Protein thermal stability, conformational dynamics and solvent properties: insights with atomistic molecular dynamics simulations

Thesis submitted to AcSIR for the award of the Degree of
DOCTOR OF PHILOSOPHY In Chemical Sciences

By

Mr. Prathit Chatterjee

Registration No. **10CC11J26043**

Research Supervisor: **Dr. Neelanjana Sengupta**

CSIR – National Chemical Laboratory,
Pune – 411008, India

December, 2015

ॐसहनाववतु सहनौ भुनक्तु सहवीर्यं करवावहै। तेजस्विनावधीतमस्तु मा विद्विशावहै।

Om, let us (both teachers and students) be protected, nourished and strengthened with knowledge. May we work together with great energy, and our intellect be sharpened. While having so, may we not hate or despise each other.

सीएसआयआर–राष्ट्रीय रासायनिक प्रयोगशाला
(वैज्ञानिक तथा औद्योगिक अनुसंधान परिषद)
डॉ. होमी भाभा मार्ग, पुणे – 411 008. भारत
CSIR-NATIONAL CHEMICAL LABORATORY
(Council of Scientific & Industrial Research)
Dr. Homi Bhabha Road, Pune - 411008. India

CERTIFICATE

This is to certify that the work presented in this thesis titled ***Protein thermal stability, conformational dynamics and solvent properties: insights with atomistic molecular dynamics simulations*** submitted by **Mr. Prathit Chatterjee** to Academy of Scientific and Innovative Research (AcSIR) in fulfillment of the requirements of the Degree of ***Doctor in Philosophy in Chemical Sciences***, embodies original research carried out under my supervision at the Physical and Materials Chemistry Division, CSIR–National Chemical Laboratory, Pune-411008, India. I further certify that this work is original and has not been submitted partly or fully, for any degree or diploma to this or to any other University. All the research materials from the other sources are duly acknowledged. Any text, illustration, table, etc., used in the thesis from other sources have been duly cited and acknowledged.

Mr. Prathit Chatterjee
(student and PhD. candidate)

Dr. Neelanjana Sengupta
(Research supervisor)
Scientist and Assistant Professor (AcSIR)
Physical and Materials Chemistry Division
CSIR- National Chemical Laboratory
Dr. Homi Bhabha Road, Pashan
Pune – 411008, Maharashtra, India

Dated: 10/12/2015
Place: Pune

DECLARATION BY RESEARCH SCHOLAR

I, hereby declare that all the computational studies embodied in this thesis **Protein thermal stability, conformational dynamics and solvent properties: insights with atomistic molecular dynamics simulations** submitted for the degree of Doctor of Philosophy in Chemical Sciences, to the AcSIR-NCL have been carried out by me at the Physical and Materials Chemistry Division, CSIR–National Chemical Laboratory, Pune-411008, India, under the supervision of **Dr. Neelanjana Sengupta**. This work is original and has not been submitted in part or full by me, for any degree or diploma to this or to any other University.

Dated: /12/2015

Mr. Prathit Chatterjee

Senior Research Fellow and PhD candidate,
Physical and Materials Chemistry Division
CSIR- National Chemical Laboratory
Dr. Homi Bhabha Road, Pashan
Pune – 411008, Maharashtra, India

Dedicated to

My Parents,

My Teachers,

and

My Late Maternal Grandparents.

Acknowledgement

It is almost impossible for a person to attain something novel without the assistance and guidance obtained from others, especially during doctoral dissertation. For me, the list of acknowledging those includes people from diverse prospects like scientists, non-academic staffs, students, friends, etc. Although, the list remains long, I am mentioning a selected few among them, whose contributions have been inevitable in channelizing my career prospects in a constructive way and inspiring me to pursue doctoral research.

First of all, I would like to thank *Dr. Neelanjana Sengupta*, my research supervisor, for all the help and support. She helped me develop my sense of reasoning, the power of discrimination and proper analyses during the course of my PhD. Moreover, she has also understood my personal issues, cooperated accordingly and strengthened me morally during my hard times to fight back against all atrocities and stand firm in tackling the problems with equal efficiency, in both professional and personal level.

I acknowledge my collaborator, *Dr. Sayan Bagchi*, for guiding and assisting me in one of my projects involving spectroscopic studies. I thank *Dr. H. V. Thulasiram* for giving me the opportunity to extend computational support in his experimental studies on protein-ligand interaction. Also, I would like to thank *Ms. Sonali Rohamare* and *Dr. Sushma Gaikwad* for collaborating with us in probing the structural propensity of proteases. The experience we have gained while collaborating with experimentalists has helped strengthen our expertise in the field.

Further, I acknowledge my doctoral advisory committee members, *Dr. Kumar Vanka*, *Dr. Sudip Roy*, *Dr. Sarika M. Bhattacharyya*, *Dr. Pankaj Doshi* and *Dr. Rajnish Kumar* for their graceful presence in all my assessments, as per the *AcSIR* norms. In addition, I also thank *Dr. Anu Raghunathan*, and *Dr. Arun Venkatnathan* (*IISER, Pune*) for appearing in my CSIR upgradation seminars. I thank *Dr. Sudip Roy* and *Dr. Ajithkumar* for coordinating the course works conducted in our times. I thank *Dr. Arnab Mukherjee* for allowing me pursue statistical mechanics course at IISER. I thank the chairpersons of *student academic office* (*SAO*), *Dr. P.A. Joy* and *Dr. C.G. Suresh*, and the present *AcSIR* coordinator *Dr. Chetan Gadgil* for all the support. I thank *Dr. Ruchi Anand*, of *IIT Bombay*, for supporting me before I joined for PhD. at NCL. I thank *Dr. Srabani Taraphdar* of *IIT Kharagpur*, for her affectionate yet careful considerations of my academic career, right from my post-graduation till date. Among other professional acquaintances, I would like to specially mention *Dr. Nayana Vaval*, and *Dr. Kavita Joshi*. I thank *Dr. Anil Kumar*, the present chairperson of Physical and Materials Chemistry Division, for his kind interactions with me. Last but not the least, I thank *Dr. Sourav Pal*. I have seen him change from the head of Physical

and Materials Chemistry, to the Director of the institute, then join *IIT Bombay* as an invited professor, during the course of my PhD. Apart from being connected to him at a personal level, I am also indebted to him for facilitating my fellowship at CSIR-NCL, which was getting delayed due to some complexities. I am also privileged to attain the advance quantum mechanics classes taken by him. I thank the hostel wardens *Dr. B.L.V. Prasad*, *Dr. Suresh Bhatt*, *Dr. C.P. Vinod*, *Dr. Mugdha Gadgil* and also my guide for managing the hostel facilities with proper care. Finally, I thank *Dr. Sulekha Hazra*, *Ms. Chandana Roy Bardhan*, *Dr. and Mrs. Ramrup Sarkar*, *Dr. Suman Chakrabarty* and *Dr. Debashree Ghosh* for all the recreational entertainments. Among the non-academic staffs, I thank *Ms. Purnima Kolhe*, *Ms. Puranik*, *Deepa* and others from *SAO* for sorting out the official and residential matters. I thank *Ms. Priyanka Saxena* for introducing myself in the *NGO* activities undertaken by *Pankhudi*, *Pune*. The little effort I had given here bore fruit in motivating a dyslexic student pass his 10^{th} standard board examinations and pursue arts further. A brief report of the same has been submitted along with my NCL-403 documents. I thank *CSIR* for funding and *CDAC*, *Pune* for providing computational facilities.

I thank all the teachers and professors who have helped me attain the place where I am today. I specially thank *Dr. A.V. Saha* (*D.Sc.*), *Dr. Prashanta Ghosh* and *Dr. Hrishikesh Chatterjee* of *Ramakrishna Mission Residential College*, *Narendrapur*, *Kolkata* for their motivating classes on fundamental chemistry; *Dr. Asis Nag*, of *St. Xavier's College*, *Kolkata*, for teaching us Physical Chemistry; and *Dr. Chandan Saha* (C*alcutta School of Tropical Medicine*) for inspiring organic classes, undertaken during my graduation times.

Now is the turn to remember my seniors, colleagues, juniors and friends with whom I have interacted throughout my PhD tenure, without which this thesis would not have been possible. First and foremost, I would want to mention my colleague and labmate *Ms. Jaya C. Jose*, with whom I had joined NCL, and worked together in specific problems. I thank my juniors *Asis*, *Sneha*, *Savan* for their support. I want to specially mention *Sneha* and *Savan* for additional help and assistance, whether it is in checking my written documents or discussing any paper or any other topics apart from what was attempted in our group meeting. I thank project trainees *Nupur Bansal* (graduate student at *Michigan State University*), *Brij Kishore Agrawal* (*IIT BHU*), *Aishwary Shivgan* (*IISER Pune*), and *Prabhakar* (PhD. student under *Dr. Thulasiram, presently at IGIB, Delhi*) for collaboratively working with me. I thank *Sumantrada*, *Bhaktidi*, *Subratada*, *Subhadeepda*, *Deeptidi*, *Debaratidi*, *Niveditadi*, *Analda*, *Mudit*, *Jitendra Bhaiya*, *Nageshwar*, *Shashikant*, *Amit Nagare*, *Saba*, *Madhulita*, *Vidhika*, *Saikatda*, *Tamasda* and others for all their support as seniors. I thank my labmates, friends and colleagues: *Late Agnimitra Banerjee*, *Shantanu* (*Kadam*), *Yuvraj*, *Mriyunjay*, *Anagha*, *Amrita*, *Deepak*, *Manoj*, *Nishamol*, *Jugal*, *Baljinder* (*Dr. Grewal!*), *Aryya* (post-doc at *University of Wisconsin*), *Achintya* (post-doc at *University of Moonstar*), *Susanta*

(post-doc at *Bar-Ilan*), *Himadri*, *Turbasu*, *Sudip*, *Manzoor*, *Chandanda*, *Prithvi Raj* (presently working at *Invictus Oncology*, *New Delhi*), *Swagata*, *Subhadip*, *Souvik*, *Mayuresh*, *Anil* (post-doc at *Bar-Ilan*), *Vaibhav*, *Xavier*, *Aishwarya*, and others for their relentless encouragement. I thank my juniors *Sayantan*, *Santu* (presently at *Visva Bharati*), *Rupa*, *Piyali*, *Saikat*, *Sutanu*, *Tamal*, *Subhrashis*, *Tapas*, *Pranab*, *Rahul*, *Samik*, *Mukunda* (graduate student in *Minnesota*), *Sarita* (*NIT Durgapur*), *Diksha* (*HCU*), *Amol* (*NISER Bhubaneshwar*), *Nilesh*, *Bappa*, *Amit*, *Vrushali*, *Niharika*, *Sunipa*, *Paulami* and others for their affectionate nature towards me. I thank my hostel mates and roommates *Priyanka*, *Subhadarshinee*, *Sujit*, *Abhik*, *Arpan*, *Tanay*, *Aniruddha*, *Arijit*, *Anirban*, *Late Subhra Sarkar*, *Chandanda*, *Subashji* (*Dr. Subash Chandra Sahoo*, presently a faculty at *Central University of Punjab*), *Suman*, *Santanu* (*Pattanayak*), *Atanu*, *Bittu*, *Atreyee*, *Monalisa*, *Soumyajyoti*, *Chayanika*, *Sudip*, *Eeshwara*, *Tharun*, *Sachin Patil*, *Manik*, *Vasudevan*, *Goudappagouda* and others for all the entertainment and cooking sessions (on off-days/ while staying at *SA-56*).

At last, I want to mention the constant support of my parents (specially my *mother*) for standing by me during hard times. I sincerely thank my *father* again for constantly encouraging me to proceed forward and not giving up. I pray for my *late maternal grandparents* to rest in peace, without whose efforts, I would not have reached where I am. I thank my *sister* for all the sibling fun, amusement and happiness, and pray for her all-round well-being.

I do not want to thank *The Almighty*, as all that is done is only because of *the unknown volition* taking place in our day-to-day lives; else we have to keep thanking Him in each and every moment. I only want to recall two lines from a song of Tagore: "***Thou hast made me endless, such is Thy pleasure. This frail vessel Thou emptiest again and again, and fillest it ever with fresh life.***"

Prathit Chatterjee

Contents

Abstract

This thesis, titled '*Protein thermal stability, conformational dynamics and solvent properties: insights with atomistic molecular dynamics simulations*' is a computational study of the structural and dynamical behavior of intrinsically disordered and globular proteins at varying thermo-physical conditions. In recent times, novel experimental methods have been developed to probe the interactions and pathways in determining the underlying causes of complexity in protein behavior. These methods are complemented to a large extent by computational methods, which are capable of providing a direct, molecular-level picture. In this work, fully atomistic molecular dynamics (MD) simulations were implemented to investigate the early conformational propensities of a disease associated amyloidogenic protein, and conformational stability of a mesophilic protein coupled to surrounding aqueous environment at thermal denaturation conditions. Structural propensities of native and a point mutant variant of the amyloidogenic protein αSynuclein (αS), responsible for diverse neurological disorders, have been elucidated. The continuing research extends this project to shed light on early nucleation mechanism of the two variants. Investigations on thermal stability of proteins and the correlation with local hydrophobicity and behavior of the surrounding aqueous environment have been studied in detail for a putative mesophilic protein; the observations have been compared with those corresponding to a thermally stable globular protein. Spectral features corresponding to different molecular and solvent properties have been calculated from high frequency simulation runs. In addition, future perspectives and ongoing work to elucidate self-assembly and nucleation of αS, and structural and dynamic properties of nano-confined aqueous systems are discussed. The thesis is divided into six chapters are described below.

Chapter 1: In the introductory chapter, the concept of protein folding and its relevance in present day biophysics have been briefly discussed. Since a part of the thesis deals with propensities of amyloidogenic proteins, the energy landscapes of these systems have been discussed vis-à-vis those of folded, globular proteins. Further, the association of proteins with surrounding solvent environment, and how solvent mediated physicochemical processes are important in biology are discussed. Investigations of solvent properties within hydrophobic nano-channels that can be considered analogous to biological ion channels are briefly elaborated.

Chapter 2: With an overview on the scope of computations in atomic level understanding, the theory behind atomistic molecular dynamics simulations, control of thermodynamic parameters in MD, analyses protocols and advanced MD techniques are been briefly elaborated.

Chapter 3: Early conformational propensities of the native and a mutated variant of the protein αS have been studied in explicit aqueous environment. The mutation was undertaken in the N-terminal domain of

the protein replacing the amino acid Alanine (Ala) with Proline (Pro). There was a transient destabilization in the helical domain due to the mutation, which was not seen in the C-terminal helix and in the native variant. Further, the destabilization was accompanied with a permanent bending in the helical domain of the protein, originating from the rupture of hydrogen bonding network due to Pro mutation.

Chapter 4: The mechanisms of early cold denaturation phenomenon considering the mesophilic protein yeast frataxin (Yfh1), whose range of thermal stability extends from 280 K to 323 K, are elucidated, and their results have been compared with the corresponding behavior of thermally stable protein, Ubq. For Yfh1, denaturation at lower temperatures occurs preferentially for the beta domains, in qualitative agreement with recent NMR experiments with Yfh1. The disruption has been associated with greater hydration propensity, plausibly due to decreased hydrophobicity of the corresponding domains, reflected from calculated IR spectra, and the dynamical and structural propensities of the surrounding hydration water layer. Ubq does not show such phenomena on the simulated time scales.

Chapter 5: The present chapter deals with probing the behavioral patterns of Yfh1 throughout its range of thermal stability and comparing those with Ubq under similar thermodynamic conditions. Signatures of decreased hydrophobicity are preferentially found for the beta domains of Yfh1 at cold denaturing conditions, along with a minor increase in their effective hydrophobicity at higher temperatures, reflected from the corresponding water density distributions. Helices show negligible changes in their hydrophobicity and significant increase in hydrophobicity at cold and hot temperatures. Increased rattling motions, elucidated from velocity autocorrelation functions (VACF) and vibrational density of states spectra, occur selectively for water surrounding the beta domains of Yfh1, with increase in temperature. Similar preferential patterns as a function of temperature are not exhibited by the structural domains of Ubq and helices of Yfh1, in its structural propensities and corresponding hydration behavior.

Chapter 6: Future perspectives in relation to continuing studies are discussed. As differences in early conformational propensities are found to exist in the two variants of αS, mechanistic differences in oligomerization propensity between the two variants are investigated. It is found that αS tends to form oligomers with greater number of monomeric units, regardless of the mutation state. Inter-monomer interactios are facilitated by the central hydrophobic core of the protein. In addition, properties of water under nano-confinement, analogous to the water flow within biological channels, are studied using artificial channels created with single-walled carbon nanotubes (SWNT) and graphene sheets. Ongoing investigations focus on delineating structural and dynamical properties of nano-confined water in presence of co-solvents with reduced polarity.

List of Figures

Figure 2.4 Structural analyses of mesophilic protein, yeast frataxin (Yfh1), at three different temperatures. a) plots of MSF as a function of the residues; b) evolution of P, and c) R_g, with simulation time.

Figure 2.5 Orientational autocorrelation function ($P_2(t)$) for the dipole moment vector of bulk water and the hydration water molecules surrounding different domains (beta, helix, $beta_H$ (or hydrophobic beta), $helix_H$ (or hydrophobic helix)) of Yfh1 at different temperatures.

Figure 2.6 Plots of *g(r)* between the backbone heavy atoms and water oxygens for the different beta strands of Yfh1.

Figure 3.1 Snapshots of the WT (upper row) and the A30P variant (lower row) taken at 10ns, 45ns and 80ns of the respective trajectories. The N-helix of WT and A30P are shown in blue, and yellow, respectively. The proline mutation in A30P is marked with CPK representation. C-helices are shown in orange.

Figure 3.2 Evolution of the backbone RMSDs of WT (*in black*) and A30P (*in red*) as a function of simulation time.

Figure 3.3 Plots of the persistence parameter *P* of the WT (*in black*) and A30P (*in red*) variants as a function of simulation time, shown for *a)* the whole peptide segments, *b)* the N-helix, *c)* the C-helix.

Figure 3.4 Plots of the persistence parameter *P* for 3 residue groups in the N-helix of the WT (*in black*) and A30P (*in red*) variants as a function of simulation time.

Figure 3.5 Plots of the persistence parameter *P* for 3 residue groups in the C-helix of the WT (*in black*) and A30P (*in red*) variants as a function of simulation time.

Figure 3.6 Histograms of bend angles θ_1 and θ_2 for the WT (*in black*) and A30P (*in red*) variants, taken over 0-20 ns, 30-50 ns and 70-90 ns.

Figure 3.7 Evolution of the bend angles θ_1 and θ_2 for the WT (*in black*) and A30P (*in red*) variants as a function of simulation time.

Figure 3.8 Plots of the backbone RMSDs for *a)* N-Helix and *b)* C-Helix as a function of simulation time for the WT (*in black*) and A30P (*in red*) variants.

Figure 3.9 Plots of the hydrogen bonding network energy as a function of simulation time for N-Helix (upper row) and C-Helix (lower row) as a function of simulation time for the WT (*in black*) and A30P (*in red*) variants. The hydrogen bonding network energy for each helix is calculated as the total non-bonded interaction energy between the backbone amide groups and the oxygen atoms of the backbone carbonyl groups.

Figure 4.1 Representative snapshots, distributions of R_g, distributions of *SASA* and contact maps. For the snapshots, helices (H1, H2) and the turn & coils are shown in gray; beta strands (S1, S2, S3, S4, S5 &S6) are shown in blue, pink, purple, green, violet and magenta respectively. For the contact maps, lower triangle denotes backbone contacts while upper triangle denotes sidechain contacts. Representative snapshots of the protein at a) T_s and b) at T_c;c) distributions of R_g at T_s *(black)* and T_c *(blue)*, d) distributions of SASA at T_s *(black)* and T_c *(blue)*, e) contact map at T_c and f) contact map at T_s.

Figure 4.2 a) Residue-wise ΔP (or $P_{Ts} - P_{Tc}$) plots; and b) residue-wise ΔC_v (or $C_{vTs} - C_{vTc}$) plots as a function of the residue numbers (N_{res}). Units of C_v are in $JK^{-1}\ Mol^{-1}$.

Figure 4.3 Radial distribution functions ($g(r)$) and interaction energy of the whole protein with its hydration layer. a) Plots of $g(r)$ between backbone carbonyl carbon atoms of protein and water oxygen atoms, b) plots of $g(r)$ between all heavy atoms and water oxygen atoms, c) distributions of interaction energy of the whole protein with its hydration layer; at T_s *(black)* and T_c *(blue)*.

Figure 4.4 2D-contour plots of P versus N_w for; a) whole protein at T_c, b) whole protein at T_s, c) beta domains at T_c, d) beta domains at T_s, e) helical domains at T_c, f) helical domains at T_s.

Figure 4.5 IR absorption spectra for the a) whole protein, b) S2, c) S4, d) S5, and e) representative helix H1; at T_s *(black)* and T_c *(blue)*.

Figure 4.6 Second-rank rotational correlation function ($P_2(t)$) for dipoles of bulk water and hydration water layer of selected domains at a) T_c and b) T_s; c) probability distributions of water densities near the surfaces of sidechains of the selected domains at T_s *(black)* and T_c *(blue)*.

Figure 4.7 Distributions of the tetrahedral order parameter (Q) for bulk water, hydrophobic beta sheet domain and hydrophobic alpha helical domain at a) T_s and b) T_c. Insets show the expanded views for the range Q = 0.5 – 0.8. P(Q) denotes the probability of the corresponding Q distributions, while $P(Q)_{max}$ corresponds to the maximum value of the corresponding probability

Figure 4.8 a) Representative snapshot at T_s, b) Representative snapshot at T_c, c) distribution of R_g at T_s *(black)* and T_c *(blue)*, and d) distribution of *SASA* at T_s *(black)* and T_c *(blue)*; of thermally stable *Ubq*.

Figure 4.9 a) $g(r)$ of Ubq between backbone carbonyl carbon atoms of the protein and water oxygen atoms at T_s *(black)* and T_c *(blue)*, b) $\Delta g(r)$ of Ubq (black straight line) and Yfh1 (black dashed line). c) $g(r)$ for the $Beta_H$ domains of Yfh1 at T_s *(black)* and T_c *(blue)*, d) $g(r)$ for the $Beta_H$ domains of Ubq at T_s *(black)* and T_c *(blue)*, e) $g(r)$ for the $Helix_H$ domains of Yfh1 at T_s *(black)* and T_c *(blue)*, f) $g(r)$ for the $Helix_H$ domains of Ubq at T_s *(black)* and T_c *(blue)*.

Figure 4.10 2D-contour plots of P versus N_w for; a) whole protein, b) beta domains, c) helical domains, at T_h; and comparison of the probability distributions of water densities near the surfaces of sidechains of d) $Beta_H$ and e) $Helix_H$, at T_s *(black)*, T_c *(blue)* and T_h *(red)*; for Yfh1.

Figure 4.11 Cumulative configurational entropy per atom of Yfh1 for the backbone heavy atoms of a) the whole protein, b) the beta domains, and c) the $beta_H$ domains; at T_s *(black)*, T_c *(blue)* and T_h *(red)*.

Figure 5.1 Sequence and secondary structure obtained from the PDB structures of a) Yfh1 (2GA5), and b) Ubq (1UBQ). Hydrophobic and non-hydrophobic amino acids are marked with green and orange, respectively. Alpha helical domains are indicates in red; 3_{10} helix in pink; beta sheets in yellow; beta bridges in tan; turns in purple; and coils in black.

Figure 5.2 Plots of $R^{/}$ and $P^{/}$ for Yfh1 and Ubq as a function of temperature, as well as representative snapshots of Yfh1 and Ubq at selected temperatures. a) evolution of $R^{/}$ of Yfh1 at

all the simulated temperatures; snapshots of Yfh1 at b) T_c, c) T_s, d) T_h; e) evolution of $P^{/}$ of Yfh1 at all the simulated temperatures; snapshots of Ubq at f) T_c, g) T_s, f) T_h.

Figure 5.3 Probability distributions of P_{wh} *vs.* N_w (obtained from the last 50 ns of simulations) for Yfh1 (upper row) and Ubq (lower row), at all the simulated temperatures.

Figure 5.4 Probability distributions of P_{beta} *vs.* N_{beta} (obtained from the last 50 ns of simulations) for Yfh1 (upper row) and Ubq (lower row), at all the simulated temperatures.

Figure 5.5 Probability distributions of P_{hlx} *vs.* N_{hlx} (obtained from the last 50 ns of simulations) for Yfh1 (upper row) and Ubq (lower row), at all the simulated temperatures.

Figure 5.6 Distributions of $P(n^*)$ (obtained from the last 50 ns of simulations) for a) $beta_H$, and b) $helix_H$ of Yfh1, at the simulated temperatures; and c) $beta_H$, and d) $helix_H$ of Ubq, at the simulated temperatures; e) $beta_{NH}$ domains of Yfh1, at temperatures T_c, T_s, T_h.

Figure 5.7 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the whole protein domains of Yfh1 and Ubq, compared with the corresponding bulk behavior, at the selected temperatures (T_c, T_s, T_h).

Figure 5.8 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the whole protein domains of Yfh1 and Ubq, compared with the corresponding bulk behavior, at the selected temperatures (T_{c1}, T_{h1}, T_{h2}).

Figure 5.9 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, $beta_H$, and $helix_H$ domains of Yfh1, compared with the corresponding bulk behavior, at the selected temperatures (T_c, T_s, T_h); differences in VDOS for beta & $beta_H$ and helix & $helix_H$ are represented by the magnified plots in the respective insets.

Figure 5.10 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, $beta_H$, and $helix_H$ domains of Yfh1, compared with the corresponding bulk behavior, at the selected temperatures (T_{c1}, T_{h1}, T_{h2}); differences in VDOS for beta & $beta_H$ and helix & $helix_H$ are represented by the magnified plots in the respective insets.

Figure 5.11 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, $beta_H$, and $helix_H$ domains of Ubq, compared with the corresponding bulk

behavior, at the selected temperatures (T_c, T_s, T_h); differences in VDOS for beta & $beta_H$ and helix & $helix_H$ are represented by the magnified plots in the respective insets.

Figure 5.12 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, $beta_H$, and $helix_H$ domains of Ubq, compared with the corresponding bulk behavior, at the selected temperatures (T_{c1}, T_{h1}, T_{h2}); differences in VDOS for beta & $beta_H$ and helix & $helix_H$ are represented by the magnified plots in the respective insets.

Figure 5.13 Plots of $R^{/}$ and $P^{/}$ for Yfh1 and Ubq as a function of temperature, as well as representative snapshots of Yfh1 and Ubq at selected temperatures, with TIP4P/2005 water model. a) evolution of $R^{/}$ of Yfh1 at all the simulated temperatures; snapshots of Yfh1 at b) T_c, c) T_s, d) T_h; e) evolution of $P^{/}$ of Yfh1 at all the simulated temperatures; snapshots of Ubq at f) T_c, g) T_s, f) T_h.

Figure 6.1 Initial setup, in implicit solvent, of the a) dimer (d_N), b) trimer (tr_N), and c) tetramer (tt_N) systems of WT αS. Partially helical monomer conformations are generated from Accelerated MD (AMD).

Figure 6.2 Evolution of the clusters from the cartesian PCA for the tt_N system. Two-dimensional distribution of the density function ΔG, on the first two principal components pC1 and pC2, has been illustrated; the cluster corresponding to the greatest population (C1) is additionally presented.

Figure 6.3 Selected plots of contact maps between the monomers for a) C1 of tr_M, b) C3 of tr_M, c) C1 of tt_N, d) C3 of tt_N.

Figure 6.4 Selected plots of cumulative configurational entropy of the monomeric units (αS_1, αS_2, αS_3, and αS_4; denoted in *black*, *red*, *green* and *blue*) in a) C1 of tr_M, and b) C3 of tt_N.

Figure 6.5 Representative snapshots of the starting structures for the water-SWNT-Mix systems. Upper solvent-reservoir contains water (yellow) and HFIP molecules. Only pure water molecules (blue) are incorporated in the lower reservoir.

Figure 6.6 *f* as a function of nanotube length for systems with different N_{HFP}.

List of Tables

List of Publications

1. **Chatterjee P** and Sengupta N, Signatures of Protein Thermal Denaturation and Local Hydrophobicity in Domain Specific Hydration Behavior: A Comparative Molecular Dynamics Study. (*under revision*)

2. Rohamare S, Gaikwad S, Jones D, Bhavnani V, Pal J, Sharma R, **Chatterjee P**, Cloning, expression and *in silico* studies of a serine protease from a marine actinomycete, (2015), *Process Biochem*. 50, 378-387.

3. **Chatterjee P**, Bagchi S and Sengupta N, The non-uniform early structural response of globular proteins to cold denaturing conditions: A case study with Yfh1, (2014), *J. Chem. Phys*. 141, 205103.

4. Jose J C, **Chatterjee P** and Sengupta N, Cross Dimerization of Amyloid-β and αSynuclein proteins in Aqueous Environment: A Molecular Dynamics Simulations Study, (2014), *PLOS One* 9, e106883

5. **Chatterjee P** and Sengupta N, Effect of the A30P mutation on the structural dynamics of micelle-bound αSynuclein released in water: a molecular dynamics study, (2012), *Eur. Biophys. J*. 41, 483 – 489.

Manuscripts to be submitted

1. **Chatterjee P**, Agrawal B and Sengupta N, Effect of hexa-fluoro-2-propanol on the dynamics and structure of water within nano-channels: a molecular dynamics simulations study.

2. **Chatterjee P**, Shivgan A and Sengupta N, Investigations of the early oligomerization propensity of the wild type and a point mutant αSyncuclein: a

3. Srivastava P L, Ramakrishnan K, Daramwar P P, **Chatterjee P**, Mantri S T, Rincy R, Kumar S, Sengupta N and Thulasiram H V, Insight for highly evolved sesquiterpene synthase involved in biosynthesis of complex and highly strained molecule.

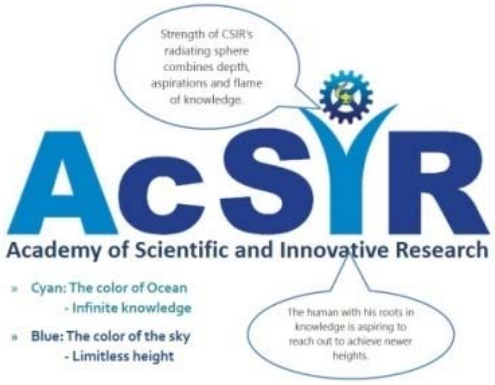
Abbreviations

αS	αSynuclein
Aβ	Amyloidβ
A30P	Alanine 30 Proline mutation
AFM	Atomic force microscopy
AMD	Accelerated molecular dynamics
CD	Circular dichroism
CHARMM	Chemistry of Harvard Molecular Mechanics
CNT	Carbon nanotube
comd	center of mass distance
DFT	Density functional theory
DNA	Deoxyribonucleic acid
DWNT	Double-walled carbon nanotube
FFCF	Frequency-frequency correlation function
FRET	Förster resonance energy transfer
GBIS	Generalized born implicit solvent
HFIP	Hexa-fluoro-2-propanol
IDP	Intrinsically disordered protein
IR	Infrared
KE	Kinetic energy
LJ	Lennard Jonnes
MD	Molecular dynamics
MSD	Mean squared displacement
MSF	Mean squared fluctuation
MWNT	Multi-walled carbon nanotube
NAC	Non-amyloid component
NAMD	Nanoscale molecular dynamics
NEMD	Non equilibrium molecular dynamics
NMR	Nuclear magnetic resonance
NMRD	Nuclear magnetic resonance dispersion

NNC	Nearest neighbor coupling
PCA	Principal Component analysis
PD	Parkinson's disease
PDB	Protein database
PME	Particle Mesh Ewald
QM/MM	Quantum and molecular mechanics
R_g	Radius of gyration
RMSD	Root mean squared deviation
RMSF	Root mean squared fluctuation
RNA	Ribonucleic acid
SASA	Solvent accessible surface area
SWNT	Single-walled carbon nanotube
TDC	Transition dipole coupling
Ubq	Ubiquitin
VACF	Velocity autocorrelation function
VDOS	Vibrational density of states
vdW	van der Waals
VMD	Visual molecular dynamics
WT	wild type variant
Yfh1	Yeast frataxin

Chapter 1

Introduction



Scientists… proceed by common sense and ingenuity. There are no rules, only the principles of integrity and objectivity, with a complete rejection of all authority except that of the fact.
- Joel H. Hildebrand

1.1 Proteins and peptides

Proteins, the building blocks of living organisms, are large macromolecules constituted of multiple amino acids. There are 20 different amino acids present in living systems, which, when connected to each other by peptide bond linkages form proteins. The biosynthesis of proteins (pictorially represented in Figure 1.1) takes place in the ribosome of the cell, where the amino acids are assembled to form a long polymeric chain of amino acids in presence of translational RNA (Ribonucleic acid). The sequence specificity of each protein encoded in DNA (deoxyribonucleic acid)[1] is transported to the ribosome of a cell by the messenger RNA. The term 'proteome' (analogous to the term 'genome' which signifies the genetic material of an organism) represents all protein molecules present in the living system. Proteomes can change with time in their overall content based on different factors, which is responsible for heterogeneous functional properties of a living system at different times.[2] Proteins can be even functionalized with different groups and/or metal-ions for specific functional properties,[3] and provide the essential machinery for carrying out biological activities such as enzymatic processes, regulation of cellular activities, signal transmission, vesicle trafficking, molecular and ion transport, binding mechanisms and developmental roles in organisms.[4-12]

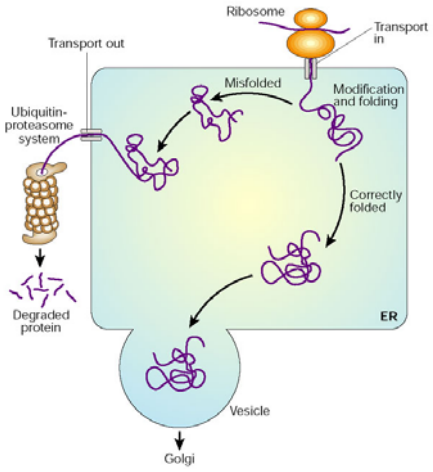


Figure 1.1 Regulation of protein biosynthesis in the ribosome prior to the accurate folding in cell cytoplasm. Misfolded sequences obtained in folding funnel are excreted out and degraded further. [Reproduced from ref. 1 with prior permission from *Nature* publishing group]

Almost 50% of the overall cell content is composed of proteins, whereas DNA and RNA only contribute 3% and 20% respectively.[13] Based on structure and function, proteins can be categorized into different types. The intracellular matrix consists of cytoskeleton proteins that

define the shape and size of living cells. Globular proteins comprise most of the proteome and have definitive three-dimensional structure and function. There are transmembrane proteins present in cell membranes with a hydrophobic exterior and hydrophilic interior. These are responsible for transport of ions and other essential components from outside to inside of a cell and vice-versa. Besides, there are hormones, which control physiological and behavioral activities; and enzymes, which catalyze different biochemical processes like digestion. Further, 20-30% of the total proteome consists of intrinsically unstructured proteins,[14] the details of which will be discussed in a later sub-section. Also, there are smaller proteins called molecular chaperones present in ribosome,[14] assisting the protein folding process. Peptides are short fragments of a protein, generally constituting of not more than fifty amino acids. They are oftentimes generated by enzymatic cleavage from the mother protein,[15] although they may be generated in a similar procedure as proteins. Therefore, proteins can either be constituted of one peptide unit[16] or a collection of polypeptides and other biological sub-units.[3] The orientations of peptide units as well as its sequences are responsible for the functionalization of proteins.

