protein partially or completely [107, 108]. Thus, some adaptive mechanisms are mandatory to continue functioning of the cell. Conditions like dessication or dehydration, high- or low-hydrostatic pressure, temperature, salinity etc. cause osmotic stress [109, 110]. Then in response to these osmotic stresses small molecules are released in the cells or tissues that are known as osmolytes [111]. Trimethylamine-N-oxide (TMAO) is one such osmolyte found in intracellular environment of certain marine organisms which is accumulated in order to counteract the effect of urea and the effect of high hydrostatic pressure [34, 112, 113].

The application of denaturing agent to aqueous protein solution results in to the unfolding of the protein. In this study we have used urea as denaturing agent which has been studied extensively through both theoretical as well as different experimental methods (as discussed in Chapter 1). Studies in support of both direct (preferential interaction of urea with the protein) and indirect (alteration of water structure which in turn affects the hydration of the protein) mechanisms of urea denaturation are reported (discussed in Chapter 1). Some recent studies put consensus views on direct interaction mechanism. However, molecular level studies on urea are still exploring new facts to provide finer explanations of denaturation mechanism [114]. In contrast to the favorable interaction of urea with protein, the protecting osmolytes have no such favorable interactions but they offer stability to the protein [115]. Folding propensity of the protein was proposed to arise due to enhanced water structure in presence of TMAO [36, 37]. Some studies are found to highlight the direct interaction between protein and TMAO molecules [36, 38, 39], whereas some others highlight the indirect interaction [33, 40-46, 116, 117]. These are mentioned in brief in Introduction chapter. Numerous studies (that considered mixture of protecting and denaturing osmolytes) are reported to establish the mechanism by which TMAO stabilizes the protein and counterbalances the effect of urea. However a unified scenario has yet to be emerged.

In the present work we have carried out replica exchange molecular dynamics simulation (REMD) simulation [21] of Trp cage in pure water, in binary urea solutions as well as

in ternary solutions of urea and TMAO in two different concentrations of TMAO in order to explore the sensitivity of protein folding/unfolding equilibrium towards the surrounding osmolytes and sort out the essential features of the intermediates formed by the protein. Because of its small size and fast kinetics, Trp cage is an extensively investigated protein in terms of computational and experimental works [15-18, 118]. We have considered REMD simulation technique over the regular molecular dynamics (MD) simulation because of the fact that for the latter technique it is usually difficult to obtain accurate canonical distributions at low and room temperature limits [21]. This is because of the fact that simulations at such temperatures tend to get trapped in one of a huge number of local or minimum energy states.

In this study we have particularly emphasized the effect of urea and TMAO on the three key stabilizing factors of Trp cage: (i) salt bridge between residues Asp9 and Arg16, (ii) the buried hydrophobic core surrounding the indole plane of residue Trp6 and (iii) orientational preference of different aromatic planes constituting the protein. Salt bridges are formed by oppositely charged residues and believed to contribute to protein stability. But, it has also been reported that it is not always mandatory to form a salt bridge to reach the stable conformation of the protein [63, 119, 120].

Comparative studies with mutated form of protein and non-mutated forms of hydrophobic core established the relationship between protein stability and hydrophobic core formation. Hydrophobic core provides tertiary interactions among the secondary structure elements/motifs (such as helix or sheet or turn or coil). Thus, the maintenance of the hydrophobic core determines the overall stability of the protein. It is well established that changes in interactions in the hydrophobic core can have effects on the protein conformation. Experimental findings along with theoretical investigations have shown that mutations in the hydrophobic core enhance instability of the protein [15, 121]. Long range hydrophobic interactions are identified between aromatic planes that contribute towards stabilization of protein. Studies involving mutations at different sites of the protein Trp cage revealed that interactions of the aromatic planes between different pair of residues contribute largely in its stabilization. For example, Andersen and co-workers [15] have shown that substitutions of proline residues at position 17 and 19 resulted in less folded structure of Trp cage (Pro19 exhibits interactions with Trp6 and Tyr3). Furthermore, hydrophobic interactions between Trp6 and Tyr3 are also found to draw attention of researchers [122-124]. The outline of this chapter is as follows. The details of the simulation protocols is provided in next section. Then we present and analyzed the results for all the systems. This is followed by summary

and conclusions of the work.

■ MODELS AND SIMULATION METHOD

REMD simulations were performed to study denaturation and counteraction of urea and TMAO respectively on the mini-protein Trp cage (PDB ID 1L2Y [18], Figure 3-1). The protein has 20 residues. At neutral pH, lysine, aspartic acid and arginine residues are charged and hence one chloride ion was added to neutralize the mini-protein.

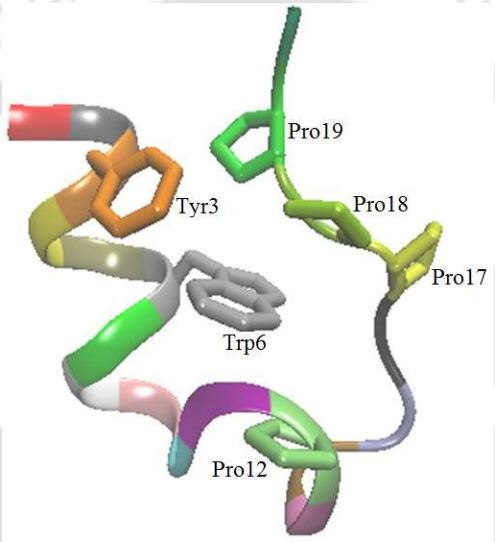


Figure 3-1. *NMR structure of Trp cage with residue-wise color identity. The aromatic planes of different residues are also shown.*

Simulations were performed using AMBER12 program [70]. The systems considered in this study are described in Table 3-1. 8M aqueous urea solution was used to cause unfolding of the protein and two different urea-TMAO mixed solutions with varying TMAO concentrations are used to examine the counteracting effect of TMAO. We have used comparatively larger systems for ternary solutions (than binary urea solution) so that urea and TMAO molecules are sufficiently solvated by water molecules. We have used Kast model [88] of TMAO rather than Osmotic model [33] since it is evident from our previous study [53] that Kast model of TMAO gives better protection against urea when used in conventional 2:1 ratio of urea:TMAO. Some other models of TMAO are also available viz. Netz model [51], Shea model [52]. These are developed on the base of Kast model to reproduce thermodynamic properties, density etc. of aqueous TMAO solution. But till date no such

force field parameters for TMAO is available that adequately describes the balance of interactions in mixture of solution. So we have restricted our study to Kast model based on our previous study.

For solvation of the protein in all the simulations TIP3P model [125] of water was used. Lennard-Jones parameters of different atomic sites of TIP3P water, urea [87] and TMAO [88] and their corresponding partial charges are considered according to the citations mentioned. For the protein, Amber ff12SB parameters [126] were used. To begin with, the initial configurations of all the systems were generated by Packmol program [89]. All non-bonded interactions were treated with 9.0 Å cutoff. SHAKE algorithm [90] was used to put constrain in all bonds involving hydrogen. Langevin dynamics [91] was used for controlling temperature and Particle mesh Ewald summation method [93] was used to treat electrostatic interactions. Before starting REMD simulation, each of the systems was subjected through some mandatory steps. At first we have carried out minimization of the Packmol generated systems through steepest descent method. Then heating was carried out gradually from 0 to 300K in terms of 5 ps steps under NVT conditions with a temperature jump of 50K. The resulting systems were then subjected to NPT conditions (1 atmospheric pressure and 300K temperature) for 300 ps and finally NVT conditions for another 400 ps. These steps ensured that Trp cage was homogenized sufficiently and the structure of Trp cage remained unperturb. The systems obtained from these steps were then used for REMD simulations.

Table 3-1. Overview of Systems[a]

System	N_W	N_U	N_T	*No.of Replicas*	*Temperature* (K)	*Density* (g/cc)	*Box Length*(Å)
PW	2640	0	0	32	280-444	0.99	43.58
PWU	2293	376	0	32	280-447	1.12	43.27
PWUT (4M)	4000	1152	576	40	288-430	1.13	64.56
PWUT (5M)	3720	1152	649	40	288-430	1.16	64.45

[a] N_W, N_U and N_T represent the number of water, urea and TMAO molecules respectively. ρ represents density of the system. P, W, U, and T refer to protein, water, urea, and TMAO respectively.

With a distribution of temperatures ranging from 280 to 450K over 32 to 40 replicas

(depending on the systems as described in Table 3-1) REMD simulations were carried out. To set up the temperature spacing between the replicas, energy distributions were obtained from a series of short (0.5 ns) constant volume MD simulations at five different temperatures (viz. 275, 325, 375, 425 and 475K). Average energy values for the corresponding temperatures were then fitted with a polynomial equation which was then solved to evaluate the temperature spacing with 20% accepted exchange probability between neighboring replicas. In REMD periodic state exchange moves are attempted in between neighboring replicas (i and j) according to the acceptance rule

$$P_{acc} = min(1, exp\left[\left(\frac{1}{k_B T_i} - \frac{1}{k_B T_j}\right)(E_i - E_j)\right]) \tag{3.1}$$

where E_i and E_j represent the energy of the replicas i and j respectively (discussed in Methodology of Chapter1). Swapping of temperature between neighboring replicas are successful when the above equation 3.1 is satisfied so that there is a sufficient overlapping of energy of the corresponding replicas. Temperature based trajectories were then sorted for each single temperature. The simulation time for each of the systems was 50 ns per replica. Almost a decade ago Paschek et al. [127] carried out REMD simulation of Trp cage for 100 ns starting from an extended conformation. They found that at least 40 ns was required for equilibration period. In this study the initial conformation of the protein is not the extended one, hence equilibration period of 50 ns is found to be sufficient. The trajectories were analyzed with AMBER ptraj and cpptraj [128] module as well as VMD[94] (Visual Molecular Dynamics package) as mentioned whenever in need of them.

■ RESULTS AND DISCUSSION

Root Mean Square Fluctuations (RMSF) and Secondary Structure

Root mean square fluctuations (RMSF) provide information about the magnitude of fluctuations of C_α-atom of each residue of the protein. Residues with high RMSF values exhibit distinct difference in flexibility from that of the native state indicating the structural change. We have calculated RMSF for both low (at 288K) and high (at $\leq$447K) temperature limits (see Figure 3-2). At high temperature limit the peptide residues for every system are found to be more flexible than the corresponding low temperature limit which is as per with the expectation. For the binary urea system (PWU) all of the residues show maximum flexibility. Action of TMAO against the denaturing action of urea is reflected from this RMSF plot with a decreased flexibility of the residues in presence of TMAO than that for binary urea system. This effect is quite prominent for low temperature limit.

At high temperature, (which is not physically relevant) the counteracting effect of TMAO against the deleterious effect of urea is not very prominent.

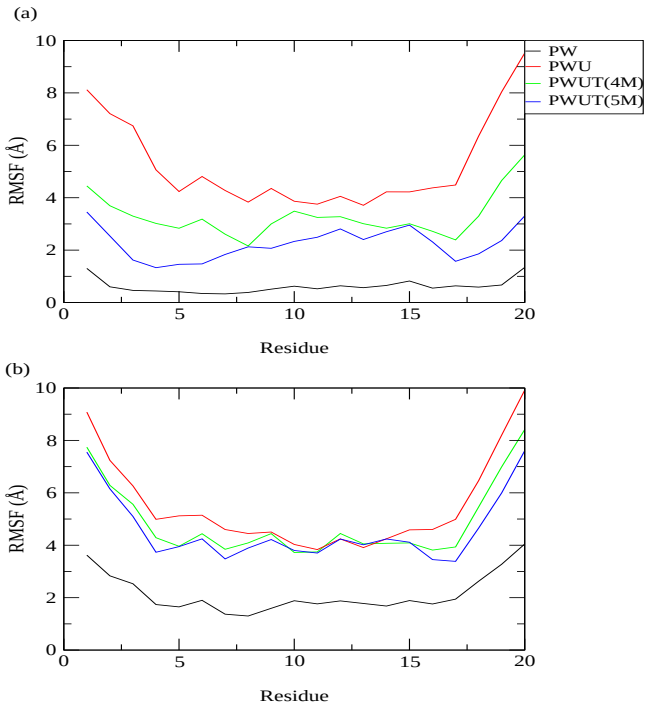


Figure 3-2. *Root mean square fluctuations (RMSF) of C_α-atoms of each of the residues of the protein at: (a) low temperature and (b) high temperature.*

Based on dictionary of protein secondary structure (DSSP) algorithm (by W. Kabsch and C. Sander [129]), secondary structure analyses are carried out at 288K temperature to find out the secondary structure elements adopted by each of the residues throughout the entire simulation path. A comparison of the DSSP plots for different systems (shown in Figure 3-3) clearly indicates that the ease of the protein to retain its native structure varies depending on the surrounding chemical environments. The native structure is absolutely retained in pure water. In binary urea solution the native structure is hardly maintained with a very high appearance of coil structure (except for the initial phase of simulation). In 4M ternary solutions of TMAO, protein's native conformation is conserved to some extent. In ternary solution with 5M TMAO, a moderately higher retention of the native structure is observed (when compared to 4M ternary TMAO solutions).

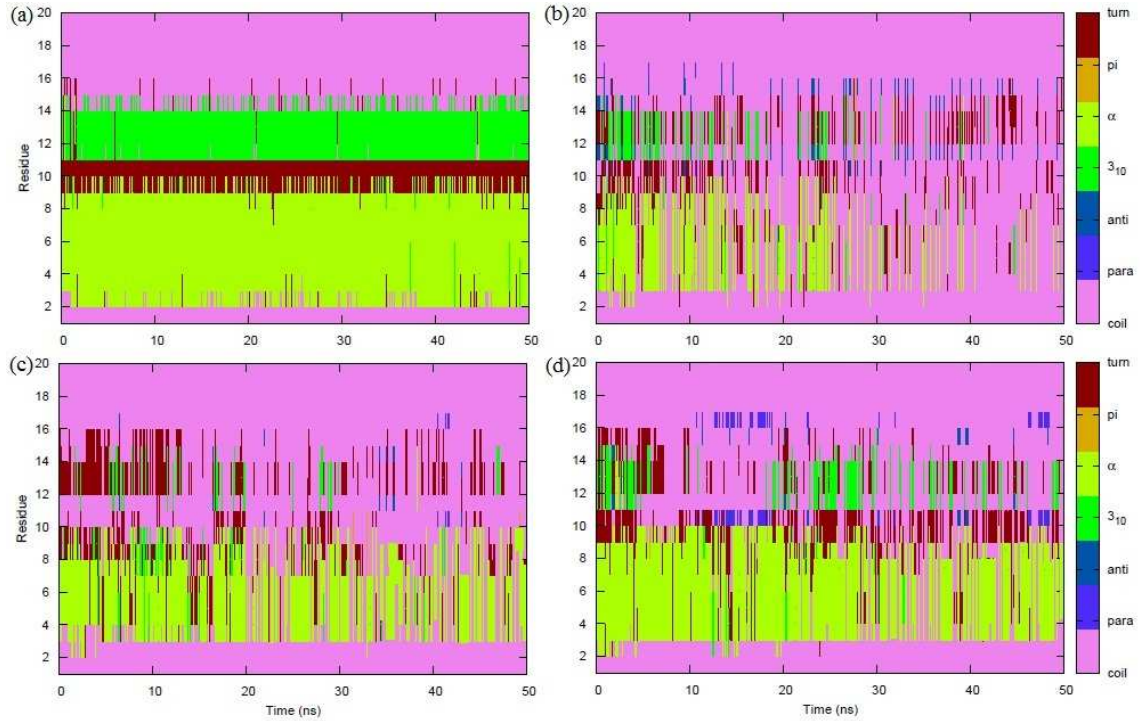


Figure 3-3. *Plot of secondary structure analysis of the protein vs. simulation time. (a), (b), (c) and (d) correspond to PW, PWU, PWUT (4M) and PWUT (5M) systems.*

Radius of Gyration

The analysis of radius of gyration (R_g) values of Trp cage can provide information about the effective size of the protein in absence/presence of osmolytes. For better understanding, the probability distribution of R_g values for every system is plotted at the two temperature limits (see Figure 3-4 (a) and (b)). Low temperature ensembles are found to exhibit smaller effective size than the high temperature ensembles. We note that our results are in accordance with the findings of Garcia and co-workers [63]. The maximum compact form is retained by the protein in pure water system. In binary of urea system, the protein attains unfolded conformation as is reflected in high value of R_g (Å). In mixed osmolyte systems, the effect of urea is reduced leading to a smaller effective size of the protein than that of binary urea system. Furthermore, the average R_g values for each temperature of all the systems are calculated. The effect of temperature and the effect of osmolytes on the protein are evident from the Figure 3-4 (c). As observed before [63], we also find that the value of R_g increases gradually with increasing temperature. The lowest average R_g value

(7.3 to 7.7 Å) is maintained for pure water system over the entire range of temperature. It indicates that the most compact conformation of the protein is attained in absence of any osmolyte.

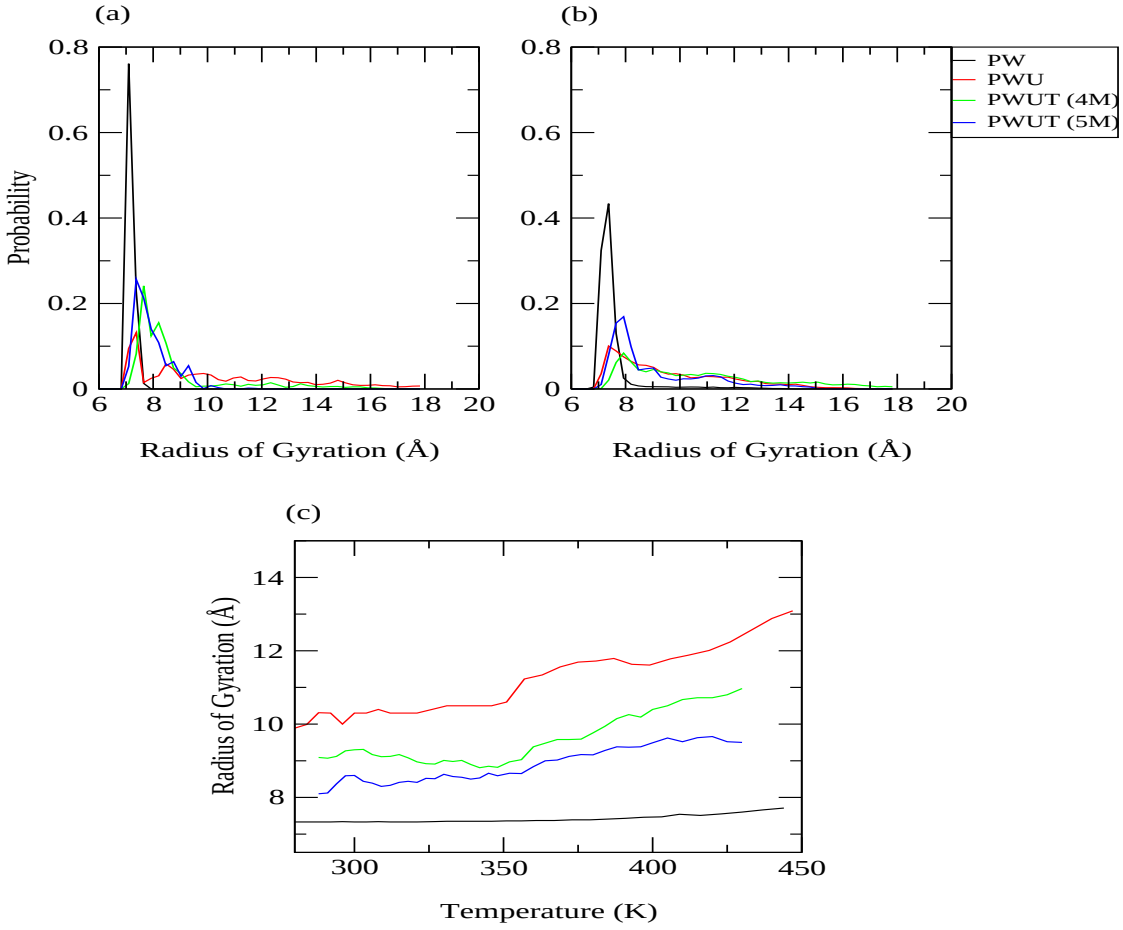


Figure 3-4. *Probability distribution of radius of gyration at (a) low and (b) high temperature limits. (c) Variation of average radius of gyration with temperature.*

Melting Curves and Free Energy Landscapes (FELs)

The use of REMD simulation helps us to construct the melting curves of Trp cage in pure water system as well as in the systems consisting of urea and TMAO (Figure 3-5 (a)). In absence of any osmolyte, the protein shows highly stable behavior upto very high temperature. The folded conformation of the protein is almost retained throughout all the temperatures except at high temperatures where a modest reduction of the folded

fraction value is obtained. The addition of urea to pure water system causes unfolding of the protein even at the lowest temperature range and more unfolding is observed as temperature increases. For the mixed ternary solutions counteracting effect of TMAO can be observed from the folding fraction values that are much higher than the binary urea solutions. Again, ternary solution with 5M TMAO shows better protection than that of 4M TMAO solution, which is quite evident from the higher folded fraction value for the former solution. Nevertheless, for 5M TMAO solution, we observe that at room temperature the protein retains approximately 80% folded conformation. The corresponding free energy changes for respective systems are also calculated [63] and plotted in Figure 3-5 (b). ΔG_u can be calculated as:

$$\Delta G_u = -RTlnK = -RTln\left[\frac{(1-x_f)}{x_f}\right] \tag{3.2}$$

where K is the equilibrium constant and x_f is the folded fraction of the protein. It is apparent that for pure water system, a sharp change is observed at ~380K, above which the ΔG_u value starts to decrease indicating loss of free energy due to appearance of a few unfolded conformations above this temperature. This is also evident from the calculation of free energy landscapes discussed below.

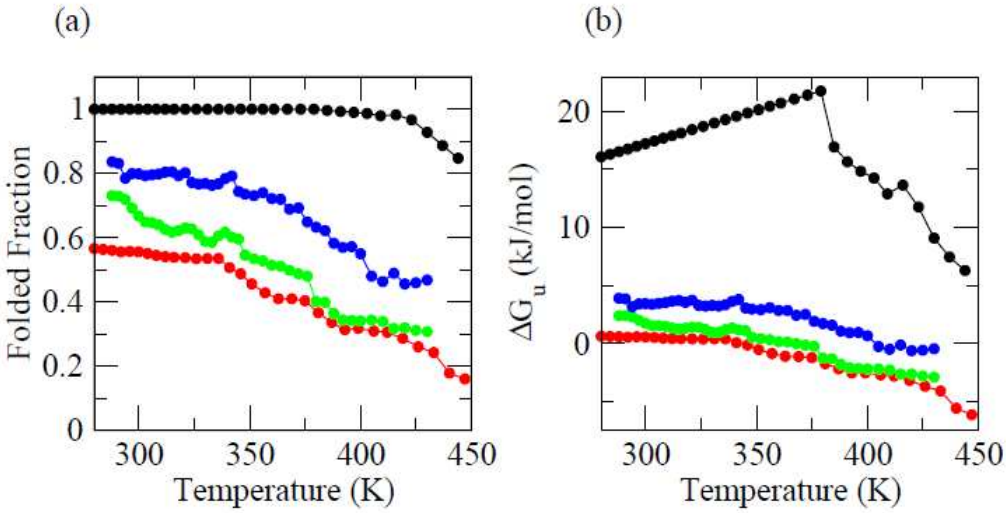


Figure 3-5. *Plot of (a) helical fraction with temperature and (b) corresponding free energy of unfolding for respective temperatures.*

Free energy landscape (FEL) has always been proved to be a useful tool to capture all possible conformations undergone by the protein [131, 132]. We have used two parameters viz. RMSD (root mean square deviation) and temperature to construct the FELs for all the systems. We cover the entire range of temperatures and considered different conformational changes adopted by the protein. Free energy is computed by using the following the equation:

$$\Delta G(V) = -k_B T[lnP(V) - lnP_{max}] \tag{3.3}$$

where k_B is the Boltzmann constant, T is the temperature, P(V) is the probability distribution of the coordinate V (RMSD in this study) and P(V) is subtracted from the maximum of the distribution, P_{max}, so that ΔG becomes zero for the free energy minimum.

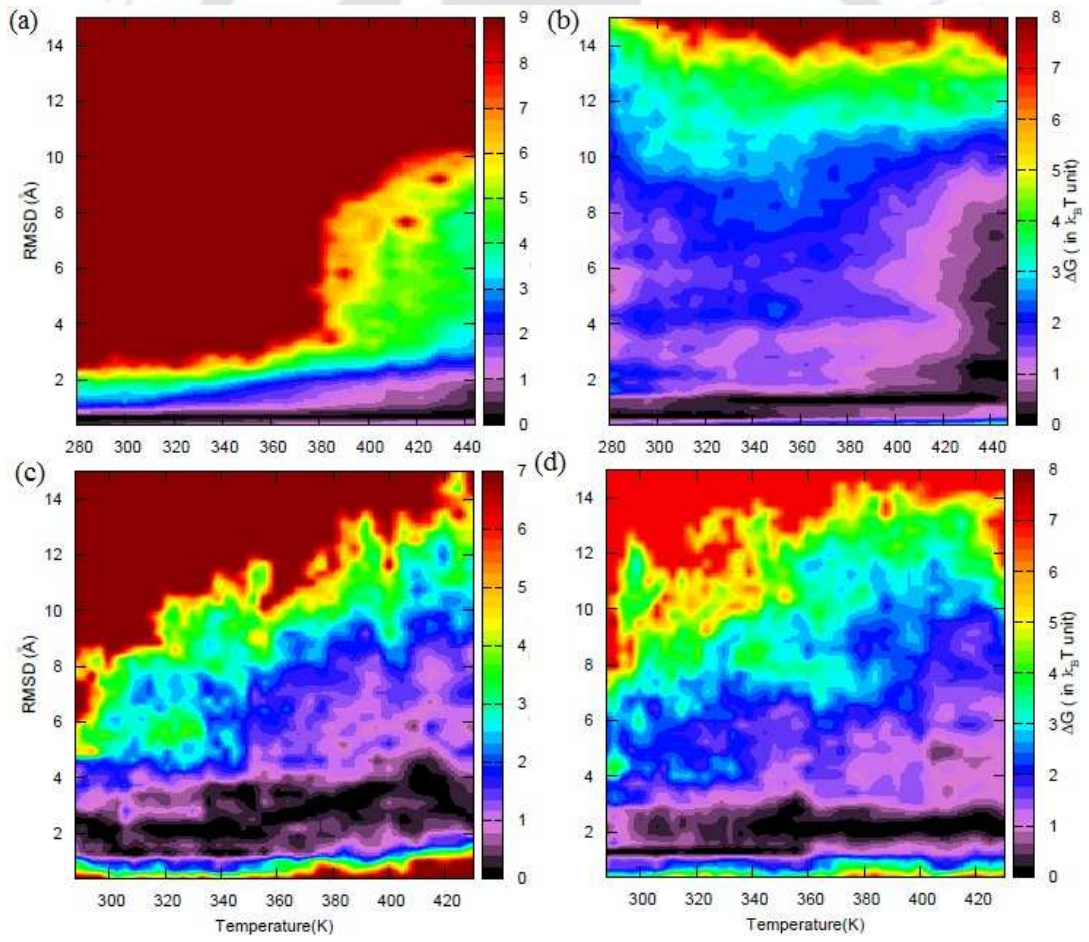


Figure 3-6. *Free energy landscapes (FELs) for systems (a) PW, (b) PWU, (c) PWUT (4M) and (d) PWUT (5M).*

In Figure 3-6, we present the FELs of the protein for all the systems as function of C_α-RMSD and temperature. Conformations with RMSD values less than 2.3 Å are considered to remain in native conformation. The ruggedness of surfaces indicate the random walks in energy spaces visited by the protein. As shown in Figure 3-6 (a), FEL of pure water samples the least variations of conformations. Above 380K temperature, however, we observe conformations with slightly higher RMSD values. For 8M binary urea solution, the FEL is found to expand from the low temperature limit (Figure 3-6 (b)). Throughout the entire temperature range, the protein adopts conformations with RMSD values ranging from 0 to 15 Å. It indicates the unfolding of the protein from the low temperature limit, which continues upto high temperature. Thus, the low appearance of native form and high appearance of largely unfolded conformations are observed in 8M urea solution. TMAO shows its protecting ability and counteract the effect of urea largely which is quite evident from the narrower distribution of RMSD values in comparison to the FEL of binary urea system. From the Figure 3-6 (c) and (d), it becomes obvious that 5M TMAO solution provides better protection in comparison to 4M TMAO solution.