1.1.1 The protein folding process

Peptides and proteins, synthesized as predominantly unstructured polymers of amino acids, typically attain a unique 3-dimensional structure associated with distinctive functional properties, after post-translational modification or specific functionalization of some residues. This phenomenon, which is perhaps the most essential transformation undergone by a protein, is known as protein folding.[17] The timescales over which an unstructured polypeptide achieves its correctly folded form varies in the range of several microseconds to a few minutes.[14] Interestingly, in 1969, Cyrus Levinthal had postulated that if a standard amino acid chain samples all the possible conformations prior to attaining the native structure, the required time for the folding process may transcend the age of the universe. The existence of finite folding timescales suggests that proteins adapt only selected intermediate states in its pathway.[18]

The attainment of the folded state within finite timescales further suggests that protein folding is not a stochastic procedure, and there have been detailed investigations to probe the driving force behind the process of protein folding.[19] As mentioned earlier, the 3-dimensional structural representation of a protein is embedded in its sequence.[20] However, understanding the pathways undertaken by a nascent protein chain on the way to its folded form, and the

heterogeneous states encountered, are critically important. Entropic studies of their folded polymeric states have revealed that conformational states with higher stability have lesser number of conformations, suggesting a "funnel shaped" free energy landscape for folding (Figure 1.2).

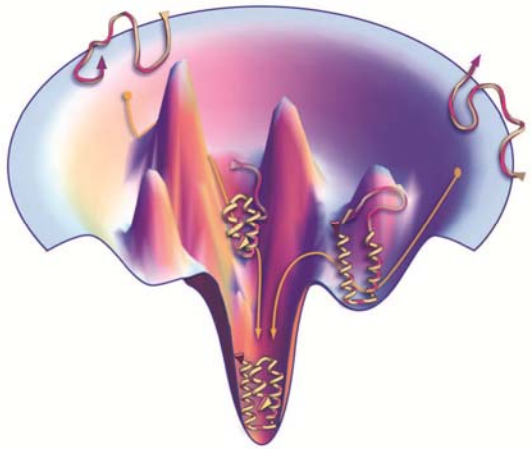


Figure 1.2 Representation of the free energy landscape of protein folding. [Reproduced from ref. 28 with prior permission from *American Association for the Advancement of Science*]

Presumably, the conformationally folded states are distinguished by thermal barriers, and the extent of conformational heterogeneity diminishes with advancement in the folding process. The energetic barriers as well as heterogeneity in folded intermediate states determine the kinetics of protein folding, and further in maintaining the dynamic demeanor of folded intermediates.[21-24] Investigations have been carried out on the kinetics of protein folding to detect the partially folded states in order to elucidate the reason behind preference of certain intermediates in the folding pathway. These involved experimental techniques such as hydrogen atom exchange, mutational analysis, fluorescence and other spectroscopic methods.[21,25-27] Most experiments have indicated that the folded state is characterized by reduced conformational entropy, and thus folding is predominantly an enthalpy driven process.[28] It is important to note that the additive downhill folding funnel shaped energy landscape discussed here is applicable only for folded globular proteins. The energetics of other proteins like intrinsically disordered proteins (IDPs) or membrane proteins are entirely different; IDPs will be briefly discussed in the next section.

The understanding of the folding process has increased substantially by the discovery of folded structures of more than 100,000 globular proteins (http://www.rcsb.org/).[29] The predominant factors that contribute to the overall energetics are i) hydrogen bonds governing the secondary structural propensity of the protein; ii) short-range interactions between the neighboring moieties; iii) steric interactions; iv) short and long range interactions between the polar or charged moieties; v) hydrophobic interactions present among the non-polar residues; vi) entropic decrease of the folding chain; vii) solvent influences.[17,28]

Several models have been proposed to characterize the protein folding pathway. One of the initial models were 'ion-paired hydrogen-bonding model' which was thought to drive the protein folding process.[17] It further proposed that folding occurs with a decrease in volume, but in reality folding is followed by a corresponding increase in volume. Later, the 'one-dimensional Ising model' supported favorability of the intra-chain hydrogen bonds over the peptide-solvent hydrogen bonds.[17] But formation of such bonds needed an overcoming of the configurational entropy of the chain and arrangement of the adjacent bonds in proper secondary structural configuration. A more recent model based on effective ion-paring is the 'polyelectrolyte model' which proposes that the unfolded polypeptide is energetically more stable than the folded state at increased salt concentration.[17] At low pH, increase in salt concentration shields the charge-repulsion more effectively for the unfolded peptide chains, as salt can penetrate more when a protein is unfolded. This will eventually promote folding of proteins in less ionic environments with oppositely charged ion-pairs in near proximity. Among the recently proposed models, the 'framework model' stresses the fact that after immediate generation of the protein, it attains nascent secondary structure, followed by a compact molten-globule form, which further develops into the formation of tertiary structure.[30] It is important to mention here that molten globules, representing previously mentioned folded intermediates, lack definitive tertiary contacts due to disordered sidechain packing, although their structural characteristics are similar to the natively folded protein.[21,22] The more prevalent 'hydrophobic collapse' model postulates the non-specific collapse driven by the hydrophobic effect as the predominant driving force of protein folding.[31]

1.1.2 Protein stability and denaturation

Globular proteins tend to denature at large deviations from physiological conditions. Denaturation can occur at temperatures higher or lower than normal physiological temperatures,

at high pressures, in presence of co-solvents and denaturing agents, with point mutations introduced at specific secondary structures, or with changes in pH of the solvent medium.[32-36] High pressure denaturation mechanism of proteins has similar characteristics to that obtained through chemical or pH induced denaturation, which involve disruption of the hydrogen bonding network with surrounding molecules for the secondary structural domains.[37] Further, theoretical investigations have suggested resemblance in the heat induced and chemical denaturation processes where both exhibit an expansion of hydrophobic core before denaturation.[38]

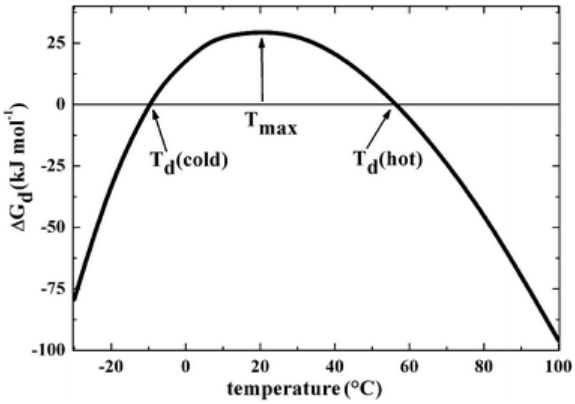


Figure 1.3 The free energy difference between folded and unfolded states of a globular protein as a function of temperature, calculated from a model study. [Reproduced from ref. 39 with prior permission from the *Royal Society of Chemistry*]

Protein thermal denaturation has been an extremely important topic, especially in addressing the mechanistic aspects of protein folding. Melting of the folded protein states occur when they deviate from the corresponding temperature of maximal stability to higher or lower temperatures.[33] The magnitude of free energy difference between the folded and unfolded states attains a maximum at the temperature of maximum stability, but correspondingly changes with increase or decrease in temperature in either direction (Figure 1.3), and is negligible at temperatures where cold and heat denaturation mechanisms are initiated.[33,39] While heat denaturation is often easily interpreted due to experimental accessibility of the thermodynamical conditions, phenomenon of cold denaturation of proteins is partially interpreted due to hindrance in obtaining the cold thermodynamical states experimentally. The hydrophobic effect, as stated before, is believed to be one of the predominant factors governing the folding of proteins and stability of its native structure.[31] Disruption of the hydrophobic effect at cold temperatures, when the entropic gain of water molecules expelled from the protein core to the bulk reduces sharply,

is considered one of the central reasons behind cold denaturation phenomena.[33,39] In the fourth and fifth chapter of this thesis, the mechanism of cold denaturation of a mesophilic protein[40,41] is attempted to be interpreted and contrasted with the corresponding heat denaturation mechanism, and further compared with the thermal response of a thermostable globular protein.

1.1.3 Protein misfolding and aggregation

It has been established that a protein's function is intricately dependent on both its structure and its essential dynamics. At normal physiological conditions, the sequence of a globular protein is inclined towards attaining its stable structure. Subtle changes in the structure arising from physico-chemical changes in the environment as well as point mutations may result in the formation of misfolded structures, and thereby annihilation of protein functions. Although the competing entropic and enthalpic factors present in the chain sequence predominantly control the thermodynamics of protein folding, these may be further influenced by externally binding peptides of relatively smaller size, known as molecular chaperones, which help maintain the delicate balance of forces responsible for protein folding.[1,14] Small proteins can refold without any assistance of external agents, but large multi-domain proteins take minutes to hours to fold. Misfolded protein conformations have been associated with several diseases, the mechanisms behind which have recently gained significant attention.[1,42,43] To avoid attaining misfolded conformation, larger sized proteins thus need molecular chaperones in order to attain fully folded stable structure efficiently within biological timescales.[14]

As stated previously, globular proteins, although prone to forming stable native structures, may have a tendency to unfold with altering thermodynamical conditions, like exposure to high temperature, changes in pH, mutations, and additions of denaturant.[44] Due to alteration in the balance of forces, and loss in external control, misfolded or partially folded states can become energetically more favorable than the native conformation. The energy landscapes of structured peptide or aggregates have further proven that the aggregated state is situated at a deeper minimum with a high energy barrier in a separate non-native funnel.[44,45] Such states typically attain aggregated cross-beta sheet structures, that are continuous parallel beta sheets arranged perpendicular to the fibril axis.[44] These beta-sheet aggregates are structurally similar for most cases and known as amyloidogenic states, resembling the structured aggregated states obtained from amyloidβ peptides. The misfolding and aggregation due to

amyloidβ are further known as amyloidosis, as the aggregated protein fragments gets similarly stained by Congo red dye as the starch (amylose).[44] Few of the predominantly studied misfolding diseases are Alzheimer's disease, Parkinson's disease, Renal Amyloidosis and Type-II diabetes, which happen due to misfolding and aggregation of the proteins amyloidβ, αSynculein, lysozyme and amylin respectively.[42,46-48] It is worth noting here that nearly 30% of proteins present in eukaryotic cells are intrinsically unstructured, and only attain stable folded states in close proximity with membranes and other macromolecules[14] as they lack proper free energy funnel arising due to ruggedness in their energy landscape.[44] Such intrinsically disordered proteins (IDPs) have specific structures and functions like signal transmission and vesicle trafficking when they are soluble in cell environments, but are akin to misfolding and further amyloidosis.

There are several proposed pathways of amyloid fibril formation.[49-51] Partially folded intermediates along the folding pathway sometimes promote pre-fibrillar oligomeric species, responsible for protein aggregation.[1] The domains containing the hydrophobic residues are believed to initiate protein aggregation.[52] Further, enthalpic association of the hydrophobic residues due to intra-peptide backbone hydrogen bonding, along with exposed side-chains to the outer surface, is considered to trigger the aggregation pathway.[50] Water molecules in close proximity to the associated hydrophobic domains are expelled out from the core, thus incorporating a dewetting mechanism in aggregation.[51] So water molecules have a significant role to play in amyloid aggregation. But, little is known about the early mechanistic aspects of the aggregation pathway in solvent environments, especially with IDPs which lack definitive structure in cell cytoplasm. Also, investigation of the mechanistic processes of soluble oligomers forming insoluble fibrils, remain elusive. Further, as stated earlier, it is yet unclear whether pre-fibrillar soluble oligomers or aggregated fibrils are the toxic species. In the third chapter of this thesis, early onset structural and dynamical propensities of the wild-type (WT) and a naturally occurring point mutant αSynculein, an intrinsically disordered protein, will be discussed in aqueous environment. The ongoing work and future perspectives on the conformational propensities of the oligomeric αSynculein will be briefly discussed in the concluding chapter.

1.1.4 Protein dynamics

Despite the correlation between the structure and biological function of proteins, the latter should not be considered as immovable and inflexible bodies. Rather, proteins can be

constitutively categorized as a "soft matter" with an extensive flexibility, such that the folded native state includes a conformational ensemble composed of rapidly fluctuating substates.[53] The corresponding functions of the protein are dependent on both the dynamical fluctuations as well as the natively folded state.[54]

The dynamics of a protein molecule can be categorized into three different time domains.[55] The atomic motions, bond vibration, formation as well as breakage of hydrogen bonding are phenomena corresponding to the fastest protein dynamics, occurring on the sub-picoseconds timescales. The immediately slower protein dynamics occurring at picoseconds to nanoseconds timescales correspond to the small-scale concerted mobility of atoms in groups, such as the rotational movements of the methyl groups and sidechains and thermally induced displacements of loops, turns and coils. The substates in the natively folded ensemble can be distinguished by these sub-microsecond timescale dynamics.[56] The slowest protein dynamics, consisting of concerted movements of larger magnitudes involving individual domains of a protein, take place on the micro- to millisecond timescale. These are responsible for differentiating the specific kinetic states responsible for varied biological processes.[55] Recent research shows that the long timescale dynamics is intricately coupled with the faster short timescale dynamics and the latter could play in controlling the collective large amplitude protein dynamics responsible for specific functional properties of proteins.[57]

1.2 Solvent influences

In physicochemical reaction pathways, the surrounding solvents have a definitive role to play in controlling the processes. The reactivity of the substances is often tuned due to viscosity and polarity of the surrounding solvent environment. Importantly, most biological systems are ubiquitously surrounded by aqueous environment, which crucially affects their behavior and functional properties.[58] Water is thus an essential component and lubricant of living systems. The properties of water in the vicinity of biological moieties is found to be distinctly different from behavior of bulk water.[58] The present thesis is based on work done with proteins in aqueous media, and therefore, the roles of solvents in different systems are briefly discussed.

1.2.1 Role of solvent in protein stability

While it is known that large scale motions of proteins are coupled with those corresponding to small amplitude motions, the surrounding solvent environment also have a role in maintaining the structure, function and dynamics of proteins. The change in solvent polarity effectively alters the structural rigidity as well as the biological functional properties of protein. A decrease in polarity rigidifies protein structure, due to lower competition for the intramolecular hydrogen bonded structure by the surrounding solvent.[59] Recent studies have even depicted a retention and rigidification of the helical content of proteins with increase in the non-polar solvent molecules in the surrounding environment.[60] On the contrary, increasing polarity causes accelerated protein denaturation, as formation of intermolecular hydrogen bonds due to solvent competes with the corresponding intramolecular hydrogen bonding.[61] Further, the viscosity of the solvent affects the chemical kinetics of the protein, in controlling the rates of enzyme catalysis and ligand binding.[62] The decrease in viscosity increases the conformational fluctuations present in proteins, which facilitates binding of ligands with proteins through the corresponding solvent.[63] Subsequently, the rates of unfolding of globular proteins can be regulated by changing the solvent viscosity of the bulk aqueous environment.[64]

Proteins undertake a majority of their biological functions in predominantly aqueous environments. Water is responsible for regulating the structural and dynamical interactions responsible for the functional properties of proteins. The water layer present in close proximity to the protein surface, is dynamically heterogeneous and differs in relaxation behavior, polarity and viscosity, from water in the bulk phase.[58,65-69] These water molecules in close vicinity to the protein surface are known as 'biological water', and are responsible for controlling the structure, and eventually the dynamical and functional properties of proteins, mediating protein-protein interaction, protein-ligand binding, etc.[58,70,71] Water near the protein surface has sometimes also been referred to as the protein 'hydration layer'.[71]

Experiments and computer simulations have together established that the relaxation rates of 'biological water' are slower than those of water in bulk phase.[58,66,68] Also, the behavior of waters around different domains of the same protein may differ due to differential anisotropic effects in protein-water interactions.[58] Consequently, the hydration layer water molecules may be tightly restrained (bound water) or weakly bound to the protein surface, the effect of which is

reflected in the dynamical deviation from bulk characteristics.[58] The slower dynamics and relaxation rates of bound water thus depend on the binding affinity to the proximal protein surface.[70] Also, bound water molecules are helpful in retaining the conformational stability of the protein molecule. Recently, it has also been shown that density distribution of water molecules in close proximity to a surface reflects the effective hydrophobicity of the corresponding surface.[72] Thus, signatures of the chemical nature of the protein surface can be found in the density fluctuation in the hydration layer.[72]

Despite overwhelming evidence of the coupling of solvent and protein behavior, the exact role of water in controlling protein structure and function remains a topic of debate. It has been argued that the essential properties of proteins are encoded in its sequence, and thus the surrounding solvent should only have a secondary, modulating role.[73] Neighboring water molecules get expelled from the inner hydrophobic core of proteins, thereby promoting hydrophobic binding in proteins and corresponding acquiring of globular structure. Reducing the aqueous interactions with non polar amino acids is undertaken with a simultaneous burial of the hydrophobic residues in the respective hydrophobic core of proteins. Proteins remain in 'glassy' state and undergo dynamical transition with increase in temperature.[74] The amplification of such behavior in hydrated environments signifies the role of water in protein dynamical behavior.[27,75,76]

1.2.2 Properties of nano-confined water

It is evident from the discussion above that structural and dynamical properties of water molecules associated with proteins and biomolecules have key differences from those present within the bulk phase.[77] Systems with water molecules entrapped within nano-pores and nano-channels are therefore relevant in explaining the mechanism of ion and water transport through water channels,[78] which are essentially transmembrane proteins.[79-81] The hydrophobic exterior of these channels interacts with the non-polar lipid chains to posit them across cell-membranes (Figure 1.4). Flow of water and selectivity of ions through the pore are regulated by the pore diameter, as well as by conserved hydrophobic residues present along the pore wall promoting quicker water transport through the channel. It is to be noted that these biological channels have widths between 10-20 Å.[82] In addition, studies of the structural and dynamical properties of water and water-like solvents in nano-confinement are useful in elucidating the basic aspects of

chemical separation and nano-electronic gating devices.[78,83-85] Thus, the flow of confined water within artificial nanochannels should have some resemblance with the transport properties of water in biological channels.

Unlike bulk water, the hydrogen bonding of water molecules in nano-confined environments are restricted to only two of the neighboring water molecules, resulting in significant reduction in their tetrahedral character.[84,86] There is a significant variability in the dynamics of water molecules in nano-channels with the dimensions of the channels and the physiochemical characteristics of materials present in it.[87-89] Further, small molecules, voltage-gated functionalized channels, charged ions and other co-solvents (with altered polarizability and permittivity) can significantly affect the properties of water molecules.[90-93] The translational and rotational dynamics have been observed to be coupled to each other in a concerted progression of water molecules in mono-lamellar flow of water within nano-channels.[84] Further, the flow of water molecules has been found to exponentially decrease with increase in length of the nano-channels.[94] In the concluding chapter, ongoing work on flow and dynamics of water molecules with differing lengths of hydrophobic nano-channels has been briefly discussed.

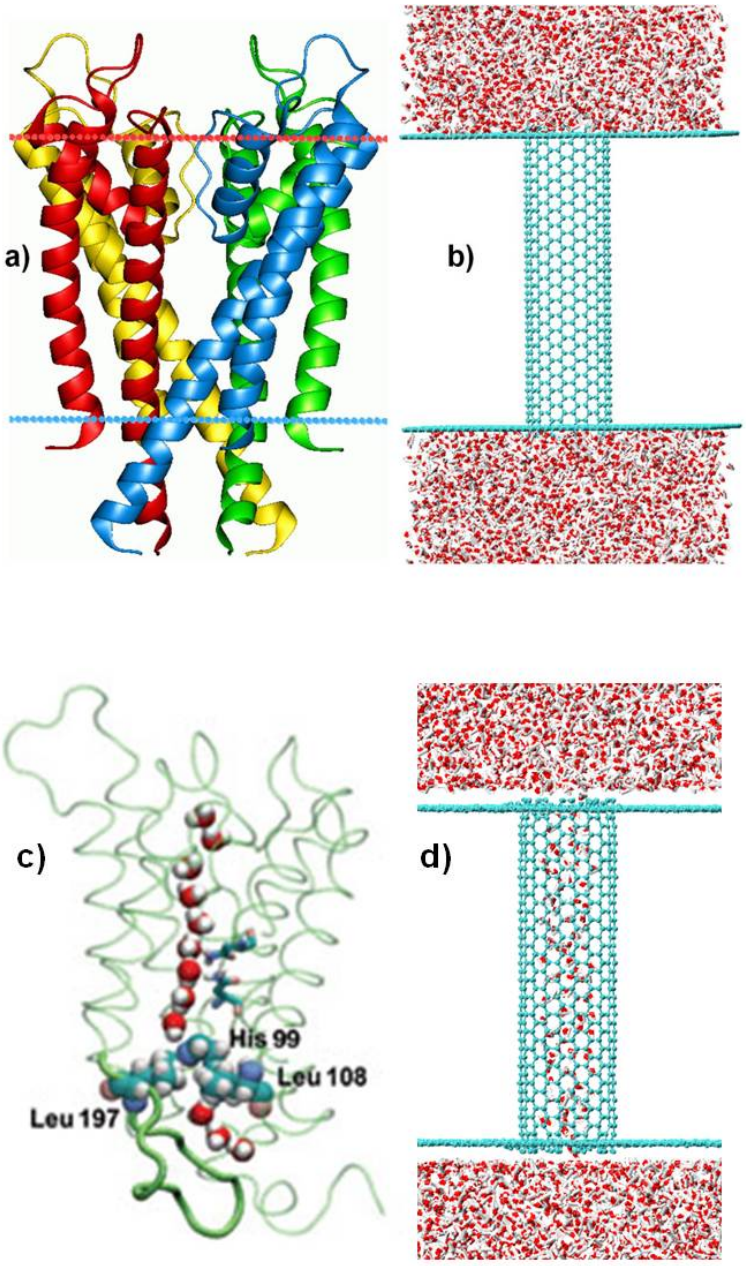


Figure 1.4: a) Penetration of the transmembrane aquaporin protein across lipid bilayer (upper and lower layers of which are signified by red and blue dots), reproduced the figure of ref. 80 from *Wikimedia*, with prior permission from *Elsevier* publishing group; b) representative model system for studying nano-confined properties of water within single-walled carbon nanotube, resembling hydrophobic confinement similar to *Figure 1.4 a*; c) snapshot of water channel within plant aquaporin, reproduced from ref. 82, with prior permission from *Nature* publishing group; d) flow of water within hydrophobic nanochannels, resembling water transport through biological ion-channels.

1.3 Bibliography

(1) Dobson, C. M. Protein folding and misfolding. *Nature*. **2003**, *426*, 884-890.

(2) Phizicky, E.; Bastiaens, P. I. H.; Zhu, H.; Snyder, M.; Fields, S. Protein analysis on a proteomic scale. *Nature*. **2003**, *422*, 208-215.

(3) Chan, N.-L.; Rogers, P. H.; Arnone, A. Crystal Structure of the S-Nitroso Form of Liganded Human Hemoglobin. *Biochemistry*. **1998**, *37*, 16459-16464.

(4) McFadden, T. B.; Akers, R. M.; Kazmer, G. W. Alpha-Lactalbumin in Bovine Serum: Relationships with Udder Development and Function. *J. Dairy Sci.* **1987**, *70*, 259-264.

(5) Schonichen, A.; Alexander, M.; Gasteier, J. E.; Cuesta, F. E.; Fackler, O. T.; Geyer, M. Biochemical Characterization of the Diaphanous Autoregulatory Interaction in the Formin Homology Protein FHOD1. *J. Biol. Chem.* **2006**, *281*, 5084-5093.

(6) Morillas, M.; Eberl, H.; Allain, F. d. r. T.; Glockshuber, R.; Kuennemann, E. Novel Enzymatic Activity Derived from the Semliki Forest Virus Capsid Protein. *J. Mol. Biol.* **2008**, *376*, 721-735.

(7) Kornberg, R. D. The molecular basis of eukaryotic transcription. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 12955-12961.

(8) Spoerner, M.; Herrmann, C.; Vetter, I. R.; Kalbitzer, H. R.; Wittinghofer, A. Dynamic properties of the Ras switch I region and its importance for binding to effectors. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4944-4949.

(9) Lee, H. J.; Kang, S. J.; Lee, K.; Im, H. Human αsynuclein modulates vesicle trafficking through its interaction with prenylated Rab acceptor protein 1. *Biochem. Biophys. Res. Commun.* **2011**, *412*, 526-531.

(10) Scholander, P. Oxygen transport through hemoglobin solutions. *Science*. **1960**, *131*, 585-590.

(11) Brown, A. M.; Birnbaumer, L. Direct G protein gating of ion channels. *Am. J. Physiol. - Heart and Circ. Physiol*. **1988**, *254*, H401-H410.

(12) Xu, X.; Cho, M.; Spencer, N. Y.; Patel, N.; Huang, Z.; Shields, H.; King, S. B.; Gladwin, M. T.; Hogg, N.; Kim-Shapiro, D. B. Measurements of nitric oxide on the heme iron and β-93 thiol of human hemoglobin during cycles of oxygenation and deoxygenation. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 11303-11308.

(13) Donald Voet; Voet, J. G. Biochemistry *Wiley: Hoboken, NJ.* **2004**, *Vol 1 3rd ed.*

(14) Hartl, F. U.; Bracher, A.; Hayer-Hartl, M. Molecular chaperones in protein folding and proteostasis. *Nature*. **2011**, *475*, 324-332.

(15) Shoji, M.; Golde, T. E.; Ghiso, J.; Cheung, T. T.; Estus, S.; Shaffer, L. M.; Cai, X. D.; McKay, D. M.; Tintner, R.; Frangione, B.; et, al. Production of the Alzheimer amyloid beta protein by normal proteolytic processing. *Science*. **1992**, *258*, 126-129.

(16) Chang, X.; Jørgensen, A. M. M.; Bardrum, P.; Led, J. J. Solution Structures of the R6 Human Insulin Hexamer. *Biochemistry*. **1997**, *36*, 9409-9422.

(17) Dill, K. A. Dominant forces in protein folding. *Biochemistry*. **1990**, *29*, 7133-7155.

(18) Levinthal, C. ARE THERE PATHWAYS FOR PROTEIN FOLDING ? *Jour. Chim. Phys.* **1968**, *65*, 44-45.

(19) Zwanzig, R.; Szabo, A.; Bagchi, B. Levinthal's paradox. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 20-22.

(20) Wolynes, P. G.; Onuchic, J. N.; Thirumalai, D. Navigating the folding routes. *Science*. **1995**, *267*, 1619-1620.

(21) Kuwajima, K. The molten globule state as a clue for understanding the folding and cooperativity of globular-protein structure. *Proteins. Struc. Func. Gen.* **1989**, *6*, 87-103.

(22) Dobson, C. M. Unfolded proteins, compact states and molten globules. *Curr. Op. Struct. Biol.* **1992**, *2*, 6-12.

(23) Ptitsyn, O. B. Molten globule and protein folding. *Adv. Prot. Chem.* **1995**, *47*, 83-229.

(24) Bychkova, V. E.; Ptitsyn, O. B. The molten globule in vitro and in vivo. *Chemtracts: Biochem. Mol. Biol.* **1993**, *4*, 133-163.

(25) Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R. The Protein Folding Problem. *Annu. Rev. Biophys.* **2008**, *37*, 289-316.

(26) Jha, S. K.; Udgaonkar, J. B. Direct evidence for a dry molten globule intermediate during the unfolding of a small protein. *Proc. Natl. Sci. U.S.A.* **2009**, *106*, 12289-12294.

(27) Tsai, A. M.; Udovic, T. J.; Neumann, D. A. The Inverse Relationship between Protein Dynamics and Thermal Stability. *Biophys. J.* **2001**, *81*, 2339-2343.

(28) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science*. **2012**, *338*, 1042-1046.

(29) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235-242.

(30) Udgaonkar, J. B.; Baldwin, R. L. NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature*. **1988**, *335*, 694-699.

(31) Agashe, V. R.; Shastry, M. C. R.; Udgaonkar, J. B. Initial hydrophobic collapse in the folding of barstar. *Nature*. **1995**, *377*, 754-757.

(32) Luan, B.; Shan, B.; Baiz, C.; Tokmakoff, A.; Raleigh, D. P. Cooperative Cold Denaturation: The Case of the C-Terminal Domain of Ribosomal Protein L9. *Biochemistry*. **2013**, *52*, 2402-2409.

(33) Privalov, P. L. Cold Denaturation of Protein. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 281-306.

(34) Pace, C. N.; Tanford, C. Thermodynamics of the unfolding of β-lactoglobulin A in aqueous urea solutions between 5 and 55^{o}. *Biochemistry*. **1968**, *7*, 198-208.

(35) Martin, S. R.; Esposito, V.; De Los Rios, P.; Pastore, A.; Temussi, P. A. Cold Denaturation of Yeast Frataxin Offers the Clue to Understand the Effect of Alcohols on Protein Stability. *J.Am. Chem. Soc.* **2008**, *130*, 9963-9970.

(36) Ravindra, R.; Winter, R. On the Temperature–Pressure Free-Energy Landscape of Proteins. *ChemPhysChem*. **2003**, *4*, 359-365.

(37) Hayakawa, I.; Linko, Y.-Y.; Linko, P. Mechanism of High Pressure Denaturation of Proteins. *LWT - Food Sci. Technol*. **1996**, *29*, 756-762.

(38) Bennion, B. J.; Daggett, V. The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Sci. U.S.A.* **2003**, *100*, 5142-5147.

(39) Graziano, G. On the molecular origin of cold denaturation of globular proteins. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14245-14252.

(40) He, Y.; Alam, S. L.; Proteasa, S. V.; Zhang, Y.; Lesuisse, E.; Dancis, A.; Stemmler, T. L. Yeast Frataxin Solution Structure, Iron Binding, and Ferrochelatase Interaction. *Biochemistry*. **2004**, *43*, 16254-16262.

(41) Pastore, A.; Martin, S. R.; Politou, A.; Kondapalli, K. C.; Stemmler, T.; Temussi, P. A. Unbiased Cold Denaturation: Low- and High-Temperature Unfolding of Yeast Frataxin under Physiological Conditions. *J.Am. Chem. Soc.* **2007**, *129*, 5374-5375.

(42) Selkoe, D. J. Folding proteins in fatal ways. *Nature*. **2003**, *426*, 900-904.

(43) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundström, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. Structure of an Intermediate State in Protein Folding and Aggregation. *Science*. **2012**, *336*, 362-366.

(44) Jahn, T. R.; Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **2005**, *272*, 5962-5970.

(45) Knowles, T. P. J.; Vendruscolo, M.; Dobson, C. M. The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell. Biol*. **2014**, *15*, 384-396.

(46) Chiti, F.; Dobson, C. M. Amyloid formation by globular proteins under native conditions. *Nat Chem Biol*. **2009**, *5*, 15-22.

(47) Karpinar, D. P.; Balija, M. B. G.; Kügler, S.; Opazo, F.; Rezaei-Ghaleh, N.; Wender, N.; Kim, H.-Y.; Taschenberger, G.; Falkenburger, B. H.; Heise, H.; Kumar, A.; Riedel, D.; Fichtner, L.; Voigt, A.; Braus, G. H.; Giller, K.; Becker, S.; Herzig, A.; Baldus, M.; Jäckle, H.; Eimer, S.; Schulz, J. B.; Griesinger, C.; Zweckstetter, M. Pre-fibrillar α-synuclein variants with impaired β-structure increase neurotoxicity in Parkinson's disease models. *The EMBO J.* **2009**, *28*, 3256-3268.

(48) Haass, C.; Selkoe, D. J. Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer's amyloid β-peptide. *Nat. Rev. Mol. Cell. Biol*. **2007**, *8*, 101-112.

(49) Khandogin, J.; Brooks, C. L. Linking folding with aggregation in Alzheimer's *β*-amyloid peptides. *Proc. Natl. Acad. Sci. U.S.A*. **2007**, *104*, 16880-16885.

(50) Lee, C.; Ham, S. Characterizing amyloid-beta protein misfolding from molecular dynamics simulations with explicit water. *J. Comput. Chem.* **2010**, *32*, 349-355.

(51) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Accounts Chem. Res.* **2012**, *45*, 83-92.

(52) Uéda, K.; Fukushima, H.; Masliah, E.; Xia, Y.; Iwai, A.; Yoshimoto, M.; Otero, D. A.; Kondo, J.; Ihara, Y.; Saitoh, T. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl. Acad. Sci. U.S.A*. **1993**, *90*, 11282-11286.

(53) Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. The energy landscapes and motions of proteins. *Science*. **1991**, *254*, 1598-1603.

(54) Valente, A. P.; Miyamoto, C. A.; L. Almeida, F. C. Implications of Protein Conformational Diversity for Binding and Development of New Biological Active Compounds. *Curr. Med. Chem.* **2006**, *13*, 3697-3703.

(55) Henzler-Wildman, K.; Kern, D. Dynamic personalities of proteins. *Nature*. **2007**, *450*, 964-972.

(56) Wand, A. J. Dynamic activation of protein function: A view emerging from NMR spectroscopy. *Nat. Struct. Mol. Biol*. **2001**, *8*, 926-931.