Effect of Osmolytes on the Salt Bridge

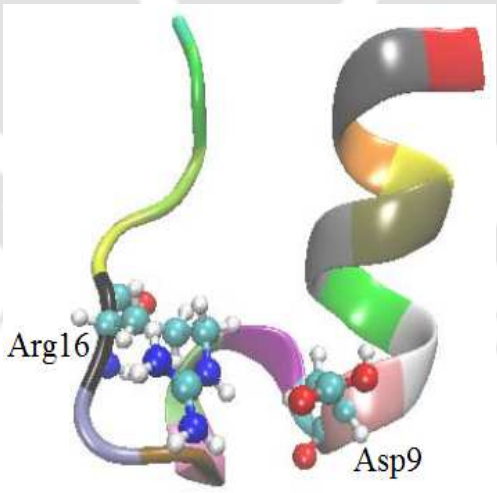


Figure 3-7. *Salt bridge forming residues Asp9 and Arg16 of Trp cage with residue-wise color identity.*

Salt-bridges are known to be important structural motifs that deliver structural stability to the protein [63]. In Trp cage, salt bridge mediated by residues Arg 16 and Asp 9 (see Figure 3-7), is reported to play crucial role in delivering overall stabilization to the protein conformation [133]. The average distance maintained by the residues Arg 16 and

Asp 9 is calculated. For a better understanding, the probability distributions of the salt bridge distances for different systems are plotted and these are shown in Figure 3-8 (a). A comparison of these two parameters for different systems reveals that in pure water the salt bridge is retained throughout the simulation time. Urea affects the salt bridge largely so that the two residues tend to fall apart from each other. When TMAO is present along with urea, the effect of urea is reduced to large extent in particular for 5M TMAO solution, which is reflected from the probability distribution of salt bridge distances.

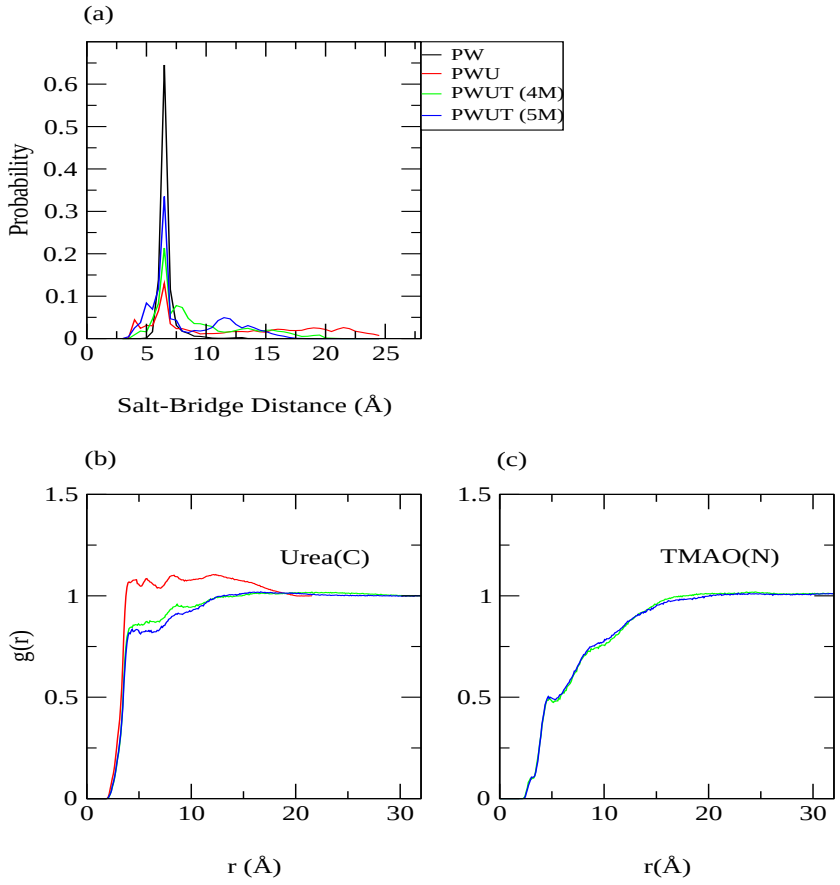


Figure 3-8. *(a) Probability distribution of salt bridge distances for all the systems at low temperature limit (288K). (b) and (c) are the radial distribution functions involving urea and TMAO central atoms and salt bridge residues.*

The above observations suggest that in binary urea solution the presence of urea molecules affects the interactions of the salt bridge forming residues. This prompts us to

examine the presence of urea molecules, if any, in the vicinity of the salt bridge. Thus, radial distribution functions (rdfs) involving urea carbon atom and the salt bridge ion pair of Trp cage for different systems are calculated (see Figure 3-8 (b)). The figure clearly indicates that in binary solution, urea molecules are very much prone to stay in the nearby position of the salt bridge. There is a distinct decrease of urea distribution in ternary solutions which is quite evident from the decreased first peak height of the rdf. This reduction of urea density may be due to replacement of some of urea molecules from this region by other solvent or cosolute molecules. Plot of rdf of TMAO central atoms around the salt bridge (see Figure 3-8 (c)) reveals the presence of a few TMAO molecules in the first solvation shell of it. Thus, TMAO molecules play a role on protein stabilization which may have arisen through the expulsion of some urea molecules and producing effect through indirect interaction which ultimately contributes to protein stability. These observations are in line with the surface rdfs obtained by Li et al. [117] for the protein HP35 in urea and mixed urea-TMAO solutions.

In support of the above observations we have further calculated spatial distribution functions [164] of urea and TMAO molecules surrounding the salt bridge and the same are shown in Figure 3-9. It is quite apparent that the density of urea molecules in the nearby position of the salt bridge in binary urea solution is much higher than that for the ternary solutions. In ternary solutions, the presence of TMAO molecules in the vicinity of the salt bridge is also quite noticeable.

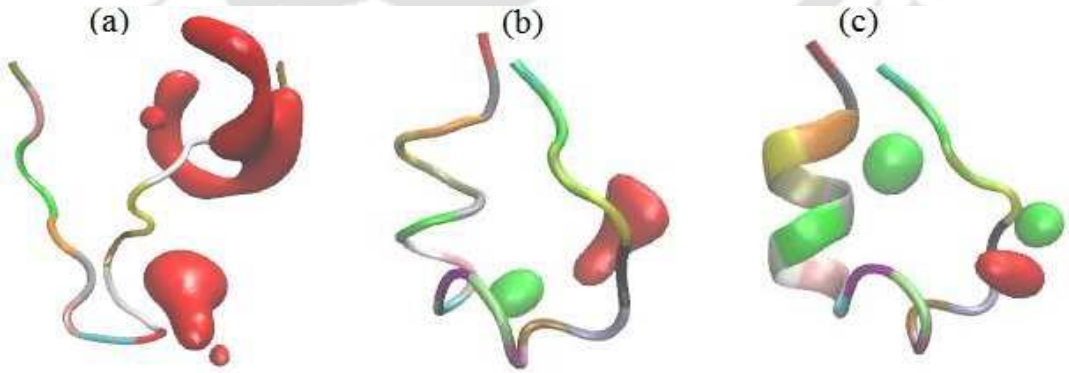


Figure 3-9. *Spatial distribution functions (sdfs) of urea (red) and TMAO (green) surrounding the salt bridge residues of systems (a) PWU, (b) PWUT (4M) and (c) PWUT (5M). Isosurface value of 3.6 is used for calculation of sdf.*

Stabilization of the Hydrophobic Core and the Effect of Osmolytes

A packed hydrophobic core inside a protein provides significant contribution to the stabilization of its native structure. Native conformation is retained as long as hydrophobic core is maintained. To better understand the role of osmolytes in stabilization and destabilization of the protein, it is important to examine their effect on the hydrophobic core. The protein Trp cage possesses a hydrophobic core, which is constituted by the residues 3(Tyr), 6(Trp), 7(Leu), 11(Gly), 12(Pro), 18(Pro) and 19(Pro). Residue 6 i.e. tryptophan contains one indole group which is located at the center of the hydrophobic core region [15, 133, 134]. Thus, the measurement of distances of the center of mass of the indole plane from the other residues of hydrophobic core for different systems can be effective in providing the effect of osmolytes on the hydrophobic core. Figure 3-10 (a) shows that for the first 30 ns of the simulation, the hydrophobic core is conserved for all the systems (i.e. native conformation is not lost). But after that, for binary urea system and ternary solution with 4M TMAO, the hydrophobic core is not conserved. However, for pure water and ternary solution with 5M TMAO the hydrophobic core is conserved. Figure 3-10 (b) shows the average number of central atom of urea, TMAO and water molecules that are present within 3.5 Å of the hydrophobic core forming residues. It can be seen that urea molecules (although small in number) are involved in solvating the considered hydrophobic residues, which is only slightly higher for binary urea solution than mixed solutions. In mixed osmolyte solutions, along with urea and water molecules, some TMAO molecules also become available to solvate the residues. It indicates that TMAO molecules replace some of the water molecules from the vicinity of the hydrophobic core forming residues and cause a reduction in the number of water molecules when compared to that for binary urea solution.

The calculation of hydration number for each of the residues contributing to the hydrophobic core (see Figure 3-10 (c)) further reveals that residue 6 i.e., tryptophan containing indole plane remain buried in native conformation (with low hydration number) in pure water system. But, for binary urea system it becomes highly hydrated. Presence of TMAO, however, prevents its exposure to water to some extent with a slightly more pronounced effect in 5M TMAO solution.

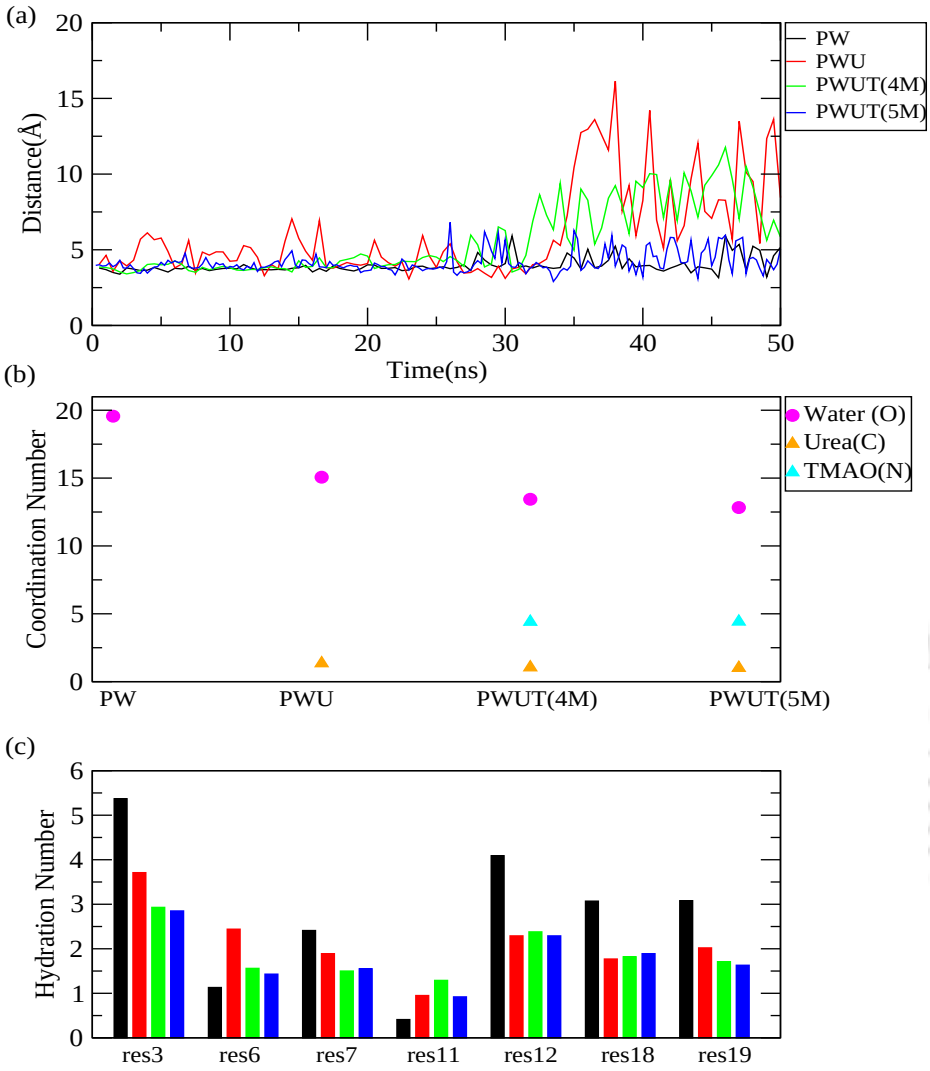


Figure 3-10. *(a) Distance of hydrophobic core forming residues from center of the core. (b) and (c) are the coordination number and hydration number of the hydrophobic core forming residues for different systems.*

Solvent accessible surface area (SASA) is also calculated for the residues constituting the hydrophobic core. The SASA values for all the systems are shown in Table 3-2. For the calculation of SASA, a probe radius of 1.4 Å (which approximates the size of a water molecule) is considered [97]. It estimates the surface area of the protein that is available for the surrounding solvent molecules. Thus high- and low-values of SASA speak about high- and low-exposure of the protein residues to solvent molecules respectively. Since, in this

study we have considered SASA only for hydrophobic core forming residues, the average value of SASA indicates that in pure water system hydrophobic core is maintained (SASA value is minimum for this case). For binary urea system, SASA value is the maximum suggesting that hydrophobic core is more exposed allowing more solvent molecules to come in contact with the protein. However, in presence of TMAO, in particular for 5M TMAO solution, the effect of urea is reduced significantly and we get a fairly low value of SASA, which is comparable with that of pure water system.

Table 3-2. SASA of the hydrophobic core for respective systems[a]

System	$SASA$(Å^2)
PW	941.38 (±20.89)
PWU	1413.65 (±33.60)
PWUT (4M)	1110.14 (±35.48)
PWUT (5M)	1027.85 (±31.82)

[a] Solvent accessible surface area (SASA) of the hydrophobic core. The standard errors are shown inside the parentheses and are calculated by dividing the total simulation runs in to five independent blocks.

Orientation of Indole Plane with respect to Other Aromatic Planes of Trp Cage

The protein Trp cage contains four proline residues (viz. residues 12, 17, 18, 19) that contain one pyrrole ring each and one phenyl ring containing residue (Tyr3). Since indole ring plane (of Trp6) occupies the center of the hydrophobic core, it will be interesting to examine the different orientations adopted by the indole plane with the other aromatic planes as osmolytes are added. To define inter-plane angle we have considered the angle between the vectors normal to the planes of the respective rings (see Figure 3-11). Thus, angle values close to 0° and 90° correspond to parallel and perpendicular orientations respectively. Center of mass (COM) distances between these planes are also calculated to verify positional displacements of the planes in presence/absence of osmolytes. For all the systems, we have calculated the normalized probability distributions of inter-plane orientational angles, P(θ), as well as inter-plane COM distances, P(Å), and plotted them against the inter-plane angle and the inter-plane COM distances respectively (see Figure 3-12). Since the protein retains most of its native conformation in pure water, the distributions of angles and distances for this case can be considered as optimal distributions. Deviations

from this optimal range indicate conformational changes that may lead to destabilization. It can be seen that the probability functions are non-uniform suggesting orientational preferences of the planes considered.

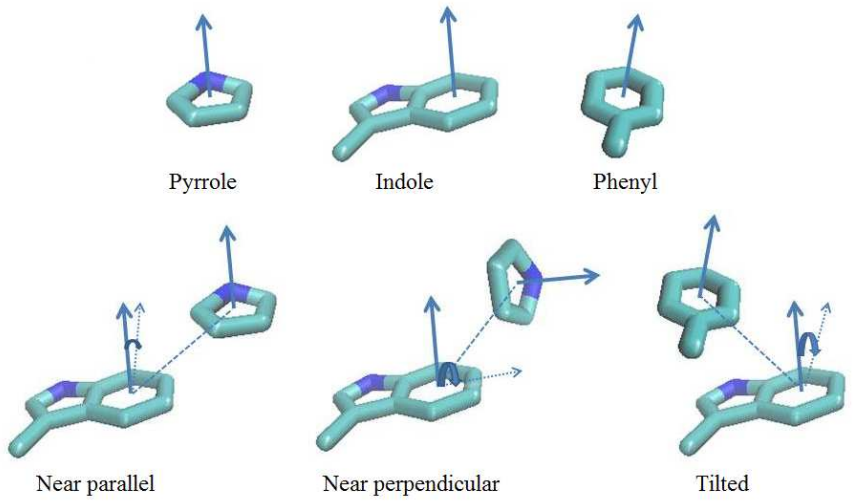


Figure 3-11. *Different planes showing the vectors perpendicular to the planes and orientational angles between two planes.*

As is shown in Figure 3-12 (a) that for pure water system, the value of P(θ) attains a maximum for $\theta \approx 25°$. This infers that the indole plane of Trp6 prefers to maintain a tilt angle of approximately 25° with phenyl ring of Tyr3. As urea is added (for binary urea system), the probability distribution is for this angle is broadened. This suggests that the extent of deviation of the tilt angle of the pair of rings has some influences on Trp cage stabilization. Remarkably, in mixed osmolyte solutions the effect of urea is neutralized to large extent as is evident from the much narrower distribution of P(θ) without changing its position of maximum. The very similar behaviors are also observed for the probability distributions of inter-plane angles between Trp6-Pro17, and Trp6-Pro18. But for the angle between planes Trp6-Pro12, the probability distribution gets broadened on addition of urea and it is not recovered in mixed osmolyte solutions.

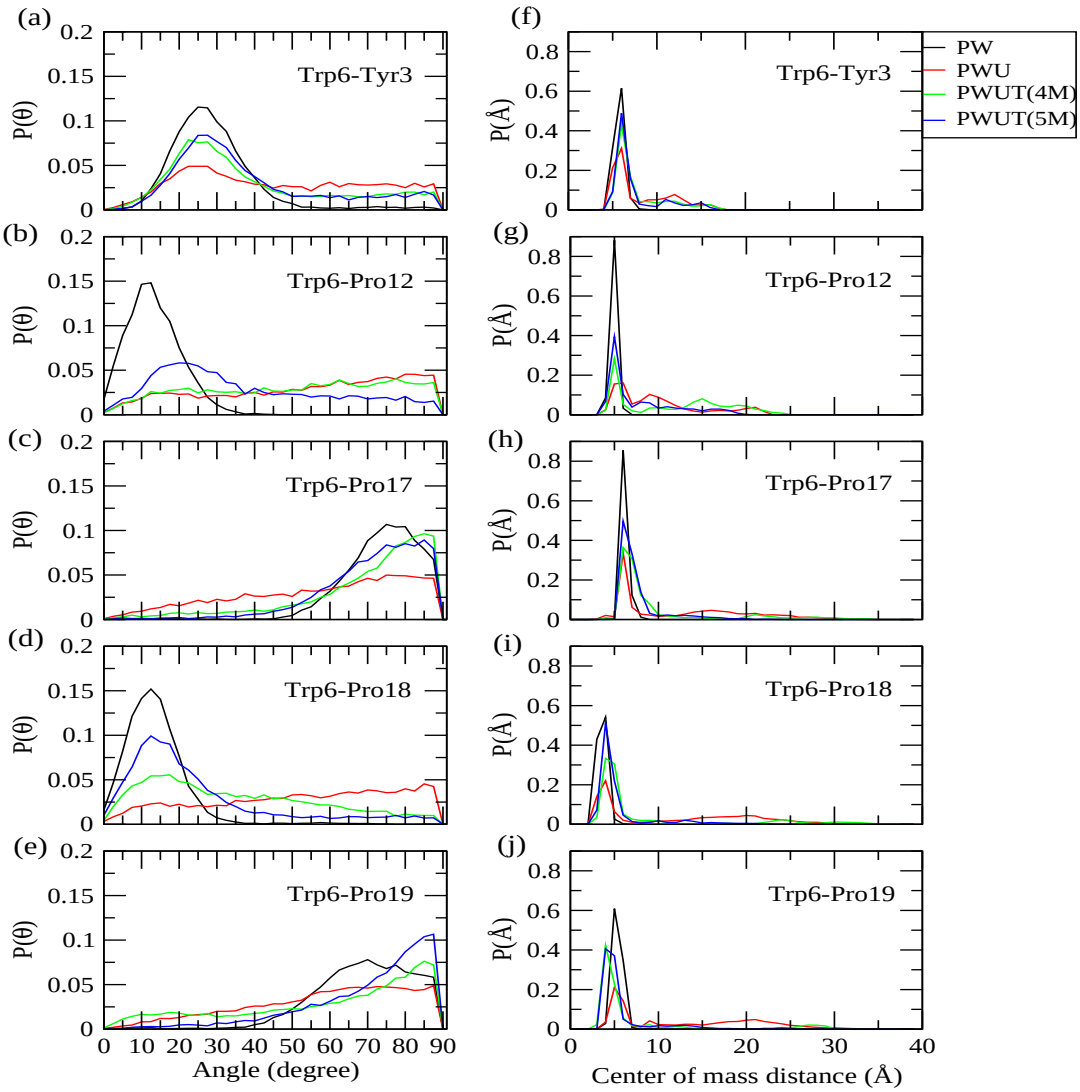


Figure 3-12. *Probability distribution of the orientational angles (left) and COM distances (right) of indole plane of Trp6 with respect to ring planes of Tyr3, Pro12, Pro17, Pro18 and Pro19.*

The protein Trp cage exhibits native-like behavior in mixed osmolyte solutions (as evident from previous discussions). Thus, it can be concluded that Trp6-Pro12 orientation does not play a significant role in the stabilization of the protein in presence of mixture of osmolytes. In this context it is worth to mention that Andersen and co-workers [15] and Yang and co-workers [133] have also found that interaction of this pair is not essential for protein stabilization (already discussed in Introduction section). They further established that if single point mutations at proline sites cause significant decrease in Trp cage stability, those proline residues contribute dominant cooperative interactions in folding of the protein.

Figure 3-12 (c) and (d) depict that Pro17 and Pro18 maintain near perpendicular and near parallel orientations with respect to the indole plane in pure water and mixed osmolyte solutions. In binary urea solutions these orientations are perturbed and positional shifting of the planes also occurs. Previous studies [15, 134] also confirmed the influence of Pro17 and Pro18 on the stability of Trp cage folded structure.

Figure 3-12 (e) indicates that no sharp orientation pattern is followed between the planes of Trp6 and Pro19 in pure water system (since the distribution is near broad). But in mixed osmolyte systems the planes exhibit a tendency to preserve perpendicular orientations which was not observed for binary urea system. Evidence of interactions for this pair is again noted in the work of Andersen and co-workers [15]. Taken together, it can be concluded that specific orientations (parallel/perpendicular/tilted) of proline rings with Trp6 and other residues may provide suitable conditions for hydrophobic interactions among the residues that lead to overall stabilization of the protein. In our study, residues Pro17 and Pro19 preserve approximately perpendicular orientations with the indole plane and Pro18 preserves approximately parallel orientations to maintain the fold stability of Trp cage, whereas Trp6-Tyr3 pair is more likely to maintain a tilt angle of approximately 25°. The probability distributions of COM distances P(Å) exhibit the similar trends as that of P(θ), as osmolytes are added (see Figure 3-12 (f)-(j)). In brief, addition of urea to pure water system decreases P(Å) value without changing the position of maximum. On the other hand, in mixed osmolyte system, the value of P(Å) increases and the effect of urea gets nullified.

Effect of TMAO on Protein-Urea Interactions

Stability of the protein conformation is governed by the nature of the surrounding osmolytes. As is evident from the literature, for denaturation action of urea, two mechanisms are proposed viz. direct and indirect (already discussed in **Chapter 1**). The direct interactions are again classified into electrostatic and vdW interactions. The interaction energy of urea with the protein via hydrogen bonds or other electrostatic interactions are considered as electrostatic interactions. On the other hand, the interaction of urea with the protein via vdW or other dispersion interactions are considered as vdW interactions. Previous studies showed that depending upon the system's environments, any one component of these two interactions may play a major role and they are not always mutually exclusive [79, 100, 101]. Some researchers [79] have found that electrostatic interaction increases population of urea near the protein surface. On the other hand, several other studies [78, 136, 137] supported the fact that dispersion interactions play a major role in

delivering urea denaturation.