(57) Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D. A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature*. **2007**, *450*, 913-916.

(58) Nandi, N.; Bhattacharyya, K.; Bagchi, B. Dielectric relaxation and solvation dynamics of water in complex chemical and biological systems. *Chem. Rev.* **2000**, *100*, 2013-2046.

(59) Ru, M. T.; Dordick, J. S.; Reimer, J. A.; Clark, D. S. Optimizing the salt-induced activation of enzymes in organic solvents: Effects of lyophilization time and water content. *Biotech. Bioengg.* **1999**, *63*, 233-241.

(60) Anderson, V. L.; Webb, W. W.; Eliezer, D. Interplay between desolvation and secondary structure in mediating cosolvent and temperature induced alpha-synuclein aggregation. *Phys. Biol.* **2012**, *9*, 056005.

(61) Mattos, C.; Ringe, D. Proteins in organic solvents. *Curr. Op. Struc. Biol.* **2001**, *11*, 761-764.

(62) Kramers, H. A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*. **1940**, *7*, 284-304.

(63) Gavish, B.; Werber, M. M. Viscosity-dependent structural fluctuations in enzyme catalysis. *Biochemistry*. **1979**, *18*, 1269-1275.

(64) Pradeep, L.; Udgaonkar, J. B. Diffusional Barrier in the Unfolding of a Small Protein. *J. Mol. Biol.* **2007**, *366*, 1016-1028.

(65) Martinez, A. V.; MaÅ‚olepsza, E.; Rivera, E.; Lu, Q.; Straub, J. E. Exploring the role of hydration and confinement in the aggregation of amyloidogenic peptides Aβ16-22 and Sup357-13 in AOT reverse micelles. *J. Chem. Phys.* **2014**, *141*, 22D530.

(66) Adrover, M.; Martorell, G.; Martin, S. R.; Urosev, D.; Konarev, P. V.; Svergun, D. I.; Daura, X.; Temussi, P.; Pastore, A. The Role of Hydration in Protein Stability: Comparison of the Cold and Heat Unfolded States of Yfh1. *J. Mol. Biol.* **2012**, *417*, 413-424.

(67) Tsai, C.-J.; Maizel, J. V.; Nussinov, R. The Hydrophobic Effect: A New Insight from Cold Denaturation and a Two-State Water Structure. *Crit. Rev. Biochem. Mol. Biol.* **2002**, *37*, 55-69.

(68) Marchi, M.; Sterpone, F.; Ceccarelli, M. Water Rotational Relaxation and Diffusion in Hydrated Lysozyme. *J. Am. Chem. Soc.* **2002**, *124*, 6787-6791.

(69) García de la Torre, J.; Huertas, M. L.; Carrasco, B. Calculation of Hydrodynamic Properties of Globular Proteins from Their Atomic-Level Structure. *Biophys. J.* **2000**, *78*, 719-730.

(70) Nandi, N.; Bagchi, B. Dielectric Relaxation of Biological Water. *J. Phys. Chem. B.* **1997**, *101*, 10954-10961.

(71) Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. Biological Water: Femtosecond Dynamics of Macromolecular Hydration. *J. Phys. Chem. B.* **2002**, *106*, 12376-12395.

(72) Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Hagan, M. F.; Chandler, D.; Garde, S. Sitting at the Edge: How Biomolecules use Hydrophobicity to Tune Their Interactions and Function. *J. Phys. Chem. B.* **2012**, *116*, 2498-2503.

(73) Jungwirth, P. Biological Water or Rather Water in Biology? *J. Phys. Chem. Lett.* **2015**, *6*, 2449-2451.

(74) Angell, C. A. Formation of Glasses from Liquids and Biopolymers. *Science*. **1995**, *267*, 1924-1935.

(75) Wood, K.; Frölich, A.; Paciaroni, A.; Moulin, M.; Härtlein, M.; Zaccai, G.; Tobias, D. J.; Weik, M. Coincidence of Dynamical Transitions in a Soluble Protein and Its Hydration Water: Direct Measurements by Neutron Scattering and MD Simulations. *J. Am. Chem. Soc.* **2008**, *130*, 4586-4587.

(76) Tarek, M.; Tobias, D. J. Role of Protein-Water Hydrogen Bond Dynamics in the Protein Dynamical Transition. *Phys. Rev. Lett.* **2002**, *88*, 138101.

(77) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2007**, *108*, 74-108.

(78) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature*. **2001**, *414*, 188-190.

(79) Murata, K.; Mitsuoka, K.; Hirai, T.; Walz, T.; Agre, P.; Heymann, J. B.; Engel, A.; Fujiyoshi, Y. Structural determinants of water permeation through aquaporin-1. *Nature*. **2000**, *407*, 599-605.

(80) Zhou, Y.; MacKinnon, R. The Occupancy of Ions in the K^+ Selectivity Filter: Charge Balance and Coupling of Ion Binding to a Protein Conformational Change Underlie High Conduction Rates. *J. Mol. Biol.* **2003**, *333*, 965-975.

(81) Takata, K.; Matsuzaki, T.; Tajika, Y. Aquaporins: water channel proteins of the cell membrane. *Prog. Histochem. Cytochem.* . **2004**, *39*, 1-83.

(82) Tornroth-Horsefield, S.; Wang, Y.; Hedfalk, K.; Johanson, U.; Karlsson, M.; Tajkhorshid, E.; Neutze, R.; Kjellbom, P. Structural mechanism of plant aquaporin gating. *Nature*. **2006**, *439*, 688-694.

(83) Choudhury, N.; Pettitt, B. M. Dynamics of Water Trapped between Hydrophobic Solutes. *J. Phys. Chem. B.* **2005**, *109*, 6422-6429.

(84) Joseph, S.; Aluru, N. R. Pumping of Confined Water in Carbon Nanotubes by Rotation-Translation Coupling. *Phys. Rev. Lett.* **2008**, *101*, 064502.

(85) Choudhury, N. Effect of surface hydrophobicity on the dynamics of water at the nanoscale confinement: A molecular dynamics simulation study. *Chem. Phys.* **2013**, *421*, 68-76.

(86) Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature*. **2005**, *437*, 640-647.

(87) Striolo, A. The Mechanism of Water Diffusion in Narrow Carbon Nanotubes. *Nano Lett.* **2006**, *6*, 633-639.

(88) Barati Farimani, A.; Aluru, N. R. Spatial Diffusion of Water in Carbon Nanotubes: From Fickian to Ballistic Motion. *J. Phys. Chem. B*. **2011**, *115*, 12145-12149.

(89) Park, J. H.; Aluru, N. R. Diffusion of water submonolayers on hydrophilic surfaces. *Applied Phys. Lett*. **2008**, *93*, 253104.

(90) Majumder, M.; Zhan, X.; Andrews, R.; Hinds, B. J. Voltage Gated Carbon Nanotube Membranes. *Langmuir*. **2007**, *23*, 8624-8631.

(91) Meng, X. W.; Wang, Y.; Zhao, Y. J.; Huang, J. P. Gating of a Water Nanochannel Driven by Dipolar Molecules. *J. Phys. Chem. B*. **2011**, *115*, 4768-4773.

(92) Corry, B. Water and ion transport through functionalised carbon nanotubes: implications for desalination technology. *Energy Environ. Sci.* **2011**, *4*, 751-759.

(93) Das, P. Effect of cosolvents on nano-confined water: a molecular dynamics study. *Nanoscale*. **2012**, *4*, 2931-2936.

(94) Su, J.; Guo, H. Effect of Nanochannel Dimension on the Transport of Water Molecules. *J. Phys. Chem. B*. **2012**, *116*, 5925-5932.

Chapter 2

Computational methodology

As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.

- Albert Einstein

2.1 Scope of computations in atomic level understanding

Computational methods are proving to be indispensable for the understanding of complex molecular systems. Quantum mechanics driven computational methodologies, such as density functional theory (DFT),[1] can probe the electronic structure of many body systems but are limited by the system sizes they can probe. On the contrary, molecular mechanical models elucidate properties of atomic systems by applying the Born-Oppenheimer approximation,[2] by ignoring the electronic motions of each atom and considering only the nuclear motions. Particles are thereby considered as charged or uncharged hard spheres. Hence, models of comparatively large systems, especially biomolecules, can be computed with these protocols. Although these models cannot elucidate changes in electronic properties occurring in the subatomic domains, they are often of immense value in providing a molecular level understanding of larger systems such as polymers, liquids, and biomacromolecules.[2] Atomistic molecular dynamics (MD) simulations is one such technique of molecular mechanics, which harnesses Newtonian mechanics for probing the structural and dynamical properties. It acts as a bridge between the microscopic events and macroscopic observables, where the atomistic underpinnings of macromolecular properties in complex physicochemical systems can be cohesively addressed within a wide range of time and length scales. Over the years, MD simulations have become an efficient tool in addressing subjects related to basic and applied interdisciplinary sciences, with notable relevance in problems of biochemistry and molecular biology (such as protein folding, lipid dynamics, docking and ligand biomolecule interactions, nucleic acids, solvent associated biomolecular dynamics, etc.).[3-7] In addition, atmospheric chemistry, nanomaterials, and polymer science are few other fields where MD simulations can provide valuable molecular insights.[8-12]

To access very large time and length scales, hybrid methods consisting of both quantum and molecular mechanics (QM/MM) are utilized for complex chemically active systems. These methods are potentially useful if the active site is prone to chemical changes; the properties of rest of the system can be addressed with molecular mechanics. Further, MD with polarizable force fields, or *ab initio* MD incorporating parameterization 'on the fly' with changes in electronic configuration, are also used.[13,14] However, even these hybrid methods are severely restrictive in terms of the time and length scales that can be addressed.

In this dissertation, the problems investigated, namely biomolecular structure, dynamics, correlation with solvent in hybrid bio/nano systems and interpretations vis-à-vis vibrational spectroscopy, are most suitable for atomistic MD simulation techniques. The following sections therefore summarize the key elements of this protocol.

2.2 Atomistic force fields

Temporal evolution of a complex system is followed with MD simulations by solving Newton's second laws of motion:

$$\vec{F} = m\vec{a} = m\frac{d\vec{v}}{dt} = m\frac{d^2\vec{r}}{dt^2} \tag{2.1}$$

A particle of mass *m*, under the effect of force $\vec{F}$, moves with acceleration $\vec{a}$, which is the time derivative of the velocity $\vec{v}$, which itself is the derivative of the coordinates $\vec{r}$. Thus the movement of each particle is dependent on the total force experienced by it to the overall system. This is obtained by the negative gradient of the potential energy ϕ of the total system at the position of the i^{th} atom

$$\vec{F}_i = -\vec{\nabla}_i V \tag{2.2}$$

The potential energy at each atom is obtained from pre-defined 'empirical' potential energy function, more commonly known as a *force field*.[15,16]

The parameters for force fields are generally obtained from an iterative process involving *ab initio* quantum mechanical calculations and experimental data.[2,17] Efforts are further made to ensure that the force field parameters are transferable, i.e. same set of parameters can be used for similar moieties in different complex systems.[2] Thus parameters obtained from smaller molecules may also be implemented in larger molecules with similar *atom type* (hybridization state) and connectivity. Thus, functions producing more accurate results eventually precedes the others, leading to the gradual development of corresponding force fields.[2] The individual terms constituting the classical force fields are discussed below.

The total interactions in an atomistic force field can be divided into intramolecular (V_1) and intermolecular (V_2) forces,

$$V_{total} = V_1 + V_2 \tag{2.3}$$

Bond stretching, angle bending and torsional potential terms are included in the intramolecular or bonded interactions for the atoms chemically linked to each other, whereas the intermolecular or non-bonded interactions include the coulombic and dispersion interactions.

$$V_1 = V_{bond} + V_{angle} + V_{torsion} \tag{2.4}$$

$$V_2 = V_{Coulomb} + V_{LJ} \tag{2.5}$$

The details of the bonded and non-bonded potentials are provided below.

2.2.1 Bonded potentials

Bond length. The bond connecting any two atoms is modeled similar to a harmonic spring. The bonding potential is stated as

$$V_{bond} = k_b\left(b - b_0\right)^2 \tag{2.6}$$

where the equilibrium bond distance is b_0, and the spring constant is k_b.

Bond angle. The bond angular potential term, is again modeled like a harmonic spring, like the bonding interaction

$$V_{angle} = k_\theta\left(\theta - \theta_0\right)^2 \tag{2.7}$$

the equilibrium angle is θ_0 and the angular force constant is k_θ.

The first two terms of the bonded potential functions present in the generalized empirical force field is associated with changes in energy due to harmonic bond stretching and angle bending. Thus both functions use Hooke's law of simple mass less spring, where bond length (or angle), from the equilibrium bond length (or angle), creates a retarding force.

Torsional potential. The potential between four atoms chemically connected to each other, consists of the torsional potential. It can be categorized into *proper* and *improper* torsional potential. The *proper torsional angle*, or the *dihedral angle*, is the angle (ϕ) made by the fourth consecutive atom of a sequentially connected chain with the plane of the first three atoms. The functional form of this potential is given as

$$V_{dihedral} = k_\phi\left(1+\cos\left(n\phi-\delta\right)\right) \tag{2.8}$$

Here, k_ϕ is the force constant, n the multiplicity, and δ the corresponding phase factor for the dihedral potential. The cosine term represents the periodicity in the torsional potential, so that a complete rotation coincides with the initial conformation.

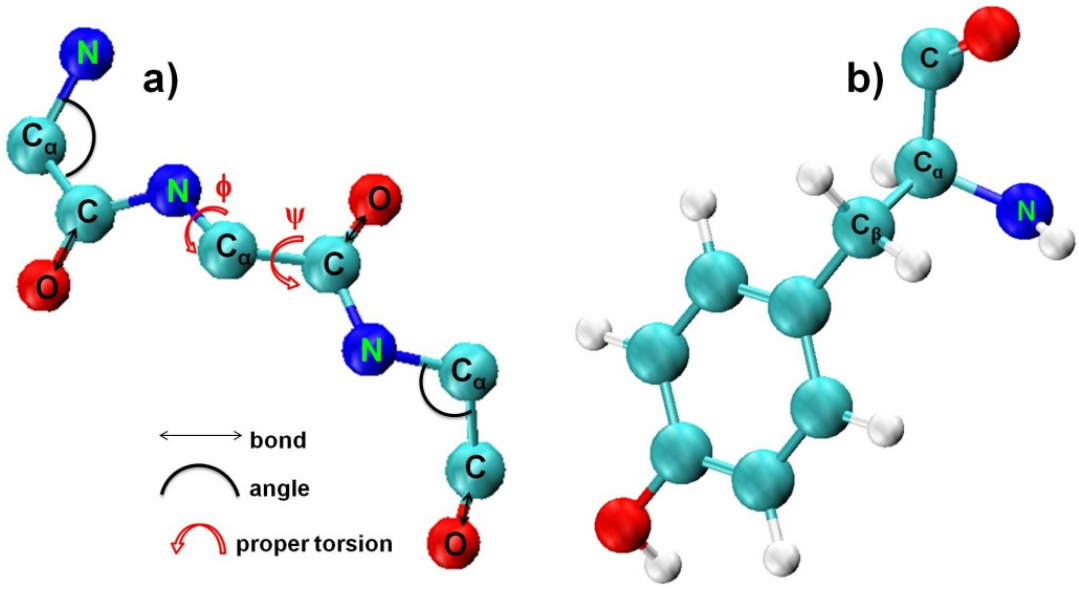


Figure 2.1 a) Representative bonds (C–O), angles (N–C_α–C), proper torsional angles ϕ (between planes C–N–C_α and N–C_α–C) and ψ (between planes N–C_α–C and C_α–C–N), in peptide backbone; b) example of improper torsional angle between C_β and the plane consisting of N–C_α–C.

The *improper torsion potential* is the potential due to improper torsion angle, which is similar in the geometric analysis as the dihedral angle. However, in this case, the fourth atom is not sequentially connected with the first three atoms, but the central atom is chemically connected to three different atoms. The improper torsion angle is the angle made by one of the lateral atoms with the plane of the other three (including the central atom). The functional form of the potential is harmonic in nature, similar to bond and angle potentials:

$$V_{improper} = k_{\phi}\left(\phi - \phi_0\right)^2 \tag{2.9}$$

Here, k_{ϕ} and ϕ_0 are respectively the improper torsional force constant and the equilibrium angle. Thus, the total torsional potential at any given atom is the summation of the proper and improper torsional potentials.

$$V_{torsion} = V_{dihedral} + V_{improper} \tag{2.10}$$

The force constants of the bonded potentials (k_b, k_θ, k_ϕ) are generally obtained experimentally or with quantum mechanical calculations.

2.2.2 Non-bonded potentials

Van der Waals interactions. The interactions between a pair of atoms due to dipole-dipole interactions, dipole-induced dipole interactions, instantaneous dipole-induced dipole interactions (London dispersion forces), are collectively known as van der Waals (vdW) interactions. The origin of these interactions lies in the reorganization of electrons surrounding an atom due to the influence of other electrons. Although changes in electronic configuration are not accounted in classical force fields, interactions between two atoms (labeled as *i* and *j*) based on corresponding Van der Waals radii are represented by the Lennard-Jones (LJ) potential in an empirical force field

$$V_{LJ} = \varepsilon\left[\left(\frac{R_{\min}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{\min}}{r_{ij}}\right)^{6}\right] \tag{2.11}$$

Here, r_{ij}, ε and R_{min} are respectively the inter-atomic distance between two atoms, the depth of LJ potential, and the location of the potential minima. The above functional form of LJ interaction can also be represented as,

$$V_{LJ} = 4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{2.12}$$

where,

$$R_{\min} = 2^{1/6}\sigma \tag{2.13}$$

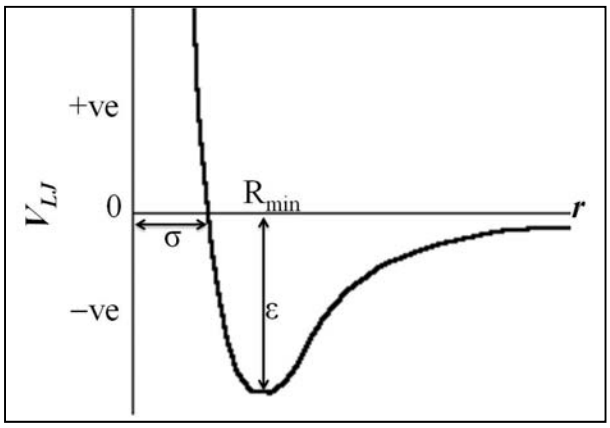


Figure 2.2 Schematic representation of Van der Waals interaction. R_{min} is the distance where the potential energy attains a minimum, ε is the corresponding well depth, and σ is the inter-atomic distance with the magnitude of the potential energy to be zero.

LJ interactions are short ranged and become repulsive at distances lesser than R_{min}, and negligible at large separation distances. Thus, the interaction distance is usually truncated after a certain distance, typically in the range of 10–12 Å.

Electrostatic interactions. Electrostatic interactions mainly constitute the non-bonded interactions between two atoms (*i* and *j*). The electrostatic interaction arising between *i* and *j* is calculated from Coulomb's law:

$$V_{Coulomb(i,j)} = k_e \frac{q_i q_j}{r_{ij}} \tag{2.14}$$

where the partial charges q_i and q_j, specified in empirical force fields for the atom types *i* and *j*, are determined from quantum mechanical derivations for the particular atom types. The separation between the two atoms is denoted as r_{ij}, and k_e is the conversion factor of the electrostatic potential in the atomic units of energy used in the force field. Unlike Van der Waals

interaction, electrostatic interactions are long-ranged. Ewald summation, reaction field and cell multiple methods are some techniques to calculate the long-range electrostatics.

2.3 Equations of motion

As stated before, the force acting on a system and its corresponding acceleration is obtained from the negative gradient of the potential energy acting on it (equation 2.2). The space-time trajectory of a particular many-bodied system is acquired with numerical integration methods. Among the algorithms used in different packages, the frequently used velocity Verlet algorithm[18] is discussed below.

The position of an atom *i* at a consecutive time '$t+\delta t$' (with time step δt), can be expressed from its initial time *t* as a Taylor expansion,

$$\vec{r}_i(t+\delta t) = \vec{r}_i(t) + \delta t \vec{v}_i(t) + \frac{\delta t^2}{2m_i}\vec{f}_i(t) \qquad (2.15)$$

where $\vec{r}_i(t)$ is the position, $\vec{v}_i(t)$ is the velocity, of the i^{th} atom at time *t*. Similarly,

$$\vec{r}_i(t) = \vec{r}_i(t+\delta t) - \delta t \vec{v}_i(t+\delta t) + \frac{\delta t^2}{2m_i}\vec{f}_i(t+\delta t) \qquad (2.16)$$

At time ($t + \delta t$), the velocity can then be obtained from the two previous equations as,

$$\vec{v}_i(t+\delta t) = \vec{v}_i(t) + \frac{\delta t}{2m_i}\left[\vec{f}_i(t) + \vec{f}_i(t+\delta t)\right] \qquad (2.17)$$

In a similar manner, equation (2.17) can then be used in determining the velocities at time ($t + 2\delta t$). It is important to note that velocity of a particular atom is one step ahead of the coordinates and they are defined from earlier simulations or from the temperature dependent Maxwell Boltzmann velocity distribution function. The fastest atomic motions in a complex chemical or biomolecular system are the bond vibrations, whose time periods are typically within a few femtoseconds. The time step (δt) should be smaller than corresponding atomic motions of fastest timescales, and hence is usually not set to be more than 2 fs in atomistic simulations. For obtaining better computational speed, oftentimes the fastest vibrational motions involving the

hydrogen atoms are constrained, thus enabling the implementation of relatively larger time steps.[19] Also, multiple time step algorithms are used within MD simulation protocols which incorporate separating the interactions depending on corresponding timescales.[20,21] Thus the slow, medium and fast time steps are incorporated for a particular motion, which increase the computational efficiency of the simulation. The dynamics of the system sampled at different time steps involve greater accuracy. These methods are suitable for probing the mechanism of macromolecule-ligand binding, intermolecular interactions, inter-residue motions, etc., the timescales of which are in the picoseconds to nanoseconds regime. But with advancement of supercomputers, phenomenon occurring at larger timescales can be efficiently addressed nowadays.[22,23]

2.4 Periodic boundary conditions

In MD simulations, the effect of boundaries should be appropriately handled in order to accomplish enhancement in speed without compromising precision. For a three-dimensional system constituting N-particles with free boundaries, the number of molecules existing on its surface is proportional to $N^{1/3}$, and they encounter forces different from those present within the bulk. The concept of periodic boundary condition has been developed in order to eliminate the surface effect. The system of reference is positioned in a unit cell, with an infinite number of images in space along all the three-dimensions. Thus, when an atom when leaves the unit cell, its mirror image enters the cell from the opposite direction (Figure 2.3). Hence, the number of particles is conserved in the unit cell throughout the process of simulation. Due to periodic boundary conditions, the minimum image condition, wherein an atom can detect only a single image of the remaining atoms present in the system, is applied. The forces are thus calculated considering the nearest neighbor (the closest image) from an atom. The computational time is thus decreased drastically while calculating the bonded and non-bonded forces between the atoms, which for the latter increase as square of the number of atoms in the system.

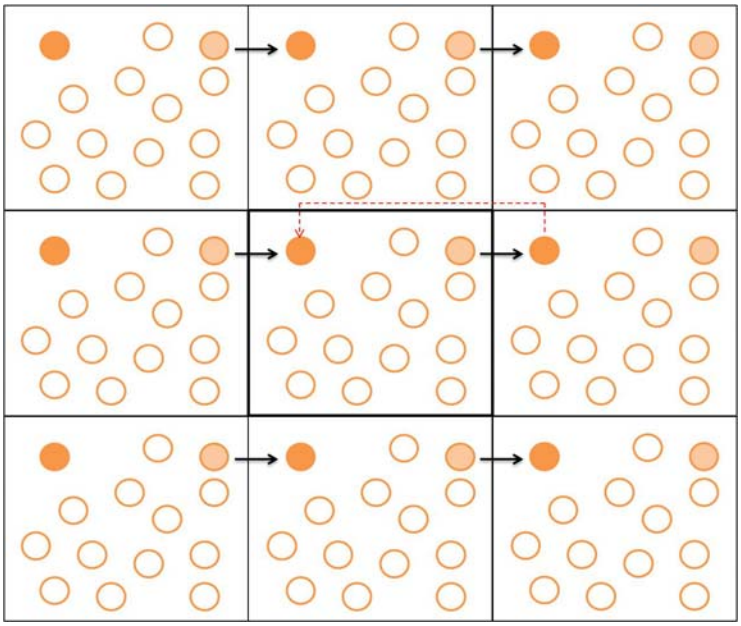


Figure 2.3 Pictorial representation of periodic boundary condition. The central simulation cell is highlighted. A particle leaving the cell enters it from the opposite direction.

2.5 Long range interactions

The calculation of the forces from bond, angle and dihedral potentials are proportional to the number of atoms present in the system and thus relatively straightforward. But, as stated in the earlier section, evaluation of non-bonded forces in MD simulations is the computationally expensive part, as it involves consideration of every pair of atoms in the system. The non-bonded dispersion (LJ) interactions decay rapidly with distance (r) and hence the potential is truncated after a certain distance. However, electrostatic interaction, which is inversely proportional to r (for two point charges) does not decay similarly. The simplification in calculating the non-bonded forces is done by applying a specific cutoff after certain distance (10 – 15 Å). The non-bonded potentials will then be calculated within the specified distance. For a system where periodic boundary condition is implemented, the cutoff should be defined in a way that the particles present in cell do not interact with their own mirror images. This corresponds to selection of the cutoff not greater than half of the distance of the cell along either of the cell dimensions. Since Coulombic interactions do not decay as rapidly as the dispersion forces, the long range electrostatic interactions are handled in a different manner, such as the reaction field method, the cell multiple method, the Ewald summation method, etc. The work presented in this

dissertation is based on simulation performed with the Ewald summation method (particle mesh Ewald).[24] This method was initially developed by Ewald while investigating the energetics of ionic crystals. Here, all the particles are allowed to interact with every other particle as well as all the mirror images present in the periodic cells. The electrostatic potential due to all charged pairs (*i* and *j*) in the reference cell, obtained from the simulation can be formulated as,

$$V = \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{2.18}$$

q_i (or q_j) is the partial charge of the particle i (or j), ε_0 the dielectric constant, N the number of particles present in the system, and r_{ij} the minimum distance between particles i and j, respectively. Additional contribution arising due to images present in the periodic cells is included in the following modified expression

$$V = \frac{1}{2}\sum_{n}\sum_{i=1}^{N}\sum_{j=1}^{N}\frac{q_i q_j}{4\pi\varepsilon_0\left(r_{ij}+n\right)} \tag{2.19}$$

In this equation, n is the vector distance of every image cell from the cell of reference. The components of each vector are integral multiples of the length (L) of the central cell ($n=n_xL$, n_yL, n_zL). The summation function in equation (2.19) converges gradually; hence the function is split into two parts, each of which converges much faster than the sum. The splitting should be done in a way such that the corresponding modified equation takes into account the dependence of the potential on r at smaller distances, as well as the slower decay of the Coulomb function at greater distances. The inverse distance can be written as,

$$\frac{1}{r} = \frac{f(r)}{r} + \frac{1-f(r)}{r} \tag{2.20}$$

The short-ranged and long-ranged Coulomb interactions are dealt with the function $f(r)$; the first term in the right hand side of eqn. (2.20) is short ranged and calculated in the real space; the second term deals with long ranged interaction, and are calculated in the reciprocal (or 'k') space. Further, every charge in Ewald summation is taken to be enveloped by a Gaussian

distribution of opposite charge but of the same magnitude. The functional form of the Gaussian charge distribution with width $\sqrt{2/\alpha}$ is,

$$\rho_i(r) = \frac{q_i \alpha^3}{\pi^{3/2}} e^{-\alpha^2 r^2} \tag{2.21}$$

The width of the neutralizing Gaussian distribution, which is responsible for faster rate of convergence of the summation function, is represented by the parameter α, and decided later based on computational efficiency. Thus the earlier expression is presently expressed as summation of the charged interactions and distributions of the neutralized charges. The functional for the real space can then be written as,

$$V = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{|n=0|} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \cdot \frac{erfc\left(\alpha \left| r_{ij} + n \right|\right)}{\left| r_{ij} + n \right|} \tag{2.22}$$

Here *erfc* is the complementary error function,

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{\infty} e^{-t^2} dt \tag{2.23}$$

Thus, the function $f(r)$ is the corresponding error function $erfc(r)$, which helps in rapid convergence of the function, and presumed to be inconsequential after a certain cut-off.

To negate the initially added neutralized Gaussian charged distribution, a second counter charge is added. For a cubical cell of side 'L', this contribution is formulated as,

$$V = \frac{1}{2} \sum_{|k \neq 0|} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{\pi . L^3} \cdot \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \cdot \frac{4\pi^2}{k^2} e^{-\frac{k}{4\alpha^2}} \cos\left(\vec{k} . \vec{r}_{ij}\right) \tag{2.24}$$

This corresponds to summation in reciprocal space which converges more rapidly for smaller α, unlike the real space sum. The present summation includes interaction of the neutralizing Gaussian charged distribution term with itself, which needs rectification by subtraction of another factor from Ewald summation function,

$$V = -\frac{\alpha}{\sqrt{\pi}} \sum_{k=1}^{N} \frac{q^2 k}{4\pi\varepsilon_0} \tag{2.25}$$

Also a different correction factor corresponding to the contribution of the surrounding medium around the simulation cell is taken into account as,

$$V_{correct} = \frac{2\pi}{L^3} \left| \sum_{i=1}^{N} \frac{q_i}{4\pi\varepsilon_0} r_i \right| \tag{2.26}$$

where ε_0 is the relative permittivity of the surrounding medium. The Ewald summation function, and the additive functional form of equations (2.22), (2.24), (2.25), (2.26); are proportional to square of the number of molecules (N^2), N being the number of molecules present in the simulated system. This is again computationally expensive for systems comprising biomacromolecules like proteins, lipids, DNA, etc. We point out here that particle mesh Ewald summation (implemented in the CHARMM force field)[24] has been formulated by Darden et. al. where the scaling factor is proportional to $N \log N$.

2.6 Control of thermodynamic states

The conformational trajectories obtained from the numerical integrations given in the earlier section are based on Newton's equations of motion. Thus, the ensemble generated from these trajectories corresponds to the constant energy, constant volume, microcanonical (NVE) ensemble. But, many studies are carried out with systems which maintain a thermal or mechanical equilibrium with the environment. These often require simulations of the system as well as the surroundings at either constant temperature (NVT) or pressure (NPT), or both. Also, analyses of processes like phase equilibria in liquid systems, gas adsorption on surfaces, etc. require regulation of chemical potential of the system, which will eventually require generation of grand canonical ensemble (with constant chemical potential and temperature) for these studies. But, in the present thesis, the scope of such studies is limited, and the focus will be given mostly to the methods formulating the control of temperature and pressure, expediting NVT or NPT simulations.

2.6.1 Temperature control

According to the *equipartition theorem*, the kinetic energy of a system, in thermal equilibrium at temperature T, is equally distributed to all the independent degrees of freedom with a magnitude of $\left(\frac{1}{2}k_bT\right)$ to get uniform gap; k_b being the Boltzmann constant. Thus, the mean kinetic energy can be estimated from its corresponding temperature. For a system with N_f degrees of freedom, its temperature T is proportional to the kinetic energy (*K.E.*) of the system as,

$$\langle K.E.\rangle = \frac{1}{2}N_f k_b T \tag{2.27}$$

Since the total kinetic energy of a system can be expressed as a summation of individual kinetic energies of the constituent atoms, the temperature for an N particle system can be obtained as

$$T = \sum_{i=1}^{N} \frac{m_i v_i^2}{k_b N_f} \tag{2.28}$$

m_i and v_i are the corresponding mass and the instantaneous velocity of the i^{th} particle of the system. Equation (2.28) shows that we need to control the velocities of the corresponding particles of a specific system in order to regulate the corresponding temperature. Evidently, we can modify the temperature by scaling the velocities proportionally; like scaling the particle velocity by a factor of $\sqrt{T_{final}/T_{initial}}$ will eventually change the temperature from $T_{initial}$ to T_{final}. One more way to modify the temperature is to reassign the particle velocities from the corresponding Boltzmann distribution function at the new temperature. However, this can lead to alteration of the dynamical behavior of the particles in the corresponding NVE ensemble.

A different method for controlling the temperature of a particular system is with *Langevin dynamics* and stochastic collisions of the system particles.[25] This is similar to random yet periodic bombardment of a thermal particle at the selected temperature. The mathematical interpretation of the Newton's second law of motion due to these collisions can thus modified as,

$$f_i - \gamma_i \frac{dr_i}{dt} + R_i = m_i \frac{d^2 r_i}{dt^2} \qquad (2.29)$$

In the left hand side, the second term indicates the frictional drag on the system dependent on its velocity, which is overcome by the additional random force represented by the third expression in the left. *Nośe-Hover* method is often implemented as an additional feature of temperature control via the Langevin method. It is regarded as an extended system, assigned a 'fictitious mass', as well as provided an additional degree of freedom, which involves implementation of a virtual thermal reservoir at the chosen temperature. The strength of coupling with the original system is controlled by the magnitude of the 'fictitious mass'. The potential and kinetic energy of the reservoir control the dynamics of the system with which it is connected to, maintaining the effective temperature of the system likewise.

2.6.2 Pressure control

MD simulations are oftentimes performed under isobaric conditions to validate results from experiments under equivalent conditions. Like the Nośe-Hover Langevin piston method for controlling the temperature, the pressure of a simulated system is controlled likewise by coupling it to a 'pressure bath'.[25] In *Berendsen's weak coupling method*,[26] the volume is considered as an added dynamical variable, and the equation, which is of the first order represents the evolution of volume. The volume is coupled to the external pressure such that changes in pressure are represented with simultaneous fluctuation in volume. However, thermodynamic variables of the observables obtained from this method cannot be correlated with real thermodynamic properties, as meaningful ensembles are not obtained from the simulated trajectories after this method is applied to normal MD protocols.

In the constant pressure Langevin Piston method, the piston has partial damping mode. The degrees of freedom in the piston are formulated with a Langevin equation, after which the equations of motion of the system are expressed as

$$\frac{dr_i}{dt} = \frac{p_i}{m_i} + \frac{1}{3V}\frac{dV}{dt} r_i \qquad (2.30)$$

$$\frac{dp_i}{dt} = f_i - \frac{1}{3V}\frac{dV}{dt}p \tag{2.31}$$

$$\frac{d^2V}{dt^2} = \frac{1}{W}\left[P(t) - P(ext)\right] - \gamma\frac{dV}{dt} + R(t) \tag{2.32}$$

P(t) and *P(ext)* are respectively the instantaneous and externally applied pressure; *W* the piston mass; γ the Langevin collision frequency; *V* the respective volume of the system; *R(t)* is the random force obtained from a Gaussian distribution; r_i is the position; p_i the momentum; f_i the force; and m_i the mass of the i^{th} particle present in the system. The mean value of *R(t)* is zero, and its variance is,

$$\langle R(0)R(t)\rangle = \frac{2\gamma k_b \delta(t)}{W} \tag{2.33}$$

When γ=0, this method reduces to the previous Berendsen's weak coupling method of pressure control. At any point of time (*t*), the instantaneous pressure can be obtained as

$$P(t) = \frac{1}{3V}\left(2K.E + \sum_i f_i r_i\right) \tag{2.34}$$

where *K.E.* denotes the total instantaneous kinetic energy, while the second summation is indicative of the system's internal virial.

2.7 Analysis of simulation trajectories

The concept of *ergodic hypothesis* states that all microstates in a system have equal accessibility at long time scales, or, the time averaged value obtained from an equilibrium quantity is equal to the value obtained from ensemble average. Thus, from trajectories obtained from equilibrated MD simulations, we can acquire many useful observables. Also, the dynamical properties and energetics can be procured from MD simulations. Herein, we review some of the trajectory analyses methods relevant in biomolecular studies.

2.7.1 Protein structural analyses

Protein structural information can be extracted from the corresponding conformational ensemble generated by MD simulation trajectories. Configuration data of proteins saved at required

frequencies provides estimates of position and velocity obtained from the simulations. Although preliminary insights of the conformational ensembles are obtained by visualizing the simulated trajectories, a more detailed description is obtained from specific elucidation of protein structures. Few of those are briefly discussed below.

Radius of gyration (or R_g) is the normal distance of a point mass from the corresponding rotational axis of a substance whose moment of inertia (MI) is same as those obtained from summation of MI of each individual components constituting the substance. It can be mathematically computed as the average root mean squared distance of each component from the center of mass of the substance. R_g can be formulated as

$$\sqrt{\frac{1}{N}\sum_{n=1}^{N}(\delta r_n)^2} \qquad (2.39)$$

where δr_n is the distance between the n^{th} component from the center of mass axis. The effective compactness of a protein can be estimated by calculating the corresponding R_g.

Root mean squared deviation (or RMSD) is the estimation of mean distances between atoms of superimposed conformations (of proteins/biomolecules). It helps in determining the possible changes in conformational propensity of proteins in course of simulations. Mathematically it can be represented as

$$\sqrt{\frac{1}{N}\sum_{i=1}^{N}\delta i^2} \qquad (2.40)$$

where δi is the distance between N equivalent pairs of atoms (usually the backbone or C_α atoms of a protein).

Root mean squared fluctuation (or RMSF) is the fluctuation of the atoms of interest (backbone or C_α atoms of a protein) with respect to reference position, usually the mean position of atoms. Unlike RMSD, which represents the mean conformational changes over all atoms, RMSF provides extent of fluctuation of each atom with respect to time. It can be mathematically represented as

$$\sqrt{\frac{1}{T}\sum_{i=1}^{T}((x_i(t)-\bar{x}_i)^2} \qquad (2.41)$$

T is the period of time over which the averaging is done for a particular atom, $\bar{x}_i$ is the reference position or the mean atomic position, and $x_i(t)$ is the corresponding position at the i^{th} time. Often mean squared fluctuation (MSF), or the square of RMSF, is plotted as a function of residue numbers to denote the extent of fluctuations present in protein backbones.

Secondary Structural Persistence (or P), initially defined in our group,[27] is a measure of the changes in secondary structure of protein with respect to a reference conformation. Generally the initial conformation (experimentally determined structure obtained from the PDB database) is taken as reference. The mathematical expression of P is,

$$P = \frac{1}{N}\sum_{n=1}^{N} e^{-(\Delta\phi/\phi_{\max})}.e^{-(\Delta\psi/\psi_{\max})} \qquad (2.42)$$

$\Delta\phi$ and $\Delta\psi$ are changes in dihedral angles ϕ and ψ with respect to the reference structure. ϕ_{max} and ψ_{max} are the maximal changes in ϕ and ψ happening in the Ramachandran diagram .

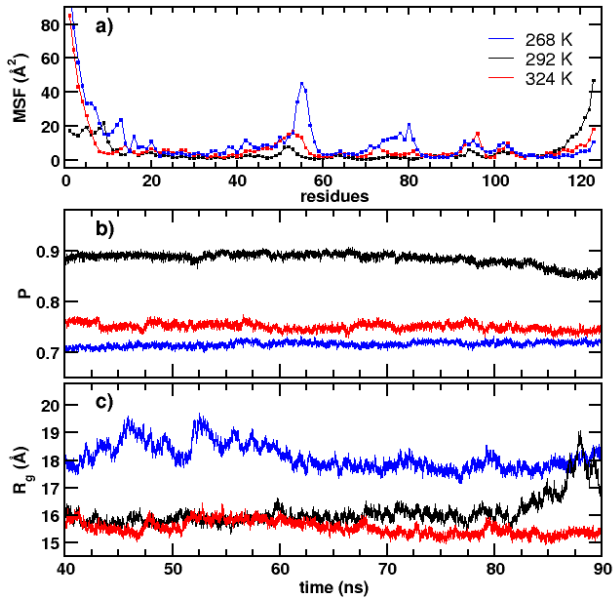


Figure 2.4 Structural analyses of mesophilic protein, yeast frataxin (Yfh1), at three different temperatures. a) plots of MSF as a function of the residues; b) evolution of P, and c) R_g, with simulation time.[28]

2.7.2 Principal component analyses

Principal component analysis (PCA) is a statistical protocol to convert a set of time dependent observations to orthogonally transformed, linearly uncorrelated variables. These variables, known as principal components, are less than or equal to the number of original variables. The

population of each linearly uncorrelated orthogonal component decreases sequentially, representing the most significant conformational ensembles in order of their population. The important modes obtained from MD simulations can be distinguished by PCA.[29-31] The theory behind PCA, wherein the original space of interdependent variables is transformed into a set of independent variables with this protocol, is described briefly in this section. We do this by considering the specific example of protein dynamics generated by MD simulation trajectories. The low dimensional protein motions, capturing the dominant modes of the protein dynamics, are obtained from the corresponding high dimensional motions. The calculation involves two essential components: i) estimation of the elements of covariance matrix (C_{ij}) of proteins, which estimates positional fluctuation of the protein molecules from the ensemble averaged structure, ii) corresponding diagonalization of the matrix to obtain the independent clusters. The $3N \times 3N$ sized covariance matrix is basically the correlation between positional deviations of the protein atoms. The translations and rotations of the molecules are completely removed from the atoms, and a minimum least square fit is implemented on them in order to estimate the effective fluctuation with respect to the reference structure. The mathematical representation of C_{ij} is,

$$C_{ij} = \frac{\left(\Delta\vec{r}_i(t).\Delta\vec{r}_j(t)\right)}{\left(\left\langle\left(\Delta\vec{r}_i(t)\right)^2.\left(\Delta\vec{r}_j(t)\right)^2\right\rangle\right)^{1/2}} \qquad (2.43)$$

where the angular brackets denote means of the enclosing quantity, $\Delta\vec{r}_i(t)$ is the displacement of the position of the i^{th} atom ($\vec{r}_i(t)$) from its corresponding mean value position. C_{ij} varies from +1 to –1, depending on correlation or anti-correlation of the i^{th} and j^{th} C_α atoms. The numerator of equation (2.43) determines the elements of 3-N dimensional covariance, denoting the ensemble of protein structures. The symmetric matrix C_{ij} can be diagonalized to obtain the corresponding Eigen vectors A,

$$A^T.C.A=\lambda \qquad (2.44)$$

The Eigen vectors signify the directions of the collective motions in configurational space, and λ denotes the corresponding Eigen values indicating the mean-squared fluctuations of the motions in the direction of these vectors. The Eigen vectors are thus the clusters of protein configurations, ranked sequentially based on the population of each of these clusters (signified by the magnitude of the eigen values) as well as on their contribution to the total mean squared fluctuation. Thus

most populated cluster, segregated from the protein conformational ensemble with PCA, is the cluster signified by the first Eigen vector, and so on. We mention here that PCA is often performed by considering a subset of the total number of protein atoms for computational efficiency. For example, covariance matrix of the backbone C_α atoms is oftentimes a sufficient representation of the overall dynamical extent.

2.7.3 Correlation functions

Correlation between two observables (x and y, for instance) obtained from MD simulation trajectories at every step (i) can be estimated by generalized correlation functions, which capture the effect of one variable on the properties of the other.

$$C_{xy} = \frac{1}{M}\sum_{i=1}^{M} x_i y_i \equiv \langle x_i y_i \rangle \tag{2.45}$$

Equation (2.45) gives a correlation of 1 or 0 or –1, depending on the observables correlated, uncorrelated or anti-correlated to each other over M data points. At certain times, mean values of the quantities x_i and y_i attain non-zero values. In such situations, the correlation function is obtained only from the fluctuating part as,

$$C_{xy} = \frac{1}{M}\sum_{i=1}^{M} \left(x_i - \langle x \rangle\right)\left(y_i - \langle y \rangle\right) \tag{2.46}$$

$\langle x \rangle$ and $\langle y \rangle$ are the mean values of the observables over M data points.

The phase space occupied by the system can evolve with time. Thus, properties as a function of time are often calculated from time dependent MD trajectories. So the correlation between two observables can also be estimated as *time correlations*, the generalized form of which is expressed as,

$$C_{xy}(t) = \frac{1}{N_0}\sum_{t_0} x_i(t + t_0) y_i(t_0) \tag{2.47}$$

At time t_0, the variable y can be correlated with variable x at a later time (t+t_0), over a total N_0 time frames, provided the variables are different from each other. If they are same, the

correlation is termed as an *autocorrelation function* (example provided in Figure 2.5) and expressed as

$$C_x(t) = \frac{1}{N_0}\sum_{t_0} x_i(t+t_0)x_i(t_0) \qquad (2.48)$$

Autocorrelation functions are measures of the retention of the properties of a particular variable with time. They can be even normalized by the function $C_x(0)$, from which decay functions of the variable can be retrieved. Examples or applications of autocorrelation function will be presented in the analyses like orientational relaxation, calculated IR spectra obtained from the Fourier transform of frequency-autocorrelation (in Chapter 4) and velocity-autocorrelation (in Chapter 5) respectively.

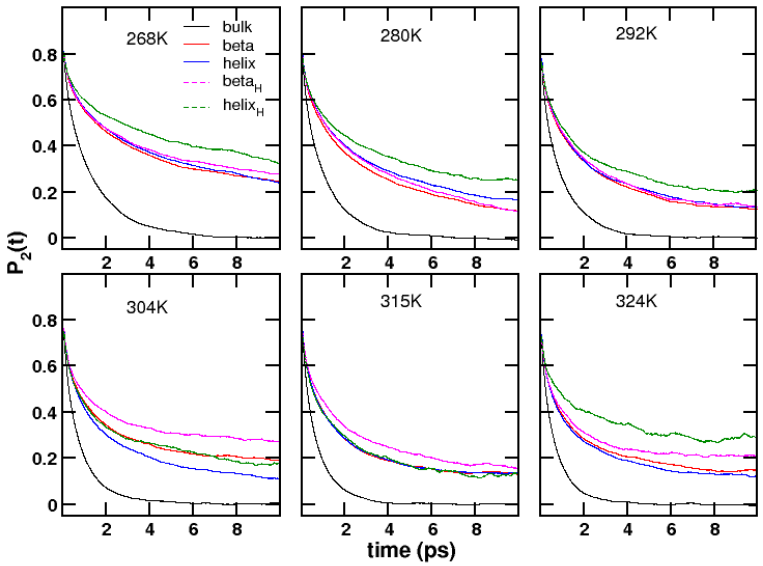


Figure 2.5 Orientational autocorrelation function ($P_2(t)$) for the dipole moment vector of bulk water and the hydration water molecules surrounding different domains (beta, helix, $beta_H$ (or hydrophobic beta), $helix_H$ (or hydrophobic helix)) of Yfh1 at different temperatures.

In addition to the temporal correlation function, where correlation between two observables are determined as a function of time, *spatial correlation functions* elucidate the spatial orientation between two random variables, eminent of which is the *radial distribution function*, (or *pair correlation function*), *g(r)* between two atom types (Figure 2.6). It describes

the variation of density or packing of an atom with respect to another, as a function of distance, and can be estimated by calculating the normalized probability of the position of the center of one atom at a particular distance with respect to the other, in three dimensional Cartesian coordinate system. Mathematically it can be represented as

$$g(r) = 4\pi r^2 \rho . dr \qquad (2.49)$$

ρ is the number density of the former atom as a function of distance, with respect to the other.

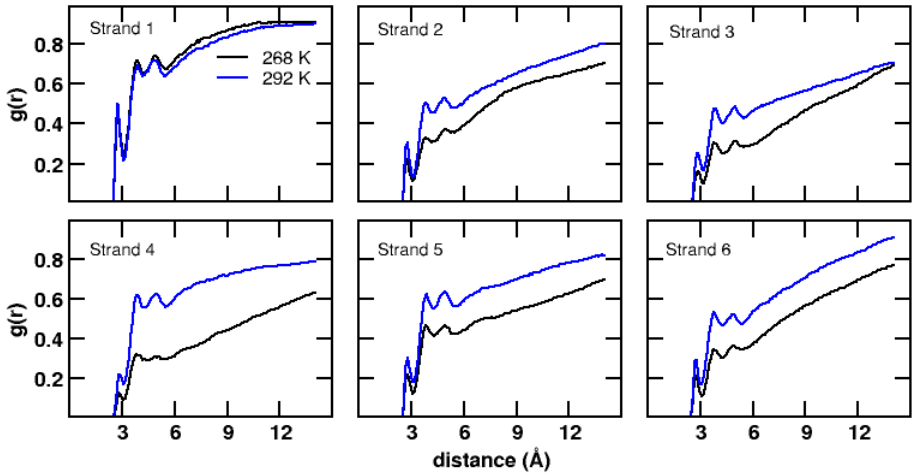


Figure 2.6 Plots of *g(r)* between the backbone heavy atoms and water oxygens for the different beta strands of Yfh1.

2.7.4 Association with spectroscopic measures

MD simulations are currently indispensable in understanding the structural properties, static and dynamic properties, thermodynamics and energetics, etc. of condensed matter from an atomic level perspective. On the other hand, spectroscopic methods deal with probing the properties of system by application of stimuli; the properties investigated as well as the results obtained depend on the nature of stimuli and the intrinsic system characteristics. In present times, MD simulations in conjunction with different spectroscopic techniques have been significantly helpful in obtaining invaluable insights for proteins, lipids, nucleic acids, etc., that were otherwise unattainable by either of the methods alone. In addition, replication of various spectroscopic measures in MD studies have recently proven to extensively determine their role

on complex biological systems, an example of which is the *Laser induced infrared pulsed spectroscopy*. Originally it has been proven to dissociate amyloid fibrils, and can also be replicated in simulation studies to computationally determine their effects in disrupting the fibrillar aggregates into smaller species.[32] The pulsed spectroscopy can be theoretically mimicked by application of a discontinuous electric field to the specified molecular aggregate, in a non-equilibrium MD (NEMD) setup. The molecule(s) of interest is thus relatively more heated; the remaining system is kept in equilibrium with the virtual thermal reservoir (as mentioned previously) to maintain the temperature of the overall system. The excess energy of the heated molecule gradually dissipates to the surrounding (solvent) environment, after excitation of the corresponding vibrational mode of proteins. Apart from reproducing different spectroscopic protocols in MD studies, there have been extensive collaborations between spectroscopic and simulation protocols to elucidate biomolecular properties. Herein, we discuss few of the important spectroscopic methods that are frequently used in conjunction with MD simulation of biomolecules.

Nuclear magnetic resonance (NMR) is a unique experimental protocol used extensively in biomolecular studies. In addition, nuclear magnetic resonance dispersion experiments (NMRD) provide valuable information about the hydration dynamics around different biomolecules.[33-35] Deuterated water, or water with ^{17}O is used, and the longitudinal and transverse relaxation rates are measured as a function of relaxation frequency, and the behavior of the hydration waters can be elicited from the relaxation properties. The measurements can be obtained for a wide range of frequency (few 100 MHz), corresponding to the dynamics obtained from sub-nanoseconds to micro-seconds timescale. The relaxation time of NMR corresponds mostly with motions on micro-seconds timescale, not overlapping with respective timescales of MD simulations. However, the dynamics of fast moving hydration water molecules on the nano- and sub-nanosecond timescales surrounding proteins and other biomolecules, captured at frequencies greater than 100 MHz, can be evidently compared with simulation results. Further, order parameters (S^2), useful in distinguishing different phases of the same substance, can be obtained from NMR experiments.[36] It helps determining the flexibility in protein secondary structure or orientation by the alignment of the N-H bonds (with isotopically labeled nitrogen (^{15}N) atoms) of backbone amide linkage (for instance), and can also be computed from MD simulations.[37-39] The functional form of S^2 can be denoted as,

$$S^2 = \frac{1}{2}\left[3\sum_{\alpha=1}^{3}\sum_{\beta=1}^{3}\left\langle \mu_\alpha . \mu_\beta \right\rangle^2 - 1\right] \tag{2.50}$$

Here, μ_α (α = 1, 2, 3) are the normalized components of the N–H vector along the three Cartesian coordinates. S^2 is calculated after a non-linear least square fit of C_α carbons of the final protein structure to the initial conformation, removing the overall rotational motion. But order parameters are mostly sensitive to the local motions of the proteins and least affected by the changes in global conformation.

As discussed in the first chapter, a number of biomolecular processes (especially for proteins) like making and breaking of non-covalent bonds, hydrogen bond formation, bond vibration, etc., occur on the pico- to nano-second timescales. These are further pronounced for smaller proteins and peptides with higher conformational flexibility, and greater interaction strength with the surrounding aqueous environment. These systems can be evidently useful in studying the formation of secondary structures during protein folding. In proteins, the intramolecular hydrogen bonds, often responsible for obtaining stable secondary structures, are responsible for conformations with well separated free energy minima. However, for small proteins and peptides, the number of such hydrogen bonds is much lesser, producing shallower free energy in conformational space, which helps exchange of conformations with each other at a much faster rate than proteins. Although, spectroscopic measures like NMR are useful in determining different structural parameters (as discussed), these protocols are relatively slower and thus inefficient to capture the indistinguishable conformational substates of proteins, giving only the time-averaged protein structure. Thus rapid conformational changes occurring in smaller peptides and proteins can be evidently captured with high resolution experimental protocols like Infrared (IR) spectroscopy, whose timescales are within the range of 1 ps. This method thus apparently freezes the conformational substates, and evidently records the rapid transition between the relative populations. In the fourth chapter, while studying the hydration behavior of denaturated protein segments in early cold denaturation mechanism, we have calculated IR spectra from high frequency MD simulation runs, based on the methodology originally formulated by Cho and co-workers.[40] It was initially developed on the amide-I mode of the N-methyl acetamide dimer. The changes in the corresponding base frequency due to influence of the surrounding medium (the remaining atoms including backbone, sidechain, and the solvent)

can be computed by the overall electrostatic potential exerted on the four atoms (C, O, N and H atoms) of the local amide-I mode. It can be mathematically represented as

$$\Delta\nu = \sum_{j=1}^{4} l_j \phi_j \tag{2.51}$$

Here, ϕ_j is the electrostatic potential on the j^{th} atom of the amide-I chromophore, l_j is the corresponding linear expansion constants. The off-diagonal elements of the amide-I modes are obtained from the nearest neighbor coupling (NNC) map.[41] The couplings due to the remaining chromophore units of the peptides are obtained from the transition dipole coupling (TDC) map.[42] The corresponding change in frequency is added to the base frequency (1707.7 cm^{-1}, for the CHO4 parameterization). The auto-correlation between the obtained frequencies, with subsequent Fourier transformation of the correlation function, will yield the motional narrowing incorporated IR frequency respectively.

The transverse, longitudinal and librational oscillations of the sequentially located water molecules have their corresponding peaks at ~50 cm^{-1}, ~200 cm^{-1}, and ~500 cm^{-1} respectively.[6,43,44] These modes of vibration are further enhanced when water molecules are restricted to confinements when present near to macromolecular solutes (proteins, DNA, nano-surfaces, etc.), and with changes in temperature. Water structural and dynamical properties in close association with biomolecular solutes are thus addressed by obtaining the corresponding spectra. Interestingly, these spectra can be calculated from Fourier transformation of velocity auto-correlation function of the oxygen and hydrogen atoms of water molecules from simulated trajectories, leading to computationally estimated vibrational density of states. The velocity auto-correlation function of the corresponding atom (O or H) of water molecule can be computed as

$$C_v(t) = \frac{\langle v_i(t).v_i(0)\rangle}{\langle v_i(0).v_i(0)\rangle} \tag{2.52}$$

Here, $v_i(0)$ and $v_i(t)$ are respective velocities of the i^{th} atom at initial time frame and time t respectively. In the fifth chapter, we have tried to correlate the changes in restricted motions of hydration water layer with corresponding changes in effective hydrophobicity of the protein

domains, by computing the velocity autocorrelation function and the corresponding vibrational density of states.

2.8 Advanced MD techniques

Unbiased MD simulation is often inadequate in overcoming large barriers (greater than several k_bT) in the energy landscapes of the complex system of interest. Implementations of advanced techniques, with the introduction of suitable biases, are therefore often incorporated in enhanced MD protocols; these enable the crossing of large barriers within finite computational times. We briefly discuss the principles of one such technique that has been useful in sampling protein conformational space.

Accelerated MD simulation (AMD). In this method, a boost energy, *E(b)*, and an acceleration parameter α, is applied to the original potential, $V(r)$, of the system of reference.[45,46] $V(r)$ can be the full potential, or a subset of the full potential. The modified potential $V^*(r)$ can be written as,

$$V(r) \geq E(b) \tag{2.53}$$

$$V^*(r) = V(r) + \Delta V(r), \quad V^*(r) < E(b) \tag{2.54}$$

$$\Delta V(r) = \frac{(E(b) - V(r))^2}{(E(b) - V(r)) + \alpha} \tag{2.55}$$

Extent of the corresponding 'acceleration' can be increased (or decreased) by regulating the parameters $E(b)$ (or α). AMD method has been successfully applied to the torsional part of the P.E. In this protocol, first, short unbiased simulations are undertaken for obtaining the dihedral potential from which $V(r)$ is determined. The magnitude of $E(b)$ is determined in a way that difference in $V(r)$ and $E(b)$ leads to 4 kcal mol^{-1} times the number of residues present in the protein.[46]

2.9 Bibliography

(1) Parr, R. G. Density functional theory. *Annu. Rev. Phys. Chem.* **1983**, *34*, 631-656.

(2) Andrew, R. L. Molecular Modelling Principles and Applications. *Pearson Education Ltd. 2nd edition*. **2001**.

(3) Berkowitz, M. L.; VÁcha, R. Aqueous Solutions at the Interface with Phospholipid Bilayers. *Accounts Chem. Res.* **2012**, *45*, 74-82.

(4) Bhattacharjee, N.; Biswas, P. Structure of hydration water in proteins: A comparison of molecular dynamics simulations and database analysis. *Biophys. Chem.* **2011**, *158*, 73-80.

(5) Bizzarri, A. R.; Cannistraro, S. Molecular Dynamics of Water at the Protein-Solvent Interface. *J. Phys. Chem. B*. **2002**, *106*, 6617-6633.

(6) Chakraborty, K.; Bandyopadhyay, S. Effect of temperature on the low-frequency vibrational spectrum and relative structuring of hydration water around a single-stranded DNA. *J. Chem. Phys.* **2015**, *142*, 015101.

(7) Wang, C.; Bradley, P.; Baker, D. Protein-Protein Docking with Backbone Flexibility. *J. Mol. Biol.* **2007**, *373*, 503-519.

(8) Das, P. Effect of cosolvents on nano-confined water: a molecular dynamics study. *Nanoscale*. **2012**, *4*, 2931-2936.

(9) Das, P.; Matysiak, S. Direct Characterization of Hydrophobic Hydration during Cold and Pressure Denaturation. *J. Phys. Chem. B*. **2012**, *116*, 5342-5348.

(10) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature*. **2001**, *414*, 188-190.

(11) Jose, J.; Sengupta, N. Molecular dynamics simulation studies of the structural response of an isolated $A\beta_{1\text{-}42}$ monomer localized in the vicinity of the hydrophilic TiO_2 surface. *Eur. Biophys. J.* **2013**, *42*, 487-494.

(12) Vieceli, J.; Roeselová, M.; Potter, N.; Dang, L. X.; Garrett, B. C.; Tobias, D. J. Molecular Dynamics Simulations of Atmospheric Oxidants at the Air-Water Interface: Solvation and Accommodation of OH and O3. *J. Phys. Chem. B*. **2005**, *109*, 15876-15892.

(13) Lamoureux, G.; Harder, E.; Vorobyov, I. V.; Roux, B. t.; MacKerell Jr, A. D. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chem. Phys. Lett.* **2006**, *418*, 245-249.

(14) Kresse, G.; Hafner, J. *Ab Initio* molecular dynamics for liquid metals. *Phys. Rev. B*. **1993**, *47*, 558-561.

(15) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *J. Am. Chem. Soc.* **1995**, *117*, 5179-5197.

(16) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400-1415.

(17) Klauda, J. B.; Venable, R. M.; Freites, J. A.; O'Connor, J. W.; Tobias, D. J.; Mondragon-Ramirez, C.; Vorobyov, I.; MacKerell, A. D.; Pastor, R. W. Update of the CHARMM All-Atom Additive Force Field for Lipids: Validation on Six Lipid Types. *J. Phys. Chem. B*. **2010**, *114*, 7830-7843.

(18) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.* **1999**, *151*, 283-312.

(19) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327-341.

(20) Grubmüller, H.; Heller, H.; Windemuth, A.; Schulten, K. Generalized Verlet Algorithm for Efficient Molecular Dynamics Simulations with Long-range Interactions. *Mol. Simul.* **1991**, *6*, 121-142.

(21) Tuckerman, M.; Berne, B. J.; Martyna, G. J. Reversible multiple time scale molecular dynamics. *J. Chem. Phys.* **1992**, *97*, 1990-2001.

(22) Sborgi, L.; Verma, A.; Piana, S.; Lindorff-Larsen, K.; Cerminara, M.; Santiveri, C. M.; Shaw, D. E.; de Alba, E.; Muñoz, V. Interaction Networks in Protein Folding via Atomic-Resolution Experiments and Long-Time-Scale Molecular Dynamics Simulations. *J. Am. Chem. Soc.* **2015**, *137*, 6506-6516.

(23) David, E. S.; Grossman, J. P.; Joseph, A. B.; Brannon, B.; Butts, J. A.; Jack, C. C.; Martin, M. D.; Ron, O. D.; Amos, E.; Christopher, H. F.; Anthony, F.; Joseph, G.; Gennette,

G.; Brian, G.; Ho, C. R.; Douglas, J. I.; Lev, I.; Jeffrey, S. K.; Richard, H. L.; Timothy, L.; Li-Siang, L.; Adam, K. L.; Chester, L.; Daniel, K.; Kenneth, M. M.; Shark Yeuk-Hai, M.; Mark, A. M.; Rolf, M.; Lawrence, J. N.; Jon, L. P.; Terry, Q.; Daniel, R.; John, K. S.; Daniele, P. S.; Schafer, U. B.; Naseer, S.; Christopher, W. S.; Jochen, S.; Ping Tak Peter, T.; Michael, T.; Horia, T.; Brian, T.; Benjamin, V.; Stanley, C. W.; Cliff, Y. Anton 2: raising the bar for performance and programmability in a special-purpose molecular dynamics supercomputer. In *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*; IEEE Press: New Orleans, Louisana, 2014.

(24) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.

(25) Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **1980**, *72*, 2384-2393.

(26) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **1984**, *81*, 3684-3690.

(27) Chatterjee, P.; Sengupta, N. Effect of the A30P mutation on the structural dynamics of micelle-bound αSynuclein released in water: a molecular dynamics study. *Eur. Biophys. J.* **2012**, *41*, 483-489.

(28) Chatterjee, P.; Bagchi, S.; Sengupta, N. The non-uniform early structural response of globular proteins to cold denaturing conditions: A case study with Yfh1. *J. Chem. Phys.* **2014**, *141*, 205103.

(29) Lin, Y.-S.; Bowman, Gregory R.; Beauchamp, Kyle A.; Pande, Vijay S. Investigating How Peptide Length and a Pathogenic Mutation Modify the Structural Ensemble of Amyloid Beta Monomer. *Biophys. J.* **2012**, *102*, 315-324.

(30) Jose, J. C.; Chatterjee, P.; Sengupta, N. Cross Dimerization of Amyloid-β and αSynuclein Proteins in Aqueous Environment: A Molecular Dynamics Simulations Study. *PLoS ONE*. **2014**, *9*, e106883.

(31) Nguyen, P. H.; Li, M. S.; Derreumaux, P. Amyloid oligomer structure characterization from simulations: A general method. *J. Chem. Phys.* **2014**, *140*, 094105.

(32) Hoang Viet, M.; Derreumaux, P.; Li, M. S.; Roland, C.; Sagui, C.; Nguyen, P. H. Picosecond dissociation of amyloid fibrils with infrared laser: A nonequilibrium simulation study. *J. Chem. Phys.* **2015**, *143*, 155101.

(33) Denisov, V. P.; Halle, B. Protein Hydration Dynamics in Aqueous Solution: A Comparison of Bovine Pancreatic Trypsin Inhibitor and Ubiquitin by Oxygen-17 Spin Relaxation Dispersion. *J. Mol. Biol.* **1995**, *245*, 682-697.

(34) Denisov, V. P.; Carlström, G.; Venu, K.; Halle, B. Kinetics of DNA hydration. *J. Mol. Biol.* **1997**, *268*, 118-136.

(35) Persson, E.; Halle, B. Cell water dynamics on multiple time scales. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 6266-6271.

(36) LeMaster, D. M. NMR Relaxation Order Parameter Analysis of the Dynamics of Protein Side Chains. *J. Am. Chem. Soc.* **1999**, *121*, 1726-1742.

(37) Smith, L. J.; Mark, A. E.; Dobson, C. M.; van Gunsteren, W. F. Comparison of MD Simulations and NMR Experiments for Hen Lysozyme. Analysis of Local Fluctuations, Cooperative Motions, and Global Changes. *Biochemistry*. **1995**, *34*, 10918-10931.

(38) Koller, A. N.; Schwalbe, H.; Gohlke, H. Starting Structure Dependence of NMR Order Parameters Derived from MD Simulations: Implications for Judging Force-Field Quality. *Biophys. J.* **2008**, *95*, L04-L06.

(39) Gu, Y.; Li, D.-W.; Brüschweiler, R. NMR Order Parameter Determination from Long Molecular Dynamics Trajectories for Objective Comparison with Experiment. *J. Chem. Theory Comput.* **2014**, *10*, 2599-2607.

(40) Ham, S.; Cho, M. Amide I modes in the N-methylacetamide dimer and glycine dipeptide analog: Diagonal force constants. *J. Chem. Phys.* **2003**, *118*, 6915-6922.

(41) la Cour Jansen, T.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. Modeling the amide I bands of small peptides. *J. Chem. Phys.* **2006**, *125*, 044312.

(42) Wang, L.; Middleton, C. T.; Zanni, M. T.; Skinner, J. L. Development and Validation of Transferable Amide I Vibrational Frequency Maps for Peptides. *J. Phys. Chem. B*. **2011**, *115*, 3713-3724.

(43) Pal, S.; Balasubramanian, S.; Bagchi, B. Dynamics of bound and free water in an aqueous micellar solution: Analysis of the lifetime and vibrational frequencies of hydrogen bonds at a complex interface. *Phys. Rev. E*. **2003**, *67*, 061502.

(44) Choudhury, N.; Pettitt, B. M. Dynamics of Water Trapped between Hydrophobic Solutes. *J. Phys. Chem. B*. **2005**, *109*, 6422-6429.

(45) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919-11929.

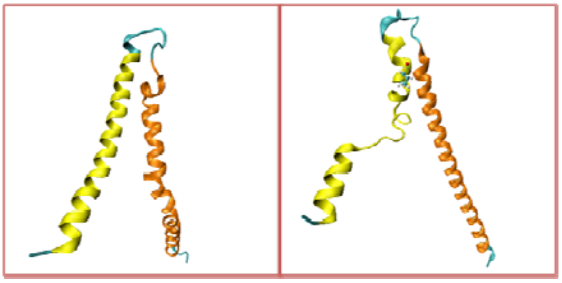
(46) Markwick, P. R. L.; McCammon, J. A. Studying functional dynamics in bio-molecules using accelerated molecular dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20053-20065.

Chapter 3

Early structural dynamics of the wild type and a mutated variant of the protein αSynuclein in explicit aqueous environment

Do not figure out big plans at first, but, begin slowly, feel your ground and proceed up and up.
- Swami Vivekananda

Abstract



Atomistic molecular dynamics simulation is used to probe the effect of A30P mutation on the structural dynamics of micelle-bound, helical αSynuclein when released in aqueous environment. On the simulated timescales, the effect of the mutation on secondary structure is restricted to local level changes close to the mutation site in the N-terminal helical domain. The changes are transient, and all residues except Lys23 recover initial structure. The local behavior due to the mutation gives rise to a global level difference in the A30P mutant in the form of a permanent kink in the N-terminal helical domain.