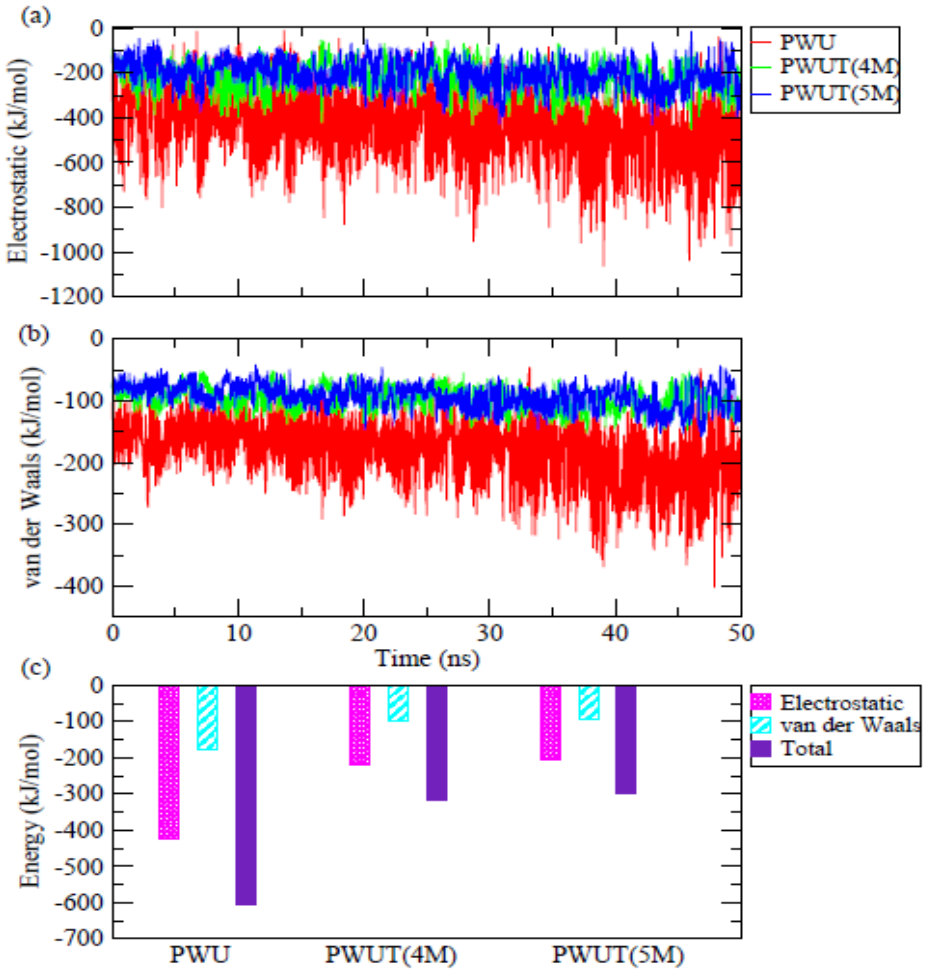


Figure 3-13. *Protein-urea (a) electrostatic, and (b) van der Waals interaction energies vs. simulation time. (c) represents the average value of electrostatic, van der Waals and total protein-urea interaction energies.*

Following earlier work [100], we have decomposed protein-urea interaction energy into two components: electrostatic and van der Waals (vdW) energy and they are plotted as a function of simulation time (see Figure 3-13 (a) and (b)). As can be seen that in binary urea system, both electrostatic as well as vdW components of protein-urea interactions are favorable with a slightly more contribution of the former than the latter. Garcia and co-workers [135] have also supported favorable interactions through electrostatic and vdW forces between protein and urea. Plot of total interaction energy i.e. the sum of the average

vdW and electrostatic energy also supports these observations (see Figure 3-13 (c)). Thus, the interaction between protein and urea is not solely governed by electrostatic energy but vdW energy also plays a significant role, and, the presence of TMAO molecules has profound effect on the interaction between urea and protein. The highly favorable energy suggests that protein-urea interaction is more feasible in binary urea system which results into unfolding of the protein. In ternary systems, TMAO reduces protein-urea interactions so that both electrostatic and van der Waals become less favorable and the native conformation of the protein is retained. In this context we note that Kokubo and co-workers [44] suggested the favorability of urea's interaction (vdW interaction) with denatured protein rather than the compact helix state. They have also argued that in mixed solution of urea and TMAO, there is more favorable vdW contribution of protein-urea interaction than protein-TMAO interaction.

Protein Solvation and Hydrogen Bond Properties

Analysis of average number of solvent molecules around the protein is carried out by considering a cutoff distance of 3.5 Å (see Table 3-3). Solution species having its heavy atom within this distance from the protein heavy atoms are considered to be inside the solvation shell of the protein. It is apparent from the Table 3-3 that the protein's solvation shell is most heavily loaded in binary urea system. The protein is found to be solvated with the minimum number of solution species in pure water system. This indicates that the most compact form of the protein is retained in pure water. This observation further confirms the idea of spontaneous folding behavior of the protein in pure water [103]. A comparison of different coordination values of binary and ternary solutions shows that with the addition of TMAO molecules, some TMAO molecules occupy the solvation shell of the protein with a significant depletion of urea molecules from it. Thus, effect of urea is reduced in ternary solutions since less number of urea molecules are allowed to exert their effect on the protein. Furthermore, as the number of TMAO molecules increases (from 4M to 5M TMAO), more number of urea molecules are expelled from the protein solvation shell.

Table 3-3. Average number of heavy atoms of water, urea and TMAO within first solvation shell[a]

System	$N_{W(O)}$	$N_{U(C)}$	$N_{U(N)}$	$N_{U(O)}$	$N_{T(C)}$	$N_{T(N)}$	$N_{T(O)}$	
PW	$288K$	74.46	$--$	$--$	$--$	$--$	$--$	$--$
PWU	$288K$	51.69	6.77	34.43	18.18	$--$	$--$	$--$
PWUT (4M)	$288K$	43.52	4.17	27.32	12.43	11.23	0.42	3.45
PWUT (5M)	$288K$	42.50	4.01	25.91	11.79	11.59	0.46	3.62

[a] $N_{W(O)}$, $N_{U(C)}$, $N_{U(N)}$, $N_{U(O)}$, $N_{T(C)}$, $N_{T(N)}$, $N_{T(O)}$ are the average number of heavy atoms of water, urea and TMAO that are present within distance of 3.5 Å from any heavy atom the protein.

For the calculation of intra- and intermolecular hydrogen bonds, following earlier work [103], we have considered the geometric crieteria of donor-acceptor distance of 3.5 Å and a minimum donor-hydrogen-acceptor angle of 120°. Average number of hydrogen bonds formed among different species are calculated and they are presented in Table 3-4. The intra-molecular protein-protein hydrogen bond maintains the residue-residue contacts and thus gives an indication about whether native conformation is conserved or not [138]. It is observed that the maximum number of intra-protein hydrogen bond is formed in pure water system and the minimum number is found for binary urea system. Ternary solutions can preserve the native conformation to some extent as is evident from the higher number of intra-protein hydrogen bonds than that for the binary urea system. In binary urea solution, there is a large number of protein-urea hydrogen bonds. This implies that direct interaction of urea with the protein contributes to the unfolding of the protein. As TMAO is introduced, there is a reduction of protein-urea (and protein-water) hydrogen bonds and a significant number of protein-TMAO hydrogen bonds is formed. Thus, introduction of TMAO results into replacement of some urea molecules from the protein surface. Increasing concentration of TMAO from 4M to 5M results into the more involvement of TMAO molecules in depletion of urea molecules from the protein surface. The small number of protein-TMAO hydrogen bond provides insufficient support for direct interaction of TMAO with the protein. Thus overall it can be concluded that TMAO replaces urea molecules from the protein surface and provides stabilization by indirectly affecting the chemical environment of the protein.

Table 3-4. Average number of hydrogen bonds[a]

System	$Temperature$	n_{HB}^{W}	n_{HB}^{U}	n_{HB}^{T}	n_{HB}^{Total}	n_{HB}^{P-P}
PW	$288K$	60.89	$--$	$--$	60.89	16.11
PWU	$288K$	42.68	35.45	$--$	78.13	9.42
PWUT (4M)	$288K$	38.25	27.09	3.04	68.38	10.75
PWUT (5M)	$288K$	37.42	26.12	3.29	66.83	11.46

[a] Average number of hydrogen bonds formed by the protein with water, urea and TMAO are presented by n_{HB}^{W}, n_{HB}^{U} and n_{HB}^{T} respectively. n_{HB}^{total} is the total number of hydrogen bond formed by the protein with the solution species. n_{HB}^{P-P} refers to the average number of intra-protein hydrogen bonds.

■ SUMMARY AND CONCLUSIONS

In this study we have used REMD sampling technique to elucidate the role of different osmolytes on the protein Trp cage and to find out the molecular level insights of the counteraction mechanism. Structural properties of the protein are revealed by analyzing RMSF, DSSP and radius of gyration plots. Melting curves are also explored to study folding propensity in different chemical environment. The overall conclusion that can be drawn from these observations is that the most and the least native form of the protein is retained in pure water and in binary urea solution respectively. Ternary mixed solution with 5M TMAO provides counteraction against action of urea to large extent so that the protein attains native like conformation largely. The plot of free energy landscapes (FELs) also supports these observations. FELs are plotted in order to depict the conformational space visited by the protein covering the whole temperature range (viz. ∼280K to ∼450K). The ruggedness of the surfaces is found to increase with temperature irrespective of all the systems and it indicates thermal induced unfolding of the protein. Highly rugged surface of binary urea solution is an indicative of urea induced unfolding that involves large number of intermediate states. The ruggedness of the surfaces are minimized to some extent in presence of TMAO revealing counteraction provided by TMAO.

Role of osmolytes in stabilization of salt bridge of Trp cage reveals that in binary urea solution, urea molecules are more prone to stay nearby the salt bridge pair which in turn helps in disruption of the salt bridge. The presence of TMAO molecules in ternary mixed

osmolyte solutions makes less availability of urea molecules around the salt bridge and thus the deleterious effect of urea is reduced. Focusing on the hydrophobic core we have found that the center of the core becomes heavily hydrated in binary urea solution. TMAO prevents to some extent the hydration of the hydrophobic core of Trp cage in ternary solutions. Thus it can be concluded that TMAO-induced stabilization of the protein Trp cage stems from both the protection of the salt bridge as well as the hydrophobic core. However, we do not observe any direct interaction of TMAO with the protein. This leads us to conclude that TMAO-induced stabilization is effective by means of indirect interaction. The investigations of orientations of the indole plane of Trp6 with respect to other aromatic planes of the protein reveal that the specific orientations of these planes provide stability to the protein. In ternary TMAO solutions, we have not found near parallel orientation of Trp6-Pro12 as an essential factor for protein stabilization, whereas near parallel orientation of Trp6-Pro18 and near perpendicular orientation of Trp6-Pro17 and Trp6-Pro19 are found to be significant contributors in stabilizing the native conformation of the protein. The calculation of total protein-urea interaction energy and the estimation of different hydrogen bond numbers indicate direct interaction between protein and urea plays a profound role. In ternary solution, TMAO replaces some urea molecules from the protein surface. More the number of TMAO molecules, more number of urea molecules are expelled from the protein surface. As a result, both electrostatic and vdW components of protein-urea interaction energy become more unfavorable in presence of TMAO (than that of binary urea system).

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Chapter 4

Effect of Non Polar Confinement on Protein Trp Cage Conformation in Aqueous Osmolyte Solutions

"In vivo, protein dynamics occur in the context of the crowded cellular milieu and in confined spaces such as the chaperonin cavity, the proteosome, the ribosome exit tunnel, the translocon, etc. When considering these factors it is reasonable to assume that proteins may experience different energy landscapes when folding in vivo than in bulk, and these differences may constitute a significant piece of the folding puzzle."

– D. Lucent, V. Vishal, V. S. Pande *Proc. Natl. Acad. Sci.* **104**, 10430 (2007)

Chaperonins assist protein folding by providing a cavity in which the newly synthesized protein can be encapsulated. Here, we investigate the effect of non polar confinement in chaperonin mediated protein folding through simulation technique. We have used REMD simulation technique to achieve efficient sampling. It helps in escaping a local minimum by exchange with a higher temperature replica. To elucidate the role of osmolytes urea and TMAO in presence of non polar confinement on the protein Trp cage is the main objective of this study. Different analyses are implemented in order to explain the effect of confinement and its consequent effect on the osmolytes in determining the folding/unfolding equilibrium of the protein. The study reveals that urea molecules exert their action on the protein more effectively beyond the presence of confinement, which is prominent from analysis of the structural properties of the protein. The protein also experiences counteracting effect of TMAO sufficiently so that native like behavior is preserved in mixed osmolyte solution.

We have enunciated on the role of salt bridge, hydrophobic core and orientational preferences of different planes of amino acid side chains in determining overall stability to the protein Trp cage. As we have already discussed in **Chapter 3** that salt bridge is not a necessary factor in determining stability of the protein. Researchers have reported stable protein conformations without maintaining the salt bridge. However, we have tried to highlight the role of salt bridge in presence of non polar confinement. Surrounding the central tryptophan residue, Trp cage possesses a hydrophobic core. We have examined the existence of the core and its influence in pertaining stability to the protein. Orientational preferences of different planes of the side chains of the protein residues with respect to central indole plane of tryptophan residue are also investigated in order to examine their role in protein stabilization. We have also analyzed average number of hydrogen bond among different species and protein-urea interaction energy in presence and absence of TMAO. These results support us to formulate the role of confinement and osmolytes in determining folding/unfolding equilibrium of the protein. Further, comparison of the results of this chapter with **Chapter 3** (i.e. without confinement) helps us to put a comparative view of the role of osmolytes in presence and absence of confinement on protein folding equilibrium.

Use of REMD to explore the effects of osmolyte solutions and non polar confinement on the protein Trp cage

Overview: In this article we demonstrate the effects of confinement and osmolytes urea and trimethylamine N-oxide (TMAO) on the conformational behavior of the protein Trp cage using replica exchange molecular dynamics (REMD) simulation. We have used a near spherical non polar confinement consisting of 2940 carbon atoms in order to mimic the role of chaperonin unit in protein folding. Encapsulated binary solution of urea exerts denaturing effect on the protein due to favorable direct interaction with the protein. Hydrophobic core of the protein is completely disrupted in presence of binary urea solution. In ternary solution, presence of TMAO makes protein-urea electrostatic interaction more unfavorable when compared to that for binary urea solution. Encapsulated ternary solution of urea and TMAO provides better protection to the protein conformation than that of without confinement. Confinement reduces the conformational space of the protein which is more pronounced for ternary mixed solution. We have observed that free energy landscapes of the protein are significantly modified when confinement is used in comparison to those without confinement.

■ INTRODUCTION

Some proteins do not attain their native states spontaneously when newly synthesized. For those proteins external assistance is required to achieve the folded states [56, 57]. Inside the cell, the process of folding is aided by molecular chaperones, that assist folding by preventing aggregation and misfolding, stabilizing the nascent polypeptides and initiate the folding process [55]. The GroEL/GroES chaperonin system of E. coli has been studied most extensively [55-57, 139]. Proteins are stabilized by many long-range interactions and tend to populate kinetically trapped folding intermediates [58]. Brinker et al. [62] suggested that GroEL/GroES chaperonin system induces a confinement effect on the protein that smoothen up the energy landscape of the protein by preventing kinetically trapped intermediates. In literature many computational and experimental studies have been reported to study protein folding/unfolding behavior under confinement in order to understand the functioning of the chaperone unit [63-68, 140-142]. These studies shed light on exploring the mechanism through which the chaperonin cage plays a passive role (by preventing the protein from aggregation), or an active role (by accelerating the rate of folding). Recent computational studies consider the effects of confinement on protein folding and stability in the context of chaperons, pores, and other cellular environments [63, 66, 67, 142]. The common observations from most of these studies are that confinement stabilizes the folded state of proteins, confined solvent also plays a crucial rule and the effects of confinement are reflected in the reduction of conformational entropy associated with protein folding, which in turn leads to the stabilization of the folded states.

The effect of denaturing and counteracting osmolyte on the protein has been a popular topic since few decades to explore the structural and dynamical information. Osmolytes show profound effect on protein solubility and stability [34, 112, 113]. In nature, osmolytes are used to overcome the denaturation process that resulted from environmental stress. For example trimethylamine-N-oxide (TMAO) accumulates in marine organisms to offset the denaturing effect of urea [109, 111]. In this work we have studied the effect of denaturing osmolyte i.e. urea and counteraction provided by TMAO molecules inside a non polar confinement. Urea is extensively studied through theoretical and experimental methods. Researchers suggest that "direct" and/or "indirect" interaction mechanisms are operative in urea-induced protein denaturation [77, 79, 114, 143, 144]. Preferential "direct" interaction of urea with the protein backbone and side chain is considered as the direct interaction mechanism. The indirect mechanism proposes that water structure is altered

due to presence of urea which in turn affects the hydration of the protein. Though the counteracting ability of TMAO molecules (against the action of urea) has been studied quite extensively but we are far from a definitive answer which can successfully explain the mode of action of TMAO on protein stabilization. For example, the enhancement of water structure in presence of TMAO is proposed to be a factor that stabilizes the protein [36, 37]. On the other hand, some studies emphasized on the direct interaction between protein and TMAO [36, 38, 39]. Bolen and co-workers [43] proposed the osmophobic interaction between protein backbone and TMAO. According to Thirumalai and co-workers [40] TMAO acts as a crowding agent that favors native conformation of the protein through excluded volume effect. Graziano proposed that a decrease in the solvent-excluded volume upon folding is responsible for TMAO stabilization [145]. Pettit and co-workers [44] coined about favorable van der Waals interactions that contribute in protein stabilization. Garcia and co-workers [33] reported that TMAO molecules are preferentially excluded from the protein surface.

Herein we have carried out Replica Exchange Molecular Dynamics (REMD) [21] simulation of Trp cage inside a non-polar confinement in pure water and in presence of osmolytes urea and TMAO. Effect of confinement and osmolytes on the protein is investigated thoroughly. The protein Trp cage has been widely acknowledged as a fast folding protein. It has a well packed hydrophobic core and typical tertiary contacts that makes it a highly stable paradigm in the field of protein research [122, 124]. Highly stable nature makes it an ideal target to study through both experimental and computational methods [15-18, 63]. We have chosen REMD simulation technique since enhanced conformational sampling is preferred for protein simulation over regular MD simulation because REMD ensures efficient sampling of the system by crossing the energetic barriers. The method is also useful in order to interpret the free energy landscape of the protein over a wide range of temperature. The details of the simulation techniques utilized in this study are described in Methodology section. In section Results and Discussions, the results are discussed. Then conclusion of the work is summarized in section Summary and Conclusions.

■ MODELS AND SIMULATION METHOD

We have performed REMD simulations of the aqueous protein solutions Trp cage (PDB ID 1L2Y [18]) inside a non polar near-spherical confinement. Three different systems were constructed with the protein comprising of pure water, a binary solution of urea (8M) and a ternary solution of mixed osmolytes urea and TMAO (8M:4M). The initial configurations of all the systems were prepared using Packmol [89]. To generate the non polar spherical

confinement, we have used CaGe software [146]. In this study, a fullerene-like ball (nearly spherical in shape) consisting of 2940 carbon atoms was used that served the purpose of a spherical confinement. All the carbon atoms of the ball were uncharged. Its parameters and topology files were generated according to GAFF force field [147]. To reduce the interaction between the different atomic sites of the protein and the fullerene ball, we have decreased the Lennard-Jones parameter ϵ of the carbon atoms of fullerene to half of its original value [148]. One chloride ion was added to each of the system to neutralize them. All systems were simulated using AMBER12 program [70] (see Table 4-1). Amber FF12SB [126] force field parameters were used for the protein. TIP3P model [125] of water was used for all the systems. Lennard-Jones parameters and corresponding partial charges of different atomic sites of TIP3P water [125], urea model [87] and TMAO model (Kast) [88] used for this study are obtained from their respective citations. We have chosen Kast model of TMAO over Osmotic model [33] of TMAO since the results of our previous study [53] demonstrated that Kast model of TMAO provides better protection against urea for the said ratio of urea:TMAO.

The Packmol generated systems were then energy minimized through 2500 steps of steepest descent method, and then heated gradually from 0 to 300K with a temperature jump of 50K (for 5 ps each) in canonical (NVT) ensemble. Thereafter a short period of 300 ps and 400 ps simulation run were carried out in NPT (isothermal-isobaric ensemble) and NVT ensemble respectively to make the systems homogenized sufficiently. With the resulting systems obtained from above steps, 60 ns REMD simulations were carried out for each of the systems. The temperature was controlled by Langevin thermostat [91]. All bonds involving hydrogen atoms were constrained by SHAKE algorithm [90]. A cutoff distance of 9 Å was used for short ranged van der Waals interactions. The Particle Mesh Ewald (PME) [93] method was employed to treat the coulombic interactions. Periodic boundary conditions were applied in all three directions. A time integration step of 2.0 fs was used.

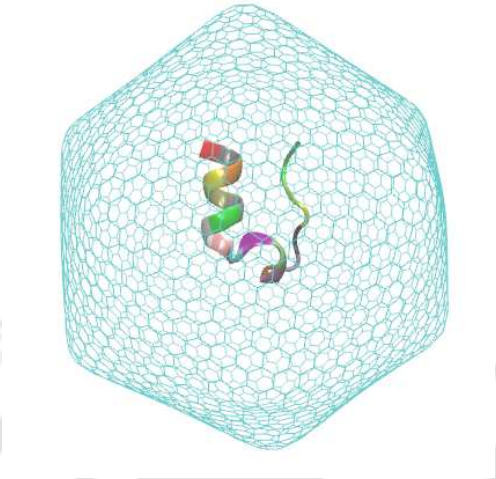


Figure 4-1. *NMR structure of Trp cage with residue-wise color identity inside the non polar spherical confinement.*

Table 4-1. Overview of Systems[a]

System	N_W	N_U	N_T	*No.of Replicas*	*Temperature* (K)	*Density* (g/cc)	*Box Length*(Å)
PW	1260	0	0	30	285-455	0.82	49.63
PWU	810	189	0	30	285-456	0.85	49.63
PWUT	800	190	90	32	284-449	0.95	49.65

[a] N_W, N_U and N_T represent the number of water, urea and TMAO molecules respectively. ρ is the density of the system. P, W, U, and T refer to protein, water, urea, and TMAO respectively.

To set up the temperature distributions for REMD simulations we have chosen a temperature range of 280 to 460K. Depending on the system sizes, 30 to 32 replicas were used for this study (as mentioned in Table 4-1). The temperatures were chosen such that the acceptance ratios of replica exchange are sufficiently large ($\sim 20\%$). Replica exchange probability is calculated according to the Metropolis equation (discussed in methodology of previous Chapter). With the progress of simulation replicas are exchanged and at the end

of simulation temperature based trajectories were sorted out for further analysis. Analyses were performed by using ptraj and cpptraj [128] program (available with AMBER) and VMD software [94] was used as and when needed.

Here we would like to mention that the experimental melting temperature of the mini protein Trp cage is 315K [149, 150] which is markedly different from those found in different computational studies [63, 151, 152]. Moreover, there is no consistency between the reported melting temperatures of Trp cage mini protein, and it varies from 360K [151] to around 450K [152] depending on the force field parameters used in different studies.

■ RESULTS AND DISCUSSION

Root Mean Square Fluctuations (RMSF) and Secondary Structure

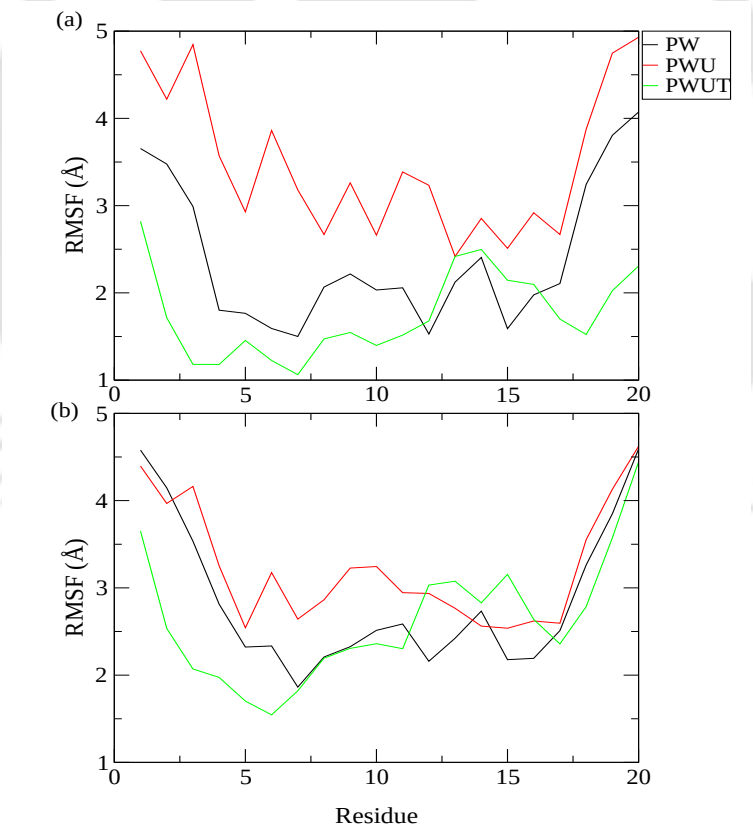


Figure 4-2. *Root mean square fluctuations (RMSF) of C_α-atoms of each of the residues of the protein at: (a) low temperature and (b) high temperature.*

The magnitude of Root Mean Square Fluctuations (RMSF) of each C_α-atom of all the residues of the protein are evaluated for different systems (see Figure 4-2). RMSF indicates the extent of structural changes from the native state of the protein. The protein in binary urea solution exhibits higher RMSFs than that in pure water and ternary mixed solution for both low ($\sim$300K) and high ($\leq$456K) temperature limits. Residues with high RMSF values indicate more degree of fluctuation from the native conformation. Thus, at high temperature we observe higher RMSF values than that at low temperature. However, in this study RMSF values are found to be restricted to some extent due to the presence of confinement. In our previous study of Trp cage in binary urea solution without confinement, (**Chapter 3**) [153] the RMSF values reach upto 10 Å even at low temperature (unlike 5 Å in this study). Thus, it can be stated that the effect of confinement is highly pronounced for all the systems. The confinement provides limited space to the protein that influences the flexibility of the protein and results in low RMSF values than that of without confinement. Low RMSF values in pure water and ternary mixed solution indicate that the protein exhibits less flexibility i.e. more stable behaviors in these systems (when compared to that for binary urea system).