3.1 Introduction

Spontaneous aggregation of αSynuclein (αS), a ~14.5 KDa protein present in neuronal cytosol, is the causative factor behind Parkinson's disease.[1] Although its physiological role is unclear, monomeric αS is involved in neurotransmitter release,[2,3] in ER/Golgi trafficking,[4] or associated with synaptic vesicles.[5,6] Residues 61 to 95 are associated with the non-amyloid component (NAC) aggregates in Alzheimer's disease plaques.[7] The monomeric form is thought to have no stable tertiary structure in aqueous environment, while it assumes helical conformations when bound to amphiphilic micelles and lipid bilayers.[8-11] *In vivo*, roughly 15% of αS is membrane bound at any given moment, and the seeding of aggregation in cytosolic αS has been found to be initiated by the membrane bound form.[12] It preferentially binds to anionic membranes but also binds with neutral membranes.[13] Since αS co-exists in aqueous media with the lipid bound form *in vivo*, it is plausible that the helix ✶ coil ✶ *β*-sheet transitions could constitute important, early components of the fibrillogenesis pathway, an understanding of which could shed more light on the importance of helical or partially folded intermediates in amyloidogenesis.[14-17]

Early-onset Parkinson's disease has been linked to three point mutations (A30P, A53T, E46K) in αS.[18-20] *In vitro* studies indicate that fibrillation rate increases with A53T and E46K,[21,22] while A30P promotes formation of prefibrillar oligomeric species.[23] NMR analysis suggests that A30P and A53T do not affect the protein's global structural properties except some minor perturbation around the A30P mutation site.[24] But analysis of C_α secondary shifts reveals that A30P mutation decreases the helical propensity and elongates the fibrillization lag phase.[24,25]

Despite numerous studies, the mechanistic role of mutations in affecting αS behavior remains incompletely understood. The A30P mutation presents a particularly interesting case as proline, usually found at helical ends, in turn regions, or at the edge strands of beta sheets, behaves as a structural disruptor when occurring mid-sequence in a protein.[26] In αS, this mutation has been found to cause several interesting alterations to the protein's intrinsic structural dynamics.[24,27] Experiments and simulations indicate that in lipid environment, it gives rise to a destabilization upstream and downstream from the mutation site,[24,27] and decreases the helical propensity and lipid binding affinity.[28-30] The strong interactions of WT αS when bound to micelles are found to induce strain in the protein structure which is released with the A30P mutation.[27,31] However, it remains to be understood whether the strain is also present when the helical micelle bound form

is released into the aqueous environment, and to what extent the A30P mutation affects the structural persistence of the WT form, hence reflecting the difference of the two variants in the early components of fibrillogenesis as discussed earlier.

Starting with the micelle-bound conformation, we have compared the early structural evolution of the monomeric WT and A30P mutant to WT αS using atomistic molecular dynamics (MD) simulations. While neither form unfolds completely within the simulation time of 90 ns, our analysis reveals key differences in their intrinsic structural dynamics, attributable to the N-terminal region. The backbone RMSD of the A30P mutant attains a high but stable value at about 20 ns, while that of the WT protein continues to increase throughout the simulation. Our torsional angle based structural persistence parameters indicate that unlike the WT, the first helical domain of the A30P undergoes a transient loss in secondary structure due to the weakening of the backbone hydrogen-bonding network. The loss in helicity is localized, occurring at least five residues away from the mutation site towards the N-terminus, but gives rise to a kink-like feature that persists even after secondary structure is recovered. Our results indicate that on release into the aqueous environment, the A30P form of the micelle bound protein releases structural strain more easily than the WT form. We have discussed the implications of our observations in the aggregation pathway of αS.

3.2 Materials and methods

3.2.1 System setup and MD simulations

The micelle-bound form of αS (PDB entry 1XQ8) was the starting structure in our study.[31] This structure consists an N-terminal helical domain (Val3-Val37), a well-ordered extended linker (Leu38-Thr44), a C-terminal helical domain (Lys45-Thr92), another linker (Gly93-Lys97) and a highly disordered tail (Asp98-Ala140).[31] The tail region (residues 96 to 140) was truncated to study the enhanced fibrillization propensity, [32,33] while selecting the helical domains. The structure was mutated at position 30 with Proline in the trans conformation to obtain the A30P form.[27] Four chloride counter ions were added to neutralize each system, which were then solvated explicitly with about 9000 TIP3P water molecules.[34] The box dimensions used for the systems measured about 60 Å, 102 Å and 46 Å. The systems were simulated using the NAMD 2.7 simulation package,[35] and the CHARMM22 all-atomic force field with CMAP correction.[36] The time step was taken to be 2 fs. Initially energy minimizations of 15000 steps were performed

based on the conjugate gradient method, following which the systems were simulated at a pressure of 1 atmosphere and a temperature of 310 K. Constant temperature was maintained using Langevin dynamics with a collision frequency of 1 ps^{-1}. The Nose-Hoover Langevin piston maintained constant pressure,[37] and SHAKE was used to constrain the lengths of hydrogen bonds.[38] Cutoffs of 12 Å were employed with smooth truncation starting at 10 Å. Long range electrostatics was calculated with Particle Mesh Ewald (PME).[39] 90 ns simulation trajectories were generated for both forms. VMD was used for visualization.[40]

3.2.2 Analyses protocol(s)

Structural Persistence Parameter

We define a structural persistence parameter *P*, to characterize the extent of secondary structure persistence of a chosen protein domain, relative to a reference structure (as depicted in *Chapter 2*). This parameter is calculated as,

$$P = \frac{1}{N_{res}} \sum_{j=1}^{N_{res}} e^{-(\Delta\phi_j / \Delta\phi_{\max})} . e^{-(\Delta\psi_j / \Delta\psi_{\max})} \qquad (3.1)$$

Here, $\Delta\phi_j$ and $\Delta\psi_j$ are, respectively, the changes in the ϕ and ψ torsional angles of the *j*'th residue over a reference structure, and the scaling factors $\Delta\phi_{\max}$ and $\Delta\psi_{\max}$ represent the maximal changes that can occur in the torsional angles in the Ramachandran diagram. The maximum value of the persistence parameter, $P = 1$, represents a conformation whose secondary structure is completely unchanged relative to the reference. For a conformation in which nearly every residue undergoes maximal change in the torsion angles, *P* is $\sim e^{-2}$. Lower values indicate lesser overall degrees of secondary structural persistence, and thus *P* is a measure of the global extent of structural disorder of the chosen domain over the reference structure.

3.3 Results and discussions

We present representative snapshots of the protein from the WT and A30P trajectories in Figure 3.1, from which emerge several aspects of the structural evolution, and some interesting differences between the 'global' behavior of WT and A30P monomers. First, the initial, micelle-bound structure does not appear to undergo extensive unfolding in water within the simulated timescale. Second, the loss in helicity appears to be transient and localized to small, contiguous domains within the N- and C-terminal helical domains. While no noticeable helicity loss is found

in the N-terminal domain (residues Val3-Val37; henceforth referred to as N-helix) in the WT trajectory, transient 'local' level helicity losses are noticed in the A30P trajectory. Closer inspection reveals that the helicity loss occurs away from the site of the mutation, towards the N-terminus (ie. 'upstream'), and is recovered at the latter part of the trajectory. For the C-terminal helical domain (Lys45-Thr92; henceforth referred to as C-helix), there appear to be transient helical losses to a smaller extent for both mutants. The snapshots also suggest that the helical domains bend and form distinct kinks. The N-helix forms distinct kinks only in the A30P trajectory, which persists after the helicity has been recovered. On the other hand, bends in the C-helix appear in both trajectories, and appear to be of a more transient nature.

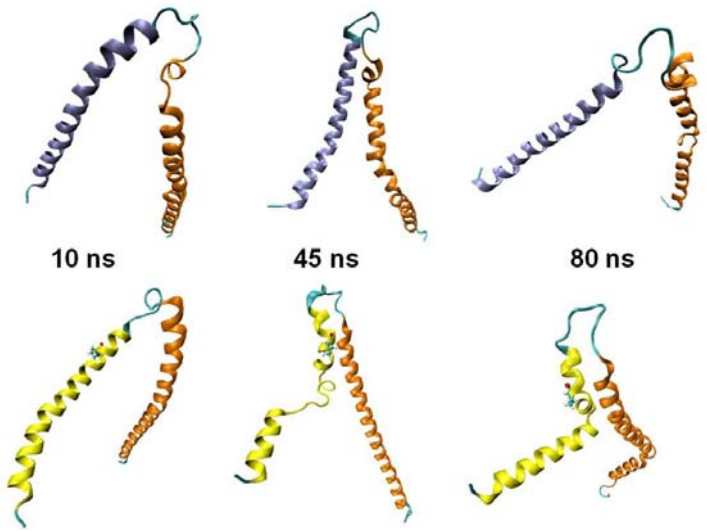


Figure 3.1 Snapshots of the WT (upper row) and the A30P variant (lower row) taken at 10ns, 45ns and 80ns of the respective trajectories. The N-helix of WT and A30P are shown in blue, and yellow, respectively. The proline mutation in A30P is marked with CPK representation. C-helices are shown in orange.

In Figure 3.2, we have plotted the root mean squared deviations (RMSD) of backbone atoms of the WT and A30P proteins obtained over the simulations, relative to the initial structure. The RMSD for the WT protein displays a steady rise over the 90 ns, indicating a gradual but steady change in the structural evolution. On the other hand, the RMSD for the A30P variant shows an initial rapid rise (by nearly 10 Å), followed by a nearly 70 ns period of marked stability. Thus on the whole, the A30P variant has an initial phase wherein it undergoes a rapid structural change, and the resultant stable structure persists for a longer time. The mean RMSD value of the WT

over the entire simulation trajectory is 10.5 (± 3.2) Å. For the A30P, the mean value over the first 20 ns is 9.7 (± 3.3) Å, and 12.4 (± 0.9) Å for the latter part.

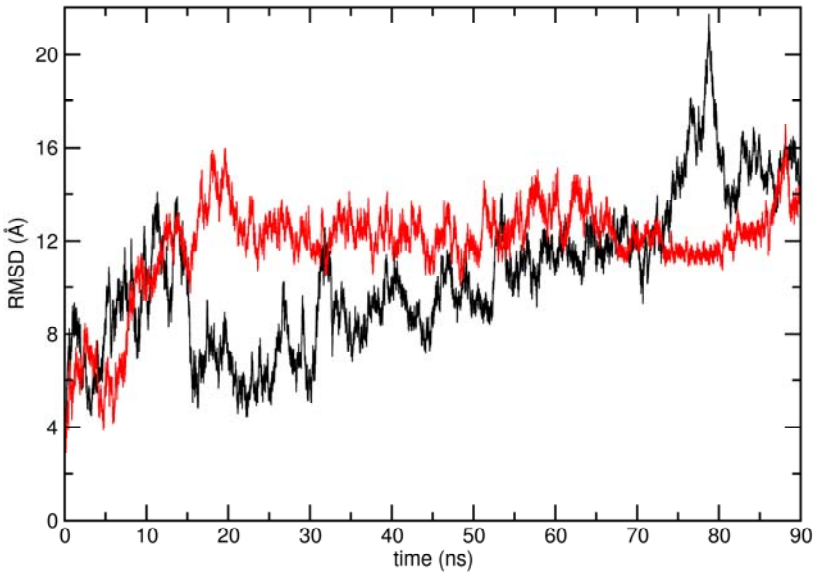


Figure 3.2 Evolution of the backbone RMSDs of WT (*in black*) and A30P (*in red*) as a function of simulation time.

To quantify the extent of structural persistence of the micelle-bound structure, and to probe the effect of the proline mutation on the structural evolution at the 'global' and 'local' scale, we have obtained the persistence parameter *P* as defined in equation 1 for various domains along the trajectories, with the original structure as the reference. In Figure 3.3, we show *P* obtained for the whole protein, and individually for the N- and the C-helix. When considering the entire protein, *P* consistently shows a high value (0.85 or more) throughout the simulation time. Thus, at the scale of the full protein, the overall structural integrity of both the whole WT and A30P systems is largely maintained. At the level of the individual helical domains, *P* remains high but reveals interesting differences between WT and A30P. For the N-helix in WT, the mean value of *P* fluctuates over the simulation is 0.93 (± 0.01). The helix in A30P, on the other hand, shows a dramatic drop between 25 and 30ns, which recovers after 65 ns. We note that the drop commences at around the same time that the RMSD value attains the high, stable value, and is coincident with the onset of the upstream helical losses and bending in the N-helix. The mean *P* values between 0-25 ns, 30-60 ns, and 70-90 ns for the A30P N-helix are 0.93 (± 0.01), 0.82 (± 0.04) and 0.91 (± 0.01), respectively. For the C-helix, *P* follows a pattern that is independent of

the N-helix in both trajectories. While *P* for this helix is high for both systems (~0.9 or more), the values are slightly higher for the A30P, but merge during the last ~25 ns. Thus, the slightly higher whole protein *P* of A30P over the WT during the first ~20 ns should be attributable to unfolding differences in the C-terminal domain; the fall and subsequent merging with the value of WT at ~25 ns is due to the significant unfolding in the N-terminal domain.

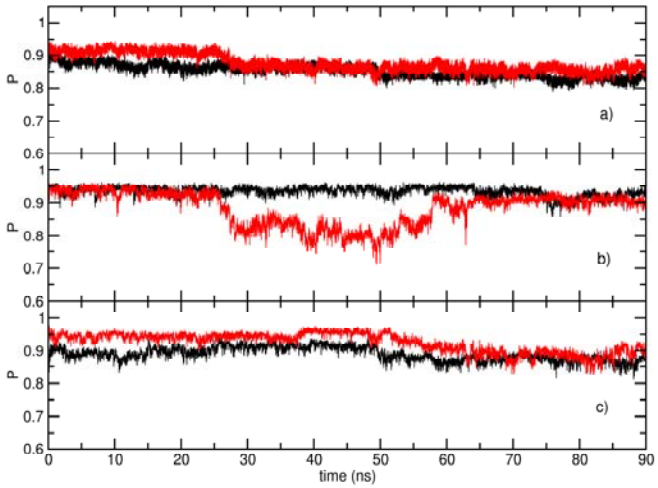


Figure 3.3 Plots of the persistence parameter *P* of the WT (*in black*) and A30P (*in red*) variants as a function of simulation time, shown for *a)* the whole peptide segments, *b)* the N-helix, *c)* the C-helix.

To elucidate the local, residue-level unfolding behavior due to mutation, we have recalculated *P* for three-residue groups in the N- and the C- helices. In Figure 3.4, we compare *P* for groups within the N-helix from the WT and A30P trajectories. Comparison with the evolution of *P* for the entire N-helix confirms that the disorder due to the mutation arises upstream, towards the N-terminus. Residue groups downstream (see $K_{32}TK_{34}$) and at the mutation site (see $A_{29}PG_{31}$) have high *P* values (~0.95) throughout the simulation. The domain-level drop in *P* observed for the entire N-helix at 25 ns is clearly found to be due to residue groups $K_{23}QG_{25}$, $A_{17}AA_{19}$, and $G_{14}VV_{16}$. Groups $V_{26}AE_{28}$, $E_{20}KT_{22}$ and $A_{11}KE_{13}$ display high but fluctuating *P* values that show no evident correlations with the domain-level trend. Notably, while *P* is recovered for $A_{17}AA_{19}$ and $G_{14}VV_{16}$, the decrease in *P* persists for $K_{23}QG_{25}$ for the remaining simulation time, which is thus the only domain that records a permanent secondary structure change due to the mutation. Further, as backbone hydrogen bonding network in alpha helices are of $i \leftarrow i+4$ nature (the arrow indicating the direction of hydrogen bond), the absence of the (NH) group at the location of the

mutation should cause a cascading weakening in structural hydrogen bonds towards the N-terminus. Individually, we find that Lys23 shows the greatest 'permanent' drop in *P* at 25 ns (by about 0.6; data not shown); *P* for other residues recovers close to 65 ns. On the other hand, 'local' *P* values in the C-terminal domain do not reveal clear differences between; transient helical unfolding from Gly67-Val74 is noticed in both the variants (Figure 3.5).

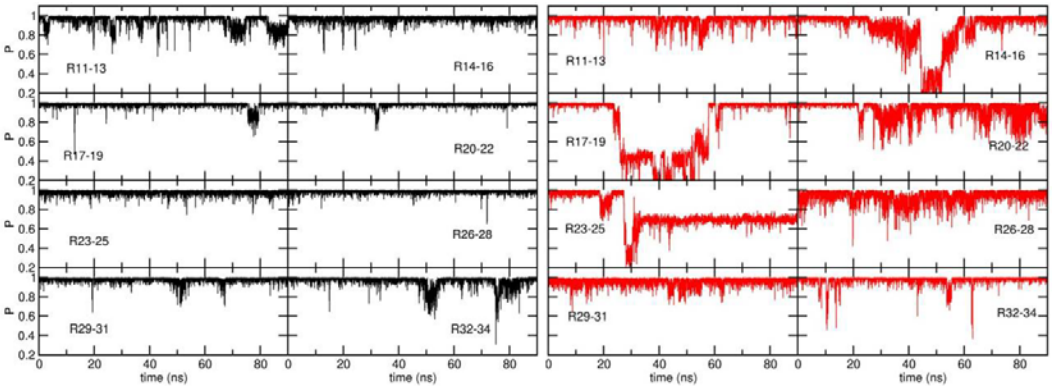


Figure 3.4 Plots of the persistence parameter *P* for 3 residue groups in the N-helix of the WT (*in black*) and A30P (*in red*) variants as a function of simulation time.

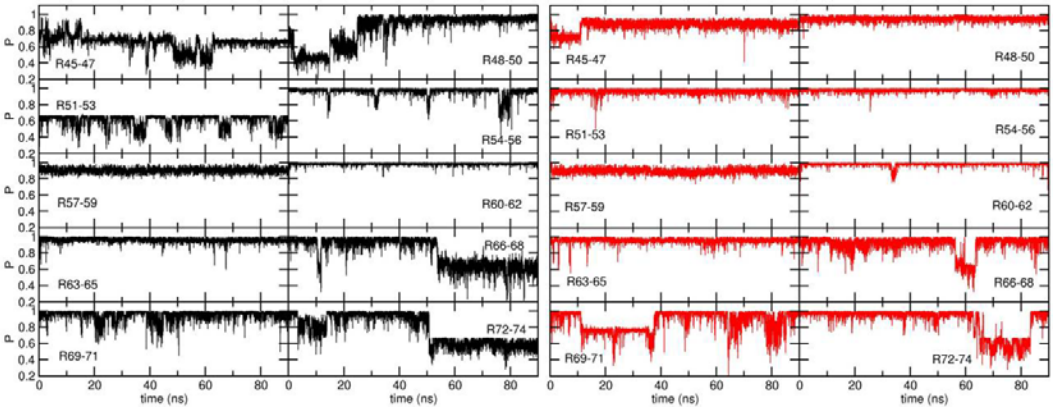


Figure 3.5 Plots of the persistence parameter *P* for 3 residue groups in the C-helix of the WT (*in black*) and A30P (*in red*) variants as a function of simulation time.

To probe the propensity for helical bending, we have evaluated the bend angles in the N-helix, θ_1 (defined as the angle at Lys23, ie. between Val3-Thr22 and Gln24-Val37) and in the C-helix, θ_2 (the angle made by the residue stretches Lys45-Val66 and Thr75-Val95, which show higher structural persistence). In Figure 3.6, we have compared histograms of θ_1 and θ_2 over segments of the simulation trajectory chosen when *P* for the N-helix in A30P did not undergo significant changes (0-20ns; 30-50 ns; and 70-90 ns). For the WT trajectory, the distribution of θ_1 is peaked

between 160° and 170°, and does not show any significant changes over the chosen time windows. However, for A30P, this distribution shows a clear shift in the peak maximum as time progresses. Within the first 20 ns, most conformations correspond to a nearly unbent N- helix, with a distinct peak in θ_1 at ~160°, and a smaller peak at ~120°. Between 30 and 50 ns, we find distinct shifts in the peak, with significant contributions from 'kinked' conformations where θ_1 is peaked at ~75° and ~115°. In the last 20 ns of the simulation, the distribution is strongly peaked at ~65°, showing a distinct tendency of the first helical domain to bend in the A30P trajectory.

The distributions for the bending angle θ_2, on the other hand do not show clear differences between the simulated WT and A30P systems. For WT, the initial peak (at ~140°) at 0-20 ns shifts towards a slightly *higher* value (~160°) between 30-50 ns. For the last 20 ns, the distribution (peaked at ~130°) is broader, showing no clear preference for either the straightened or the bent form. In the case of A30P, the first 20 ns correspond to two distinct peaks (at ~80° and ~160°). However, between 50 and 70 ns, the peak at ~80° is lost, signifying a decrease in the bent structures. For the last 20 ns, the distribution actually contracts around the peak of ~140°, also demonstrating a final preference for neither straight nor bent structure. Therefore, we find no clear correlation between the trends of θ_1 and θ_2.

We point out here that for the A30P trajectory, the initial drop and subsequent stability in θ_1 (Figure 3.7) is reminiscent of the temporal pattern of the backbone RMSD. Thus, the kink formation in the first helical domain is the main structural effect of the mutation at the 'global' level. This is confirmed by evaluating the backbone RMSDs individually for the N- and the C-helix (Figure 3.8).

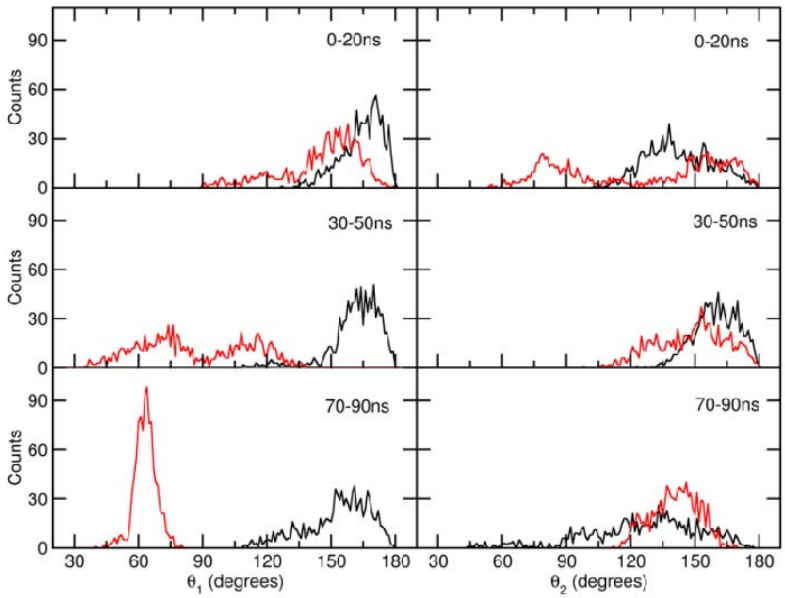


Figure 3.6 Histograms of bend angles θ_1 and θ_2 for the WT (*in black*) and A30P (*in red*) variants, taken over 0-20 ns, 30-50 ns and 70-90 ns.

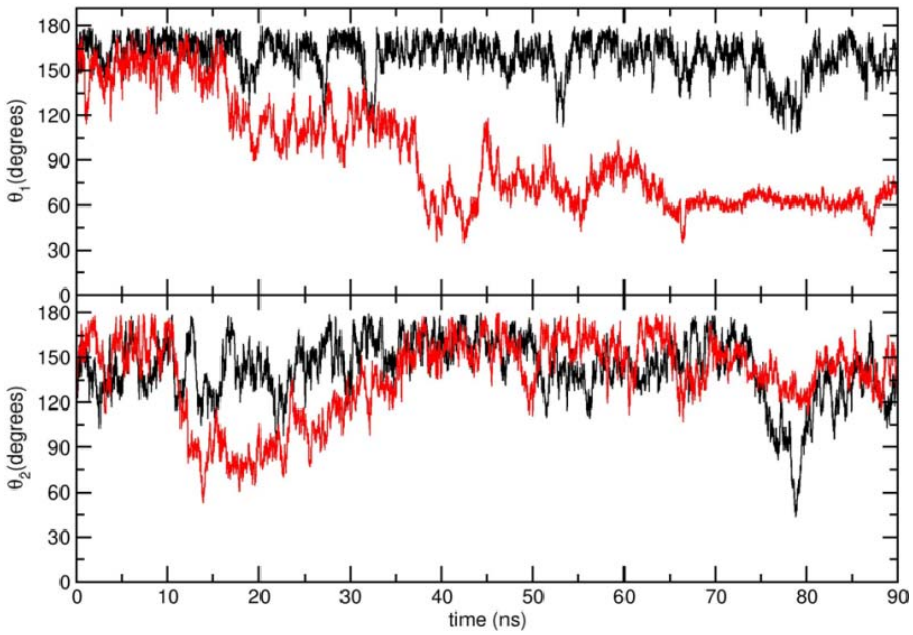


Figure 3.7 Evolution of the bend angles θ_1 and θ_2 for the WT (*in black*) and A30P (*in red*) variants as a function of simulation time.

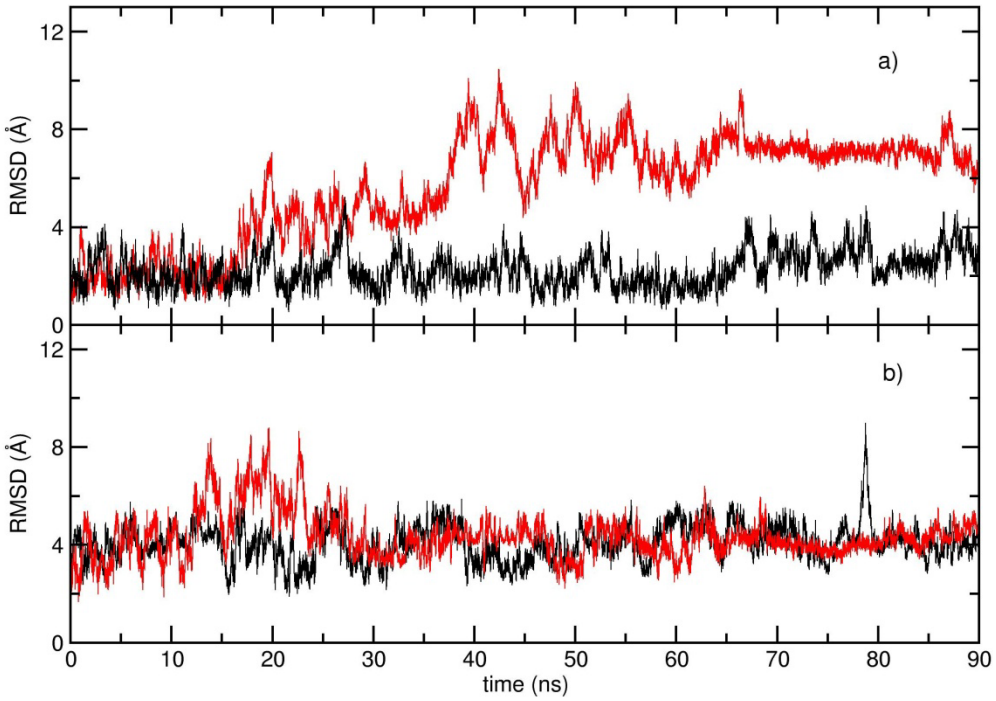


Figure 3.8 Plots of the backbone RMSDs for *a)* N-Helix and *b)* C-Helix as a function of simulation time for the WT (*in black*) and A30P (*in red*) variants.

3.4 Conclusions

We have reported the effect of A30P mutation on the early structural dynamics of the micelle bound structure of αS when released in aqueous environment. Globally, both the WT and the mutant forms show high levels of structural persistence on nanosecond timescales. However, the A30P mutation is found to cause partial unfolding in the N-terminal domain. The loss of helicity in this domain is transient, and all residues except Lys23, located seven residues upstream from the site of the mutation, fully regain their helical conformation. We find that the sudden drop in the persistence parameter for the N-helix of A30P is nearly commensurate with the onset of a kink-like feature. The kink formation is stable, and does not reverse with the recovery of overall helicity for most residues. We find that the sharp increase in backbone RMSD of the A30P trajectory at ~20 ns is commensurate with a drop in θ_1. Unlike the persistence parameter, the starting value of θ_1 is not recovered. Thus, the higher, stable RMSD in the A30P variant in the latter part of the trajectory is due to the formation of a stable kink in the N-helix. This is confirmed by comparing the backbone RMSD plots of the individual helices with θ_1.

It is noted here that micelle-bound monomeric αS is reported to be a strained structure,[27] and the Proline mutation has been shown to play a key role in the strain release of the A30P form by either upstream or downstream helical destabilization.[24,27] Our MD simulation results suggest that in an aqueous environment, the transient structural disorder upstream from the mutation site and the permanent kink formation in the N-helix of the A30P could be a key mechanism for strain release. In comparison, the extent of strain released in the C-helix should be comparable in both mutant varieties. Continuing studies in our group will focus on the interplay of various thermodynamic factors behind the structural changes and associated strain release. As a preliminary analysis, we have compared temporal patterns of the energies of backbone hydrogen-bonding network of the N- and C-helices for both trajectories (Figure 3.9). The data shows that the A30P mutation transiently weakens the hydrogen-bonding network associated with the structure of the N-helix; the magnitude of the destabilization from 25 to 60 ns is greater than 400 kcal mol^{-1}. Comparison with the *P* for the N-helix suggests that entropic effects play an important role in the helical unfolding. We note that the hydrogen-bonding network in the C-helix remains unaffected, reiterating that structural dynamics of the two domains are independent within the simulated timescales.

The destabilization of backbone hydrogen-bonding network in the A30P mutant could be one of the key reasons leading to the recovery of the helical structure on the simulated timescales. However, since all mutant forms of monomeric αS are experimentally found to be intrinsically unstructured in water, the helical reformation itself should be transient when longer timescales are considered. Therefore, the unfolding pathway of the protein may involve populations of structures with varying degrees of helicity. Longer trajectories (possibly up to the micro- or the millisecond time regime) will be required to reveal in detail, the interplay between various thermodynamic and kinetic factors that eventually lead to total unfolding of the peptide in aqueous environment. Further, longer simulations involving multiple monomers will be able to address how 'global' conformational features arising due to the mutations and encountered along the unfolding pathway can modulate the self-assembly process. For example, the kink in of the A30P trajectory may affect the relative proximities, and thereby the interaction strengths, of key residues that participate in the self-assembly mechanism.

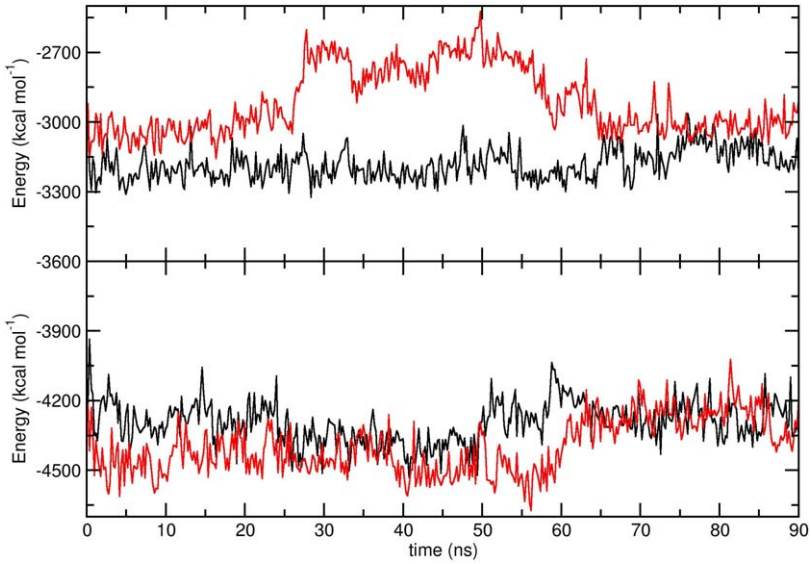


Figure 3.9 Plots of the hydrogen bonding network energy as a function of simulation time for N-Helix (upper row) and C-Helix (lower row) as a function of simulation time for the WT (*in black*) and A30P (*in red*) variants. The hydrogen bonding network energy for each helix is calculated as the total non-bonded interaction energy between the backbone amide groups and the oxygen atoms of the backbone carbonyl groups.

Single point mutations in the amino acid sequence of αS are responsible for significant alterations to the protein's structural behavior and self-assembly. A clearer, molecular level picture of the underlying differential structural dynamics of the protein's mutant forms is an essential requirement for understanding their complex aggregation behavior. Our results, pertaining to key differences between the intrinsic structural dynamics of the A30P variant and the WT form, are therefore relevant in describing the very early steps of the self assembly of αS.

3.5 Bibliography

(1) Iwai, A.; Masliah, E.; Yoshimoto, M.; Ge, N.; Flanagan, L.; Rohan de Silva, H. A.; Kittel, A.; Saitoh, T. The precursor protein of non-A*β* component of Alzheimer's disease amyloid is a presynaptic protein of the central nervous system. *Neuron*. **1995**, *14*, 467-475.

(2) Lotharius, J.; Brundin, P. Pathogenesis of parkinson's disease: dopamine, vesicles and α-synuclein. *Nat Rev Neurosci*. **2002**, *3*, 932-942.

(3) Norris, E. H.; Giasson, B. I.; Lee, V. M. Y.; Gerald, P. S. *α*Synuclein: Normal Function and Role in Neurodegenerative Diseases. In *Curr*. *Top*. *Dev*. *Biol.*; Academic Press, 2004; Vol. Volume 60; pp 17-54.

(4) Cooper, A. A.; Gitler, A. D.; Cashikar, A.; Haynes, C. M.; Hill, K. J.; Bhullar, B.; Liu, K.; Xu, K.; Strathearn, K. E.; Liu, F.; Cao, S.; Caldwell, K. A.; Caldwell, G. A.; Marsischky, G.; Kolodner, R. D.; LaBaer, J.; Rochet, J.-C.; Bonini, N. M.; Lindquist, S. *α*Synuclein Blocks ER-Golgi Traffic and Rab1 Rescues Neuron Loss in Parkinson's Models. *Science*. **2006**, *313*, 324-328.

(5) Cookson, M. R. The biochemistry in Parkinson's Disease. *Annual Review of Biochemistry*. **2005**, *74*, 29-52.

(6) McLean, P. J.; Kawamata, H.; Ribich, S.; Hyman, B. T. Membrane Association and Protein Conformation of *α*Synuclein in Intact Neurons. *J*. *Biol*. *Chem*. **2000**, *275*, 8812-8816.

(7) Uéda, K.; Fukushima, H.; Masliah, E.; Xia, Y.; Iwai, A.; Yoshimoto, M.; Otero, D. A.; Kondo, J.; Ihara, Y.; Saitoh, T. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc*. *Natl. Acad*. *Sci*. *U.S.A*. **1993**, *90*, 11282-11286.

(8) Uversky, V. N.; Li, J.; Fink, A. L. Evidence for a Partially Folded Intermediate in *α*Synuclein Fibril Formation. *J*. *Biol*. *Chem*. **2001**, *276*, 10737-10744.

(9) Davidson, W. S.; Jonas, A.; Clayton, D. F.; George, J. M. Stabilization of *α*Synuclein Secondary Structure upon Binding to Synthetic Membranes. *J*. *Biol*. *Chem*. **1998**, *273*, 9443-9449.

(10) Chandra, S.; Chen, X.; Rizo, J.; Jahn, R.; Südhof, T. C. A Broken α-Helix in Folded *α*Synuclein. *J. Biol. Chem.* **2003**, *278*, 15313-15318.

(11) Eliezer, D.; Kutluay, E.; Bussell Jr, R.; Browne, G. Conformational properties of αSynuclein in its free and lipid-associated states. *J. Mol. Biol.* **2001**, *307*, 1061-1073.

(12) Lee, H.-J.; Choi, C.; Lee, S.-J. Membrane-bound αSynuclein Has a High Aggregation Propensity and the Ability to Seed the Aggregation of the Cytosolic Form. *J. Biol. Chem.* **2002**, *277*, 671-678.

(13) Mihajlovic, M.; Lazaridis, T. Membrane-bound structure and energetics of α-synuclein. *Proteins: Struct. Funct. Bioinform*. **2008**, *70*, 761-778.

(14) Abedini, A.; Raleigh, D., P. . A role for helical intermediates in amyloid formation by natively unfolded polypeptides? *Phys. Biol*. **2009**, *6*, 015005.

(15) Dobson, C. M. Protein misfolding, evolution and disease. *Trends Biochem. Sci*. **1999**, *24*, 329-332.

(16) Hauser, C. A. E.; Deng, R.; Mishra, A.; Loo, Y.; Khoe, U.; Zhuang, F.; Cheong, D. W.; Accardo, A.; Sullivan, M. B.; Riekel, C.; Ying, J. Y.; Hauser, U. A. Natural tri- to hexapeptides self-assemble in water to amyloid *β*-type fiber aggregates by unexpected *α*-helical intermediate structures. *Proc. Natl. Sci. U.S.A.* **2011**, *108*, 1361-1366.

(17) Kelly, J. W. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* **1998**, *8*, 101-106.

(18) Polymeropoulos, M. H.; Lavedan, C.; Leroy, E.; Ide, S. E.; Dehejia, A.; Dutra, A.; Pike, B.; Root, H.; Rubenstein, J.; Boyer, R.; Stenroos, E. S.; Chandrasekharappa, S.; Athanassiadou, A.; Papapetropoulos, T.; Johnson, W. G.; Lazzarini, A. M.; Duvoisin, R. C.; Di Iorio, G.; Golbe, L. I.; Nussbaum, R. L. Mutation in the αSynuclein Gene Identified in Families with Parkinson's Disease. *Science*. **1997**, *276*, 2045-2047.

(19) Krüger, R.; Kuhn, W.; Müller, T.; Woitalla, D.; Graeber, M.; Kösel, S.; Przuntek, H.; Epplen, J. T.; Schöls, L.; Riess, O. Ala30Pro mutation in the gene encoding alpha-synuclein in Parkinson's disease. *Nat. gen*. **1998**, *18*, 106-108.

(20) Zarranz, J. J.; Alegre, J.; Gómez-Esteban, J. C.; Lezcano, E.; Ros, R.; Ampuero, I.; Vidal, L. d.; Hoenicka, J.; Rodriguez, O.; Atarés, B. a.; Llorens, V. n.; Gomez Tortosa, E.; del Ser, T.; Muňoz, D. G.; de Yebenes, J. G. The new mutation, E46K, of alpha-synuclein causes Parkinson and Lewy body dementia. *Ann. neurol*. **2004**, *55*, 164-173.

(21) Conway, K. A.; Harper, J. D.; Lansbury, P. T. Accelerated in vitro fibril formation by a mutant *α*synuclein linked to early-onset Parkinson disease. *Nat. Med*. **1998**, *4*, 1318-1320.

(22) Greenbaum, E. A.; Graves, C. L.; Mishizen-Eberz, A. J.; Lupoli, M. A.; Lynch, D. R.; Englander, S. W.; Axelsen, P. H.; Giasson, B. I. The E46K Mutation in *α*Synuclein Increases Amyloid Fibril Formation. *J. Biol. Chem.* **2005**, *280*, 7800-7807.

(23) Conway, K. A.; Lee, S.-J.; Rochet, J.-C.; Ding, T. T.; Williamson, R. E.; Lansbury, P. T. Acceleration of oligomerization, not fibrillization, is a shared property of both Î±-synuclein mutations linked to early-onset Parkinson's disease: Implications for pathogenesis and therapy. *Proc. Natl. Sci. U.S.A.* **2000**, *97*, 571-576.

(24) Bussell, R.; Eliezer, D. Effects of Parkinson's Disease-Linked Mutations on the Structure of Lipid-Associated *α*Synuclein. *Biochemistry*. **2004**, *43*, 4810-4818.

(25) Bussell, R.; Eliezer, D. Residual Structure and Dynamics in Parkinson's Disease-associated Mutants of *α*Synuclein. *J. Biol. Chem.* **2001**, *276*, 45996-46003.

(26) Barlow, D. J.; Thornton, J. M. Helix geometry in proteins. *J. Mol. Biol.* **1988**, *201*, 601-619.

(27) Ulmer, T. S.; Bax, A. Comparison of Structure and Dynamics of Micelle-bound Human *α*Synuclein and Parkinson Disease Variants. *J. Biol. Chem.* **2005**, *280*, 43179-43187.

(28) Jensen, P. H.; Nielsen, M. S.; Jakes, R.; Dotti, C. G.; Goedert, M. Binding of *α*Synuclein to Brain Vesicles Is Abolished by Familial Parkinson's Disease Mutation. *J. Biol. Chem.* **1998**, *273*, 26292-26294.

(29) Perlmutter, J. D.; Braun, A. R.; Sachs, J. N. Curvature Dynamics of α-Synuclein Familial Parkinson Disease Mutants. *J. Biol. Chem.* **2009**, *284*, 7177-7189.

(30) Jo, E.; Fuller, N.; Rand, R. P.; St George-Hyslop, P.; Fraser, P. E. Defective membrane interactions of familial Parkinson's disease mutant A30P *α*synuclein. *J. Mol. Biol.* **2002**, *315*, 799-807.

(31) Ulmer, T. S.; Bax, A.; Cole, N. B.; Nussbaum, R. L. Structure and Dynamics of Micelle-bound Human *α*Synuclein. *J. Biol. Chem.* **2005**, *280*, 9595-9603.

(32) Crowther, R. A.; Jakes, R.; Spillantini, M. G.; Goedert, M. Synthetic filaments assembled from C-terminally truncated *α*Synuclein. *FEBS Lett*. **1998**, *436*, 309-312.

(33) Hoyer, W.; Cherny, D.; Subramaniam, V.; Jovin, T. M. Impact of the Acidic C-Terminal Region Comprising Amino Acids 109−140 on *α*Synuclein Aggregation in Vitro. *Biochemistry*. **2004**, *43*, 16233-16242.

(34) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys*. **1983**, *79*, 926-935.

(35) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys*. **1999**, *151*, 283-312.

(36) Mackerell, A. D. Empirical force fields for biological macromolecules: Overview and issues. *J. Comput. Chem*. **2004**, *25*, 1584-1604.

(37) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys*. **1995**, *103*, 4613-4621.

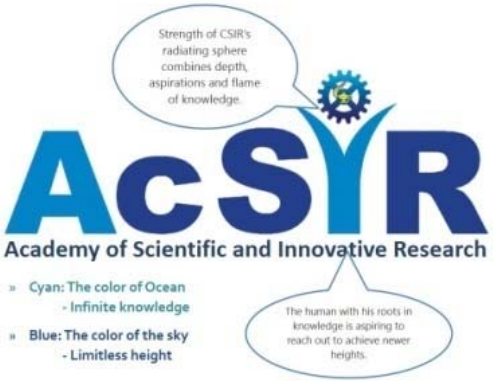
(38) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys*. **1977**, *23*, 327-341.

(39) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys*. **1995**, *103*, 8577-8593.

(40) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics*. **1996**, *14*, 33-38.

Chapter 4

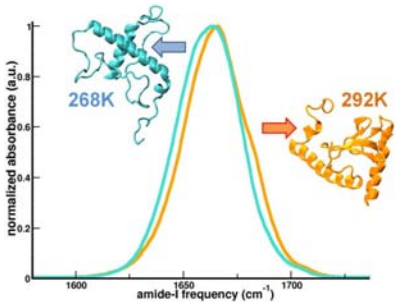
The non uniform early structural response of Yfh1 to cold denaturing conditions probed with atomistic MD simulations and calculated infrared absorption spectra



The universe is asymmetric and I am persuaded that life, as it is known to us, is a direct result of the asymmetry of the universe or of its indirect consequences.

- Louis Pasteur

Abstract



The mechanism of cold denaturation in proteins is often incompletely understood due to limitations in accessing the denatured states at extremely low temperatures. Using atomistic molecular dynamics simulations, we have compared early (nanosecond timescale) structural and solvation properties of yeast frataxin (Yfh1) at its temperature of maximum stability, 292 K (T_s), and the experimentally observed temperature of complete unfolding, 268 K (T_c). Within the simulated timescales, discernible 'global' level structural loss at T_c is correlated with a distinct increase in surface hydration. However, the hydration and the unfolding events do not occur uniformly over the entire protein surface, but are sensitive to local structural propensity and hydrophobicity. Calculated infrared absorption spectra in the amide-I region of the whole protein show a distinct red shift at T_c in comparison to T_s. Domain specific calculations of IR spectra indicate that the red shift primarily arises from the beta strands. This is commensurate with a marked increase in solvent accessible surface area per residue for the beta-sheets at T_c. Detailed analyses of structure and dynamics of hydration water around the hydrophobic residues of the beta-sheets show a more bulk water like behavior at T_c due to preferential disruption of the hydrophobic effects around these domains. Our results indicate that in this protein, the surface exposed beta-sheet domains are more susceptible to cold denaturing conditions, in qualitative agreement with solution NMR experimental results.

4.1 Introduction

The attainment of the uniquely folded, functional form of a globular protein at physiological conditions results from a complex interplay of numerous short and long-range interactions involving the amino acid sequence and the surrounding solvent.[1] As the structure of a nascent polypeptide evolves in a complex energy landscape, the loss in configurational entropy is countered by increasingly favorable interactions between particular regions of its sequence. Alterations to the thermo-chemical environment of the folded protein can disrupt the delicate balance of forces maintaining the stability of the folded form, and may result in partial or complete unfolding or misfolding events.[2,3] In recent times, a large number of experimental techniques as well as several theoretical and computational methods have been employed, sometimes in conjunction with each other, to delineate the factors responsible for the stability of the folded protein at physiological conditions.[1,4-7]

Globular proteins have been known to lose their natively folded form at temperatures that are either markedly higher or lower than physiological temperatures.[8] While denaturation brought about by either condition shares some common aspects such as loss in native contacts and increased thermal heat capacity, it is widely accepted that there are key differences underlying the thermodynamics and kinetics of 'heat' and 'cold' denaturation.[9-13] Unlike heat denaturation, cold denaturation has been described as an irreversible process.[14] It is recognized that a clearer understanding of the origins of cold denaturation process in proteins could lead to an improved understanding of the balance of forces that lead to protein folding and stability.

Under normal physiological conditions, the hydrophobic effect is a key factor underlying structural compaction incurred by a nascent, random polypeptide on the way to acquiring the folded native form.[15-17] It has been suggested that the origins of cold denaturation lie in the disruption of the hydrophobic effect at sufficiently low temperatures,[8,18,19] when the entropic costs of solvating the apolar protein core reduces sharply.[17] In apomyoglobin, a globular protein, it has been shown that a temperature drop from 310 K to 278 K results in an increase in the number of solvent contacts and isothermal compressibility, particularly for apolar regions.[19] It has been suggested that the thermodynamic advantage in terms of translational entropy gain of water molecules associated with the large decrease in solvent accessible surface area during folding is lost at lower temperatures.[20,21] Further, although cold denaturation has been largely considered to be non-cooperative,[22,23] recent NMR experiments showing similar structural loss at

distal sites[24] suggest that there may be exceptions. We also mention here that cold protein denaturation is analogous to polymer unfolding resulting from hydrophobic hydration as in both the scenarios, unfolding occurs by heating or cooling as well as by applying pressure like the globular proteins.[25]

The cold denaturation process in proteins is relatively less well understood compared to heat denaturation, in part due to experimental difficulties in accessing the denatured states below the freezing point of the aqueous solvent.[26] Indirect experiments that raise the cold denaturation temperature via cosolutes, pH changes, site mutations, or use supercooled water to prevent solvent freezing, have been used extensively to study this apparently anomalous process, although concerns regarding alterations to the inherent energy landscape arising from these modifications persist.[24,27-33] Recently, however, it was discovered that the onset of cold denaturation occurs at temperatures above the freezing point of water in yeast frataxin (Yfh1), the yeast orthologue of the human frataxin protein,[34] in the C-terminal domain of the HHR23A protein,[35] in the I98A variant of CTL9,[36] etc. These discoveries have provided systems where the cold denatured states may be accessed naturally, without biasing the solvent environment or the protein structure. This discovery therefore offers novel paradigms for understanding cold denaturation in proteins, and may consequently lead to substantial improvement in the current understanding of protein folding and stability.

Since the cold denaturation effect is accessible for Yfh1 above the melting point of water, there have been several experimental as well as theoretical approaches to address this particular problem.[24,30,34,37-40] It was found that at the temperatures at which heat and cold denaturation are initiated (304 and 280 K respectively),[34] the two denatured states exhibit similar conformational characteristics.[38] However, residual structural characteristics, compactness and the extent of hydration of these states diverge as the temperatures of complete unfolding are approached.[14] In heat and cold denatured Yfh1, the beta-sheet domains has a greater propensity to unfold when compared with the helical domains, and the N-terminal helical domain is more resistant to unfolding than the C-terminal helical domain. However, it was demonstrated that the protein-water hydrogen bond lifetimes increase sharply at cold denaturing conditions. Very recent results have indicated that influx of solvent molecules into the hydrophobic core of Yfh1 arising due to electrostatic repulsion between the similarly charged helices could be a plausible reason for its cold denaturation.[39]

In this study, using MD simulations, we have attempted to correlate the nature of protein surface hydration with the earliest structural response of Yfh1 to cold denaturing conditions. While fully atomistic, sub-microsecond simulations are insufficient to unravel the complete unfolding pathway; they demonstrate the non-uniform structural response of the different domains of Yfh1 to the thermal conditions. For a clearer picture of the solvation effects, we have used the simulated trajectories to generate infrared (IR) absorption spectra corresponding to the backbone amide-I vibrations at both the temperature of maximum stability (292 K, or T_s) and the experimentally reported temperature of complete unfolding (268 K, or T_c),[34] based on the frequency maps by Cho et al.[41] It is observed at the lower temperature that the beta sheet domains are susceptible to spectral red shifts. Further, based on our results of the structure, dynamics and density fluctuations of the hydration layer water molecules of the different secondary structural domains of Yfh1, we see that the water molecules pertaining to the hydration layer of the beta sheets are more bulk-like in nature, resulting from the structural loss in the hydrophobic residues giving rise to an enhancement in solvation. These results indicate that the red shift in the IR spectra might be dominated by the hydration of the hydrophobic residues residing in the beta sheet domains. Our results thus demonstrate that cold denaturation first originates with locally enhanced hydrophobic solvation, in agreement with the currently accepted views. Interestingly, a protein with a much lower cold denaturation temperature, Ubiquitin, does not undergo any significant unfolding at T_c. Further, the early cold denaturation behavior of Yfh1 is fundamentally different from the early heat denaturation behavior observed at 324 K (or T_h), which is close to the experimentally reported temperature (323 K) of complete heat denaturation.[14,34] The early heat denaturation is not confined to the beta-sheet domains, and is accompanied by a strengthening of the hydrophobic effect on the protein surface.

4.2 Materials and methods

4.2.1 System setup and MD simulations

Initial coordinates for the 123-residue yeast frataxin (Yfh1) protein molecule were obtained from the PDB database (PDB entry 2GA5), reported *via* solution NMR studies conducted at 298 K and 1 atm pressure.[42] The structure, comprised of two helical domains (residues 19 to 43, or H1, and residues 107 to 121, or H2), seven beta strands (residues 50-54, or S1; 57-62, or S2; 66-71, or S3; 78-83, or S4; 87-94, or S5; 97-100, or S6; 105-106, or S7) and inter-connecting, loops is depicted in Figure 4.1 a. The secondary structure was elucidated via the STRIDE algorithm[43] as

implemented in the Visual Molecular Dynamics (VMD).[44] We note here that the temperature of maximum stability (292 K) is lower than the temperature at which this structure was reported. Simulations were performed with the CHARMM22 force field with the CMAP corrections[45,46] and the NAMD 2.9 simulation package.[47] The protein was solvated with the TIP3P[48] water model, such that the edges of the resulting simulation box were 15 Å from the extreme ends of the protein. 15 sodium counter ions were used to neutralize the system, after which energy minimization was first performed for 15000 steps using conjugate gradient method. Orthorhombic periodic boundary conditions were used, and long range electrostatics was calculated with particle mesh Ewald (PME).[49] A cutoff of 12 Å was used for non-bonded interactions with smoothing starting at 11 Å. Simulations were generated at the isothermal-isobaric (NPT) ensemble. Langevin dynamics with a collision frequency of 1 ps^{-1} was used to maintain constant temperature, and the Nose´–Hoover Langevin piston[50] maintained constant pressure. The SHAKE algorithm[51] was used to constrain the lengths of bonds involving hydrogen atoms. A time step of 2 fs was used, and independent simulation trajectories, each exceeding at least 100 ns, were generated at 292 K (T_s), 268 K (T_c) and 324 K (T_h), saving snapshots every 2 ps. To validate and compare the force field effects at T_c, a thermally stable globular protein ubiquitin (Ubq) (PDB entry 1UBQ)[52] has been simulated at T_s and T_c with the same protocols. For orientational auto-correlation analyses, snapshots were saved every 100 fs. For the calculation of the IR absorption spectra, selected snapshots were simulated with similar protocol but using a 1 fs time step and the trajectories saved every 6 fs. VMD was used for visualization.[44]

4.2.2 Analyses protocol(s)

Structural persistence

In earlier studies,[53,54] we have quantified the extent of structural destabilization in proteins relative to a reference structure with the structural persistence, *P*, previously defined in *Chapter 2*, *3*. In this study, we have compared the overall as well as residue-wise extents of structural change at T_c and T_h over T_s. Thus, the mean (ϕ, ψ) values for every residue obtained over last 50 ns at T_s were used as the reference values. The persistence was calculated for the whole protein (P_{Full}), beta domains (P_{Beta}), and helical domains (P_{Helix}) as well as for individual residues (P_{res}). P_{res} is calculated in a similar manner as P_{Full}, but the summation and averaging over time is done for a particular residue instead of the full protein.

Heat capacity

We have calculated the heat capacity at constant volume, C_v as,

$$C_v = \frac{\langle \Delta E \rangle^2}{RT^2} \tag{4.1}$$

Here $\langle \Delta E \rangle^2$ is the standard deviation of the internal energy of the selected system, R is the universal gas constant and T is the absolute temperature.

Rotational relaxation

We have computed the estimates of rotational relaxation times of bulk and surface water (within 4 Å) from the mentioned simulations via time correlation of the second Legendre polynomial (the second-rank spin correlation function),

$$P_2(t) = \left\langle \frac{3\cos^2\theta(t) - 1}{2} \right\rangle \tag{4.2}$$

$\theta(t)$ being the rotational angle of a molecule fixed unit vector at time t, which for our calculation happens to be the dipole moment unit vector, giving almost identical results when OH vector unit is considered for the calculation. $P_2(t)$ does not have a simple exponential decay, but can be fitted to a stretched exponential using the Kohlrausch-Williams-Watts (KWW) function,[55,56]

$$P_2(t) \equiv e^{-(t/\tau)^{\beta}} \tag{4.3}$$

Here, the τ and β ($0<\beta<1$) denote the orientational relaxation time and the stretched exponential parameter, respectively. An increase in τ indicates a slowdown in the relaxation rate, and a lowering of β is indicative of increased heterogeneity in the dynamical behavior.

Water density fluctuation

We have calculated the density fluctuation of water[57] considering a very thin probe volume (a product of solvent accessible surface area (*SASA*) and 3Å) from the surface of the selected domains. For obtaining the water density fluctuation, we obtained $P(n^*)$ which is the probability

distribution of the water number density divided by the mean density (n^*), then normalized by the maximum probability. The distribution is then fitted to the Gaussian form,

$$P(n^*) = e^{-\frac{(n^*-1)^2}{\alpha}} \tag{4.4}$$

Tetrahedral order parameter

We have calculated the tetrahedral order parameter (Q)[58,59] of water molecules (considering only the water oxygen atoms as potential hydrogen bonding sites) within the hydration shell of the protein as,

$$Q = 1 - \frac{3}{8}\sum_{j=1}^{3}\sum_{k=j+1}^{4}\left(\cos\theta_{j,k} + \frac{1}{3}\right)^2 \tag{4.5}$$

$\theta_{j,k}$ represents the angle formed between the O-O vectors i,j and i,k from the oxygen atom (i) with its nearest neighbors (j and k; $\leq$ 4). The order parameter value of 1 resembles a perfect tetrahedral structure of water resembling its ice-state. The average values of Q range from 0 to 1.Water molecules lying between the first and the second solvation shell of the protein surface being maximally ordered,[60] we have considered the water molecules which lie within 2.5 and 4 Å of the protein heavy atoms (see Figure 4.3 b).

Cumulative configurational entropy

We have calculated the configurational entropy for the backbone heavy atoms of the protein Yfh1 as well as of its specific domains of interest for the whole trajectory at temperatures T_c, T_s and T_h by Schlitter's method.[61] The initial frame in each simulated trajectory has been used as the respective reference, to remove the translational as well as the rotational motions with respect to the center of mass of the systems. From the Schlitter's method, the absolute entropy can be approximately obtained as

$$S_{abs} < S = \frac{1}{2}k_b \ln\det\left[1 + \frac{k_b T e^2}{\hbar^2} M^{\frac{1}{2}}\sigma M^{\frac{1}{2}}\right] \tag{4.6}$$

where k_b being the Boltzmann's constant, $\hbar$ the Planck's constant divided by 2, e the Euler's number, M the the mass matrix of 3N dimension containing N atomic masses of the system and σ is the covariance matrix. The elements in the covariance matrix can be denoted as

$$\sigma_{ij} = \left\langle \left(x_i - \left\langle x_i \right\rangle\right)\left(x_j - \left\langle x_j \right\rangle\right) \right\rangle \tag{4.7}$$

where, x_i and x_j are the Cartesian coordinates of the selected atoms.

Calculation of IR absorption spectra

For the calculation of the IR absorption spectra of the amide-I modes, four representative snapshots of the simulations were selected from the maxima of the N_w vs. P_{Full} plots (Figures 4.4 a and b). These representative snapshots were equilibrated for 1 ns (using a 2 fs time step), followed by a high frequency production run for 204 ps (1 fs time step) for both T_s and T_c. Configuration data were saved at every 6 fs during each high frequency simulation to use them for the calculation of the IR spectra.

For the frequency calculation of each amide-I mode, the CHO4 parameterization[41] was used which is based on the expansion of the vibrational frequency as a linear combination of the electrostatic potential computed at the coordinates of the four atoms C,O, N and H of each amide bond. For a particular peptide bond, the electrostatic potential calculated at the coordinate of a particular atom contained contribution from all the other atoms including side-chains, backbone and water (counter ions were excluded).[62] A model that has been used previously to model the absorption of amyloid fibrils[62] defined the backbone atoms that contributed to the total electrostatic potential. Multiplying the calculated potentials at C, O, N, and H of a particular chromophore with the corresponding expansion coefficients,[41] along with a suitable conversion factor, the shifts in frequencies in the units of cm^{-1} were obtained and added to the base frequency of 1707.7 cm^{-1}.[41] For obtaining the off-diagonal elements in the system Hamiltonian, the coupling between the nearest neighbors was determined from the nearest neighbor coupling (NNC) map.[63] The couplings between all other peptide units were approximated using transition dipole coupling (TDC).[64] We would like to mention that as CHO4 parameterization excludes the chromophores pertaining to the imides,[41] the chromophores succeeded by Proline have been excluded in our present calculation. The IR absorption spectra were calculated as Fourier

transforms of the corresponding linear response function,[65] where the mean frequency of any amide-I mode is the ensemble averaged normal mode transition frequency of the amide-I mode. Using the second order cumulant expansion, the linear response function can be written in terms of frequency-frequency correlation functions (FFCF).[65] FFCF of each chromophore was calculated using a 9.6ps window over the whole trajectories. The cross correlation between two different amide-I modes was assumed to be negligibly small as shown before.[65-67] For the calculation of the IR spectrum of separate beta domains, we have considered the block diagonal amide-I frequencies as the coupling terms for these residues were negligible. As mentioned before, FFCF were separately calculated for different beta sheet and alpha helical domains to obtain the domain specific IR absorption spectra.

4.3 Results and discussions

Structural effects at T_c

We begin by comparing the overall structural characteristics of the conformations of Yfh1 obtained at the cold denaturing conditions (T_c) with those obtained at the temperature of maximum stability (T_s). In Figures 4.1 a and b, we have presented representative snapshots of the protein conformations obtained at the end of the simulations carried out at T_c and T_s. Preliminary inspection of the structures indicates a marked extent of structural loss at T_c, particularly in the beta sheet domains. In Figures 4.1 c and d, we present distributions of the protein's radius of gyration (R_g), and the protein's total solvent accessible surface area (*SASA*, calculated with the VMD tool using a probe volume of 1.8 Å radius)[68] collected over the last 50 ns of the simulated trajectories. At T_s and T_c, the mean values of R_g are 16.11 (±0.47) Å and 18.13 (±0.54) Å, respectively, while the mean values of *SASA* are 11115.6 (±420.8) Å^2 and 12861.4 (±376.3) Å^2, respectively. Thus, R_g and *SASA* increase by about 12.4% and 15.7% respectively at T_c relative to those at T_s. The increase in R_g indicates that the protein responds to the sharp lowering of temperature with a significant loss in compactness within nanosecond timescales. Further, the loss in compactness is also reflected by a distinct increase in the exposed protein surface and its solvent accessibility. In Figures 4.1 e and f, we have presented the sidechain and backbone inter-residue contact maps from the trajectories at T_c and T_s. The decrease in structural compactness at T_c is reflected in significant losses in the probability of inter-residue contacts. Particularly, the nearest neighbor contacts that characterize the presence of the beta sheet strands (strand 2 with strand 3; strand 4 with strand 5) at T_s are completely lost at T_c. The contacts largely retained at T_c

are those that constitute the inter-residue contacts due to the helical domains. We would like to mention here that T_s is slightly lower (6 K) than the temperature of the solution NMR experiments (298 K), and hence some structural rearrangement occurs at this temperature relative to the original PDB structure.

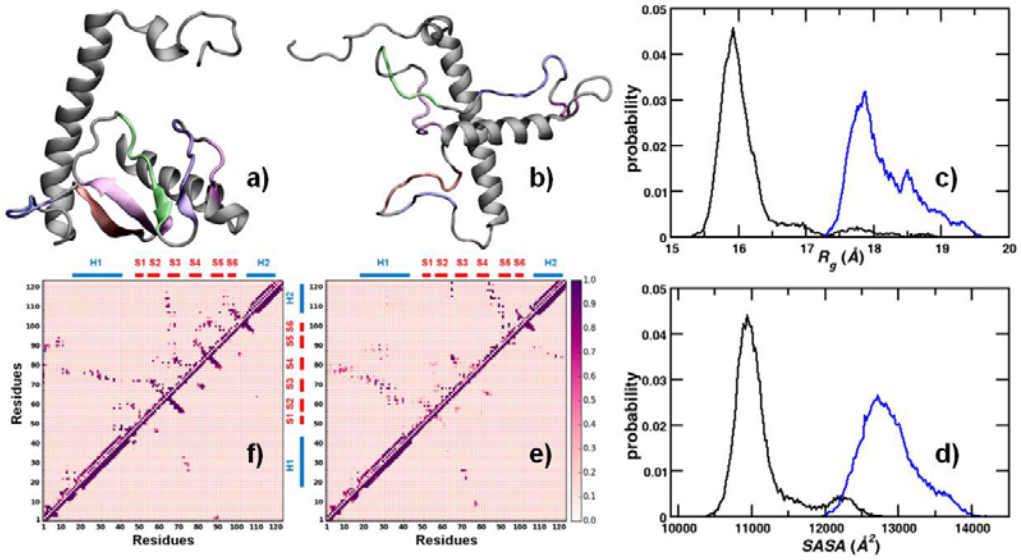


Figure 4.1 Representative snapshots, distributions of R_g, distributions of *SASA* and contact maps. For the snapshots, helices (H1, H2) and the turn & coils are shown in gray; beta strands (S1, S2, S3, S4, S5 &S6) are shown in blue, pink, purple, green, violet and magenta respectively. For the contact maps, lower triangle denotes backbone contacts while upper triangle denotes sidechain contacts. Representative snapshots of the protein at a) T_s and b) at T_c;c) distributions of R_g at T_s *(black)* and T_c *(blue)*, d) distributions of SASA at T_s *(black)* and T_c *(blue)*, e) contact map at T_c and f) contact map at T_s.

For greater insights into the structural effects brought about by the cold denaturing conditions, we calculated the full as well as residue-wise values of the structural persistence and the heat capacity. The mean values of P_{Full} shifted from 0.88 at T_s to 0.72 at T_c, while the mean value of C_v shifted from 17.34 kJ K^{-1} mol^{-1} at T_s to 24.35 kJ K^{-1} mol^{-1} at T_c. The lowering of P_{Full} and increase in C_v at T_c show that the loss in compactness at this temperature is accompanied by an increase in the structural disorder, and thus reflects an onset of structural unfolding. In Figure 4.2, we have plotted the mean residue-wise values of ΔP_{res} and ΔC_v, the difference in the values of P_{res} and C_v from T_s to T_c. Positive values of ΔP_{res} and negative values of ΔC_v reflect higher local disorder at T_c over T_s. It is interesting to note that most residues belonging to the beta strands display large shifts in the values of P_{res} and C_v, reflecting the loss in secondary structure

in these regions at T_c as evidenced in Figures 4.2 a and b. In comparison, the magnitude of ΔP_{res} and ΔC_v are negligible in most residues belonging to the helical domains, reflecting the structural stability of the helical domains at T_c within the simulated timescales.

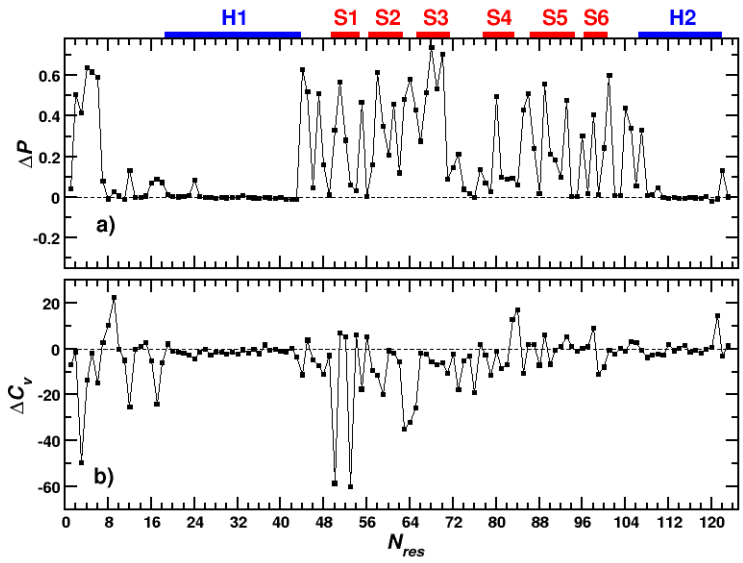


Figure 4.2 a) Residue-wise ΔP (or $P_{Ts} - P_{Tc}$) plots; and b) residue-wise ΔC_v (or $C_{vTs} - C_{vTc}$) plots as a function of the residue numbers (N_{res}). Units of C_v are in $JK^{-1}\ Mol^{-1}$.

Hydration at T_c

As mentioned earlier, insights into the protein cold denaturation process can be obtained *via* studies of polymer unfolding at low temperatures.[25] Coarse-grained simulations of homopolymers show an increase in the first solvation peak of the water radial distribution function, or $g(r)$, relative to the polymer surface. Further, specifically in the case of Yfh1, the protein-water hydrogen bonds have a longer lifetime at the cold denaturing condition.[14] These results suggest that the association of water with the protein surface strengthens at cold denaturing conditions. In Figures 4.3 a and b, we compare water-protein $g(r)$ obtained over the last 50 ns of the simulation trajectories at T_c and T_s. The $g(r)$ has been calculated for the backbone carbonyl carbon and water oxygen (Figure 4.3 a) and for all heavy atoms and water oxygen (Figure 4.3 b); the peak positions of the first solvation layer occur at 3.8Å and 2.8Å respectively. The peak heights for the *g(r)* with the backbone carbonyl atoms occur at 0.507 and 0.622 at T_s and T_c, respectively, showing an increase of ~23% at T_c over T_s. The height of the first *g(r)* peak obtained for the heavy atoms increases by about 14% from T_s to T_c. Thus, the

radial distribution functions describe a distinguishable enhancement in the protein's surface hydration due to protein cold denaturation conditions within the sub-microsecond timescales.