For all the systems we have also examined residue based secondary structure analysis of the protein as a function of time at 300K temperature (shown in Figure 4-3). For this, protocols according to DSSP algorithm (Dictionary of Secondary Structure of Proteins) of W. Kabsch and C. Sander [129] are followed. The reference secondary structure states of different residues of the protein are derived from the definitions produced by DSSP [130]. Both in pure water as well as in mixed osmolyte solution the helical structure of the protein is maintained throughout the entire simulation run with occasional local deviations from α-helix to 3_{10}-helix and turn. But in binary urea solution, the helical structure is noticeably lost and the appearance of coil and turn conformations take place. It is noteworthy that 3_{10}-helix of the protein is highly preserved only in pure water. Urea induced unfolding (Figure 4-3 (b)) and TMAO induced counteraction (Figure 4-3 (c)) of the protein can be clearly figured out from these observations. It can also be depicted that the protein maintains maximum native like conformation in pure water system.

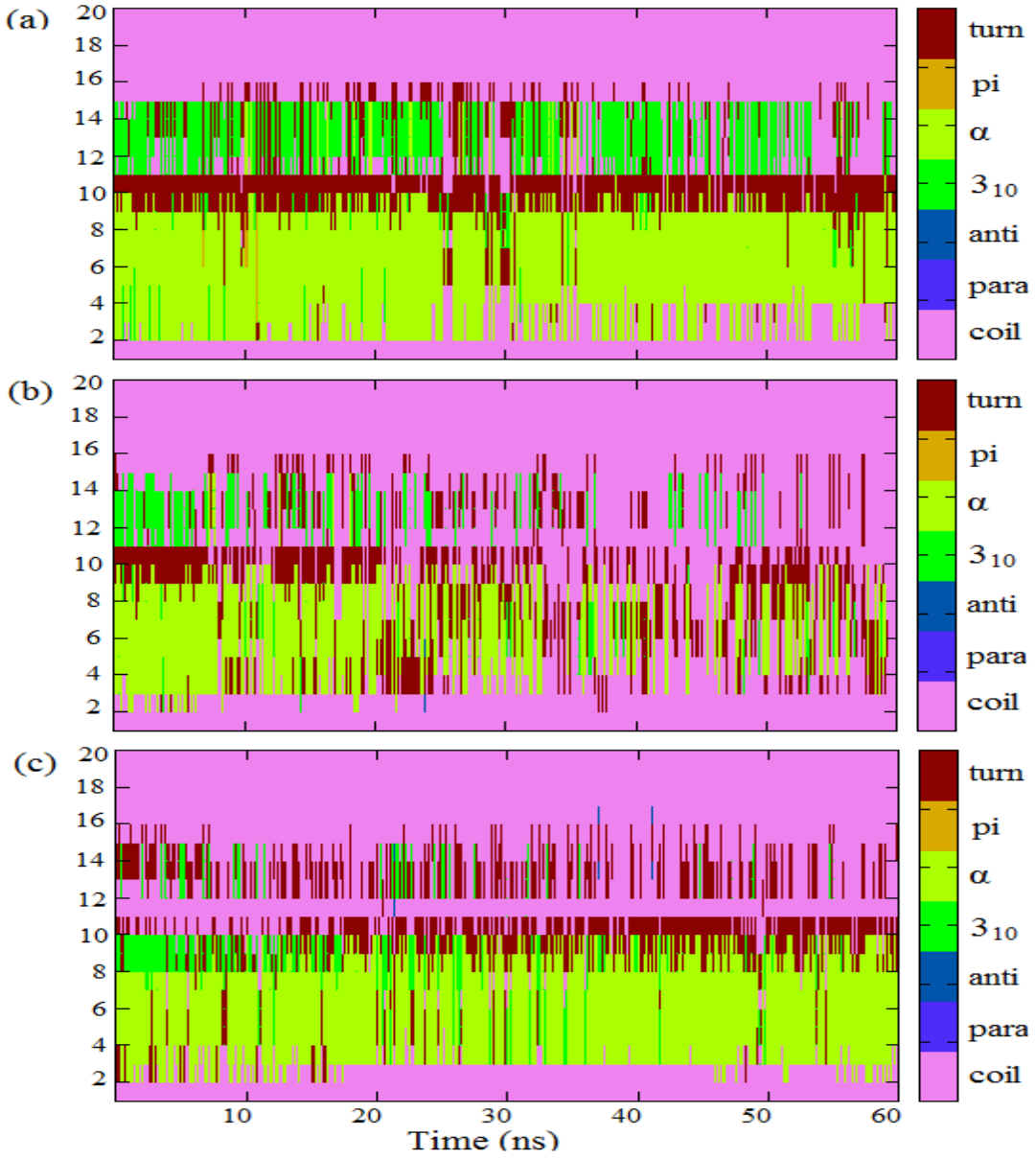


Figure 4-3. *DSSP plots of the protein for (a) PW, (b) PWU and (c) PWUT systems as a function of simulation time.*

Radius of Gyration

We have plotted the probability distributions of radius of gyration (R_g) of the protein at the two temperature limits for all the systems (see Figure 4-4). The calculations of R_g provide us information about the effective size of the protein. We observe that the protein attains the most compact form in pure water system which is followed by that

in ternary mixed solution. For binary urea solution the distribution curve is found to be expanded over a wide range of R_g values indicating the presence of unfolded conformations of the protein in that solution. A comparison of the distribution curves at low and high temperature reveals that the protein acquires more extended conformation at high temperature than that at low temperature for all the systems, which is as per the expectations. However, because of the confinement induced limited conformational space, the R_g values do not exceed beyond 14 Å. In contrast, our previous simulation study [153] of Trp cage in binary solution of urea (without confinement **Chapter 3**) showed the appearance of finite probability of R_g values upto 18 Å.

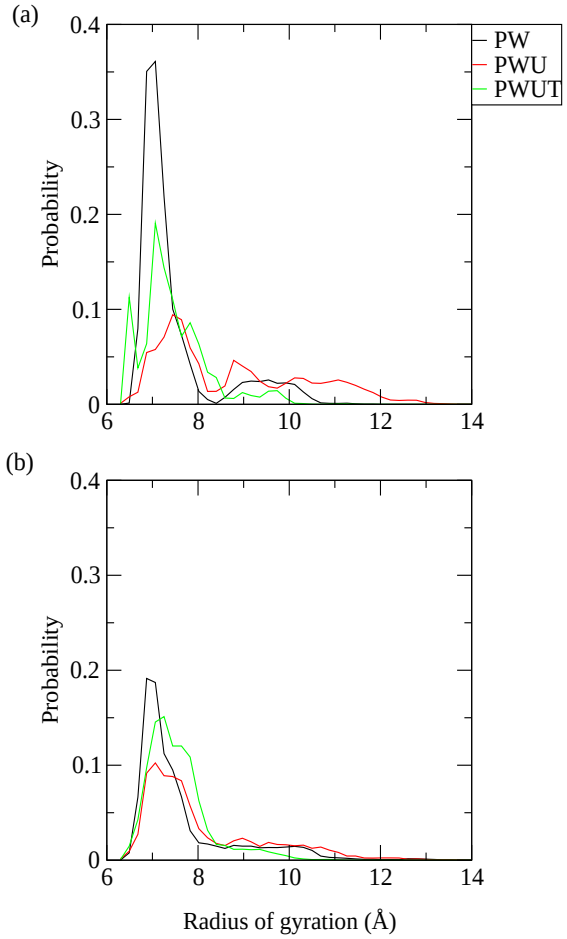


Figure 4-4. *Probability distributions of radius of gyration of the protein at (a) low and (b) high temperature limits.*

Melting Curves and Free Energy Landscapes (FELs)

In order to examine the temperature dependent conformational changes of the protein Trp cage, melting curves are plotted when it (the protein) is surrounded by different chemical environments (see Figure 4-5). It is observed that the folded fractions decrease with the rise of temperature for all the systems. Maximum folded fraction is retained for pure water system, whereas for binary urea system retention of folded fraction is minimum. In presence of mixed osmolytes, a significant retention of folded fraction is observed. This implies that the counteracting behavior of TMAO against the deleterious action of urea is successful in presence of confinement. When we compare these melting curves with those without confinement [153], it becomes quite clear that the counteracting ability of TMAO is more pronounced in the former than that of the latter. Figure 4-5 (b) represents free energy of unfolding (ΔG_u) of the protein for respective temperatures (assuming two-state unfolding equilibrium) [154]. ΔG_u is defined as follows:

$$\Delta G_u = -RT[(1 - x_{folded})/x_{folded}] \tag{4.1}$$

x_{folded} is the fraction of folded state, R is the molar gas constant and T is the replica temperature. For binary urea solution ΔG_u is found to be negative, whereas for pure water and ternary mixed solution ΔG_u shows positive values. It indicates that the folding propensity of the protein is higher in pure water and mixed osmolyte solution than that for binary urea solution.

The free energy landscape (FEL) analysis is an efficient method to evaluate stable/unstable behavior of the protein [131, 132]. Different conformations corresponding to different basins of FEL provide insights into the structural transition pathway followed by the studied protein. In this study, we have used RMSDs (root mean square deviations) and temperature of each replica to evaluate the FEL for all the systems (see Figure 4-6). Probability distributions of the RMSDs are calculated first and then free energy is calculated according to following equation:

$$\Delta G(V) = -k_B T[lnP(V) - lnP_{max}] \tag{4.2}$$

where k_B denotes the Boltzmann constant, T is the temperature of respective replica. P(V) represents the probability distribution of RMSDs which is subtracted from its maximum value i.e. P_{max}, so that we get zero value of ΔG for the free energy minimum.

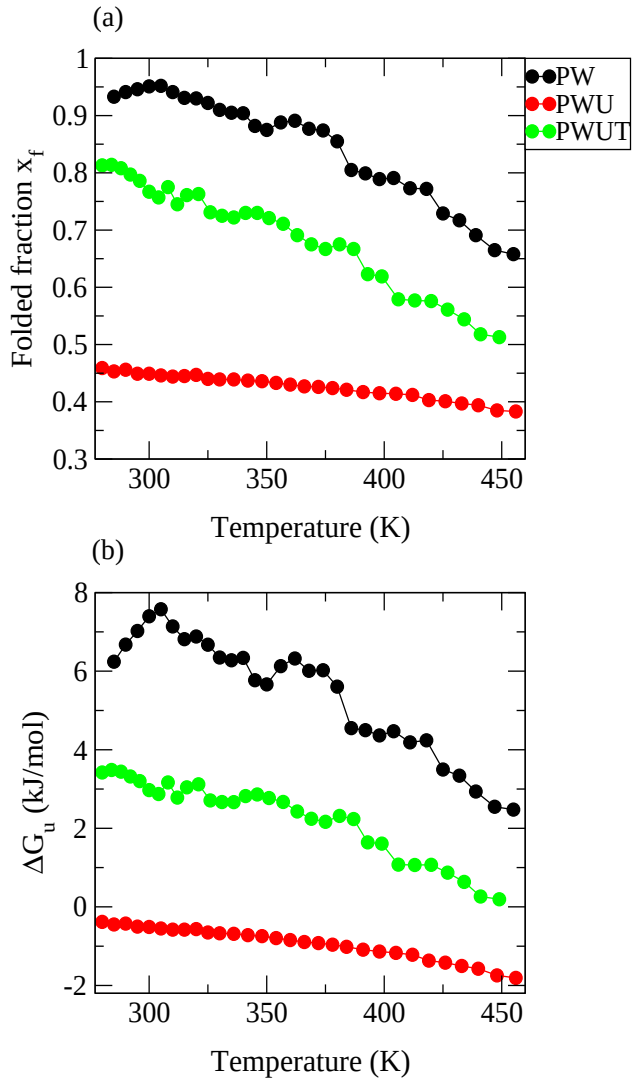


Figure 4-5. *(a) Folded fractions of the protein versus temperature and (b) corresponding free energy of unfolding of the protein (for each replica temperature).*

As presented in Figure 4-6 (a), in pure water the native conformations (RMSDs $\leq$ 1.5 Å) with the lowest energy state appear as the most dominant one over the entire range of temperatures. However, the overall conformational space is found to be extremely narrow in presence of mixture of osmolytes (Figure 4-6 (c)). This explains the combinatorial role of TMAO counteraction and the effect of confinement in delivering stability to the protein. The highly rugged energy surface for binary urea solution depicts that the maximum number of unfolded conformations of the protein are achieved for this case (Figure 4-6 (b)).

Small free energy basins that appear at the range of 3-5 Å do not contribute significantly towards native conformations (because of large RMSDs). In comparison to our previous REMD study on Trp cage without confinement [153], the FEL of the protein is found to be smoothen up in presence of confinement for the mixed osmolyte solution.

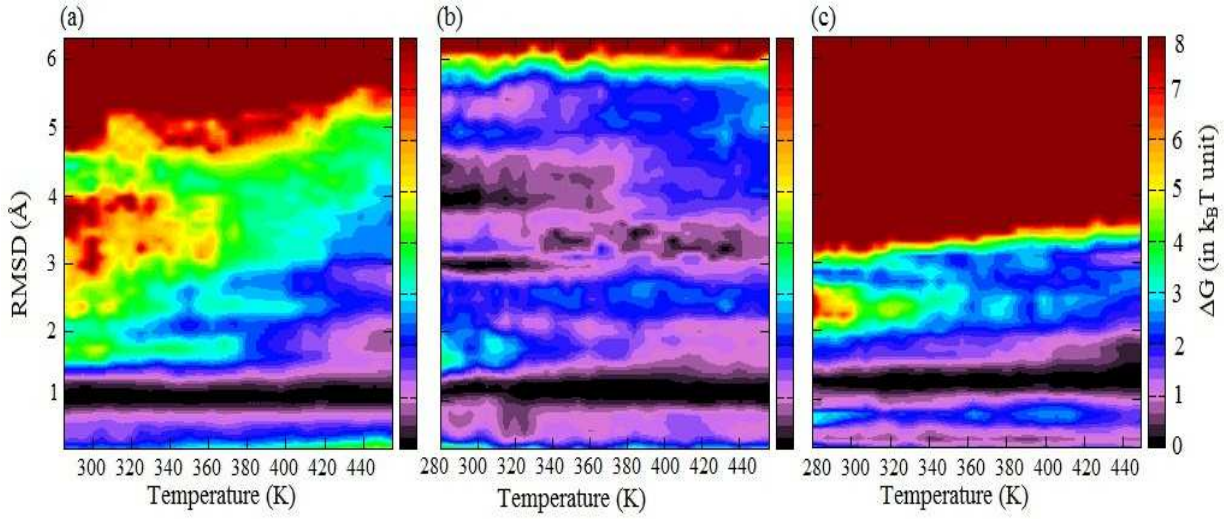


Figure 4-6. *Free energy landscapes (FELs) for systems (a) PW, (b) PWU, and (c) PWUT.*

Dihedral Angle Principal Component Analysis (dPCA)

To further characterize the conformational changes of Trp cage, dihedral angle principal component analysis (dPCA) [98] is performed. We have used the backbone dihedral angles (ϕ, ψ) of C_α atoms of the protein backbone. 'Carma: a molecular dynamics analysis program' [99] is used to construct dPCA. This method employs calculation of the covariance matrix σ_{ij}, which provides information on the two point correlations of the system and is defined as:

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \tag{4.3}$$

where x_1,...,x_{2N} are the sine- and cosine-transformed dihedral angles of the protein backbone (having N angles) and $\langle ... \rangle$ denotes the average over all sampled conformations. Diagonalization of the matrix σ_{ij} removes the instantaneous linear correlations among the variables and produces into 2N number of eigenvectors, v(i), and eigenvalues, λ_i, such that λ_1 represents the largest eigenvalue. The principal components $V_i = v(i)x$ are used to construct the free energy surface following above mentioned equation 3. In this case P(V) is the

probability distribution of the molecular system along the principal components V_i. In this study the first two principal components are used to represent the free energy surface.

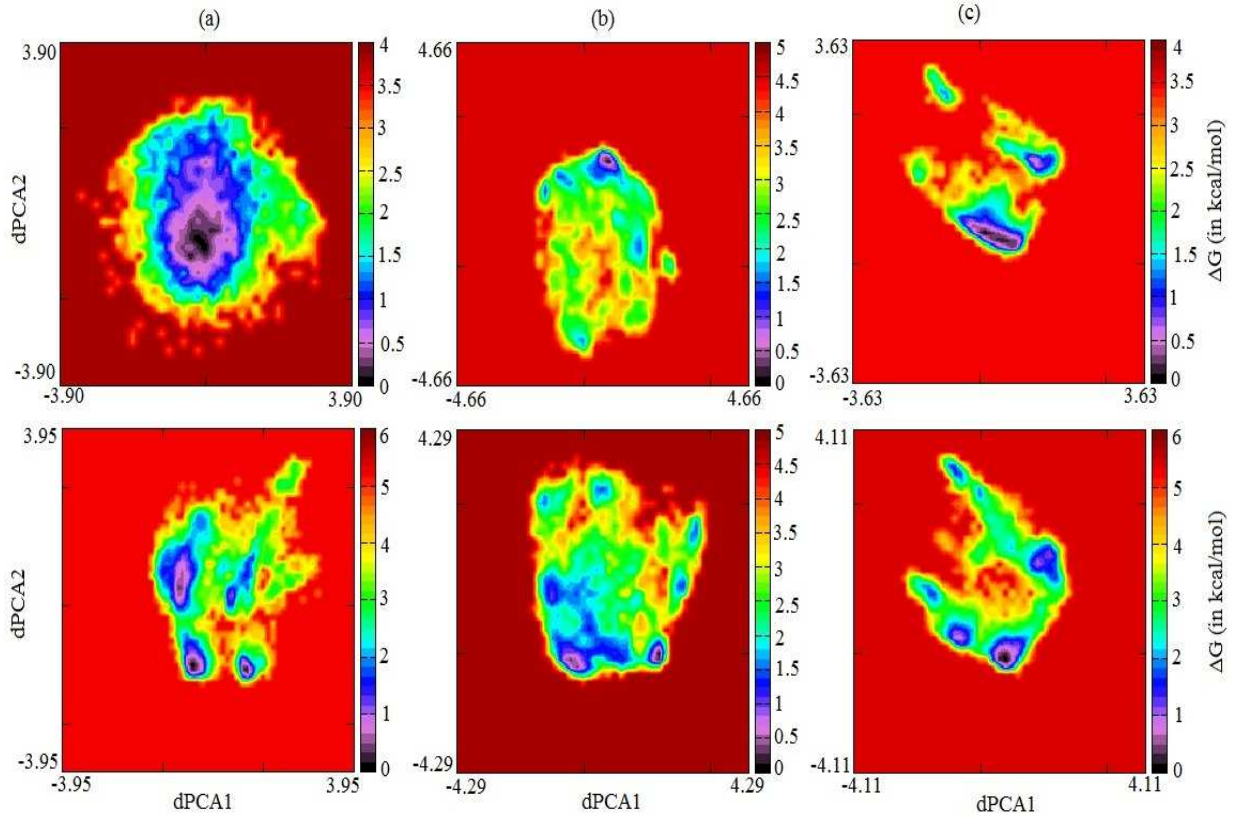


Figure 4-7. *Plot of dihedral angle principal component analysis (dPCA) for systems (a) PW, (b) PWU and (c) PWUT. The top three plots and bottom three plots represent dPCA at low temperature and at high temperature respectively.*

In Figure 4-7, we present dPCA for each of the systems at two temperature limits. At low temperature, the protein shows the formation of a single major cluster of native like conformation in pure water system. On the other hand, at high temperature limit, three different conformational clusters are observed (two of them appear with a small energy barrier between them). Figure 4-7 (b) depicts that the appearance of the native conformation diminishes in presence of urea for both the temperature limits. Moreover, at high temperature, large number of conformational clusters are observed. It implies that urea causes more denaturation of the protein at high temperature. For ternary solutions (Figure 4-7 (c)), the appearance of one major and one minor cluster that appear at low temperature suggesting that the effect of urea is counteracted by TMAO to large extent.

Density Change Across the Fullerene Ball

To visualize the effect of confinement we have calculated the density of the protein Trp cage (C_α atoms considered) and water for all the systems (see Figure 4-8 (a)) from the center of mass (COM) of the fullerene ball towards its surface. It is noticed that for all the systems density change of the protein (solid line) and water (line with diamond symbol) differs depending on the surrounding environments. In binary urea solution the protein is more prone to occupy peripheral region of the confinement rather than the central region. Conversely in mixed osmolyte solution the protein shows least tendency to occupy the peripheral region of the confinement. For pure water system the protein is found in the region 10-22 Å from the center of the fullerene ball and in mixed osmolyte solution this optimal range is found to be 5-21 Å. Density profiles of the confined water in different systems show that in pure water there is a sharp decrease of density in the region close to the confinement surface. This observation is in consistent with previously reported works that confirm about dewetting nature of water in the vicinity of the hydrophobic surfaces [63, 155]. Density profiles of water for binary urea solution and mixed ternary solution are very similar. Water densities are found to be gradually decreasing from the center to the surface of the fullerene ball. The calculations of probability distributions of COM distances between the protein and the fullerene ball also confirm above observations (see Figure 4-8 (b)). It can be seen that the protein maintains the least distance from the COM of the fullerene in ternary mixed osmolyte solutions and in binary urea solution the COM distance is maximum. It points about the possibility of the protein to interact with the fullerene ball in binary urea solution. It may also contribute to the decreased folding behavior of the protein in binary urea solution.

Figure 4-9 (a) represents the probability distribution of the protein C_α atoms within a distance of 4.5 Å from the surface of the non polar confinement. It shows that different residues of the protein interact differently with the fullerene surface. In binary urea solution all the residues of the protein show very high probability to occupy the region near the surface. In pure water system only residues 8, 10, 15, and 16 show exceptional high probability to occupy the region near the surface. In mixed ternary solution we observe that residues 2 to 8 (i.e. the α-helix) show least probabilities to occupy the region within 4.5 Å from the surface of fullerene. These observations are in accordance with the density profiles of the protein in different solutions (Figure 4-8 (a)).

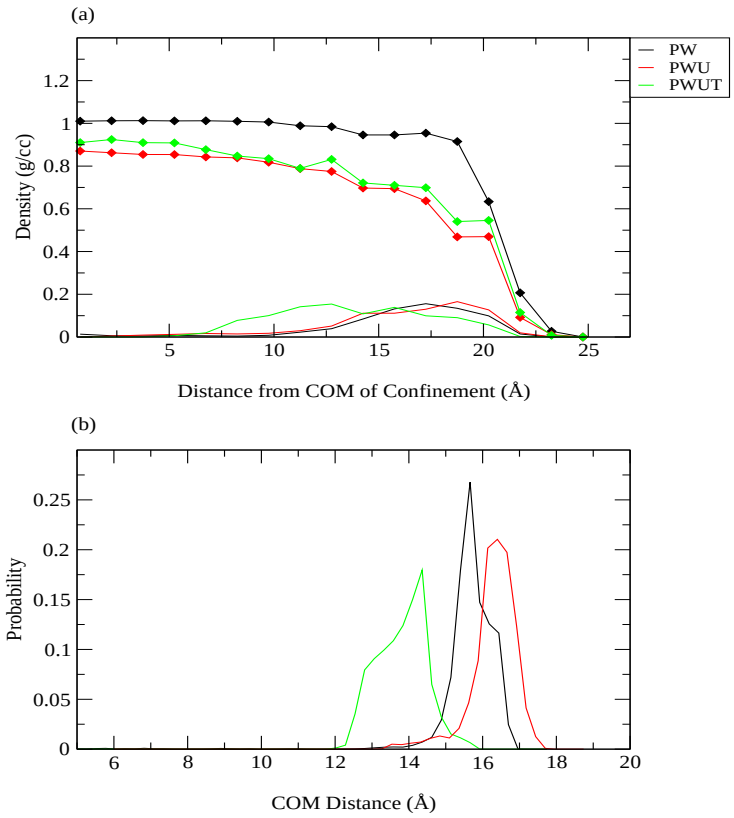


Figure 4-8. *(a) Density of water (line with diamond symbol) and protein (solid line) as a function of distance from COM of the spherical confinement. (b) represents probability distributions of COM distances of the protein as a function of distance from the center of the spherical confinement for different systems.*

Further, we have plotted radial distribution functions (rdfs) involving central atoms of urea and TMAO and the fullerene surface (see Figure 4-9 (b)). It shows that central atom of urea (for both binary and mixed osmolyte solution) appears at position nearer to the fullerene surface than the central atom of TMAO. We also observe that the rdf of urea carbon in mixed osmolyte solution shows enhanced peak height than that in binary urea solution. These imply that in binary urea solution, the urea molecules show less tendency to solvate fullerene surface (than that in ternary solution), thus they get more access to solvate the protein. Hence, the deleterious effect of urea on the protein is highly manifested. But in case of mixed ternary solution the scenario changes due to presence of TMAO molecules. TMAO molecules show a moderately high tendency to solvate the fullerene surface and they also solvate urea molecules (as is evident from hydrogen bond calculations). This

results into a situation where the protein experiences less exposure to urea molecules in ternary solution. Hence, the denaturing effect of urea is further diminished on the protein and the native conformation is preserved in mixed ternary solution. Thus, the confinement plays a significant role in protein stabilization by exerting its effect on the protein, water and osmolytes as well.

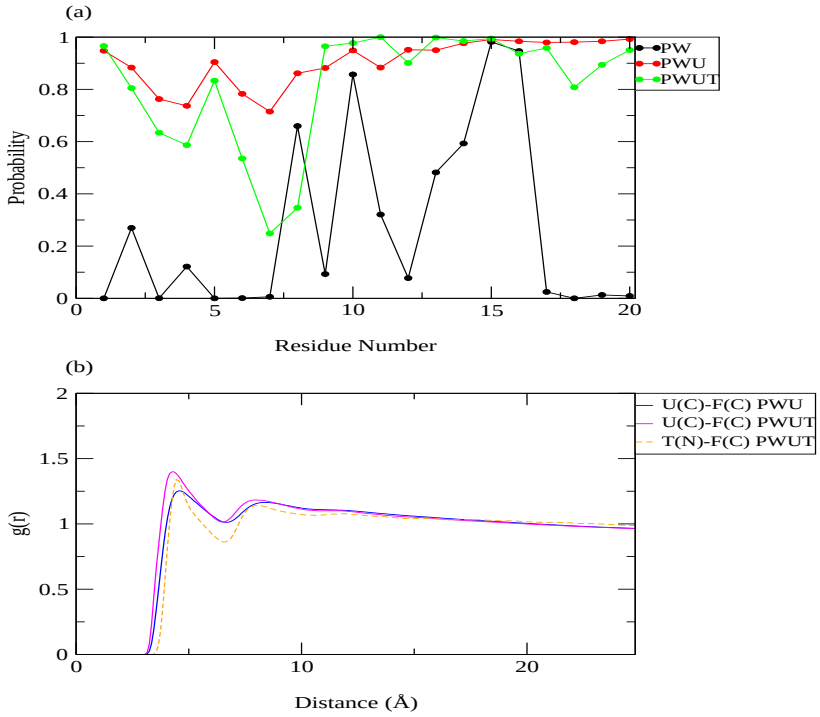


Figure 4-9. *(a) Probability of finding the C_α-atoms of the protein residues within 4.5 Å from the fullerene surface. (b) Rdf of central atoms of urea and TMAO as a function of distance from the fullerene surface.*

Role of Salt Bridge, Hydrophobic Core and Orientational Preferences of Aromatic Planes in Stabilization of Trp Cage

Previous studies identified three different factors that play significant role in stabilization of the protein Trp cage. These are: (i) salt bridge between residues Arg16 and Asp9 [63, 133], (ii) hydrophobic core constituted by residues Tyr3, Leu7, Gly11, Pro12, Pro18 and Pro19 surrounding the central residue Trp6 [15, 133], (iii) orientational preference of indole plane (of Trp6) with respect to other aromatic planes of the protein [153]. In Figure 4-10 (a), we have shown the probability distributions of the distance between the salt bridge forming residue pair Arg16 and Asp9 for the three different systems. The plot shows that

the salt bridge is rarely preserved in pure water and binary urea solution. Although in ternary mixed osmolyte solution the salt bridge exists to some extent. Thus the plot indicates that the existence of salt bridge is not a necessary factor for protein stabilization. This finding is in line with Garcia's work [63] which reported that Trp cage can retain its native form in presence as well as in absence of the salt bridge. Figure 4-10 (b) depicts the distance maintained by the hydrophobic core forming residues from the center of the core i.e. Trp6.

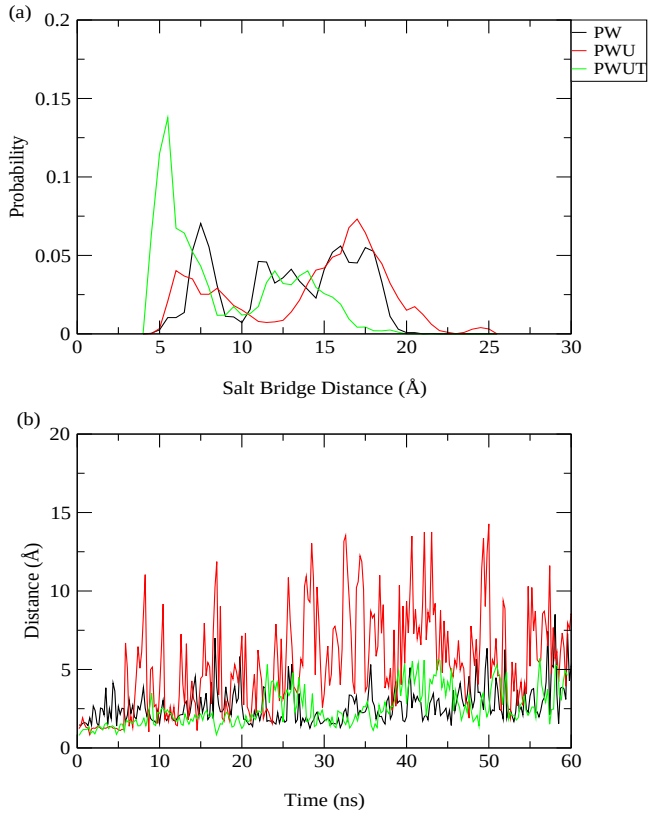


Figure 4-10. *(a) Probability distributions of the salt bridge distances for different systems. (b) Plot of distance of the hydrophobic core forming residues from the center of the core.*

We observe that for pure water and mixed osmolyte system, the hydrophobic core of the protein is properly maintained (distances are limited to 5 Å). But for binary urea solution, the hydrophobic core is disrupted (distances show fluctuations upto 15 Å). Maintenance of the hydrophobic core supports that native conformation of the protein is conserved, whereas disruption of the core indicates loss of the native conformation. Thus it leads us to conclude that maintenance of the hydrophobic core (in pure water and mixed

osmolyte solutions) plays a major role in stabilizing the protein.

For explaining the orientational preference of indole plane of Trp6 with respect to the other aromatic planes of Trp cage, we define the inter-plane angle as the angle between the vectors normal to the respective planes. Hence the angles close to 0° and 90° correspond to parallel and perpendicular orientations between the planes respectively. Normalized probability distributions of inter-plane orientational angles and COM distances are denoted by P(θ) and P(Å) respectively. In Figure 4-11, we have plotted P(θ) and P(Å) against inter-plane orientational angle and COM distances respectively. Peaks of the plots of P(θ) and P(Å) reveal the preferred orientational angle and COM distances adopted by the pair of planes. Discussions of previous sections indicate that the most native conformation of the protein Trp cage is retained in pure water, hence probability distributions for this system can be considered as the reference standard. A comparative view of P(θ) values of each of the residue pairs for the three systems reveal that the orientational preferences do not play any precise role in stabilizing the protein. We observe that a proper tilt angle of 25° for planes of Trp6-Tyr3 is maintained only for mixed osmolyte system, though COM distances between the planes of this residue pair maintain a distance within 10 Å for both pure water and mixed osmolyte solution. In our previous study of Trp cage without confinement [153], we have shown that for both pure water and mixed osmolyte solution tilt angle of 25° is maintained for the planes of Trp6-Tyr3. For residue pair Trp6-Pro12, the planes are found to maintain parallel orientations only in case of pure water system. For binary urea solution and mixed osmolyte solution, no preferred orientation is observed. This observation is in line with that obtained for without confinement [153]. Corresponding distributions of COM distances show that the planes move slightly more away in binary urea solution. For residues Trp6 and Pro17, near perpendicular orientations of the aromatic planes are observed for mixed osmolyte solution, but no sharp pattern is observed for pure water and binary urea solutions. Here we note that the orientational preference for protein-water system is dissimilar with that of without confinement study [153]. For the planes of residue pairs Trp6-Pro18 and Trp6-Pro19, we observe near parallel and near perpendicular orientations respectively for pure water and mixed osmolyte solutions (similar observations with those we have shown for without confinement [153]). Distributions of COM distances for each pair of planes depict that for binary urea solution they move to longer distances indicating unfolding of the protein. Thus, from these observations we propose that the orientational preference of different aromatic planes (with respect to indole plane of residue Trp6) has little role in Trp cage stabilization in non polar confinement.

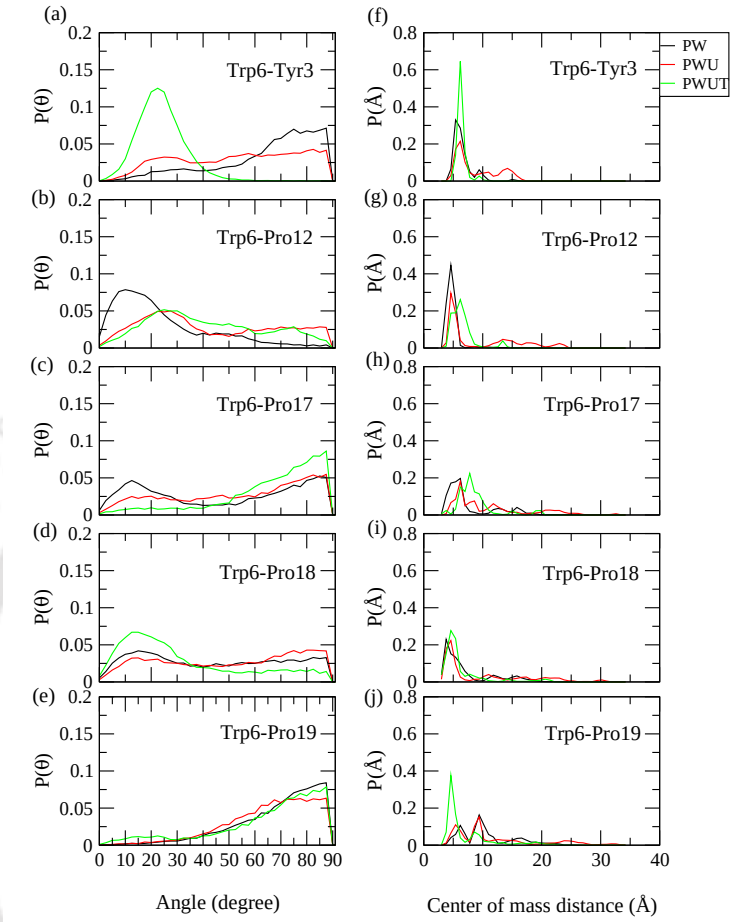


Figure 4-11. *Probability distributions of different orientational angles (left) and COM distances (right) of indole plane with other aromatic planes.*

In support of the above observations we have further calculated spatial distribution functions of urea and TMAO molecules surrounding the protein and the same are shown in Figure 4-11. It is quite apparent that the density of urea molecules in the nearby position of the protein in binary urea solution is much higher than that for the ternary solutions. In ternary solutions, the presence of TMAO molecules in the vicinity of the protein is also quite noticeable.

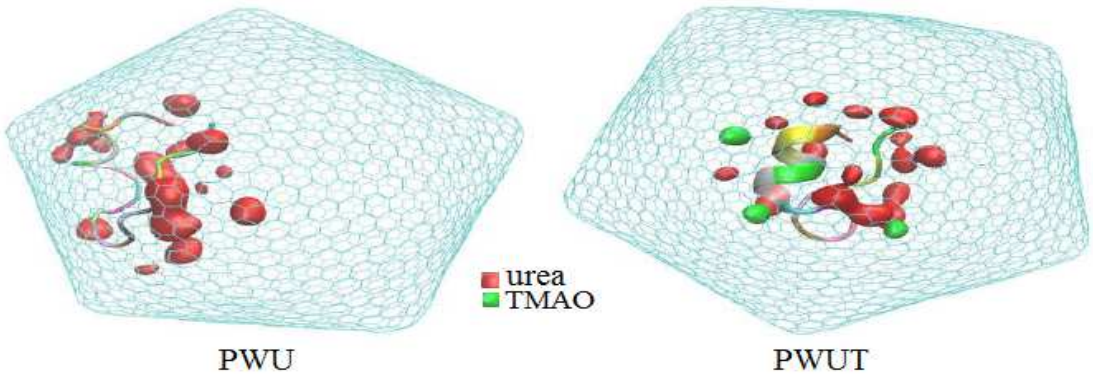


Figure 4-12. *Spatial distribution functions (sdfs) of urea (red) and TMAO (green) surrounding the protein residues of systems PWU and PWUT.*

Effect of TMAO on Protein-Urea Interactions

We have already discussed in **Chapter 1** that the denaturation action of urea is accomplished by either direct or indirect mechanism. In support of direct interaction mechanism two forms of interactions are recognized viz. electrostatic and van der Waals (vdW) interactions. Hydrogen bonds or other electrostatic interactions contribute to electrostatic interaction and vdW or other dispersion interactions contribute to vdW interactions. Literature reveals that any one of these two interactions can be dominant over the other and these interactions are not necessarily mutually exclusive. Some simulation studies support the evidence of direct interaction mechanism whereas some other support the vdW interaction mechanism [78, 79, 100, 101, 136, 137].

Similar to the study by Paul and Paul [100], we have also decomposed the protein-urea interaction energies in terms of electrostatic and vdW energy components and they are plotted as a function of time (see Figure 4-13 (a) and (b)). Figure 4-13 (a) shows that protein-urea electrostatic interactions are more favorable for binary urea solution than mixed ternary solution. Figure 4-13 (b) shows that vdW interaction energy components are similar for both the systems (slightly more favorable for binary urea solution) and less pronounced than the electrostatic component. Average values for the electrostatic, vdW and total energies (Figure 4-13 (c)) also support these observations. Presence of TMAO in mixed osmolyte solution, thus, results into decreased protein-urea interaction. Moreover, in presence of TMAO molecules, the electrostatic component of protein-urea interaction is affected more (i.e. more unfavorable) than that of vdW component. In this regard we note

that the previous study by Kokubo and co-workers [44] suggested about urea's favorable vdW interaction with unfolded protein than the folded one. They also found that in the mixed solution of urea and TMAO, vdW component is more favorable for both folded and unfolded protein, but electrostatic component is significantly unfavorable. This results into a higher free energy system where unfolded protein has significantly low solubility. Thus the folded protein gets stabilized in mixed solution.

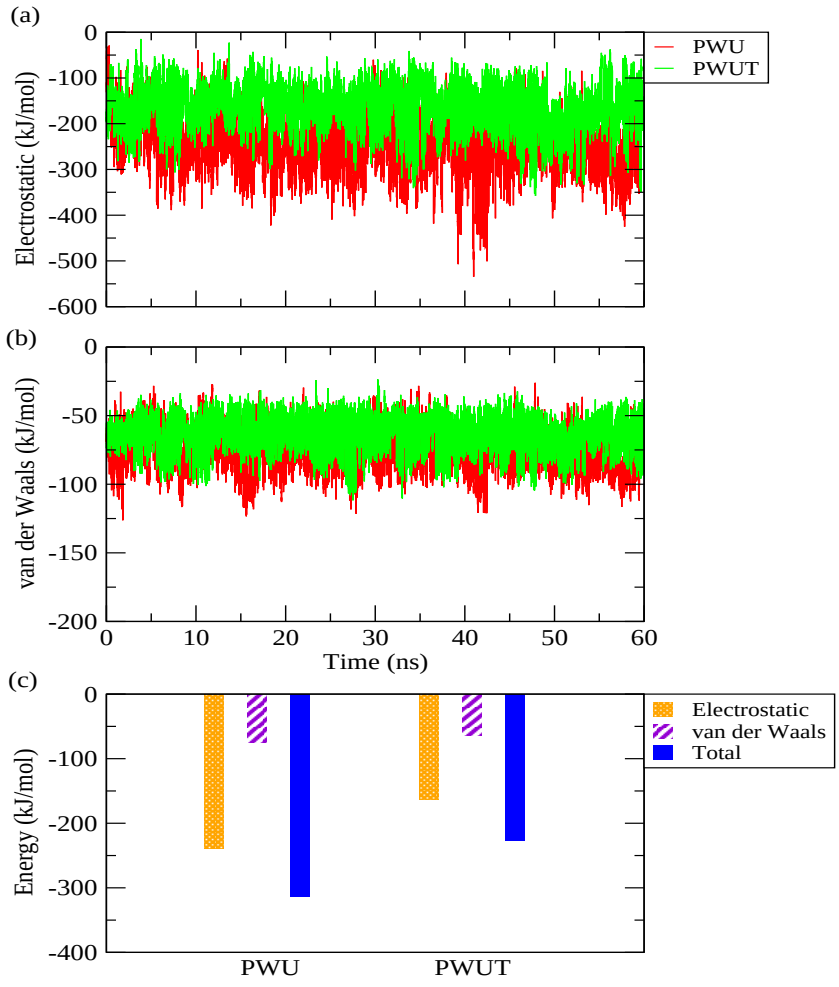


Figure 4-13. *(a) Electrostatic, and (b) van der Waals energies of protein-urea interaction as a function of simulation time. (c) Average values of electrostatic, van der Waals and total protein-urea interaction energies.*

Hydrogen Bond Properties

We have estimated the average number of intra- and inter-molecular hydrogen bonds among different species (see Table 4-2). We have used the geometric crieteria of donor-acceptor distance cutoff 3.5 Å and donor-hydrogen-acceptor angle cutoff 120° [103]. Considering the intra-protein hydrogen bonds first, we notice that in pure water system the average value is approximately 13. Addition of urea (i.e. in binary urea solution) causes the breaking of some of this type of hydrogen bonds which is regained to large extent in ternary mixed osmolyte osmolyte solution as its value reaches to 10.56. Retention of more number of intra-molecular hydrogen bonds of the protein ascertains the native conformation of the protein [138]. Thus counteraction of TMAO against denaturation action of urea is reflected from this observation. In binary urea solution large number of protein-urea hydrogen bonds are formed that are decreased in ternary mixed solution and a small number of protein-TMAO hydrogen bonds are also observed for this system. Altogether it can be stated that urea interacts through direct interaction with the protein in binary solution of urea, but in ternary mixed solution, protein-urea interactions are diminished as TMAO molecules occupy nearby positions of the protein.

Table 4-2. Average number of hydrogen bonds[a]

System	*Temperature*	n_{HB}^{W}	n_{HB}^{U}	n_{HB}^{T}	n_{HB}^{Total}	n_{HB}^{P-P}	$n_{\mathrm{HB}}^{urea-tmao}$
PW	$285K$	59.07	$--$	$--$	59.07	12.95	$--$
PWU	$285K$	42.75	27.73	$--$	70.48	8.57	$--$
PWUT	$284K$	47.63	19.61	2.72	69.96	10.56	0.46

[a] Average number of hydrogen bonds formed by the protein with water, urea and TMAO are presented by n_{HB}^{W}, n_{HB}^{U} and n_{HB}^{T} respectively. n_{HB}^{total} is the total number of hydrogen bond formed by the protein with the solution species. n_{HB}^{P-P} refers to the average number of intra-protein hydrogen bonds. $n_{\mathrm{HB}}^{urea-tmao}$ is the average number of hydrogen bonds formed between urea and TMAO per urea molecule.

■ SUMMARY AND CONCLUSIONS

By the application of REMD simulations, we have demonstrated the effect of denaturing and counteracting osmolytes on the protein inside a near spherical non polar confinement. The purpose of the use of encapsulation is to mimic the biological folding machinery chaperonin unit to explore its role on protein stabilization. Radius of the spherical

confinement is 2.3 Å which resembles with that of chaperonin unit. The profound effect of the confinement is realized when we compare the results with our previous study on Trp cage in bulk i.e. without confinement. We have found that the protein experiences decreased conformational space in presence of confinement due to limited volume provided by the confinement and the same is confirmed from R_g and RMSF plots. Highly unfolded behavior of the protein is observed in binary urea solution which is evident from FEL and melting curves. FEL also explores successful counteraction of TMAO. The decomposition of protein-urea interaction energy in to electrostatic and van der Waals components reveal that TMAO makes the former contribution less favorable in ternary solution.

In mixed ternary osmolyte solution the protein tend to occupy central part of the confinement whereas, in binary urea solution the protein is more prone to occupy peripheral region of the spherical ball. Density profiles of water for different systems indicate that for pure water system, there is dewetting of the hydrophobic fullerene surface. This observation is in consistent with previous studies [155, 63]. We have also found that the interaction behavior of different residues of the protein with the confinement surface changes depending on the osmolyte used. In binary urea solution all residues of the protein show high probability to occupy region in vicinity of the fullerene surface. Urea molecules in mixed osmolyte solution show higher extent of occupation near to the fullerene surface than that in binary urea solution. It suggests that urea molecules get more access to solvate the protein in binary urea solution to exert denaturing action. In mixed osmolyte solution, presence of TMAO molecules changes the solvation behavior of urea molecules. TMAO molecules solvate the fullerene surface as well as the urea molecules, so that we get an enhanced solvation of urea molecules near the fullerene surface. This makes urea molecules less available to solvate the protein and henceforth denaturing effect of urea is diminished in mixed osmolyte solution. Thus, confinement induces an environment that favors stable state of the protein in mixed solution whereas unstable state is favored in binary urea solution.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Chapter 5

Effect of Polar Confinement and Osmolyte Solutions on Folding/Unfolding Equilibrium of Protein Trp Cage

"It has been recognized that hydrogen bonds restrain protein molecules to their native configurations, and I believe that as the methods of structural chemistry are further applied to physiological problems it will be found that the significance of the hydrogen bond for physiology is greater than that of any other single structural feature."

– Linus C. Pauling, Nature of the Chemical Bond and the Structure of Molecules and Crystals (1939)

In this chapter we have explored the effect of polar confinement and osmolyte solutions on the folding/unfolding processes of the protein Trp cage. In vivo, folding and unfolding of protein occur in highly crowded cellular environment as well as inside folding machinery like chaperonins. To characterize the effect of chaperonin cavity or the crowded environment different approaches are developed. Considering the protein inside a spherical cavity is one such approach. This study was performed in order to examine the role of polar spherical confinement on protein stabilization and put forward a comparative view with non polar confinement (discussed in **Chapter 4**). For efficient conformational sampling of the protein we have utilized REMD simulation technique (like previous chapters). The effect of osmolytes urea and TMAO on protein destabilization/stabilization in presence of polar confinement is investigated throughout the study. Different structural properties of the protein are discussed. Role of salt bridge, hydrophobic core and orientational preferences of different planes of amino acid side chains are also implemented. Energy landscapes elucidate the conformational dynamics of the protein. Hydrogen bond analysis sheds light on the mechanisms of actions of urea and TMAO. Density profiles of water and protein across the spherical confinement are calculated to visualize their distributions inside the confinement. The study reveals that binary urea solution exerts denaturation on the protein in presence of polar confinement. But extent of counteraction decreases than in comparison to non polar confinement. For pure water system also it is observed that stabilizing effect due to polar confinement decreases than that of non polar confinement.

Use of REMD to explore the effects of osmolyte solutions and polar confinement on Trp cage

Overview: Folding/unfolding processes of the protein Trp cage in presence of osmolytes inside a polar confinement is investigated using replica exchange molecular dynamics (REMD) simulation. A near spherical fullerene ball consisting of 2940 carbon atoms (charged atoms) is used as a polar confinement. Urea exerts its action profoundly on the protein causing denaturation. Counteraction of TMAO is also observed in ternary solution of urea and TMAO. It is found that TMAO exerts its action in ternary mixed urea-TMAO solution by (i) removing some of the urea-TMAO hydrogen bonds, (ii) preserving the angles and distances between the aromatic planes of the residues Pro17, Pro18 and Pro19 with the indole ring of Trp6 and (iii) retaining the hydrophobic core of the protein Trp cage.

■ INTRODUCTION

Protein folding and unfolding are the fundamental processes that occur in cellular environments and are difficult to characterize in detail. $20\% - 40\%$ of the cytoplasmic volume is occupied by cellular macromolecules [156-160]. The crowded cytoplasmic environments significantly influence the biological processes including protein folding/unfolding equilibrium. The effect of crowding environment or the effects arising due to a chaperonin cavity on the protein can be mimicked by considering the protein within a spherical cavity [156, 157]. Chaperones are known as essential units that assist folding of newly synthesized proteins in vivo [56, 57]. GroEL/GroES is the most common example of chaperonin system that facilitates folding of a wide range of bacterial and eukaryotic proteins [55-57]. A key aspect of the chaperonin is the hydrophobic regions of the cavity that acts as the binding sites for exposed hydrophobic regions of non native protein and the cavity becomes hydrophilic upon ATP/GroES binding [54, 59, 60]. Combined effects of hydrophobic and hydrophilic regions of GroEL/GroES accomplish folding of protein. Brinker et al. [62] showed that GroEL/GroES chaperonin smoothen up the energy landscape of the protein and prevents kinetically trapped intermediates. In literature, numerous experimental techniques and theoretical methods have been reported addressing the effects of confinement on protein stabilization [63-67, 156, 161]. Tian et al. [63] addressed the effects of polar and non polar confinement on protein folding/unfolding equilibrium. They have demonstrated that non polar confinement stabilized the folded state while the polar confinement destabilized the protein. Simulations of coarse-grained proteins showed that the effect of confinement is dependent on the size of the confining volume relative to the size of the protein [67-69]. These studies show confinement-induced stabilization of the folded state of the protein. It is also established that confined solvent performs a crucial rule in determining the stability of the protein. Confinement results in reduction of the conformational entropy associated with protein folding. These factors in turn render stabilization of the protein. These studies inspired us to carry out a study of protein folding/unfolding processes in presence of binary urea and ternary solutions of urea and TMAO in polar confinement.

In literature, numerous efforts have been reported considering protecting and denaturing osmolytes to explore their mechanisms of protein stabilization through both experimental and computational methods. Depending on their nature, osmolytes play crucial role in protein's solubility and stability and function [34, 112, 113]. In this study we have emphasized on actions of osmolytes urea and TMAO. For urea denaturation, two

mechanisms are commonly suggested, namely, direct and indirect [77, 79, 114, 143, 144]. Direct interaction proposes preferential interaction of urea with the protein through hydrogen bond or dispersion interactions. According to indirect mechanism, urea alters the water structure which in turn affects protein hydration. However, recent advances on this osmolyte put consensus view on direct mechanism of urea denaturation. Alternatively, protecting osmolyte like TMAO does not show any favorable interaction with the protein. Different views of protein stabilization by TMAO have been put forwarded. However, a generalized mechanism of counteraction by TMAO has yet to be emerged. The mechanism of TMAO counteraction can be broadly categorized as (i) Direct effects and (ii) Indirect effects of TMAO [46]. In support of direct mechanism, work by Kokubo and co-workers [44] showed that favorable van der Waals (vdW) interaction of TMAO with the protein results in stabilization of the native protein. Bolen and co-workers [43] suggested that osmophobic interaction between protein backbone and TMAO contributes towards protein stabilization. In support of indirect mechanism various studies came up with different views, e.g. enhancement of water structure in presence of TMAO [36, 37], excluded volume effect of TMAO by acting as a crowding agent [40], decrease in solvent-excluded volume upon folding in presence of TMAO [145], preferential exclusion of TMAO from the protein surface and decreased self-interaction [33], mutual exclusion of urea and TMAO from the protein surface and enhancement of self-aggregation [45], TMAO-induced removal of urea from hydrophobic surface of the protein [117] etc. These studies support indirect mechanism of TMAO counteraction.

In this work, we have carried out Replica Exchange Molecular Dynamics (REMD) simulation [21] of Trp cage inside polar confinement in presence and absence of osmolytes urea and TMAO. Effect of confinement on the denaturing and counteracting ability of osmolytes and subsequent result on protein folding is investigated. As mentioned in **Chapter 3 and 4**, REMD simulation technique is chosen over classical MD since the former provides enhanced conformational sampling, which is useful to explore the protein free energy landscape to understand folding/unfolding equilibrium in a much better manner.

The rest of the chapter is organized as follows. Simulation details are provided in section models and simulation method. Then results are discussed thereafter and the last section includes the conclusion of the work.