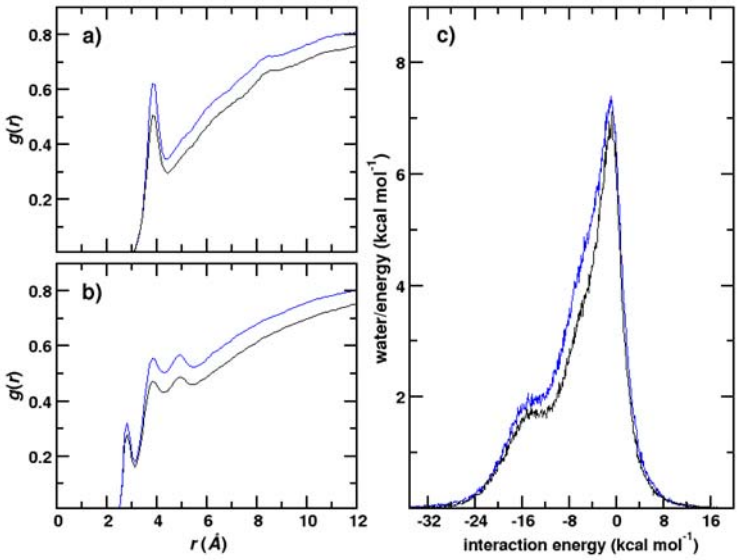


Figure 4.3 Radial distribution functions ($g(r)$) and interaction energy of the whole protein with its hydration layer. a) Plots of $g(r)$ between backbone carbonyl carbon atoms of protein and water oxygen atoms, b) plots of $g(r)$ between all heavy atoms and water oxygen atoms, c) distributions of interaction energy of the whole protein with its hydration layer; at T_s *(black)* and T_c *(blue)*.

Earlier reports have shown that the average interaction strength of the protein with the water molecules in its first hydration layer decreases when the native protein partially denatures to the molten-globule like state.[69] However, protein cold denaturation is mechanistically distinct from thermal and chemically induced denaturation that lead to molten globule forms of proteins. Further, as suggested by several recent findings and corroborated by the protein-water radial distribution functions of Yfh1, cold denaturation is accompanied by increased favorability of the protein-hydration layer interactions. For greater insight into the energetics of the protein-water interaction, we present in Figure 4.3 c, distributions of the interactions of each water molecule within the hydration layer of the protein surface, with the protein surface. For our calculations, the hydration layer includes those water molecules that lie within 4 Å of the heavy atoms of the protein. The data have been normalized such that the average hydration number equals the integrals of the distributions. The distribution is broadened at T_c relative to T_s, and the mean protein-hydration layer interaction strength shifts from –5.7 kcal mol^{-1} at T_s to –5.9 kcal mol^{-1} at T_c; showing that the interaction becomes energetically more favorable at the lower temperature.

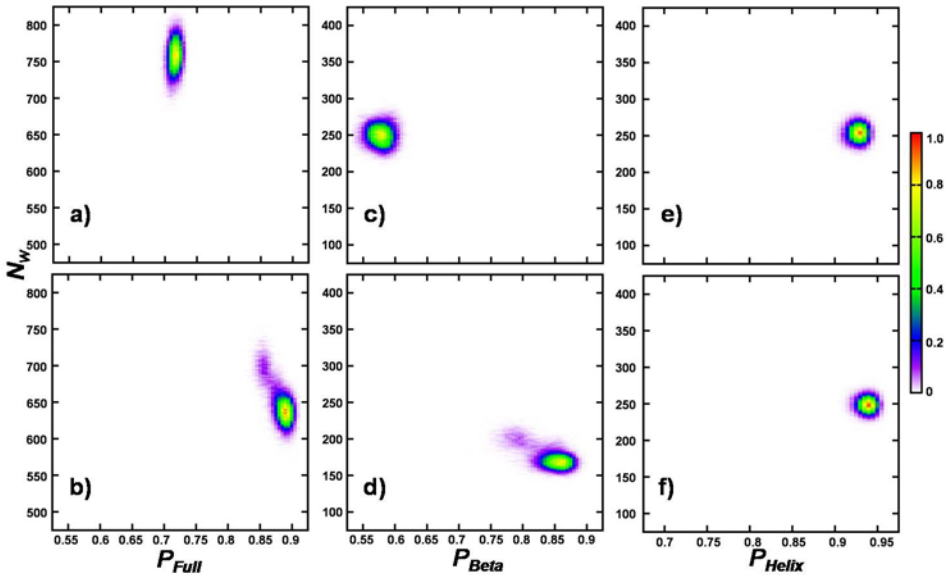


Figure 4.4 2D-contour plots of P versus N_w for; a) whole protein at T_c, b) whole protein at T_s, c) beta domains at T_c, d) beta domains at T_s, e) helical domains at T_c, f) helical domains at T_s.

In Figures 4.4 a, b we have compared distributions of the hydration layer number, N_w, and the overall structural persistence, P_{Full}, obtained at each snapshot over the last 50 ns simulation at T_s and T_c. The shift in the mean position from (647, 0.88) at T_s to (760, 0.72) at T_c shows that the increase in hydration is directly correlated with the overall structural loss and unfolding of the protein. When similar distributions are obtained for the beta strands (Figures 4.4 c, d) and the helical domains (Figures 4.4 e, f), we find that the shifts in (N_w, P_{Full}) occur predominantly due to structural loss and enhanced hydration of the beta strands. Within the timescales of interest, for the helical domains, the structures, and commensurately, the level of hydration at T_c, are largely similar to that at T_s. The mean positions of (N_w, P_{Beta}) are (175, 0.85) and (252, 0.58) at T_s and T_c, respectively, while the mean positions of (N_w, P_{Helix}) are (251, 0.94) and (256, 0.93) at T_s and T_c, respectively.

IR absorption spectra

In the recent years, IR spectra, in conjunction with computer simulations, has proven to be a powerful technique for obtaining valuable insights into structural dynamics of proteins and peptides and how such properties are influenced by solvent and thermodynamic conditions.[62,64,65,70-74] We have calculated representative IR absorption spectra of the amide-I

modes for the whole protein **(**Figure 4.5 a**)**, as well as beta **(**Figure 4.5 b, c, d**)** and helical (Figure 4.5 e) domains, as described in *Methods*. The high sensitivity of the backbone carbonyl to changes in hydration at T_c (Figure 4.3 a) suggests that the amide-I mode is a suitable IR probe for studying the effects of cold denaturation. The calculated absorption spectrum for the whole protein at T_c shows a distinct red shift when compared with the calculated spectrum at T_s, although significant differences are not observed in the peak shape. The peak maximum at T_c has a frequency of about 6 cm^{-1} lower than that at the higher temperature. It is well established that hydrogen bonding to carbonyls (C=O) produces a red-shift in the amide-I vibrational frequency.[75] The observed red shift in the amide-I spectrum at T_c indicates more extensive hydrogen bonding of the solvent water molecules to the C=O groups of the protein backbone at cold denaturation conditions. As observed earlier, a loss of secondary structure upon unfolding at T_c increases the solvent accessibility of the protein. A higher *SASA* at T_c allows more solvent water) molecules to form hydrogen bonds with the peptide units, manifested in the overall red shift of the calculated IR spectrum.

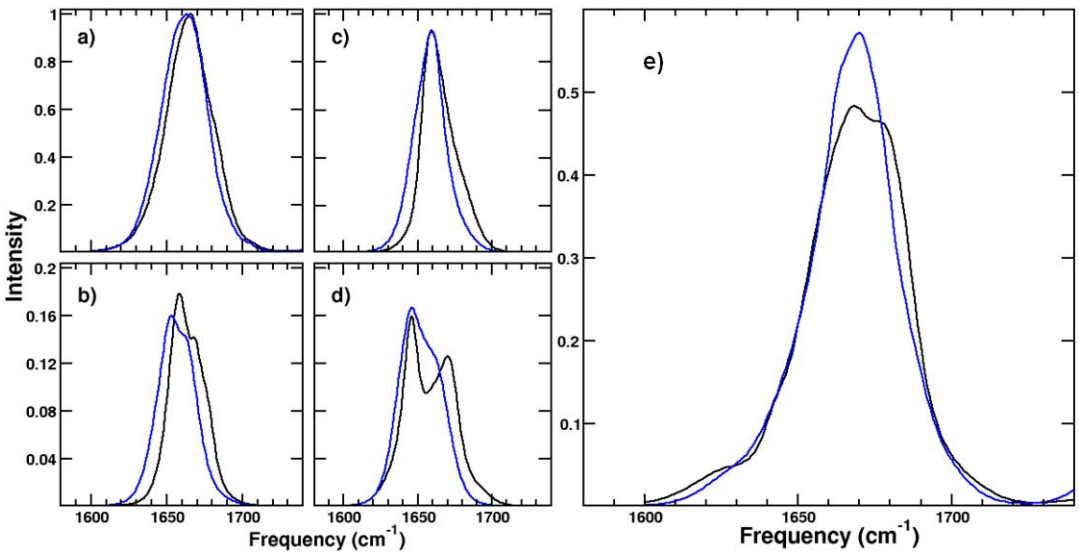


Figure 4.5 IR absorption spectra for the a) whole protein, b) S2, c) S4, d) S5, and e) representative helix H1; at T_s *(black)* and T_c *(blue)*.

As pointed out earlier, the hydration effects at T_c are not uniform over the entire protein within the timescale relevant for this study; the beta sheet domains that are more prone to structural unfolding at cold denaturing conditions are more hydrated. This observation

corroborates previous results.[14,39] In order to understand the local hydration effects upon loss of structure, we have calculated the IR absorption spectra for representative beta (S2, S4 and S5) and helical (H1) domains. A detailed inspection using IR spectra of these individual beta-sheet domains show similar trends of spectral red shifts as observed for the full protein; whereas a red shift is not observed in the calculated IR spectra of the representative helical domain of the protein upon lowering the temperature. The IR spectra calculated for the individual domains indicate that the overall red shift in the full protein amide-I IR spectrum on the nanosecond timescale primarily arises from the beta sheet regions due to increased hydration at the protein backbone in these domains. Other factors that might cause spectral red-shifts of the amide-I modes upon lowering of temperature are increase in hydrogen bond strength[76,77] and temperature dependence of the solvent dielectric properties.[78,79] Figure 4.4 suggests greater solvent accessibility for the helical domains than that of the beta sheet regions at T_s. If increasing hydrogen bond strength at T_c was the dominant factor resulting in the spectral red-shifts, a larger red-shift would be expected for the helical domains. When compared with the infrared absorption spectra at T_s, our calculations indicate that the beta sheet amide-I modes are red-shifted upon lowering of temperature (T_c) while that of the helical domains remain mostly unaltered. This suggests that the increased hydration and the local structural loss in the beta strands are manifested in the observed red-shift of the calculated spectra. For the helical domain, which largely retains the structure at T_c, solvent accessibility of the protein backbone does not significantly produce alterations in the calculated IR spectra.

Microscopic effects of hydrophobic hydration

Our results thus far demonstrate the greater susceptibility of the beta-sheet domains of Yfh1 to unfold on the nanosecond timescales commensurate with enhanced local hydration in those domains. Herein, we attempt to probe the extent of local structural and dynamical perturbation to the hydration layer water molecules in the vicinity of the unfolded domains. We first point out that the introduction of cosolutes into bulk water perturbs the hydrogen bonding network in a size dependent manner.[17] While smaller hydrophobic solutes do not rupture the overall hydrogen bonding network of bulk water, introduction of solutes of larger sizes causes significant disruption in the water hydrogen bonding network resulting in a phase separation with water. Recent research shows that the resulting entropic penalty is countered by the strengthening of the

water hydrogen bonds in the vicinity of the solute.[17,80] Thus, compared to bulk water, the hydration layer of non-polar moieties display slower rotational relaxation and enhanced tetrahedrality of water in the first hydration shell.[81,82]

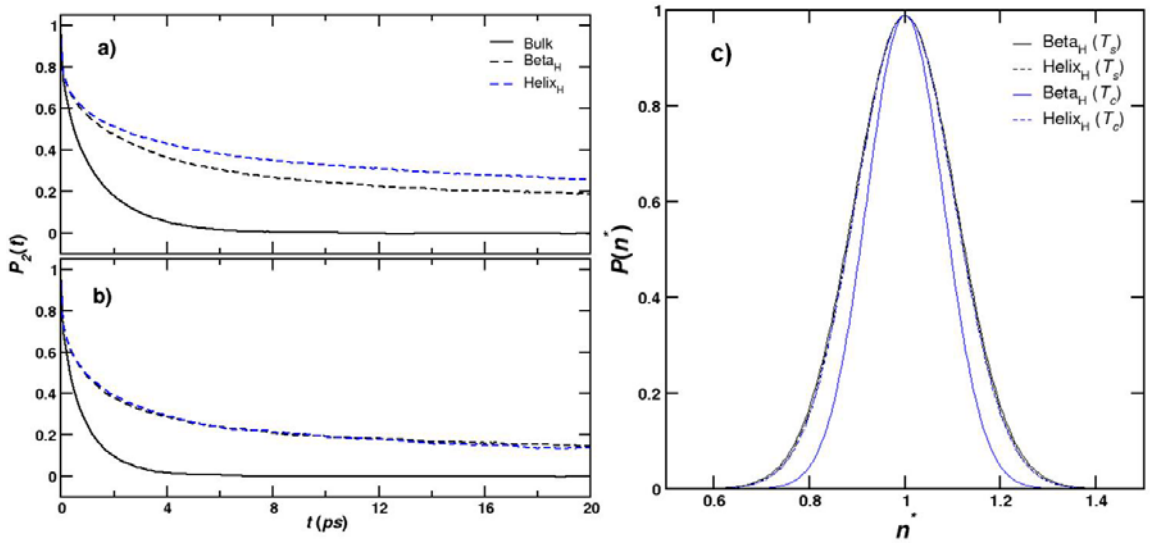


Figure 4.6 Second-rank rotational correlation function ($P_2(t)$) for dipoles of bulk water and hydration water layer of selected domains at a) T_c and b) T_s; c) probability distributions of water densities near the surfaces of sidechains of the selected domains at T_s *(black)* and T_c *(blue)*.

To evaluate the changes in the hydration water dynamics due to the temperature change, we have compared the orientational correlation functions $P_2(t)$ of the hydration water molecules in the hydrophobic residues[83] belonging to the helical ('Helix$_H$') and beta-sheet ('Beta$_H$') domains as shown in Figure 4.6 (a, b). In agreement with previous reports,[14] the decays are found to slow down at T_c compared to T_s, and the water present in the hydration layer show slower relaxation than bulk water at the same temperature.[55,69,84-87] Interestingly, however, while $P_2(t)$ for water molecules around Beta$_H$ are comparable with those of Helix$_H$ at T_s, they decay much faster than the latter at T_c. In other words, the orientational dynamics of the water molecules hydrating the Beta$_H$ show a relative shift towards bulk-like behavior at T_c. In Table 4.1, we report the values of the parameters τ and β from fits of $P_2(t)$ to Eqn. 4.3. The τ, β for Helix$_H$ and Beta$_H$ are comparable at T_s. At T_c, τ increases 3.5 times for Helix$_H$ in comparison to the corresponding value at T_s; however, the increase is only 1.9 times for Beta$_H$. We point out here that lower values of β are associated with increased dynamical heterogeneity of hydration water molecules.[69] While the β values observed for the hydration layer water molecules are comparable, we note some subtle differences. The β values corresponding to the hydration layers

of $Helix_H$ and $Beta_H$ are nearly identical at T_s. However, at T_c, the value of β is discernibly higher for $Beta_H$ suggesting a subtle shift towards bulk like character of the water molecules hydrating these regions.

Table 4.1 Dynamical data for bulk and hydration water molecules around the selected domains at T_s and T_c. τ (orientational time) and β (extent of heterogeneity) are obtained from the stretched exponential fit of the dipolar second-rank relaxation; α values are obtained from fits of the water density distribution $P(n^*)$ to Eq. 4.4 for the sidechains of selected domains.

Domains	τ		β		α	
	T_c	T_s	T_c	T_s	T_c	T_s
Bulk	0.96	0.66	0.74	0.74	-	-
$Helix_H$	8.51	2.44	0.28	0.31	2.20	2.20
$Beta_H$	4.42	2.34	0.32	0.3	1.3	2.3

The IR spectra and the observed dynamical behavior of water molecules hydrating $Beta_H$ suggest a decrease in hydrophobicity for these domains at T_c. Recently, it has been shown that increase in width of the Gaussian distribution of water-density within a small probe volume in close proximity to a surface is proportional to increase in surface hydrophobicity.[57] We have compared the water density fluctuation $P(n^*)$ within a probe volume as described in *Methods*, for $Beta_H$ and $Helix_H$ at T_s and T_c (Figure 4.6 c). The corresponding α values (see Eqn. 4.4) are tabulated in Table 4.1. While $Beta_H$ at T_s, and $Helix_H$ at both the temperatures have comparable distributions, $Beta_H$ at T_c clearly have a narrower distribution than the others, indicating distinctly reduced hydrophobicity at this temperature for these domains.

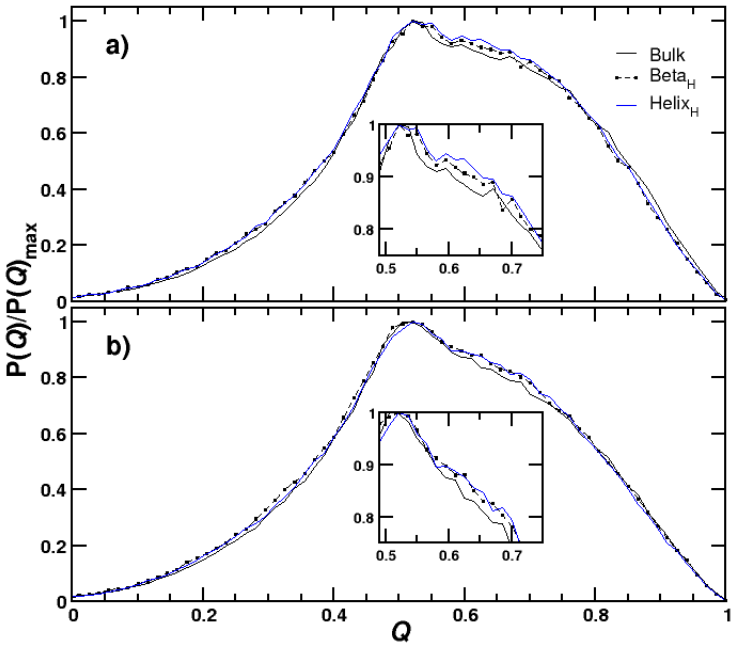


Figure 4.7 Distributions of the tetrahedral order parameter (Q) for bulk water, hydrophobic beta sheet domain and hydrophobic alpha helical domain at a) T_s and b) T_c. Insets show the expanded views for the range Q = 0.5 – 0.8. P(Q) denotes the probability of the corresponding Q distributions, while P(Q)$_{max}$ corresponds to the maximum value of the corresponding probability.

Recent simulations suggest that the dynamical slowdown of water molecules in the vicinity of hydrophobic solutes is commensurate with an increase in their ordering, reflected in an increase in their tetrahedrality.[81,82] Although these studies involved polarizable water models, we note here that the TIP3P water model has yielded useful insights into the structural behavior of water at different thermo-chemical conditions.[88,89] In Figure 4.7, we have compared distributions of the tetrahedral order parameter, Q,[59] of bulk water with that of water surrounding Helix$_H$ and Beta$_H$, at T_c and T_s. We observe a slight enhancement in tetrahedrality in agreement with recent experimental reports.[80] Further, compared to bulk water, the distributions for the hydration layer water in the vicinity of the hydrophobic residues display a marginal shift to higher Q values (particularly distinct in the range of Q = 0.5 to 0.8). Interestingly however, subtle differences in the tetrahedrality distributions for the hydration water molecules of Helix$_H$ and Beta$_H$ are observed between T_c and T_s. While at T_s, the distributions corresponding to Beta$_H$ and Helix$_H$ overlap, at T_c, it is shifted to lower values and closer to the corresponding distribution for bulk water (evident between Q = 0.5 to 0.8). Thus, at T_c, Beta$_H$ has a marginally lower order in comparison with Helix$_H$.

Behavior of a thermally stable protein at T_c

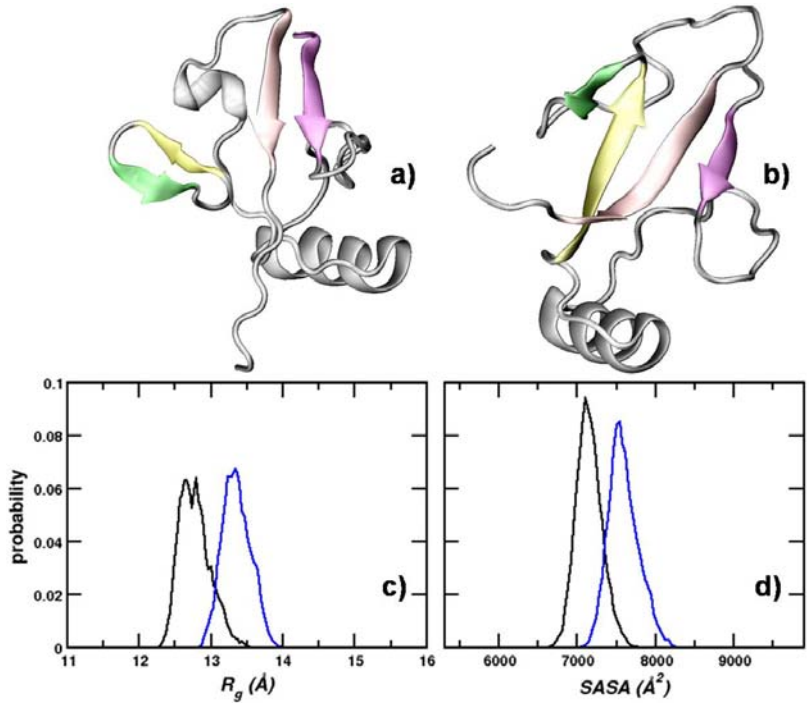


Figure 4.8 a) Representative snapshot at T_s, b) Representative snapshot at T_c, c) distribution of R_g at T_s *(black)* and T_c *(blue)*, and d) distribution of *SASA* at T_s *(black)* and T_c *(blue)*; of thermally stable *Ubq*.

Most folded proteins exhibit a markedly wider range of thermal stability compared to Yfh1. In order to gain preliminary insights into the higher susceptibility of Yfh1 to cold denaturing conditions, we compared its nanosecond timescale behavior with that of Ubiquitin (Ubq) at T_c. Ubq is known to display high thermal stability[87] and to cold denature at temperatures as low as 238 K.[23,90] Visual inspection of snapshots of Ubq from simulations produced at T_s (292 K) and T_c (268 K), presented in Figures 4.8 a and b, respectively, indicate that the structural integrity of this protein is maintained at T_c. Contrary to the structural effects in Yfh1, the beta strands of Ubq are mostly found to be nearly intact at T_c. In Figures 4.8 c and d, we have compared distributions of the R_g and *SASA* of Ubq at T_s and T_c. The distributions are found to shift only marginally to higher values at T_c, and are unlike the distinct shifts observed in Yfh1 (see Figure 4.1). The mean R_g and *SASA* increase by 4.5 and 6.1%, respectively, and are thus significantly lower than the corresponding changes in Yfh1. Further, no differences were found between the values of the structural persistence, P_{Full}, evaluated at T_c and T_h using the initial structure as reference. We

emphasize here that the structural stability of Ubq observed at T_c indicates that atomistic force fields are capable of effectively capturing cold denaturation processes in proteins.

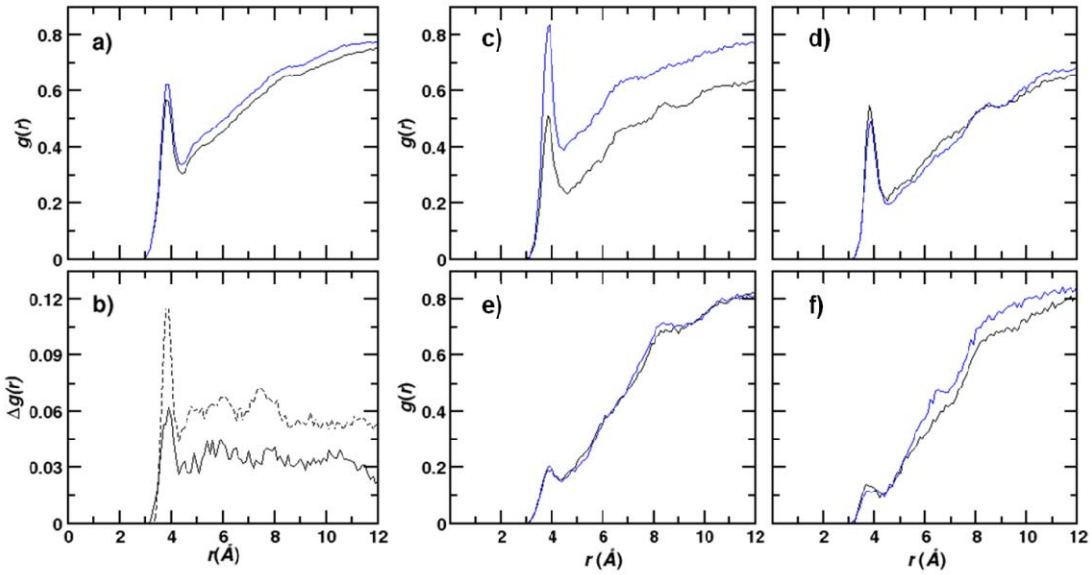


Figure 4.9 a) $g(r)$ of Ubq between backbone carbonyl carbon atoms of the protein and water oxygen atoms at T_s *(black)* and T_c *(blue)*, b) $\Delta g(r)$ of Ubq (black straight line) and Yfh1 (black dashed line). c) $g(r)$ for the Beta_H domains of Yfh1 at T_s *(black)* and T_c *(blue)*, d) $g(r)$ for the Beta_H domains of Ubq at T_s *(black)* and T_c *(blue)*, e) $g(r)$ for the Helix_H domains of Yfh1 at T_s *(black)* and T_c *(blue)*, f) $g(r)$ for the Helix_H domains of Ubq at T_s *(black)* and T_c *(blue)*.

In Figure 4.9 a, we have compared the $g(r)$ calculated for the backbone carbonyl carbon atom and water oxygens, at T_s and T_c for Ubq, and find a relatively smaller increase in the height of the first solvation peak upon temperature lowering, compared to Yfh1. This becomes eminently clear when the difference $g(r)$ plots of the two proteins are compared (Figure 4.9 b). It is also noteworthy that while the protein-water $g(r)$ for the Beta_H domains of Yfh1 shows a sharp increase from T_s to T_c, they remain virtually unchanged in the case of Ubq (Figure 4.9 c and d). Also there is relatively little or no distinction of the $g(r)$ peaks of Helix_H for both Yfh1 and Ubq from T_s to T_c (Figure 4.9 e and f). This comparison shows that the increased hydration of Yfh1 reflected in the $g(r)$ data at T_c can be partly attributed due to its structural unfolding, while Ubq remains largely unaffected in its structural and hydration propensity.

Characteristic differences between early heat and cold denaturation of Yfh1

The data presented thus far show that cold denaturation in Yfh1 is associated with the weakening of the hydrophobic effect around key domains, particularly around the beta strands. Protein heat denaturation, in contrast, is generally attributed to increased conformational disorder with the

increase in thermal energy[12]. In Figure 4.10 a, we present the distributions of N_w and P_{Full} for the whole protein at T_h. The corresponding data for the beta strands and the helical domains are presented in Figures 4.10 b and c, respectively. A comparison with Figure 4.10 indicates that overall, within the relevant timescales, the degree of structural unfolding for the entire protein is comparable at T_c and T_h. However, the distribution obtained for the beta sheet domains indicate relatively lesser beta sheet unfolding at T_h compared to T_c. Interestingly, unlike the behavior at T_c, a marginal degree of helical unfolding is observed at T_h. Marked differences are further observed in the nature of surface hydration. Unlike the scenario at T_c, the unfolding is not accompanied by an increase in surface hydration at T_h, and in fact, the mean hydration number reduces from T_s to T_h. To understand how the hydrophobic effect is altered at T_h, we obtained the water density fluctuations, $P(n^*)$, around key hydrophobic domains at T_h (Figures 4.10 d and e). Unlike the behavior at T_c, a broadening of the distributions is observed for the key domains, indicating a marginal enhancement of the hydrophobic effect at T_h. Difference in the extent of disorder induced by the effects of heat and cold denaturation are further apparent when the cumulative configuration entropies, obtained by employing Schlitter's method,[61,91,92] are compared between the temperatures (see Figure 4.11). The final values of the entropies at 90 ns are listed in Table 4.2. When the entire protein is considered, the configurational entropy shows a clear increase with temperature till about 55 ns, beyond which the entropy at T_c marginally exceeds that at T_s. However, this trend is not maintained when the beta-sheet domains, especially the hydrophobic residues belonging to these domains, are considered. For these domains, the entropy at T_c is noticeably greater than corresponding values at T_h at early times, and the values converge towards the end of the simulation trajectories. This analysis clearly demonstrates the association between the weakening of the hydrophobic effect observed at T_c in the beta-sheet and their local unfolding. The unfolding at T_h, on the other hand, arises due to enhanced configurational disorder attributable to increased thermal energy.

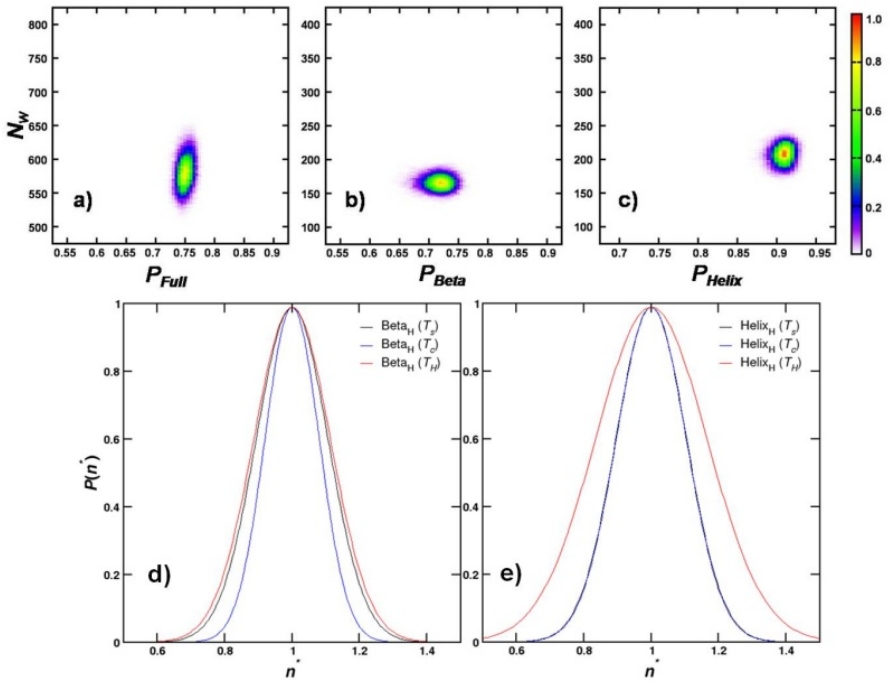


Figure 4.10 2D-contour plots of P versus N_w for; a) whole protein, b) beta domains, c) helical domains, at T_h; and comparison of the probability distributions of water densities near the surfaces of sidechains of d) Beta$_H$ and e) Helix$_H$, at T_s *(black)*, T_c *(blue)* and T_h *(red)*; for Yfh1.

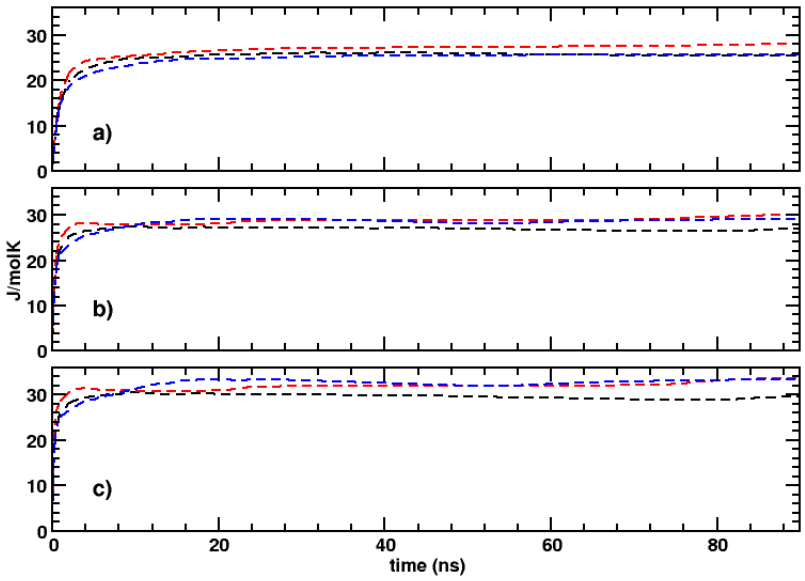


Figure 4.11 Cumulative configurational entropy per atom of Yfh1 for the backbone heavy atoms of a) the whole protein, b) the beta domains, and c) the beta$_H$ domains; at T_s *(black)*, T_c *(blue)* and T_h *(red)*.

Table 4.2 Tabulated final values of the configurational entropy per atom for the backbone heavy atoms of the whole protein (S_{Full}), beta-sheets (S_{Beta}), and hydrophobic residues of the beta-sheets (S_{BetaH}). Unit of entropy are in J/molK.

Temperature	S_{Full}	S_{Beta}	S_{BetaH}
T_c	25.72	29.04	33.36
T_s	25.57	26.95	29.27
T_h	28.00	29.99	33.64

4.4 Conclusion

We have used fully atomistic MD simulations to investigate the early aspects of the cold denaturation process of the protein, Yfh1 that is reported to cold denature at a temperature of 268 K (T_c). The protein's behavior at T_c has been compared to that at the temperature of maximum stability, 292 K (T_s). In addition, the main unfolding characteristics have been compared with those observed at 324 K (T_h). The analyses of the simulations detect a decrease in the globular compactness followed by increased surface hydration at T_c. The decrease in mean values of the residue-wise persistence parameter and increase in heat capacity (as evident from ΔP and ΔC_v plots) show considerable native structural loss in the beta domains in accordance with the recent NMR studies, while helices remain structurally unperturbed within the simulated timescales. The overall decrease in persistence parameter at T_c is simultaneously accompanied by an increased number of hydration water molecules compared to that at T_s, the contribution of which is mostly coming from the beta domains. Calculated IR absorption spectra show a distinct red shift at T_c for the entire protein compared to that at T_s. The structural losses of the beta-sheet domains are reflected by red shifts of the IR spectra of the beta-sheet domains at T_c relative to T_s. Interestingly, we observe that the hydrophobic residues of the beta domains have the major contribution towards the percentage increase in the mean *SASA*/residues (Table 4.3) at T_c.