■ MODELS AND SIMULATION METHOD

We have chosen a nearly spherical fullerene ball consisting of 2940 carbon atoms and half of these carbon atoms are charged randomly with +0.15e and rest half are charged with -0.15e in such a manner that the fullerene ball does not carry any net charge. This fullerene ball produces the effect of a polar confinement surrounding the protein Trp cage (PDB ID 1L2Y [18]) immersed in different osmolyte solutions. We have used CaGe software [146] to generate the fullerene ball. REMD was used for carrying out the simulations of three different aqueous solutions of the protein consisting of pure water, denaturing osmolyte (i.e. binary aqueous solution of 8M urea) and mixture of denaturing and counteracting osmolytes (i.e. ternary aqueous solution of 8M urea and 4M TMAO) using AMBER12 simulation program [70]. An overview of the systems are shown in Table 5-1. Packmol [89] was employed for generating the initial configurations of the systems. One chloride ion was added to each of the systems for neutralization since the protein has one unit of overall positive charge. GAFF force field [147] parameters were used for different carbon atoms of the fullerene ball. Following the work by Waghe et al. [148], we have also decreased the Lennard Jones parameter 'ϵ' of carbon atoms of the fullerene ball to half of their original values so that the interaction of the carbon atoms of the fullerene ball with other species gets reduced. We have considered Amber FF12SB [126] force field parameters for the protein. For solvation TIP3P model [125] of water was used. Smith of Urea [87] and Kast model of TMAO [88] were used for this study. We prefer to use Kast model of TMAO rather than the Osmotic model [33] since our previous study [53] had found that the former model provides better counteraction (than the latter) against the action of urea when the ratio of urea:TMAO is the conventional and experimentally relevant one i.e. 2:1. In literature, some other models of TMAO are also reported viz. Netz model [51] and Shea model [52] and that are modified version of the original Kast model. These models were developed in order to reproduce some experimental thermodynamic properties of aqueous TMAO solution. However, no such model is available till date that can adequately describe the balance of interactions in mixture of solutions. These observations lead us to choose Kast model of TMAO over the other models

For all the systems, at first we have carried out energy minimization by steepest descent method. Then the systems were heated gradually from 0K to 300K (in steps of 50K) in canonical (NVT) ensembles. The systems obtained from these steps were then equilibrated under isothermal-isobaric conditions (NPT) for 300 ps and this is followed by

400 ps of simulation run in NVT ensemble. We have carried out REMD simulations of each of the systems spanning a temperature range of 284-455K for 60 ns. For each systems 32 replicas were used. Simulation protocols are similar to those followed for **Chapter 4**. Note that, the melting temperature of a protein is the temperature at which its folded and unfolded states are expected to be sampled equally. The experimental melting temperature of the protein used in this study i.e., Trp cage is reported to be 315K [149, 150] which is found to be dissimilar with those obtained from different computational studies [63, 151, 152]. Computational studies show different melting temperatures for Trp cage that vary from 360K to 450K depending on the force field parameters used.

Table 5-1. Overview of Systems[a]

System	N_W	N_U	N_T	*No.of Replicas*	*Temperature* (K)	*Density* (g/cc)	*Box Length*(Å)
PW	1545	0	0	32	284-451	0.86	50.19
PWU	945	210	0	32	284-451	0.91	49.50
PWUT	800	190	90	32	284-444	1.01	48.76

[a] N_W, N_U and N_T represent the number of water, urea and TMAO molecules respectively. ρ is the density of the system. P, W, U, and T refer to protein, water, urea, and TMAO respectively.

■ RESULTS AND DISCUSSION

Root Mean Square Fluctuations (RMSF) and Secondary Structure

The Root Mean Square Fluctuations (RMSFs) of protein measure the positional fluctuations of the constituting residues and thus give an idea of conformational changes adopted by the protein in different chemical environments. High RMSFs indicate that the protein has experienced high degree of fluctuations. The terminal residues of a protein generally exhibit high RMSF values. RMSFs of each of the C_α-atom of the protein for three different systems are shown for both at low (300K) and high ($\leq$451K) temperature limits (see Figure 5-1). A comparison of Figure 5-1 (a) and (b) indicates that for all the systems at high temperature the protein has undergone more fluctuations than low temperature which is as per the expectation. Moreover, at low temperature the protein shows the highest fluctuation for binary urea solution i.e. more conformational changes of the protein have taken

place in that solution. For protein-water system and in ternary mixed osmolyte solution, the RMSFs show a close behavior of each other indicating a very similar conformational behaviors shown by the protein for these two systems. However, at high temperature RMSFs for all the three systems are very much similar. Correlating this study with our previous study [162] (where we have used a non polar confinement) we find that that polar confinement induces more instability to the protein conformation than that of the non polar confinement. At low temperature, in non polar confinement the RMSFs for all the residues of the protein in both protein-water system and mixed ternary solution remain below 2 Å whereas in polar confinement RMSFs exceed 2 Å. This observation leads us to propose that under polar confinement the protein attains more fluctuations than that in non polar confinement.

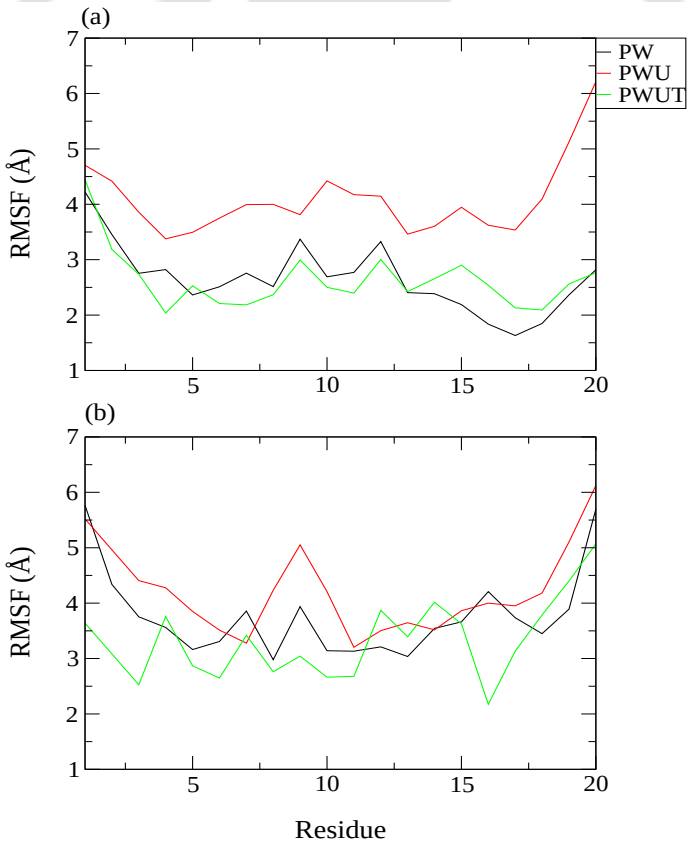


Figure 5-1. *Root mean square fluctuations (RMSF) of C_α-atoms of each of the residues of the protein at (a) low temperature and (b) high temperature.*

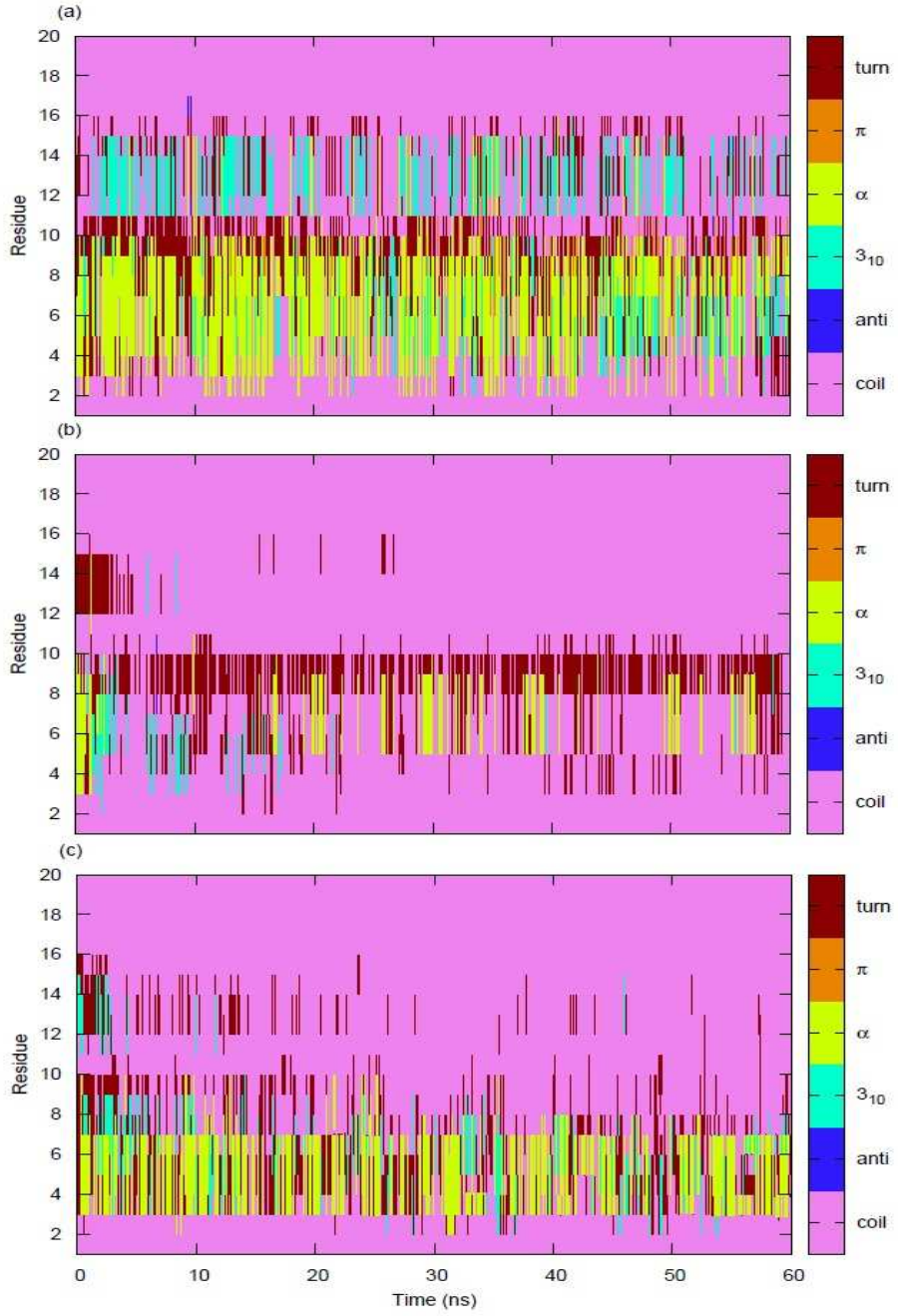


Figure 5-2. *Plot of secondary structure analysis of the protein as a function of simulation time. (a), (b) and (c) correspond to PW, PWU and PWUT systems respectively.*

To detect the changes in the protein's secondary structure in different systems, we have followed the protocol known as DSSP (Dictionary of Secondary Structure of Proteins) [129]. We have plotted the secondary structural elements as a function of simulation time

for the protein at 300K for all the systems (see Figure 5-2). It is observed that in pure water large extent of native conformation of the protein is retained than that of the other two systems (see Figure 5-2 (a)). In binary urea solution, the protein mostly acquires coil and turn conformations for the entire simulation path with a very weak appearance of α-helix conformation (see Figure 5-2 (b)). For the ternary osmolyte solution (Figure 5-2 (c)), we observe frequent appearances of coil and turn conformations in the helical region.

Radius of Gyration

We calculate the radius of gyration (R_g) which measures the effective size of the protein. At low temperature limit, the probability distributions of the R_g values of the protein Trp cage for all the systems provide a comparative view of its sizes in different chemical environments (see Figure 5-3). It is observed that the protein in binary urea solution has undergone the maximum conformational changes as is evident from a much wider distribution of R_g values. The peak positions of the distributions for pure water and ternary mixed solution are found appearing at around 8 Å which indicates a compact conformation of the protein in these two systems. However, a comparison of these distributions with our previous study of non polar confinement [162] reveals that the effective size of the protein is less compact in polar confinement.

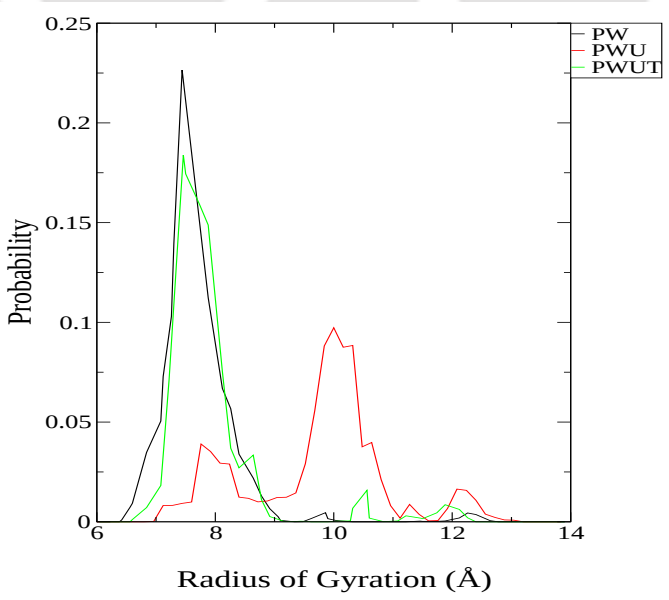


Figure 5-3. *Probability distributions of radius of gyration of the protein in different systems.*

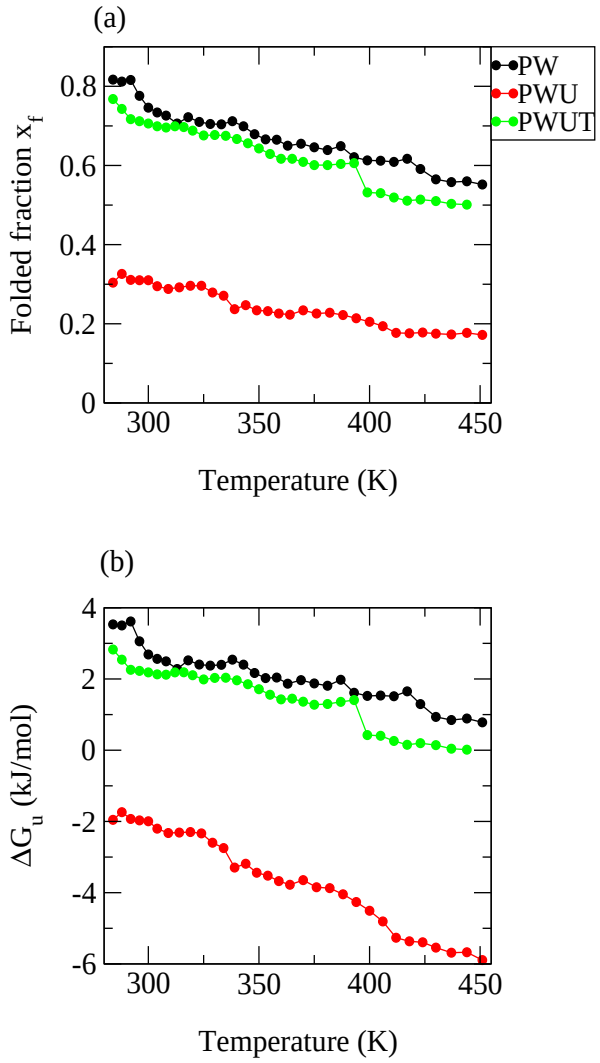


Figure 5-4. *(a) Folded fractions of the protein versus temperature and (b) corresponding free energy of unfolding of the protein at every temperatures for the respective systems.*

To find out the conformational behavior of the protein that varies with temperature we have plotted melting curves of the protein in absence and presence of osmolytes (see Figure 5-4). The folded fraction of the protein depends on the temperature as well as on the surrounding environments. Figure 5-4 (a) shows that the maximum folded fraction of the protein is retained in pure water system. For the entire temperature range the folded fraction values for ternary mixed solution is very similar to that of pure water

system. It indicates that in ternary mixed solution, TMAO exerts counteraction against the denaturing effect of urea. For binary urea solution, the folded fraction of the protein goes down markedly portraying urea-induced denaturation. Moreover, a comparison of the melting curves of the present study with those obtained for non polar confinement (as shown in our previous study [162]) reveals that the folded fractions of the protein show diminished values for polar confinement than non polar confinement. This observation simply indicates that in presence of polar confinement the protein becomes less stable than in presence of non polar one. In Figure 5-4 (b), we have shown the free energy of unfolding of the protein (ΔG_u) for every temperature. ΔG_u is obtained as:

$$\Delta G_u = -RT[(1 - x_{folded})/x_{folded}] \tag{5.1}$$

Here, x_{folded} is the folded fraction of the protein, R is the gas constant and T represents replica temperature. The negative values of ΔG_u (even at low temperature) for binary urea solution ascertain that unfolding is highly favorable in presence of urea. In case of pure water and mixed osmolyte systems high folding propensity of the protein is indicated by positive ΔG_u values at low to moderate (or entire) temperature range.

The free energy landscapes (FELs) of the protein are explored since they provide information regarding the random walk in the energy space performed by the protein. This in turn helps to characterize the structural transition pathway undergone by the protein. FELs can be constructed as a function of suitable reaction coordinate that incorporates most of the changes attained by the protein. In order to examine the conformational behavior of the protein with the rise of temperature, we have used RMSDs (root mean square deviations) and temperature of each replica for the construction of FELs. For this purpose we have calculated probability distributions of the RMSDs of the protein at every replica temperature and then FELs are constructed according to the following equation (Eq. 3):

$$\Delta G(V) = -k_B T[lnP(V) - lnP_{max}] \tag{5.2}$$

Here, k_B is the Boltzmann constant and T refers to the replica temperature. P(V) is the probability distribution of RMSDs and P_{max} is the maximum of the distribution. P_{max} is subtracted from P(V) so that ΔG becomes zero for the free energy minimum. A comparison of the FELs for different systems (see Figure 5-5 (a) and (c)) shows that the protein continues to maintain native-like structure in pure water and in presence of mixed osmolytes (RMSDs around 2 Å). But in binary urea solution, the native conformation is poorly preserved (see Figure 5-5 (b)). For the entire range of temperature the protein attains

highly unfolded conformations with RMSDs exceeding 3 Å. More unfolding of the protein is observed with the increase of temperature for all the systems which is as per the expectation. Furthermore, the folded basins of the FELs for pure water and mixed ternary solution resemble each other. It suggests that the native conformation of the protein is preserved in presence of mixture of osmolytes and acts as a corroborative evidence of what we discussed above. This again confirms that TMAO counteracts the denaturing effect of urea even inside the polar confinement. When we compare the FELs with those obtained for non polar confinement [162], we can clearly observe that for all the systems, the protein experiences more stabilizing effect under non polar confinement than polar one.

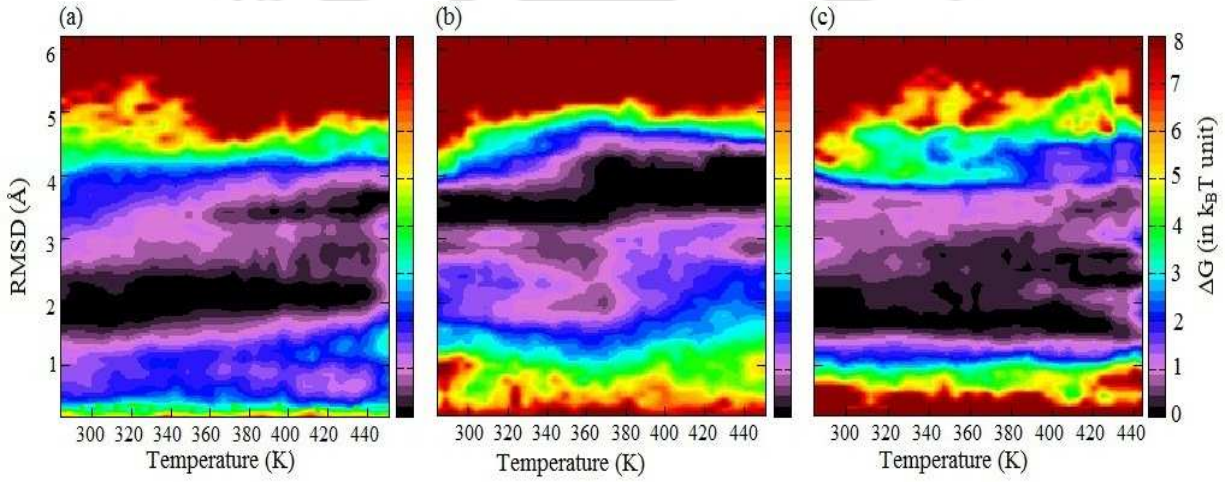


Figure 5-5. *Free energy landscapes (FELs) for systems (a) PW, (b) PWU and (c) PWUT.*

Dihedral Angle Principal Component Analysis (dPCA)

To characterize the conformational transitions of Trp cage, we have performed dihedral angle principal component analysis (dPCA) [98] by using 'Carma: a molecular dynamics analysis program' [99]. The dihedral angles (ϕ, ψ) of C_α atoms of the protein backbone are considered. At first a covariance matrix, σ_{ij}, is constructed. σ_{ij} is defined as:

$$\sigma_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle \tag{5.3}$$

where x_1,...,x_{2N} are the sine- and cosine-transformed dihedral angles (N is the number of angles) and $\langle ... \rangle$ represents the average of the samples. Then diagonalization of the covariance matrix results into 2N number of eigenvectors, v(i), and eigenvalues, λ_i. Finally, the probability distributions of the first two principal components, $V_i = v(i)x$ are used to set

up the free energy surface. Only the first two eigenvectors are chosen because they represent the most dominant collective motions that are sufficient to describe the conformational transitions of the protein.

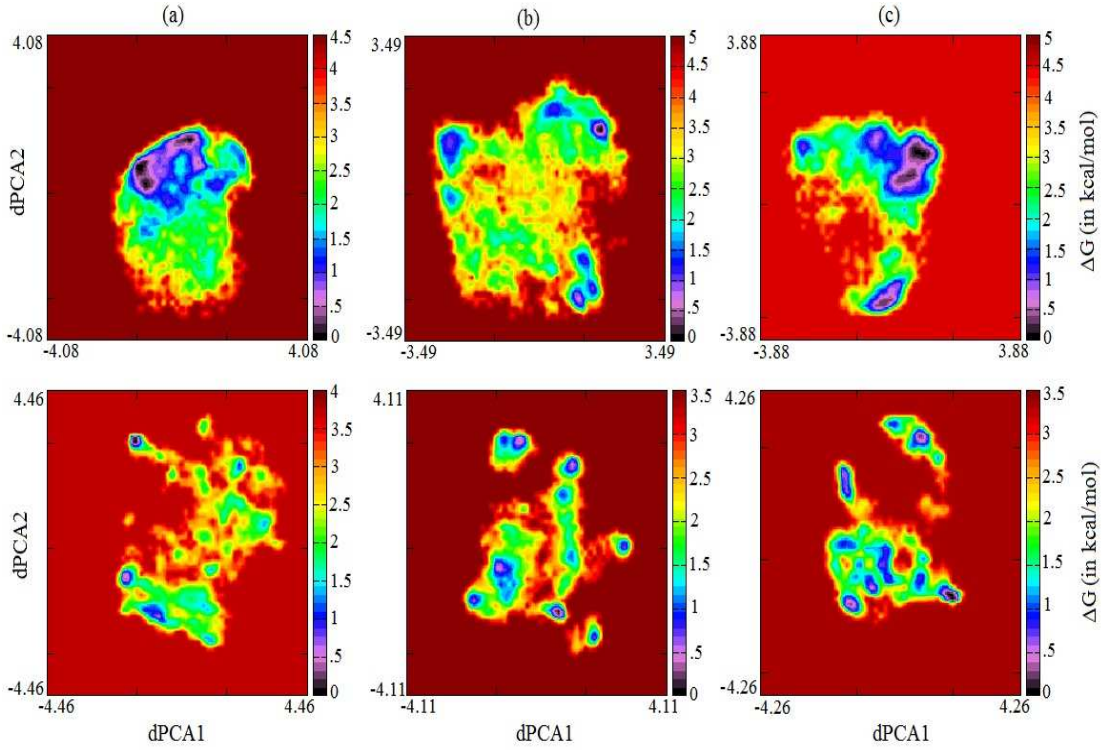


Figure 5-6. *Plot of dihedral angle principal component analysis (dPCA) for systems (a) PW, (b) PWU and (c) PWUT at low (top) and high (bottom) temperature limits.*

For pure water system at low temperature (see Figure 5-6 (a)), two major native like conformational clusters are observed that are separated by a small energy barrier. It implies that the transition between these two clusters are permissible. At high temperature a single minimum indicates weak appearance of native like conformation of the protein. For binary urea solution, at low temperature (Figure 5-6 (b)), high free energy basins separate the minima from other conformations. At high temperature no minimum is observed. For mixed osmolyte solution (Figure 5-6 (c)), at low temperature, two closely spaced local minima that allow conformational transitions are observed. Moreover, another isolated minimum is also observed with high energy barrier. It portrays that the native conformation of the protein is preserved in this system. However, at high temperature native conformation diminishes. It is observed that at high temperature the scenario is

more or less similar for all the systems. It indicates that the protein prefers non-native conformations at high temperature in presence of polar confinement. This is as per our expectation. Nevertheless, since we are interested in, mostly, the counteracting action of TMAO against urea's action on the protein at physically relevant temperature, our focus is mainly on the findings at low temperature limit (∼300K).

Density Change Across the Fullerene Ball

Density of the protein (C_α-atoms of the residues considered) across the confinement i.e. from the center of mass (COM) of the fullerene ball towards its surface is calculated to examine the location of the protein inside the confinement. Density of water across the confinement is also calculated. In Figure 5-7, the densities of the protein (solid lines) and water (lines with diamond symbol) are shown for all the systems. We have observed that there is a sharp enhancement of densities of water for all the systems near the surface of the confinement. This observation is in line with previously reported results [155, 63]. This finding further supports that water wets the hydrophilic surface indicating favorable water-fullerene interaction. Hydration layer that clings to the charged fullerene surface may help in the exclusion of the protein atoms from the fullerene surface. However, charged side chains of the protein may compete with water and penetrate the hydration layer [63]. It is observed that for all the systems the protein occupy a region 5-20 Å from the center of the confinement. It indicates that irrespective of the nature of the solution components, the protein is located within an optimal range of 5-20 Å.

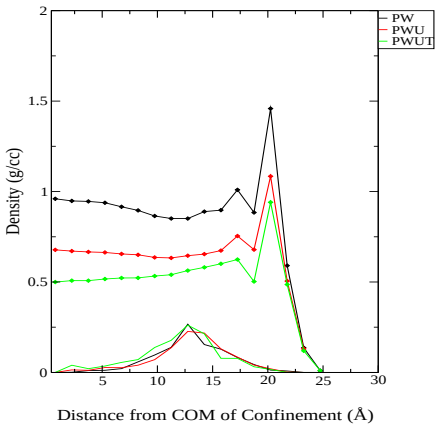


Figure 5-7. *Density of water (lines with diamond symbol) and protein (solid line) as a function of distance from COM of the spherical confinement.*

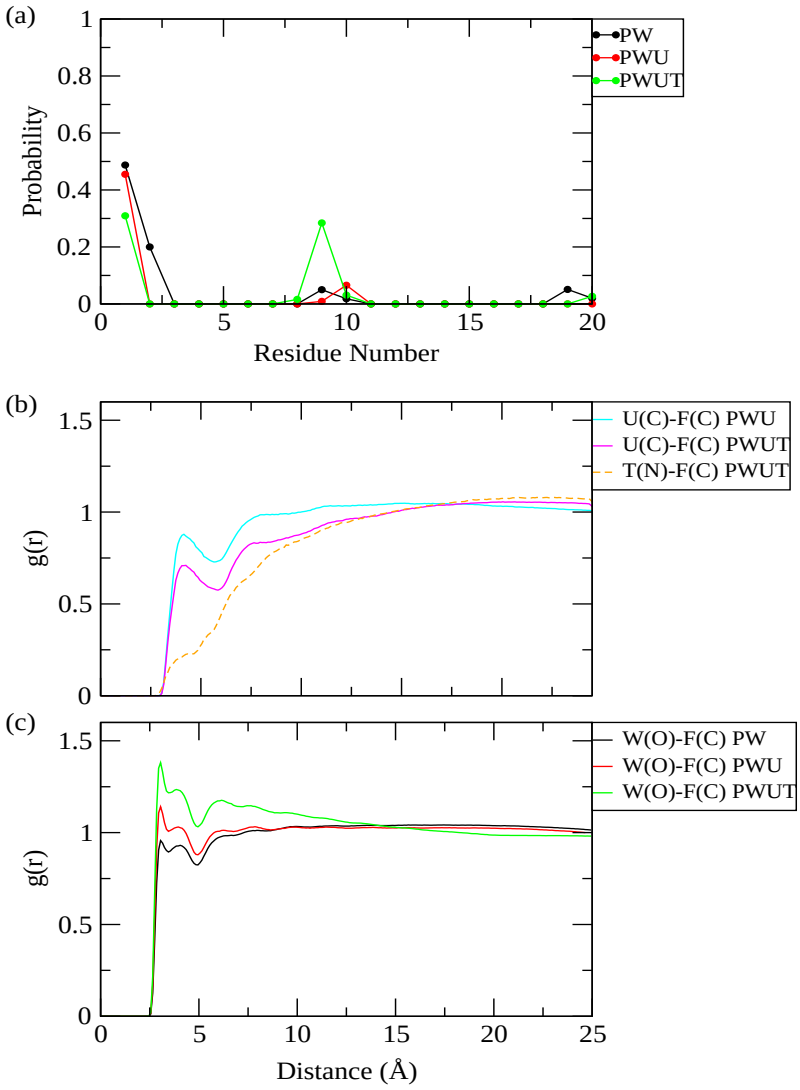


Figure 5-8. *(a) Probability of finding the C_α-atoms of different residues of the protein within 4.5 Å from the fullerene surface. (b) Radial distribution functions involving central atoms of urea and TMAO with the fullerene surface. (c) refers to the radial distribution functions involving water oxygen and the fullerene surface.*

We have calculated the probability distribution of the C_α atoms of different protein residues within a distance of 4.5 Å from the surface of the confinement (see Figure 5-8 (a)). It is apparent that all the residues of the protein Trp cage show very small tendency to occupy the region near the surface for all the systems except the residue 1. However, the residues 8, 9 and 10 show small probabilities to remain near to the surface. We have plotted

rdfs (radial distribution functions) involving central atoms of urea (TMAO, and water) and the fullerene surface (see Figure 5-8 (b) and (c)). It is observed that urea molecules in binary solution show more tendency to stay nearby the surface than in mixed osmolyte solution. This implies that the protein residues that remain available near the surface, get solvated by urea molecules in binary solution. In case of ternary mixed solution, the presence of TMAO reduces the urea density near the fullerene surface. Moreover, the fullerene surface region is also hydrated significantly in mixed osmolyte solution.

Role of Salt Bridge, Hydrophobic Core and Orientational Preferences of Aromatic Planes in Stabilization of Trp Cage

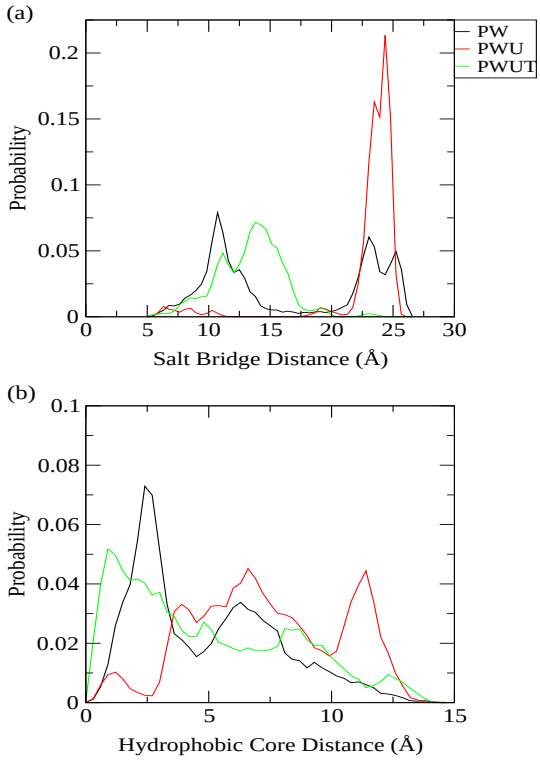


Figure 5-9. *(a) Probability distributions of the distances between the salt bridge forming residues for different systems. (b) Probability distributions of the hydrophobic core forming residues versus the distance from the center of the hydrophobic core.*

Above findings suggest that urea denatures protein Trp cage and TMAO nullifies the deleterious effect of urea on protein conformation without revealing the atomistic details of urea's action (on protein) and the counteraction by TMAO. Thus, it would be interesting to concentrate on these aspects in atomistic details. To emphasize on the dif-

ferent stabilizing factors of the protein Trp cage we have analyzed the role of salt bridge, hydrophobic core and the orientational preferences of aromatic planes of different residues of Trp cage. In the native conformation of the protein there exists a salt bridge between residues Arg16 and Asp9 and according to the results of the previous findings this salt bridge plays an important role in stabilization of the protein Trp cage [63, 119, 120]. Contrary to these findings it has also been reported that it is not always necessary to maintain the salt bridge for the stabilization of protein conformations [63, 119]. Nevertheless, in this study the probability distribution of the distance maintained by the salt bridge forming residues (i.e. Arg16 and Asp9) is calculated (see Figure 5-9 (a)). For binary urea system the salt bridge pair keeps the maximum distance from each other. For pure water system and ternary mixed solution the peaks of distributions at positions 10 and 15 Å respectively indicate insignificant contribution of the salt bridge pair on the stabilization of Trp cage.

Further, the hydrophobic core of Trp cage is formed by residues Tyr3, Leu7, Gly11, Pro12, Pro18 and Pro19 surrounding the central residue Trp6 [133, 15]. This hydrophobic core is believed to play significant role on the stabilization of the protein Trp cage [15, 121]. This dictates us to calculate the probability distributions of the distances maintained by hydrophobic core forming residues from the center of the core and the same is shown in Figure 5-9 (b). A comparison of the distributions of the three systems reveals that the hydrophobic core is mostly preserved in pure water and ternary mixed solution, but not in binary urea solution. Preservation of the hydrophobic core indicates that native like conformation is maintained in pure water and ternary solution. Thus it can be concluded that the retention of the hydrophobic core in those two systems contributes to the stability of protein Trp cage.

Moreover, the protein Trp cage possesses different aromatic planes surrounding the central indole plane of Trp6. As reported previously [153, 162], the interactions among the planes may play key role in determining the overall structure and conformation of the protein. Interactions between these planes are feasible when they maintain a distance within 7 Å and the best fitted tilt angle for a particular orientation [163]. Motivated by this, in this work also we have emphasized on the orientational preferences of aromatic ring of residue Tyr3 and pyrrole rings of residues Pro12, Pro17, Pro18, Pro19 with respect to indole plane of Trp6. We have also analyzed the distributions of COM distances for the respective residue pairs. For the calculation of orientational preferences we have considered the vectors normal to the aromatic planes of the concerned residues. The angles between these vectors close to 0° and 90° correspond to parallel and perpendicular orientations of the planes respectively.

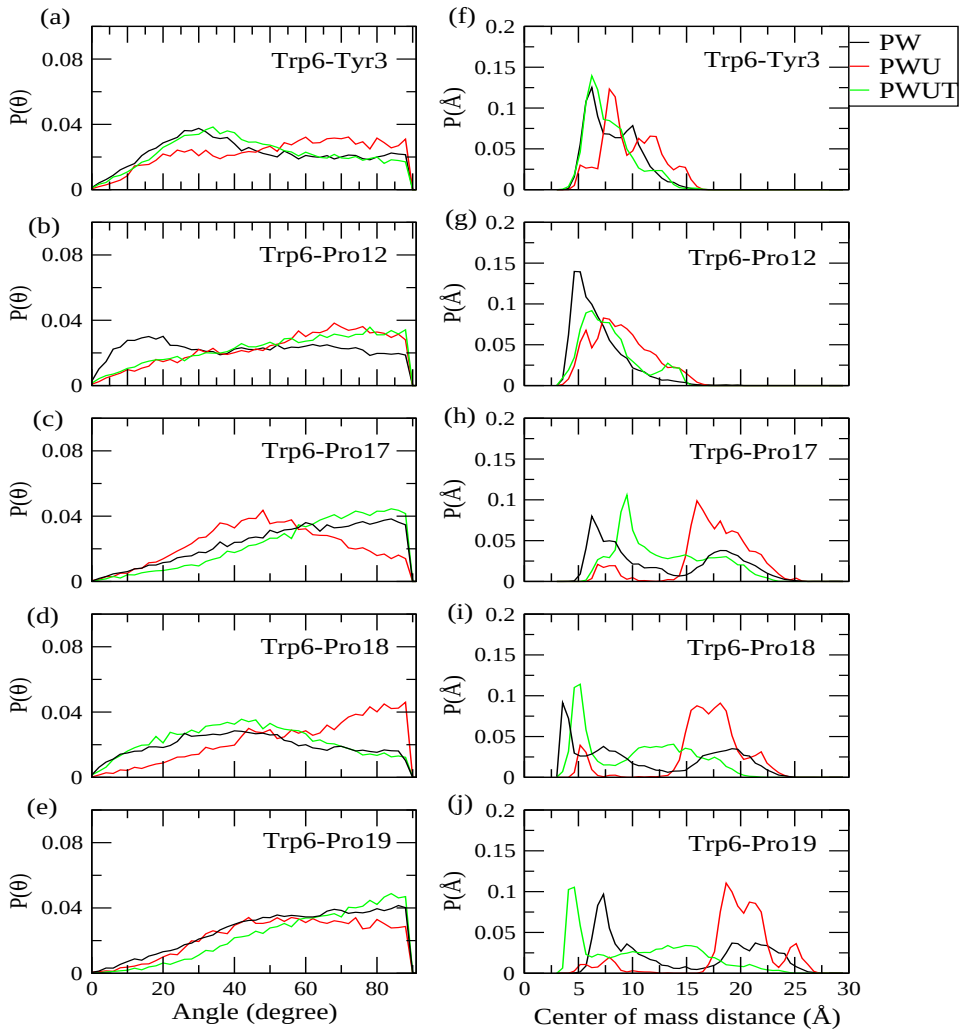


Figure 5-10. *Probability distributions of different orientational angles (left) and COM distances (right) of indole plane with other aromatic planes.*

In the Figure 5-10 (a)-(j), P(θ) and P(Å) denote normalized probability distributions of the angles between the planes and COM distances of the planes respectively. They are plotted against inter plane angles and COM distances respectively. The peaks of the distributions indicate the most preferred angle or distance between the respective pair of planes. Since the protein exhibits most native like behavior in pure water system, the distributions of P(θ) and P(Å) for this system are considered as reference distributions. It is observed that for the planes of residue pair Trp6-Tyr3, a tilt angle close to 30° is maintained for pure water and ternary mixed solution. For binary urea system, no preferential orientation for this pair is observed. The planes of residue pair Trp6-Pro12 prefer parallel

orientation with respect to each other in pure water system whereas for other two systems perpendicular orientations are preferred. For Trp6-Pro17 and Trp6-Pro19, both in pure water and mixed ternary solution near perpendicular orientations are preferred, but for binary urea solution the maximum of the distribution appears around 50°. For Trp6-Pro18, the distributions of P(θ) in pure water and mixed ternary solution are very similar and they prefer to be oriented at around 40°. On the other hand, the same in binary urea solution prefers perpendicular orientation. Distributions of COM distances reveal that for pure water and ternary mixed solution the planes are more likely to remain within 10 Å. But in case of binary urea solution, the planes mostly occupy distant positions except for the pairs Trp6-Tyr3 and Trp6-Pro12. In brief, the interactions between the planes of residues Pro17, Pro18 and Pro19 with the indole plane of Trp6 contribute mainly in retaining the native conformation of Trp cage in pure water and in mixed osmolyte solution. A comparison of these findings with our previous study carried out for non polar confinement [162] reveals that in both the studies the preferential orientations of the studied pairs of planes are similar in pure water and ternary mixed osmolyte systems. However, in non polar confinement the peak heights of these distribution functions are higher than that for polar confinement. This suggests that the protein in polar confinement shows less tendency to retain its preferred orientations, (when compared to that of non polar confinement) i.e. folding propensity becomes less pronounced in polar confinement. Furthermore, a comparison of the COM distance plots reveals that the planes maintain a closer distance in non polar confinement than in polar one. Thus it leads us to conclude that the polar confinement induces some more destabilization to the protein than that of the non polar one.

Hydrogen Bond Properties

In order to get an insight about the effect of hydrogen bonds formed between protein and different solution species on the protein stability we have estimated the average number of intra- and inter-molecular hydrogen bonds formed by the protein at 300K temperature (see Table 5-2). For the calculation of hydrogen bonds the geometric crieteria are set. If the distance between the donor and acceptor atoms falls within 3.5 Å and the minimum donor-hydrogen-acceptor angle is maintained at 120° [103], then the two molecules are considered to be hydrogen bonded. It is apparent that the maximum number of intra-protein hydrogen bond is maintained in pure water system. For ternary mixed solution also a considerable number of intra-protein hydrogen bond is preserved, whereas in binary urea solution the intra-protein hydrogen bond number gets reduced profoundly. It suggests that more native like conformation of the protein is favored in both pure water

and ternary mixed solution. But in binary urea solution the native conformation is not maintained. Further, the addition of urea to protein-water system causes a significant reduction in the protein-water hydrogen bond numbers. This is because of the fact that some of the protein-water hydrogen bonds are getting replaced by protein-urea hydrogen bonds. Now, as TMAO is added to binary urea solution, though a small reduction in the protein-water hydrogen bonds is observed but remarkably, a significant decrease in the protein-urea hydrogen bond number is noticed. This suggests that TMAO molecules replace some of the urea molecules from the surface of the protein. It results in reduction of protein-urea hydrogen bonds in ternary mixed solutions. We also observe a large number of protein-urea hydrogen bonds in binary urea solutions. This causes an enhancement in the total number of hydrogen bonds formed by the protein with the solution species in binary urea solution. It supports the idea that the unfolded conformation of the protein gets more access to form hydrogen bonds with surrounding species. However, these hydrogen bonds are decreased significantly in ternary mixed solution. Now considering protein-TMAO hydrogen bonds we find that a small number of TMAO molecules are also engaged in hydrogen bonding interactions with the protein. The above facts, suggest that TMAO's action (on urea-conferred protein denaturation) arises mainly due to removal of some of the urea molecules from the protein surface by TMAO molecules.

Table 5-2. Average number of hydrogen bonds[a]

System	n_{HB}^{water}	n_{HB}^{urea}	n_{HB}^{tmao}	n_{HB}^{total}	n_{HB}^{P-P}
PW	59.88	−−	−−	59.88	10.78
PWU	44.58	33.32	−−	77.90	3.62
PWUT	38.29	24.18	2.56	65.03	9.88

[a] n_{HB}^{water}, n_{HB}^{urea} and n_{HB}^{tmao} are the average number of hydrogen bonds formed by the protein with water, urea and TMAO respectively. n_{HB}^{total} is the total number of hydrogen bond formed by the protein with the solution species. n_{HB}^{P-P} represents average number of intra-protein hydrogen bonds.

In support of the above observations we have further calculated spatial distribution functions of water, urea and TMAO molecules surrounding the protein and the same are shown in Figure 5-11. Density of water is found to decrease gradually from pure water to binary and ternary solution. It is quite apparent that the density of urea molecules (near

the protein) in binary urea solution is much higher than that for ternary solutions. In ternary solutions, the presence of TMAO molecules is also quite noticeable.

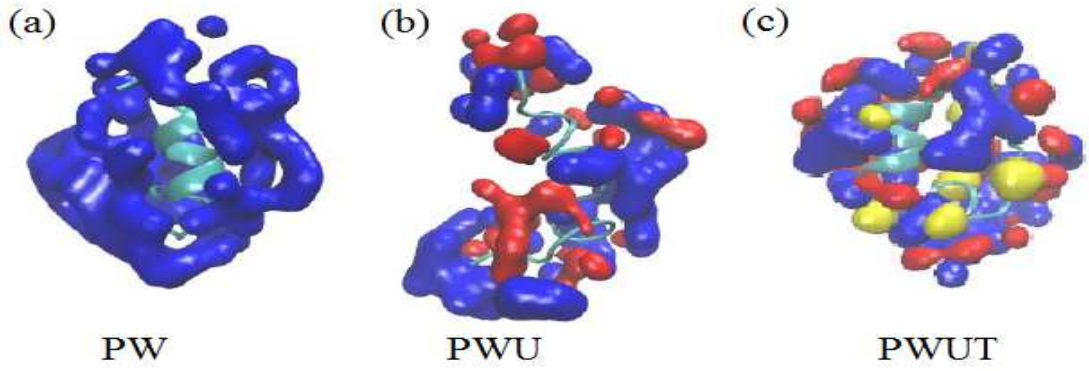


Figure 5-11. *Spatial distribution functions (sdfs) of water (blue), urea (red) and TMAO (yellow) surrounding the protein in systems (a) PW, (b) PWU and (c) PWUT*

■ SUMMARY AND CONCLUSIONS

We have studied the effect of polar confinement on the denaturing and counteracting actions of osmolytes urea and TMAO on the protein. It is hound that the protein experiences decreased conformational space due to confinement. We also observe that the protein attains highly unfolded behavior in binary urea solution. Direct interaction of urea with the protein is confirmed by the hydrogen bond analysis that results into denatured conformation. The calculations of the protein's folded fractions (obtained from melting curves) and energy landscapes indicate about the stabilizing and destabilizing effect of TMAO and urea respectively. Role of urea on protein denaturation becomes more efficient in presence of polar confinement when compared to that for non polar confinement.

Density profiles of water across the confinement reveal that water molecules wet the hydrophillic surface of the confinement in all the systems. Moreover, the small probabilities of finding the residues of the protein within a distance of 4.5 Å from the fullerene surface indicate that the fullerene surface is not enriched by the protein. The calculations of rdfs involving central atom of urea and fullerene surface show that density of urea becomes less near the surface in mixed osmolyte solution than that in binary urea solution. We have also found that the optimal range of occurrence of the protein inside polar confinement is 5-20 Å for all the systems. Hydrophobic core of the protein is found to contribute in

protein stabilization, however the poor existence of salt bridge rules out its contribution in Trp cage stabilization. Analysis of orientational preferences of different pairs of planes of the protein shows that Pro17 and Pro19 prefer near perpendicular orientations with Trp6 in pure water and mixed solution, whereas Pro18 shows no orientational preference. Tyr3 prefers a tilt angle of 30° for pure water and mixed solution. Thus altogether it can be stated that TMAO-induced stabilization of the protein Trp cage in polar confinement is achieved due to removal of some of the urea molecules from the surface of the protein by TMAO molecules and TMAO-induced retention of angles and distances between planes of Pro17, Pro18 and Pro19 with indole plane of Trp6. Moreover, the preservation of the hydrophobic core of Trp cage in mixed osmolyte solution also contributes significantly in protein stabilization. Since no evidence of direct interaction of protein with TMAO is observed this study points towards indirect mechanism of TMAO-induced stabilization of the protein Trp cage.

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Chapter 6

Summary and Our View on Osmolyte and Confinement Induced Stabilization of Protein

"Nature's patterns sometimes reflect two intertwined features: fundamental physical laws and environmental influences. It's nature's version of nature versus nurture."

– Brian Greene

The aim of this work is to gain an insight and a deeper understanding of the effect of confinement on protein folding/unfolding process in presence and absence of osmolytes. To explore the mechanisms of protein stabilization and destabilization by TMAO and urea respectively over a wide range of temperatures we have investigated different structural properties of the protein. Energy landscapes and interactions with the solution species of the proteins are explored as well. Replica exchange molecular dynamics simulation technique has been used which has helped us to achieve enhanced sampling and overcome energetic barriers at low temperature.

We started the investigations by classical MD simulation of a small protein β-hairpin. Study of literature for counteraction behavior of TMAO against action of urea showed that different models of TMAO are available. We have chosen Kast model and Osmotic model. Conformational behaviors of the protein were studied by RMSD and SASA of the protein in presence and absence of osmolytes. The study reveals that Kast model of TMAO is capable of offsetting the denaturing effect of urea in conventional 2:1 ratio of urea:TMAO whereas Osmotic model provides maximum counteraction against action of urea when used in 1:2 ratio. The study also found that electrostatic and vdW components of protein-urea interaction become more unfavorable in presence of TMAO. To explain this fact it was suggested that TMAO molecules replace some of the urea molecules from the solvation shell of the protein. In support of this, we have shown rdfs involving central atoms of urea and TMAO and heavy atoms of the protein. The calculations of average number of heavy atoms of urea and TMAO around the protein also confirms the fact that there is depletion in the number of heavy atoms of urea in the solvation shell of the protein for mixed osmolyte system. From the calculations of intra- and intermolecular hydrogen bonds we have found that in absence of any osmolyte the number of intra-protein hydrogen bond is maximum. This is in consistent with the idea of spontaneous folding behavior of the proteins when they are surrounded by aqueous environment [103]. Minimum number of intra-protein hydrogen bonds indicate unfolding of the protein under the influence of urea. Urea molecules deliver denaturation through formation of considerable number of hydrogen bonds with the protein. It is also found that the addition of TMAO decreases urea-water hydrogen bonds due to the formation of considerable amount of TMAO-water hydrogen bonds. Remarkably, the formation of urea-TMAO hydrogen bond reveals that in mixed osmolyte systems TMAO molecules are also solvated by urea molecules. This in turn makes urea molecules less available to solvate the protein to deliver denaturing effect. A further investigation of hydrogen bonds formed by the terminal residues (2 and 15) of

the protein with urea showed that these hydrogen bonds are more affected in binary urea solution and the same remains protected in presence of TMAO. Thus the formation of hydrogen bond by the terminal residues with urea plays a vital role in determining overall conformation of the protein. The overall conclusion from the study can be given as follows. Urea delivers its action through direct interaction with the protein and noticeably through the breaking of terminal hydrogen bonds that links the two sheets in parallel fashion while TMAO invokes an environment surrounding the protein, which is deficient in water and urea molecules and hence the folded form of the protein is preserved.

To obtain atomic-level picture of the protein's conformational changes REMD simulation of the protein Trp cage was carried out. To explain the molecular mechanism of actions of urea and TMAO different structural properties are analyzed first. To be more specific on the denaturing and counteracting mechanisms of urea and TMAO on the protein Trp cage we have emphasized on the three key stabilizing factors that cause large impact on the protein's stability viz. salt bridge, hydrophobic core and orientational preferences of different planes of amino acid side chains. Role of salt bridge on determining stability of the protein has been a matter of dispute. It is not always a necessary factor for protein stabilization. Our study reveals that (from the calculations of probability distributions of distances between salt bridge forming residues and occurrence of urea molecules near the salt bridge) urea molecules affects the interactions of the salt bridge forming residues. Urea molecules are very much prone to occupy region near the salt bridge. However, urea density decreases in ternary mixed solution due to replacement of some of the urea molecules from this region by other solvent or cosolute molecules. Thus, TMAO molecules contribute on protein stabilization which may have arisen through the expulsion of some urea molecules from the region near the salt bridge. Trp cage possesses a well packed hydrophobic core. This invokes us to examine the existence of this core and the role of TMAO in the protection of the hydrophobic core. We observe that the core is conserved in pure water, but in presence of urea the core becomes heavily hydrated. Presence of TMAO, however, prevents its exposure to water to some extent. This leads us to propose that TMAO counteracts against the action of urea by providing protection to the hydrophobic core of the protein. Interactions between side chains of the protein residues are regarded as a decisive factor for protein stabilization. Thus, we have studied preferential orientations of different aromatic planes of different amino acid residues and how they are affected with the change of chemical environments surrounding the protein. We have found that near parallel orientation of Trp6-Pro12 is not an essential factor for the protein stabilization,

whereas near parallel orientation of Trp6-Pro18 and near perpendicular orientation of Trp6-Pro17 and Trp6-Pro19 are found to be significant contributors in stabilizing the native conformation of the protein. We have calculated the intra- and inter-protein hydrogen bonds and the effect of TMAO on protein-urea interaction energy in order to highlight the stabilizing effect of TMAO to compensate the deleterious effect of urea. We found that direct interaction between protein and urea causes denaturation of the protein. In ternary solution, TMAO replaces some of the urea molecules from the protein surface. More the number of TMAO molecules, more number of urea molecules are expelled from the protein surface. As a result, both electrostatic and vdW components of protein-urea interaction energy become more unfavorable in presence of TMAO (than that of binary urea system). In the study, we do not observe any direct interaction of TMAO with the protein. This leads us to conclude that TMAO-induced stabilization is effective by means of indirect interaction.

In cellular environment proteins are subjected to both crowding and confinement. How the confinement exerts its effects on protein folding/unfolding process and interactions of protein with its surrounding species have been investigated by use of both non polar and polar confinements. Different structural properties were analyzed to explain the effect of confinement on the protein and consequent effect on the osmolytes in determining the folding/unfolding process of the protein. The study disclosed that the protein experiences a decreased conformational space in presence of confinement (both non polar and polar) due to limited volume provided by the confinement than in comparison to without confinement. Highly unfolded behavior of the protein is observed in binary urea solution (both in non polar and polar confinement), which is evident from FEL and melting curves. FEL also explores successful counteraction of TMAO. A comparison of non polar and polar confinement indicates that the latter induces some destabilizing effect on the protein. The decomposition of electrostatic and van der Waals components of protein-urea interaction energy reveals that TMAO makes the contribution of electrostatic energy less favorable in mixed ternary solution. The role of the salt bridge, the existence of the hydrophobic core and its influence in pertaining stability to the protein and the orientational preferences of different aromatic planes of the protein residues with respect to central indole plane of tryptophan residue are also investigated in order to examine their roles in protein stabilization. For both the confinements, we have found that the existence of salt bridge is not a necessary factor for protein stabilization. This finding is in line with Garcia's work [63] which reported that Trp cage can retain its native form in presence as well as in absence of

the salt bridge. Maintenance of the hydrophobic core (in pure water and mixed osmolyte solutions) plays a vital role in stabilizing the protein for both non polar and polar confinements. Orientational preferences considering non polar confinement show that the residue pair Trp6-Pro12 maintains parallel orientation only in case of pure water system similar to that obtained for without confinement. But, the orientations of residue pairs Trp6-Pro17 for pure water system is dissimilar with that of without confinement. For the aromatic planes of residue pairs Trp6-Pro18 and Trp6-Pro19, we observe near parallel and near perpendicular orientations respectively for pure water and mixed osmolyte solutions that is similar with those we have found for without confinement. In case of polar confinement, Pro17 and Pro19 prefer near perpendicular orientations with Trp6 in pure water and mixed solution, whereas Pro18 shows no orientational preference.

In mixed ternary osmolyte solution the protein tend to occupy central part of the non polar confinement whereas, in binary urea solution the protein is more prone to occupy peripheral region of the spherical ball. The optimal range of occurrence of the protein inside polar confinement is 5-20 Å for all the systems. Density profiles of water for different systems indicate that for pure water system, there is dewetting of the non polar fullerene surface, whereas water molecules wet the polar surface of the confinement for all systems. These observations are in consistent with previous studies [155, 63]. We have also found that the interaction behavior of different residues of the protein with the non polar confinement surface changes depending on the osmolyte used. In binary urea solution all residues of the protein show high probability to occupy region in vicinity of the non polar fullerene surface. Urea molecules in mixed osmolyte solution show higher extent of occupation near to the fullerene surface than that in binary urea solution. It suggests that urea molecules get more access to solvate the protein in binary urea solution to exert denaturing action. In mixed osmolyte solution, the presence of TMAO molecules changes the solvation behavior of urea molecules. TMAO molecules solvate the fullerene surface as well as the urea molecules, so that we get an enhanced solvation of urea molecules near the fullerene surface. This makes urea molecules less available to solvate the protein and henceforth denaturing effect of urea is diminished in mixed osmolyte solution. Confinement induces an environment that favors stable state of the protein in mixed solution whereas unstable state is favored in binary urea solution.

Altogether it can be stated that TMAO stabilization of protein inside confinement is achieved due to removal of urea molecules from the surface of the protein by TMAO molecules. TMAO-induced retention of angles and distances between aromatic planes of

Trp cage and the preservation of the hydrophobic core in mixed osmolyte solution contribute significantly in protein stabilization. These result in reduced effect of urea denaturation on the protein. Since no evidence of direct interaction of protein with TMAO is observed this study points towards indirect mechanism of TMAO stabilization.

Further studies can be expanded with the simulation of the same system under polar confinement with different spatial charge distribution. How different spatial charge distribution of polar confinement can effect on the protein and osmolytes can be compared. Using different sizes of the spherical confinement, a size-dependent comparative study can also be performed. The study can also be extended to various concentrations of urea and TMAO. This will help to make comparisions over a range of concentrations. Further, the study can also be repeated with recently developed models of urea and TMAO. Use of some other denaturant or counteracting osmolyte (except than urea and TMAO) or salt inside polar and non polar confinement can be another extension of this study. A further challenging extension of this work would be to study the effect of cold denaturation on the same systems.

Bibliography

1. C. M. Dobson, A. Sali, M. Karplus *Angew. Chem.* **37**, 868 (1998).
2. C. M. Dobson *Nature* **426**,884 (2003).
3. C. B. Anfinsen *Science* **181** 223 (1973).
4. V. Daggett *Chem. Rev.* **106**, 1898 (2006).
5. C. M. Dobson *Nature* **418**, 729 (2002).
6. A. L. Fink *Fold. Des.* **3**, 9 (1998).
7. S. Prakash, A. Matouschek *Trends Biochem. Sci.* **29**, 593 (2004).
8. F. J. Blanco, M. A. Jimenez, A. Pineda, M. Rico, J. Santoro, J. L. Nieto *Biochemistry* **33**, 6004 (1994).
9. R. M. Hughes, M. L. Waters *Curr. Opin. Struct. Biol.* **16**, 514 (2006).
10. B. Zagrovic, E. J. Sorin, V. Pande *J. Mol. Biol.* **313**, 151 (2001).
11. Q. Shao, L. Yang, Y. Q. Gao *J. Chem. Phys.* **130**, 195104 (2009).
12. P. H. Nguyen, G. Stock, E. Mittag, C. Hu, M. Suan *Proteins: Struct., Funct., Bioinf.* **61**, 795 (2005).
13. G. Bussi, F. L. Gervasio, A. Laio, M. Parrinello *J. Am. Chem. Soc.* **128**, 13435 (2006).
14. M. Bonomi, D. Branduardi, F. L. Gervasio, M. Parrinello *J. Am. Chem. Soc.* **130**, 13938 (2008).
15. B. Barua, J. C. Lin, V. D. Williams, P. Kummler, J. W. Neidigh, N. H. Andersen *Protein Eng. Des. Sel.* **21**, 171 (2008).
16. Z. Ahmed, I. Beta, A. Mikhonin, S. Asher *J. Am. Chem. Soc.* **127**, 10943 (2005).
17. R. Day, D. Paschek, A. E. Garcia *Proteins: Struct., Funct., Bioinf.* **78**, 1889 (2010).
18. J. W. Neidigh, R. M. Fesinmeyer, N. H. Anderson *Nat. Struct. Biol.* **9**, 425 (2002).

19. A. Bille, B. Linse, S. Mohanty, A. Irback *J. Chem. Phys.* **143**, 175102 (2015).

20. C. A. English, A. E. Garcia *J. Phys. Chem. B* **119**, 7874 (2015).

21. Y. Sugita, Y. Okamoto *Chem. Phys. Lett.* **314**, 141 (1999).

22. D. R. Canchi, D. Paschek, A. E. Garcia *J. Am. Chem. Soc.* **132**, 2338 (2010).

23. P. H. Ancey, M. E. Clark, S. C. Hand, R. D. Bowlus, G. N. Somero *Science* **217**, 1214 (1982).

24. K. N. Barton, M. M. Buhr, J. S. Ballantyne *Am. J. Physiol.* **276**, 397 (1999).

25. L. J. Borowitzka, A. D. Brown *Arch. Microbiol.* **96**, 3752 (1974).

26. R. D. Bowlus, G. N. Somero *J Exp. Zool.* **208**, 137 (1979).

27. R. Zangi, R. H. Zhou, B. J. Berne *J. Am. Chem. Soc.* **131**, 1535 (2009).

28. H. S. Frank, F. Franks *J. Chem. Phys.* **48**, 4746 (1968).

29. B. J. Bennion, V. Daggett *Proc. Natl. Acad. Sci.* **100**, 5142 (2003).

30. A. Caballero-HerreraA, K. Nordstrand, K. D. Berndt, L. Nilsson *Biophys. J.* **89**, 842 (2005).

31. E. P. O'Brien, R. I. Dima, B. Brooks, D. Thirumalai *J. Am. Chem. Soc.* **129**, 7346 (2007).

32. D. K. Klimov, J. E. Straub, D. Thirumalai *Proc. Natl. Acad. Sci.* **101**, 14760 (2004).

33. D. R. Canchi, P. Jayasimha, D. C. Rau, G. I. Makhatadze, A. E. Garcia *J. Phys. Chem. B* **116**, 12095 (2012).

34. A. Wang, D. W. Bolen *Biochemistry* **36**, 9101 (1997).

35. S. Gopal, J. C. Ahluwalia *J. Chem. Soc. Faraday Trans.* **89**, 2769 (1993).

36. B. J. Bennion, V. Daggett *Proc. Natl. Acad. Sci.* **101**, 6433 (2004).

37. J. Hunger, K. J. Tielrooij, R. Buchner, M. Bonn, H. J. Bakker *J. Phys. Chem. B* **116**, 4783 (2012).

38. P. Attri, P. Venkatesu, M. J. Lee *J. Phys. Chem. B* **114**, 1471 (2010).

39. F. Meersman, D. Bowron, A. K. Soper, M. H. J. Koch *Biophys. J.* **97**, 2559 (2009).

40. S. S. Cho, G. Reddy, J. E. Straub, D. Thirumalai *J. Phys. Chem. B* **115**, 13401 (2011).

41. M. R. Hilaire, R. M. Abaskharon, F. Gai *J. Phys. Chem. Lett.* **6**, 2546 (2015).

42. J. Maa, I. M. Pazosa, F. Gai *Proc. Natl. Acad. Sci.* **111**, 8476 (2014).

43. D. W. Bolen, I. V. Baskakov *J. Mol. Biol.* **310**, 955 (2001).

44. H. Kokubo, Y. C. Hu, B. M. Pettitt *J. Am. Chem. Soc.* **133**, 1849 (2011).

45. P. Ganguly, T. Hajari, J.-E. Shea, N. F. van der Vegt *J. Phys. Chem. Lett.* **6**, 581 (2015).

46. P. Ganguly, N. F. van der VegN. F. van der Vegt, J.-E. Shea *J. Phys. Chem. Lett.* **7**, 3052 (2016).

47. S. Paul, G. N. Patey *J. Phys. Chem. B* **110**, 10514 (2006).

48. P. Venkatesu, M. Lee, H. Lin *J. Phys. Chem. B* **113**, 5327 (2009).

49. S. Paul, G. N. Patey *J. Am. Chem. Soc.* **129**, 4476 (2007).

50. S. Paul, G. N. Patey *J. Phys. Chem. B* **111**, 7932 (2007).

51. E. Schneck, D. Horinek, R. R. Netz *J. Phys. Chem. B* **117**, 8310 (2013).

52. L. Larini, J. -E. Shea *J. Phys. Chem. B* **117**, 13268 (2013).

53. G. Borgohain, S. Paul *J. Phys. Chem. B* **120**, 2352 (2016).

54. R. M. Vabulas, S. Raychaudhuri, M. Hayer-Hartl, F. U. Hartl *Cold. Spring. Harb. Perspect. Biol.* **2**, a004390 (2010).

55. D. Thirumalai, G. H. Lorimer *Annu. Rev. Biophy. Biomol. Struct.* **30**, 245 (2001).

56. F. U. Hartl *Nature* **381**, 571 (1996).

57. F. U. Hartl, M. Hayer-Hartl *Science* **295**, 1852 (2002).

58. F. U. Hartl, A. Bracher, M. Hayer-Hartl *Nature* **475**, 324 (2011).

59. T. J. Piggot, R. B. Sessions, S. G. Burston *Biochemistry* **51**, 1707 (2012).

60. N. K. Tyagi, W. A. Fenton, A. L. Horwich *Proc. Natl. Acad. Sci.* **106**, 20264 (2009).

61. J. L. England, D. Lucent, V. S. Pande *J. Am. Chem. Sos.* **130**, 11838 (2008).

62. A. Brinker, G. Pfeifer, M. J. Kerner, D. J. Naylor, F. U. Hartl, M. Hayer-Hartl *Cell* **107**, 223 (2001).

63. J. Tian, A. E. Garcia *J. Am. Chem. Soc.* **133**, 15157 (2011).

64. D. Thirumalai, D. K. Klimov, G. H. Lorimer *Proc. Natl. Acad. Sci.* **100**, 11195 (2003).

65. H. X. Zhou *Arch. Biochem. Biophys.* **469**, 76 (2008).

66. D. Lucent, C. D. Snow, C. E. Aitken, V. S. Pande *PLoS Comput. Biol.* **6**, e1000963 (2010).

67. D. K. Klimov, D. Newfield, D. Thirumalai *Proc. Natl. Acad. Sci.* **99** 8019 (2002).

68. J. Mittal, R. B. Best *Proc. Natl. Acad. Sci.* **105**, 20233 (2008).

69. N. Rathore, T. A. Knotts, J. J. D. Pablo. *Biophys. J.* **90**, 1767 (2006).

70. D. A. Case, T. A. Darden, T. E. Cheatham, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, et al. *AMBER 12*; University of California: San Francisco, (2012).

71. J. W. Ponder, D. A. Case *Adv. Prot. Chem.* **66**, 27 (2003).

72. J. E. Lennard-Jones *Proc. R. Soc. London, Ser. A* **106**, 463 (1924).

73. J. D. van der Waals *Verhandelingen der Koninklijke Akademie der Wetenschappen* **1**, 1 (1893).

74. C. A. Coulomb *Collection de ḿemoires relatifs à la physique*, Gauthier-Villars, 569 (1884).

75. L. Verlet *Phys. Rev.* **159**, 98 (1967).

76. R. W. Hockney *Meth. Comp. Phys.* **9**, 136 (1970).

77. G. A. P. Oliveira, J. L. Silva *Proc. Natl. Acad. Sci.* **112**, E2775 (2015).

78. L. Hua, R. Zhoua, D. Thirumalai, B. J. Berne *Proc. Natl. Acad. Sci.* **105**, 16928 (2008).

79. A. Das, C. Mukhopadhyay *J. Phys. Chem. B* **113**, 12816 (2009).

80. L. B. Sagle, K. Cimatu, V. A. Litosh, Y. Liu, S. C. Flores, X. Xin Chen, B. Yu, P. S. Cremer *J. Am. Chem. Soc.* **133**, 18707 (2011).

81. A. Panuszko, P. Brudziak, J. Zielkiewicz, D. Wyrzykowski, J. Stangret *J. Phys. Chem. B* **113**, 14797 (2009).

82. Z. A. Levinea, L. Larinia, N. E. LaPointe, S. C. Feinstein, J. -E. Shea *Proc. Natl. Acad. Sci.* **112**, 2758 (2015).

83. H. Wei, Y. Fan, Y. Q. Gao *J. Phys. Chem. B* **114**, 557 (2010).

84. Q. Shao, Y. Q. Gao *J. Chem. Theory Copmut.* **6**, 3750 (2010).

85. J. M. Wang, P. Cieplak, P. A. Kollman *J.Comput. Chem.* **21**, 1049 (2000).

86. H. J. C. Berendsen, J. R. Griger, T. P. Straatsma *J. Phys. Chem.* **91**, 6269 (1987).

87. L. J. Smith, H. J. C. Berendsen, W. F. V. Gunsteren *J. Phys. Chem. B* **108**, 1065 (2004).

88. K. M. Kast, J. Brickmann, S. M. Kast, R. S. Berry *J. Phys. Chem. A* **107**, 5342 (2003).

89. L. Martnez, R. Andrade, E. G. Birgin, J. M. Martnez *J. Comput. Chem.* **30**, 2157 (2009).

90. J. P. Ryckaert, G. Ciccotti, H. J. C. Berendsen *J. Comput. Phys.* **23**, 327 (1977).

91. P. H. Hunenberger *Adv. Polym. Sci.* **173**, 105 (2005).

92. H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola, J. R. Haak *J. Chem. Phys.* **81**, 3684 (1984).

93. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. J. Pedersen *J. Chem. Phys.* **103**, 8577 (1955).

94. W. Humphrey, A. Dalke, K. Schulten *J. Molec. Graphics* **14**, 33 (1996).

95. A. Grossfield, D. M. Zuckerman *Annu. Rep. Comput. Chem.* **5**, 23 (2009).

96. H. W. Hatch, F. H. Stillinger, P. B. Debenedetti *J. Phys. Chem B* **118**, 7761 (2014).

97. A. Shrake, J. A. Rupley *J. Mol. Biol.* **79**, 351 (1973).

98. Y. Mu, P. H. Nguyen, G. Stock *Proteins* **58**, 45 (2005).

99. N. M. Glykos *J. Comput. Chem.* **27**, 1765 (2006).

100. S. Paul, S. Paul *J. Phys. Chem. B* **119**, 10975 (2015).

101. M. C. Stumpe, H. Grubmuller *J. Am. Chem. Soc.* **129**, 16126 (2007).

102. R. Sarma, S. Paul *J. Chem. Phys.* **139**, 034504 (2013).

103. R. Sarma, S. Paul *J. Phys. Chem. B* **117**, 9056 (2013).

104. A. Kuffel, J. Zielkiewicz *J. Chem. Phys.* **133**, 135102 (2010).

105. J. Rosgen, R. Jackson-Atogi *J. Am. Chem. Soc.* **134**, 3590 (2012).

106. D. Lambert, D. E. Draper *J. Mol. Biol.* **370**, 993 (2007).

107. K. Takai, K. Nakamura, T. Toki, U. Tsunogai, M. Miyazaki, J. Miyazaki, H. Hirayama, S. Nakagawa, T. Nunoura, K. Horikoshi *Proc. Natl. Acad. Sci.* **105**, 10949 (2008).

108. D. R. Canchi, A. E. Garcia *Annu. Rev. Phys. Chem.* **64**, 273 (2013).

109. R. Jaenicke *Eur. J. Biochem.* **202**, 715 (1991).

110. K. Heremans *Annu. Rev. Biophys. Biol.* **11**, 1 (1982).

111. D. Bolen, G. Rose *Annu. Rev. Biochem.* **77**, 339 (2008).

112. T. O. Street, D. W. Bolen, G. D. Rose *Proc. Natl. Acad. Sci.* **103**, 13997 (2006).

113. T. Y. Lin, S. N. Timasheff *Biochemistry* **33**, 12695 (1994).

114. D. Cui, S. Ou, S. Patel *J. Phys. Chem B* **119**, 164 (2015).

115. N. Kishore, N. Kumar *Biophys. Chem.* **189**, 16 (2014).

116. A. Chand, P. Chettiyankandy, S. K. Pattanayak, S. Chowdhuri *J. Mol. Liq.* (2016), http://dx.doi.org/10.1016/j.molliq.2016.05.044.

117. Y. Yang, Y. Mu, W. Li *Phys. Chem. Chem. Phys.* **18**, 22081 (2016).

118. W. Xu, Y. Mu *Biophys. Chem.* **137**, 116 (2008).

119. C. A. Jimenez-Cruz, G. I. Makhatadze, A. E. Garcia *Phys. Chem. Chem. Phys.* **13**, 17056 (2011).

120. J. Juraszek, P. G. Bolhuis *Proc. Natl. Acad. Sci.* **103**, 15859 (2006).

121. M. W. van der Kamp, V. Daggett *J. Mol. Biol.* **404** 732 (2010).

122. A. Halabis, W. Zmudzinska, A. Liwo, S. Oldziej *J. Phys. Chem. B* **116**, 6898 (2012).

123. H. Meuzelaar, K. A. Marino, A. Huerta-Viga, M. R. Panman, L. E. J. Smeenk, A. J. Kettelarij, J. H. van Maarseveen, P. Timmerman, G. Bolhuis, S. Woutersen *J. Phys. Chem. B* **117**, 11490 (2013).

124. P. Hudaky, P. Straner, V. Farkas, G. Varadi, G. Toth, A. Perczel *Biochemistry* **47**, 1007 (2008).

125. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein *J. Chem. Phys.* **79**, 926 (1983).

126. W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Jr. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman *J. Am. Chem. Soc.* **117**, 5179 (1995).

127. D. Paschek, H. Nymeyer, A. E. Garcia *J. Struct. Biol.* **157**, 524 (2007).

128. D. R. Roe, T. E. III Cheatham *J. Chem. Theory Comput.* **9**, 3084 (2013).

129. W. Kabsch, C. Sander *Biopolymers* **22**, 2577 (1983).

130. D. T. Jones *J. Mol. Biol.* **292**, 195 (1999).

131. J. Bryngelson, J. Onuchic, N. Socci, P. Wolynes *Proteins: Struct., Funct., Genet.* **21**, 167 (1995).

132. P. Weinkam, J. Zimmermann, F. E. Romesberg, P. G. Wolynes *Acc. Chem. Res.* **43**, 652 (2010).

133. X. Wu, G. Yang, Y. Zu, Y. Fu, L. Zhou, X. Yuan *Comput. Theor. Chem.* **973**, 1 (2011).

134. S. Kannan, M. Zacharias *PLoS One* **9**, e88383 (2014).

135. D. R. Canchi, D. Paschek, A. E. Garcia *J. Am. Chem. Soc.* **132**, 2338 (2010).

136. Z. Yang, P. Xiu, B. Shi, L. Hua, R. Zhou *J. Phys. Chem. B* **116**, 8856 (2012).

137. F. Rodriguez-Ropero, N. F. A. van der Vegt *J. Phys. Chem. B* **118**, 7327 (2014).

138. G. Vogt, S. Woell, P. Argos *J. Mol. Biol.* **269**, 631 (1997).

139. A. L. Horwich, F. E. Chapman, G. W. Farr *Ann. Rev. Cell. Dev. Biol.* **23**, 115 (2007).

140. F. Takagi, N. Koga, S. Takada *Proc. Natl. Acad. Sci.* **100**, 11367 (2003).

141. A. I. Jewett, J. Shea *J. Mol. Biol.* **363**, 945 (2006).

142. F. Georgescauld, K. Popova, A. J. Gupta, A, Bracher, J. R. Engen, M. Hayer-Hartl, F. U. Hartl *Cell* **157**, 922 (2014).

143. M. Auton, D. W. Bolen *Proc. Natl. Acad. Sci.* **102**, 15065 (2005).

144. R. Zangi, R. Zhou, B. J. Berne *J. Am. Chem. Soc.* **131**, 1535 (2009).

145. G. Graziano *Phys. Chem. Chem. Phys.* **14**, 13088 (2012).

146. G. O. Brinkmann, D. Friedrichs, S. Lisken, A. Peeters, N. V. Cleemput *MATCH Commun. Math. Comput. Chem.* **63**, 533 (2010).

147. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case *J. Comp. Chem.* **25**, 1157 (2004).

148. A. Waghe, J. C. Rasaiah, G. J. Hummer *Chem. Phys.* **117**, 10789 (2002).

149. A. T. Iavarone, J. H. Parks *J. Am. Chem. Soc.* **127**, 8606 (2005).

150. W. W. Streicher, G. I. Makhatadze *Biochemistry* **46**, 2876 (2007).

151. Z. Gattin, S. Riniker, P. J. Hore, K. H. Mok, W. F. v. Gunsteren *Prot. Sci.* **18**, 2090 (2009).

152. A. E. Garcia, K. Y. Sanbonmatsu *Proc. Natl. Acad. Sci.* **99**, 2782 (2002).

153. G. Borgohain, S. Paul *(Manuscript Under Revision).*

154. G. H. Zerze, B. Uz, J. Mittal *Prot.: Struct., Funct., Bioinf.* **83**, 1307 (2015).

155. M. Sakurai, H. Tamagawa, K. Ariga, T. Kunitake, Y. Inoue *Chem. Phys. Lett.* **289**, 567 (1998).

156. M. Li, W. Xu, J. Z. H. Zhang, F. Xia *J. Mol. Model.* **20**, 2530 (2014).

157. J. S. Rao, M. D. Smith, L. Cruz *J. Phys. Chem B* **118**, 3517 (2014).

158. M. S. Cheung, D. Klimov, D. Thirumalai *Proc. Natl. Acad. Sci.* **102**, 4753 (2005).

159. M. S. Cheung, D. Thirumalai *J. Phys. Chem. B* **111**, 8250 (2007).

160. H. X. Zhou *Biochemistry* **43**, 2141 (2004).

161. H. Zhang, Y. Lv, T. Tan, D. Van Der Spoel *J. Phys. Chem B* **120**, 477 (2016).

162. G. Borgohain, S. Paul *J. Mol. Liq.* **231**, 174 (2017).

163. R. Anjana, M. K. Vaishnavi, D. Sherlin, S. P. Kumar, K. Naveen, P. S. Kanth, K. Sekar *Bioinformation* **8**, 1220 (2012).

164. I. M. Svishchev, P. G. Kusalik *J. Chem. Phys.* **99**, 3049 (1993).

List of Publications

1. Gargi Borgohain and Sandip Paul, Model Dependency of TMAOs Counteracting Effect Against Action of Urea: Kast Model versus Osmotic Model of TMAO *J. Phys. Chem. B.* **120**, 2352-2361 (2016).

2. Gargi Borgohain and Sandip Paul, Use of Molecular Dynamics Simulation to Explore Structural Facets of Human Prion Protein with Pathogenic Mutations *Biophys. Chem.*, **213**, 32-39 (2016).

3. Gargi Borgohain and Sandip Paul, Temperature Mediated Switching of Protectant-Denaturant Behavior of Trimethylamine-N-Oxide and Consequences on Protein Stability from Replica Exchange Molecular Dynamics Simulation Study *Mol. Simul.*, **43**, 52-64 (2016).

5. Gargi Borgohain and Sandip Paul, Effect of Non Polar Confinement on Protein Trp Cage Conformation in Aqueous Osmolyte Solutions *J. Mol. Liq.*, **231**, 174-184 (2017).

6. Gargi Borgohain and Sandip Paul, Folding/Unfolding of Protein Trp cage in Aqueous Osmolyte Solutions Under Polar Confinement (Manuscript Under Revision).

4. Gargi Borgohain and Sandip Paul, Atomistic Level Understanding of the Stabilization of Protein Trp Cage in Denaturing and Mixed Osmolyte Solutions (Manuscript Under Revision).

7. Gargi Borgohain and Sandip Paul, Molecular Mechanism of Dodine Conferred Denaturation of λ-Repressor and Counteraction of Trehalose from Molecular Dynamics Simulation Study (Manuscript Under Revision).

8. Gargi Borgohain and Sandip Paul, (Manuscript Submitted).

भारतीय प्रौद्योगिकी संस्थान गुवाहाटी
Indian Institute of Technology Guwahati

Conferences Attended

1. Presented a poster on "Investigation of effect of urea and TMAO on folding/unfolding of protein" in the conference " Theoretical Chemistry Symposium (TCS-2012)" held at Department of Chemistry, IIT Guwahati.

2. Presented a poster on "REMD simulation of protein in presence of osmolytes" in the conference "Dynamics of Complex Chemical and Biological Systems (DCCBS-2014)" held at IIT Kanpur.

3. Presented a poster on "Temperature mediated Switching of Protectant-Denaturant Behavior of TMAO and Consequences on Protein Stability" in the conference "Theoretical Chemistry Symposium (TCS-2014)" held at Department of Chemistry, NCL-Pune.

4. Presented a poster entitled "Use of molecular dynamics simulation to explore structural facets of human prion protein with pathogenic mutations" in the conference "Contemporary Developments in Chemical Sciences (CDCS-2015)" held at Tezpur university.