Table 4.3 SASA/residues (of the side-chains) of Yfh1 for the hydrophobic (All_H) and non-hydrophobic domains (All_{NH}) of the whole protein, along with that of the hydrophobic ($Beta_H$) and non-hydrophobic ($Beta_{NH}$) beta sheet domains, calculated for the initial snapshot ($SASA_i$), and for the last 50 ns of simulations at T_c ($SASA_c$) and T_s ($SASA_s$).

Domains	$SASA_i$	$SASA_c$	% change (T_c)	$SASA_s$	% change (T_s)
All_H	117.9	146.4	24.2	137.7	16.8
All_{NH}	147.5	159.9	8.4	145.6	-1.3
Beta	93.2	148.5	59.3	120.4	29.2
$Beta_H$	135.6	201.3	48.4	163.9	20.9
$Beta_{NH}$	153.3	197.7	29.0	169.3	10.4

To obtain further insights into how hydrophobic hydration at T_c could be associated with the observed denaturation of the beta sheet domains, we investigated the dynamical and structural behavior of the hydration water at the beta sheet domains. In comparison to the bulk water behavior at the two temperatures, we find that while water molecules around $Helix_H$ are relatively more ordered at T_c with a slower rotational relaxation in accordance with the recent experimental and theoretical findings, water molecules around $Beta_H$ are having a comparatively faster decay as well as lesser structural ordering. Also, water density fluctuations within a thin probe volume suggest a lesser fluctuation for $Beta_H$ at T_c.

The denaturation behavior of Yfh1 at T_c is not observed in the case of Ubq, a protein associated with a cold denaturation temperature of 238 K. This observation underscores the validity of empirical force fields in capturing the essential elements of protein cold denaturation phenomena. Further detailed studies will be required to reveal how the stabilizing internal topologies of various folded proteins compete with the interactions arising from the solvent as the temperature is systematically lowered. Our results appear to lend credence to the 'non-cooperative' nature of cold protein denaturation that has emerged from some recent

studies.[22,23,26] The present study opens the possibilities of future experimental and theoretical investigations of the temperature effects pertaining to the varied structural and dynamical aspects of mesophilic proteins and peptide sequences.

4.5 Bibliography

(1) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science*. **2012**, *338*, 1042-1046.

(2) Dobson, C. M. Protein folding and misfolding. *Nature*. **2003**, *426*, 884-890.

(3) Jahn, T. R.; Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **2005**, *272*, 5962-5970.

(4) Michalet, X.; Weiss, S.; Jäger, M. Single-Molecule Fluorescence Studies of Protein Folding and Conformational Dynamics. *Chem. Rev.* **2006**, *106*, 1785-1813.

(5) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. Molecular Simulation of ab Initio Protein Folding for a Millisecond Folder NTL9(1−39). *J. Am. Chem. Soc.* **2010**, *132*, 1526-1528.

(6) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundström, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. Structure of an Intermediate State in Protein Folding and Aggregation. *Science*. **2012**, *336*, 362-366.

(7) Baiz, C. R.; Lin, Y.-S.; Peng, C. S.; Beauchamp, K. A.; Voelz, V. A.; Pande, V. S.; Tokmakoff, A. A Molecular Interpretation of 2D IR Protein Folding Experiments with Markov State Models. *Biophys. J.* **2014**, *106*, 1359-1370.

(8) Privalov, P. L. Cold Denaturation of Protein. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 281-306.

(9) Tamura, A.; Kimura, K.; Takahara, H.; Akasaka, K. Cold denaturation and heat denaturation of Streptomyces subtilisin inhibitor: (I). CD and DSC studies. *Biochemistry*. **1991**, *30*, 11307-11313.

(10) Griko Yu, V.; Kutyshenko, V. P. Differences in the processes of beta-lactoglobulin cold and heat denaturations. *Biophys. J.* **1994**, *67*, 356-363.

(11) Matysiak, S.; Debenedetti, P. G.; Rossky, P. J. Role of Hydrophobic Hydration in Protein Stability: A 3D Water-Explicit Protein Model Exhibiting Cold and Heat Denaturation. *J. Phys. Chem. B. 116*, 8095-8104.

(12) Dias, C. L.; Ala-Nissila, T.; Karttunen, M.; Vattulainen, I.; Grant, M. Microscopic Mechanism for Cold Denaturation. *Phys. Rev. Lett.* **2008**, *100*, 118101.

(13) Yoshidome, T.; Kinoshita, M. Hydrophobicity at low temperatures and cold denaturation of a protein. *Phys. Rev. E.* **2009**, *79*, 030905.

(14) Adrover, M.; Martorell, G.; Martin, S. R.; Urosev, D.; Konarev, P. V.; Svergun, D. I.; Daura, X.; Temussi, P.; Pastore, A. The Role of Hydration in Protein Stability: Comparison of the Cold and Heat Unfolded States of Yfh1. *J. Mol. Biol. 417*, 413-424.

(15) Agashe, V. R.; Shastry, M. C. R.; Udgaonkar, J. B. Initial hydrophobic collapse in the folding of barstar. *Nature*. **1995**, *377*, 754-757.

(16) Dill, K. A. Dominant forces in protein folding. *Biochemistry*. **1990**, *29*, 7133-7155.

(17) Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature*. **2005**, *437*, 640-647.

(18) Tsai, C.-J.; Maizel, J. V.; Nussinov, R. The Hydrophobic Effect: A New Insight from Cold Denaturation and a Two-State Water Structure. *Crit. Rev. Biochem. Mol. Biol.* **2002**, *37*, 55-69.

(19) Lopez, C. F.; Darst, R. K.; Rossky, P. J. Mechanistic Elements of Protein Cold Denaturation. *J. Phys. Chem. B.* **2008**, *112*, 5961-5967.

(20) Graziano, G. On the molecular origin of cold denaturation of globular proteins. *Phys. Chem. Chem. Phys. 12*, 14245-14252.

(21) Riccio, A.; Graziano, G. Cold unfolding of β-hairpins: A molecular-level rationalization. *Proteins. 79*, 1739-1746.

(22) Babu, C. R.; Hilser, V. J.; Wand, A. J. Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation. *Nat. Struct. Mol. Biol.* **2004**, *11*, 352-357.

(23) Pometun, M. S.; Peterson, R. W.; Babu, C. R.; Wand, A. J. Cold Denaturation of Encapsulated Ubiquitin. *J. Am. Chem. Soc.* **2006**, *128*, 10652-10653.

(24) Adrover, M.; Esposito, V.; Martorell, G.; Pastore, A.; Temussi, P. A. Understanding Cold Denaturation: The Case Study of Yfh1. *J. Am. Chem. Soc.* **2010**, *132*, 16240-16246.

(25) Das, P.; Matysiak, S. Direct Characterization of Hydrophobic Hydration during Cold and Pressure Denaturation. *J. Phys. Chem. B.* **2012**, *116*, 5342-5348.

(26) Whitten, S. T.; Kurtz, A. J.; Pometun, M. S.; Wand, A. J.; Hilser, V. J. Revealing the Nature of the Native State Ensemble through Cold Denaturation. *Biochemistry*. **2006**, *45*, 10163-10174.

(27) Pace, C. N.; Tanford, C. Thermodynamics of the unfolding of *β*-lactoglobulin A in aqueous urea solutions between 5 and 55^{o}. *Biochemistry*. **1968**, *7*, 198-208.

(28) Chen, B. L.; Schellman, J. A. Low-temperature unfolding of a mutant of phage T4 lysozyme. 1. Equilibrium studies. *Biochemistry*. **1989**, *28*, 685-691.

(29) Chen, B. L.; Baase, W. A.; Schellman, J. A. Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. *Biochemistry*. **1989**, *28*, 691-699.

(30) Martin, S. R.; Esposito, V.; De Los Rios, P.; Pastore, A.; Temussi, P. A. Cold Denaturation of Yeast Frataxin Offers the Clue to Understand the Effect of Alcohols on Protein Stability. *J. Am. Chem. Soc.* **2008**, *130*, 9963-9970.

(31) Davidovic, M.; Mattea, C.; Qvist, J.; Halle, B. Protein Cold Denaturation as Seen From the Solvent. *J. Am. Chem. Soc.* **2008**, *131*, 1025-1036.

(32) Li, Y.; Shan, B.; Raleigh, D. P. The Cold Denatured State Is Compact but Expands at Low Temperatures: Hydrodynamic Properties of the Cold Denatured State of the C-terminal Domain of L9. *J. Mol. Biol*. **2007**, *368*, 256-262.

(33) Agashe, V. R.; Udgaonkar, J. B. Thermodynamics of denaturation of barstar: evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry*. **1995**, *34*, 3286-3299.

(34) Pastore, A.; Martin, S. R.; Politou, A.; Kondapalli, K. C.; Stemmler, T.; Temussi, P. A. Unbiased Cold Denaturation: Low- and High-Temperature Unfolding of Yeast Frataxin under Physiological Conditions. *J. Am. Chem. Soc.* **2007**, *129*, 5374-5375.

(35) Buchner, G. S.; Shih, N.; Reece, A. E.; Niebling, S.; Kubelka, J. Unusual Cold Denaturation of a Small Protein Domain. *Biochemistry*. **2012**, *51*, 6496-6498.

(36) Luan, B.; Shan, B.; Baiz, C.; Tokmakoff, A.; Raleigh, D. P. Cooperative Cold Denaturation: The Case of the C-Terminal Domain of Ribosomal Protein L9. *Biochemistry*. **2013**, *52*, 2402-2409.

(37) Ascolese, E.; Graziano, G. On the cold denaturation of globular proteins. *Chem. Phys. Lett.* **2008**, *467*, 150-153.

(38) Oshima, H.; Yoshidome, T.; Amano, K.-i.; Kinoshita, M. A theoretical analysis on characteristics of protein structures induced by cold denaturation. *J. Chem. Phys.* **2009**, *131*, 205102-205111.

(39) Sanfelice, D.; Puglisi, R.; Martin, S. R.; Di Bari, L.; Pastore, A.; Temussi, P. A. Yeast Frataxin Is Stabilized by Low Salt Concentrations: Cold Denaturation Disentangles Ionic Strength Effects from Specific Interactions. *PLoS ONE*. **2014**, *9*, e95801.

(40) Aznauryan, M.; Nettels, D.; Holla, A.; Hofmann, H.; Schuler, B. Single-Molecule Spectroscopy of Cold Denaturation and the Temperature-Induced Collapse of Unfolded Proteins. *J. Am. Chem. Soc.* **2013**, *135*, 14040-14043.

(41) Ham, S.; Cho, M. Amide I modes in the N-methylacetamide dimer and glycine dipeptide analog: Diagonal force constants. *J. Chem. Phys.* **2003**, *118*, 6915-6922.

(42) He, Y.; Alam, S. L.; Proteasa, S. V.; Zhang, Y.; Lesuisse, E.; Dancis, A.; Stemmler, T. L. Yeast Frataxin Solution Structure, Iron Binding, and Ferrochelatase Interaction. *Biochemistry*. **2004**, *43*, 16254-16262.

(43) Heinig, M.; Frishman, D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res*. **2004**, *32*, W500-W502.

(44) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics*. **1996**, *14*, 33-38.

(45) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* **1998**, *102*, 3586-3616.

(46) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem*. **2004**, *25*, 1400-1415.

(47) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys*. **1999**, *151*, 283-312.

(48) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(49) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.

(50) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613-4621.

(51) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys*. **1977**, *23*, 327-341.

(52) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of ubiquitin refined at 1.8 Å resolution. *J. Mol. Biol.* **1987**, *194*, 531-544.

(53) Chatterjee, P.; Sengupta, N. Effect of the A30P mutation on the structural dynamics of micelle-bound αSynuclein released in water: a molecular dynamics study. *Eur. Biophys. J*. **2012**, *41*, 483-489.

(54) Jose, J.; Sengupta, N. Molecular dynamics simulation studies of the structural response of an isolated A$\beta_{1\text{-}42}$ monomer localized in the vicinity of the hydrophilic TiO_2 surface. *Eur. Biophys. J*. **2013**, *42*, 487-494.

(55) Marchi, M.; Sterpone, F.; Ceccarelli, M. Water Rotational Relaxation and Diffusion in Hydrated Lysozyme. *J. Am. Chem. Soc.* **2002**, *124*, 6787-6791.

(56) Abseher, R.; Schreiber, H.; Streinhauser, O. The Influence of a Protein on Water Dynamics in Its Vicinity Investigated by Molecular Dynamics Simulation. *PROTEINS.* **1996**, *25*, 366-378.

(57) Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Hagan, M. F.; Chandler, D.; Garde, S. Sitting at the Edge: How Biomolecules use Hydrophobicity to Tune Their Interactions and Function. *J. Phys. Chem. B*. **2012**, *116*, 2498-2503.

(58) Chau, P. L.; Hardwick, A. J. A new order parameter for tetrahedral configurations. *Molecular Phys*. **1998**, *93*, 511-518.

(59) Errington, J. R.; Debenedetti, P. G. Relationship between structural order and the anomalies of liquid water. *Nature*. **2001**, *409*, 318-321.

(60) Bhattacharjee, N.; Biswas, P. Structure of hydration water in proteins: A comparison of molecular dynamics simulations and database analysis. *Biophys. Chem.* **2011**, *158*, 73-80.

(61) Schlitter, J. r. Estimation of absolute and relative entropies of macromolecules using the covariance matrix. *Chem. Phys. Lett.* **1993**, *215*, 617-621.

(62) Falvo, C.; Zhuang, W.; Kim, Y. S.; Axelsen, P. H.; Hochstrasser, R. M.; Mukamel, S. Frequency Distribution of the Amide-I Vibration Sorted by Residues in Amyloid Fibrils Revealed by 2D-IR Measurements and Simulations. *J. Phys. Chem. B.* **2012**, *116*, 3322-3330.

(63) la Cour Jansen, T.; Dijkstra, A. G.; Watson, T. M.; Hirst, J. D.; Knoester, J. Modeling the amide I bands of small peptides. *J. Chem. Phys.* **2006**, *125*, 044312.

(64) Wang, L.; Middleton, C. T.; Zanni, M. T.; Skinner, J. L. Development and Validation of Transferable Amide I Vibrational Frequency Maps for Peptides. *J. Phys. Chem. B.* **2011**, *115*, 3713-3724.

(65) Choi, J.-H.; Lee, H.; Lee, K.-K.; Hahn, S.; Cho, M. Computational spectroscopy of ubiquitin: Comparison between theory and experiments. *J. Chem. Phys.* **2007**, *126*, 045102.

(66) Ham, S.; Hahn, S.; Lee, C.; Kim, T.-K.; Kwak, K.; Cho, M. Amide I Modes of α-Helical Polypeptide in Liquid Water: Conformational Fluctuation, Phase Correlation, and Linear and Nonlinear Vibrational Spectra. *J. Phys. Chem. B.* **2004**, *108*, 9333-9345.

(67) Hahn, S.; Ham, S.; Cho, M. Simulation Studies of Amide I IR Absorption and Two-Dimensional IR Spectra of β Hairpins in Liquid Water. *J. Phys. Chem. B.* **2005**, *109*, 11789-11801.

(68) Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A*. **2001**, *105*, 9954-9960.

(69) Sengupta, N.; Jaud, S.; Tobias, D. J. Hydration Dynamics in a Partially Denatured Ensemble of the Globular Protein Human α-Lactalbumin Investigated with Molecular Dynamics Simulations. *Biophys. J.* **2008**, *95*, 5257-5267.

(70) Bagchi, S.; Boxer, S. G.; Fayer, M. D. Ribonuclease S Dynamics Measured Using a Nitrile Label with 2D IR Vibrational Echo Spectroscopy. *J. Phys. Chem. B.* **2012**, *116*, 4034-4042.

(71) Bagchi, S.; Nebgen, B. T.; Loring, R. F.; Fayer, M. D. Dynamics of a Myoglobin Mutant Enzyme: 2D IR Vibrational Echo Experiments and Simulations. *J. Am. Chem. Soc.* **2010**, *132*, 18367-18376.

(72) Zanetti Polzi, L.; Amadei, A.; Aschi, M.; Daidone, I. New Insight into the IR-Spectra/Structure Relationship in Amyloid Fibrils: A Theoretical Study on a Prion Peptide. *J. Am. Chem. Soc.* **2011**, *133*, 11414-11417.

(73) Daidone, I.; Aschi, M.; Zanetti-Polzi, L.; Di Nola, A.; Amadei, A. On the origin of IR spectral changes upon protein folding. *Chem. Phys. Lett.* **2010**, *488*, 213-218.

(74) Welch, W. R. W.; Keiderling, T. A.; Kubelka, J. Structural Analyses of Experimental ^{13}C Edited Amide I' IR and VCD for Peptide β-Sheet Aggregates and Fibrils Using DFT-Based Spectral Simulations. *J. Phys. Chem. B.* **2013**, *117*, 10359-10369.

(75) Bagchi, S.; Charnley, A. K.; Smith Iii, A. B.; Hochstrasser, R. M. Equilibrium Exchange Processes of the Aqueous Tryptophan Dipeptide. *J. Phys. Chem. B.* **2009**, *113*, 8412-8417.

(76) Manas, E. S.; Getahun, Z.; Wright, W. W.; DeGrado, W. F.; Vanderkooi, J. M. Infrared Spectra of Amide Groups in α-Helical Proteins: Evidence for Hydrogen Bonding between Helices and Water. *J. Am. Chem. Soc.* **2000**, *122*, 9883-9890.

(77) Walsh, S. T. R.; Cheng, R. P.; Wright, W. W.; Alonso, D. O. V.; Daggett, V.; Vanderkooi, J. M.; DeGrado, W. F. The hydration of amides in helices; a comprehensive picture from molecular dynamics, IR, and NMR. *Protein Sci*. **2003**, *12*, 520-531.

(78) Amunson, K. E.; Kubelka, J. On the Temperature Dependence of Amide I Frequencies of Peptides in Solution. *J. Phys. Chem. B.* **2007**, *111*, 9993-9998.

(79) Kaminský , J.; Bouř, P.; Kubelka, J. Simulations of the Temperature Dependence of Amide I Vibration. *J. Phys. Chem. A*. **2011**, *115*, 30-34.

(80) Davis, J. G.; Gierszal, K. P.; Wang, P.; Ben-Amotz, D. Water structural transformation at molecular hydrophobic interfaces. *Nature*. **2012**, *491*, 582-585.

(81) Galamba, N. Water's Structure around Hydrophobic Solutes and the Iceberg Model. *J. Phys. Chem. B.* **2013**, *117*, 2153-2159.

(82) Galamba, N. Water Tetrahedrons, Hydrogen-Bond Dynamics, and the Orientational Mobility of Water around Hydrophobic Solutes. *J. Phys. Chem. B.* **2014**, *118*, 4169-4176.

(83) Moon, C. P.; Fleming, K. G. Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10174-10177.

(84) Bizzarri, A. R.; Cannistraro, S. Molecular Dynamics of Water at the Protein–Solvent Interface. *J. Phys. Chem. B.* **2002**, *106*, 6617-6633.

(85) Bandyopadhyay, S.; Chakraborty, S.; Balasubramanian, S.; Pal, S.; Bagchi, B. Atomistic Simulation Study of the Coupled Motion of Amino Acid Residues and Water Molecules around Protein HP-36: Fluctuations at and around the Active Sites. *J. Phys. Chem. B.* **2004**, *108*, 12608-12616.

(86) Pal, S.; Bandyopadhyay, S. Importance of Protein Conformational Motions and Electrostatic Anchoring Sites on the Dynamics and Hydrogen Bond Properties of Hydration Water. *Langmuir*. **2013**, *29*, 1162-1173.

(87) Pizzitutti, F.; Marchi, M.; Sterpone, F.; Rossky, P. J. How Protein Surfaces Induce Anomalous Dynamics of Hydration Water. *J. Phys. Chem. B.* **2007**, *111*, 7584-7590.

(88) Agarwal, M.; Kushwaha, H. R.; Chakravarty, C. Local Order, Energy, and Mobility of Water Molecules in the Hydration Shell of Small Peptides. *J. Phys. Chem. B*. **2009**, *114*, 651-659.

(89) Bhattacharjee, N.; Biswas, P. Structure of hydration water in proteins: A comparison of molecular dynamics simulations and database analysis. *Biophys. Chem*. **2011**, *158*, 73-80.

(90) Ibarra-Molero, B.; Makhatadze, G. I.; Sanchez-Ruiz, J. M. Cold denaturation of ubiquitin. *BBA-Proteins Struct. M.* **1999**, *1429*, 384-390.

(91) Schäfer, H.; Mark, A. E.; van Gunsteren, W. F. Absolute entropies from molecular dynamics simulation trajectories. *J. Chem. Phys.* **2000**, *113*, 7809-7817.

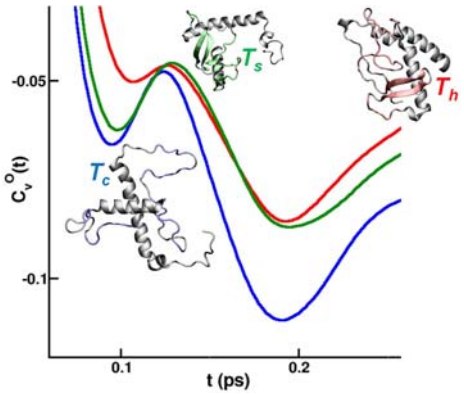
(92) Sinha, S. K.; Chakraborty, S.; Bandyopadhyay, S. Secondary Structure Specific Entropy Change of a Partially Unfolded Protein Molecule. *Langmuir*. **2010**, *26*, 9911-9916.

Chapter 5

Onset of protein thermal denaturation in selective domains is reflected in the local hydration behavior: insights from a comparative molecular dynamics study

All truths are easier to understand once they are discovered; the point is to discover them.
- Galilieo Galilei

Abstract



We investigate, using atomistic molecular dynamics, the association of surface hydration accompanying local unfolding in the mesophilic protein Yfh1 at a series of thermal conditions spanning its cold and heat denaturation temperatures. The results are benchmarked against the thermally stable protein, Ubq, and behavior at the maximum stability temperature. Local unfolding in Yfh1, predominantly in the beta sheet regions, is in qualitative agreement with recent solution NMR studies; corresponding Ubq unfolding is not observed. Interestingly, all domains, except the beta sheet domains of Yfh1, show increased effective surface hydrophobicity with rise in temperatures, as reflected by the density fluctuations of the hydration layer. Velocity autocorrelation functions (VACF) of water oxygens within the hydration layers, and the corresponding vibrational density of states (VDOS) are used to characterize alteration in dynamical behavior accompanying the temperature dependent local unfolding. Enhanced caging effects accompanying transverse oscillations of the water molecules are found to occur with the increase in temperature preferentially for the beta sheet domains of Yfh1. Helical domains of both proteins exhibit similar trends in VDOS with changes in temperature. This work demonstrates the existence of key signatures of the local onset of protein thermal denaturation in solvent dynamical behavior.

5.1 Introduction

Structural stability of soluble proteins at physiological conditions arises from a complex superposition of several short and long ranged forces.[1] Importantly, the balance of the forces is uniquely modulated by the solvent environment, and major disruptions to either the character of the solvent, or to the nature of its interactions with the protein can induce protein denaturation, misfolding and even subsequent aggregation.[2-4] The critical influence exerted by the existing physico-chemical conditions on protein stability has been elicited recently with calorimetric, circular-dichroism, nuclear magnetic resonance (NMR), atomic force microscopy (AFM), spectroscopic studies, etc.[5-12] Further, thermodynamical models and atomistic computer simulations have provided major insights into how electrostatic, dispersion and hydrogen bonding effects are synchronized in order to maintain the conformational stability of these proteins.[13-17]

Protein denaturation can be induced by drastic physical changes in temperature and pressure,[18-20] and also by alterations to chemical conditions brought about by pH changes, addition of chemical denaturants, etc.[21-28] Interestingly, the mechanism of high pressure denaturation has been shown to have some resemblance to denaturation initiated by urea and guanidine hydrochloride, particularly in the hydrogen bonding rupture with the surrounding environment.[29] Such observations are corroborated by theoretical investigations that show that both chemical and thermal denaturation processes are accompanied by an expansion of the hydrophobic core followed by an enhanced solvation.[30]

Protein *thermal* denaturation has been under intense scrutiny recently for its vast relevance in applications ranging from therapeutics and food industry to cryo-preservation of drugs (of biotechnological relevance), etc.[31-36] A severe increase or decrease in temperatures over the physiological regime is known to bring about ‘heat’ or ‘cold’ denaturation in proteins, commensurate with major structural alterations and increased conformational disorder.[2] The free energy difference between folded and unfolded states is close to zero over a small range of temperature close to the physiological temperature (temperature of ‘maximum stability’), and increases with changes in temperature in either direction.[2,37] Interestingly, despite similarities such as reduced native contacts, enhanced conformational entropy and heat capacity, the cold and heat denatured protein states have been demonstrated to be not uniform, and indeed, there

are profound mechanistic differences in the two processes.[38-42] Some studies suggest that protein cold denaturation is a non-cooperative process,[38,43,44] whereas more recent studies have depicted cooperativity in the cold denaturation mechanism.[45,46] Further, although temperature induced denaturation has been depicted to be reversible in nature,[47,48] recent studies indicate important differences between the kinetics of heat and cold denaturation; while heat denatured states can revert back to the folded form, cold denaturation results in trapping of the unfolded state.[49] Also, the hydration propensity is entirely different between the two processes, with markedly increased hydration accompanying the cold denaturation, but not the heat induced denaturation.[49,50] Incidentally, the hydrophobic effect, or the unfavorable solvation of apolar moieties in aqueous environment that can result in their assembly and aggregation, is considered one of the key driving forces for the folding of globular proteins.[17,51,52] The temperature dependence of the hydrophobic effect, known for the last few decades,[53-55] has been directly observed in recent spectroscopic experiments.[56,57] A weakening of the hydrophobic effect at lower temperatures, resulting in a substantially lower entropic penalty for hydrating the hydrophobic core, has been thought to be one of the key factors behind the cold denaturation mechanism.[2,50,58] Compared to heat denaturation, the cold denaturation process is more difficult to interpret due to the experimental difficulties in attaining the denatured state.[2,59] This difficulty is magnified by the fact that most proteins undergo cold denaturation at temperatures that are several tens of Kelvin lower than the freezing point of water.[2,36,44]

It is further interesting to note that large variations exist at the onset of temperature for cold and heat denaturation in various proteins,[28,60-62] and several studies have attempted to understand the roles played by the protein's sequence and native topology in its thermal stability.[28,45,63] The serendipitous discovery of the cold denaturation of yeast frataxin (Yfh1), the yeast orthologue of the human frataxin protein, at temperatures higher than the freezing point of water, has facilitated uninhibited studies of the protein cold denaturation, without resorting to solvent supercooling or the use of cosolvents to achieve denatured state.[62-64] Yfh1 is experimentally reported to undergo complete cold denaturation at 268 K and heat denaturation at 323 K.[49,62] Thus, the unfolding mechanism of this protein has been studied very recently using a variety of experimental techniques including solution NMR, circular dichroism (CD), Förster resonance energy transfer (FRET), as well as computational and theoretical methods.[26,45,49,63-68] The stability curve of the protein could be determined unambiguously in normal and biased

conditions. The beta sheets were found to denature at both the denaturation conditions, whereas the helices retained residual secondary structural propensity, particularly at the cold denaturation conditions. Further, the protein was relatively more hydrated at cold unfolding conditions. Using atomistic simulations, we recently demonstrated the structural response of Yfh1 to cold denaturation conditions at 268 K within the nanosecond timescales.[69] Unfolding within those timescales was non-uniform and heterogeneous, with the beta sheets being far more prone to unfolding than the helical domains, in qualitative agreement with experimental observations.[45] Further, we observed a close correlation between the extent of unfolding and the surface hydration of hydrophobic residues, corroborating the idea that cold denaturation originates due to weakening of the hydrophobic effect at low temperatures. Substantial differences in the unfolding and the nature of hydration were found between Yfh1 and Ubiquitin (Ubq),[69] a globular protein whose thermal denaturation occurs at extreme temperatures.[43,60]

Emerging experimental, theoretical and computational studies have revealed that the structural, dynamical and folding behavior of proteins and other biomolecules is closely coupled to that of its surrounding aqueous environment.[13,16,40,49,50,58,66,70-74] Signatures of biomolecular structure and function can be found in the behavior of the 'hydration layer', or the water molecules in close proximity to the biomolecular surface, which has aptly been termed 'biological water'.[75] The coupling of the hydration layer with the biomolecular surface results in sharp deviation from the structural and dynamical characteristics of bulk water. Interestingly, the deviation from bulk behavior is also observed for water in the vicinity of hydrophobic molecules, with the extent of deviation distinctly correlated with the extent of hydrophobicity.[57,76,77] Interestingly, we found in our earlier work that the weakening of hydrophobic effect and local unfolding of the beta sheet domains of Yfh1 at 268 K were accompanied by a small but distinct shift towards bulk-like character of the hydration water.[69]

Our previous results indicate that key signatures of the local thermal unfolding should be present in the hydration layer behavior. Particularly, at 268 K, the exhibited bulk like behavior in the dynamic properties (water density fluctuation and rotational relaxation) of the hydration layer water molecules are indicative of a weakening of effective hydrophobicity of apolar residues of the beta domains. These strongly suggest that the surrounding aqueous solvent contains key signatures of protein thermal denaturation. In order to derive a clear picture of the nature of the

hydration layer and its coupling to the early local thermal unfolding within the protein, we herein report investigations of the behavior of the hydration layer of Yfh1 at a series of six temperatures spanning 324 K (T_h) to 268 K (T_c). For comparison, we have performed similar investigations of the hydration layer of Ubq, which is reported to be thermally stable in the said temperature range.[43,60] We point out that several recent studies have observed a correlation between the hydrophobicity of macromolecular solutes and the properties of the surrounding aqueous environment.[40,41,50,78] Herein, we have focused on the association between the dynamical behavior of the hydration layer of Yfh1 with the hydrophobic effect associated with specific structural domains of the protein, at a range of temperatures spanning T_c to T_h. These behavior have been compared with corresponding data obtained from simulations of the thermally stable protein, Ubq, at the same conditions. The beta sheets of Yfh1 are completely lost at cold temperatures, partially retained at high temperatures and substantially preserved at 292 K (T_s), the temperature of maximum stability, while the helices remain nearly intact throughout the entire temperature range. In comparison, for Ubq, changes in secondary structural propensity with temperature are nominal. Further, the loss in secondary structure for the beta sheet domains of Yfh1 is accompanied with an increased hydration at cold denaturing temperatures. Investigations of the dynamical behavior of the hydration layer water molecules from water density distributions, velocity autocorrelation and vibrational density of states depicts domain wise differential responses to temperature and alterations in effective hydrophobicity in Yfh1, but not in the thermally stable Ubq.

5.2 Materials and methods

5.2.1 System Setup and MD simulations

Initial coordinates of the protein Yfh1, consisting of 123 amino acid residues (PDB ID 2GA5), were obtained from solution NMR studies performed at 298 K temperature and 1 atm pressure.[79] From the secondary structure determination of the initial coordinates of Yfh1, based on the STRIDE[80,81] algorithm as implemented in VMD,[82] we had earlier found that it has two alpha helices (19-43 and 107-121), and seven beta strands (50-54, 57-62, 66-71, 78-83, 87-94, 97-100, and 105-106) joined by inter-connecting loops.[69] The sequence and domain-wise information of Yfh1 is produced in Figure 5.1 a from the PDB database (http://www.rcsb.org/) using the STRIDE algorithm.[81,83] The protein was solvated to about 15 Å in either of its extremities with the TIP3P water model,[84] followed by an addition of 15 Na^+ ions to neutralize the system.

Initially a conjugate gradient based energy minimization was performed for the system at 0 K. Subsequently, simulations were run at six different temperatures ranging from T_c to T_h (268 K (or T_c), 280 K (or T_{c1}), 292 K (or T_s), 304 K (or T_{h1}), 315 K (or T_{h2}), and 324 K (or T_h)); with the CMAP correction incorporated CHARMM22 force field,[85,86] and NAMD 2.9 simulations package.[87] Coincidentally, T_{c1} and T_{h1} are the corresponding temperatures of the onset of cold and heat denaturation of Yfh1.[62] Orthorhombic periodic boundary conditions were used for each of the systems. The non-bonded interactions were calculated up to the assigned cutoff of 12 Å, with the smoothing starting at 11 Å. The long range electrostatics was calculated with particle mesh Ewald (PME).[88] Isothermal-isobaric ensemble (NPT) simulations were generated for each of the systems for 90-100 ns, with a time step of 2 fs, and configuration data saved every 2 ps. For obtaining the low frequency vibrational density of states, trajectories with time step 0.5 fs and configuration data saved every 2 fs were generated for each of the systems after equilibration. Nosé–Hoover Langevin Piston method,[89] and Langevin dynamics with a collision frequency of 1 ps^{-1} were used for maintaining the constant pressure and temperature respectively. The lengths of bonds involving hydrogen atoms were constrained using the SHAKE algorithm.[90] For Ubq (a 76 amino acid residue protein), the initial coordinates were obtained from X-ray crystallographic structure (PDB id 1UBQ)[91], whose secondary structural domains (reproduced with similar protocols as mentioned earlier) [81,83] consisted of an alpha helix (23-34) and five beta strands (2-6, 12-16, 41-45, 48-49, and 66-71); this is shown in Figure 5.1 b. Simulations of Ubq were generated with similar protocols at the same temperatures as mentioned before. Simulations obtained from the last 50 ns of simulations were used for the preliminary analyses, whereas high frequency simulations (after equilibrated simulation run) were utilized in obtaining the calculations like vibrational density of states (VDOS), obtained from the corresponding velocity auto-correlation function (VACF). VMD was used for visualization.[82] In addition, the conformations obtained at each temperature at the end of the trajectories were solvated in the TIP4P/2005 water model[92], equilibrated for a few ns, and the last 3 ns used to compare structural properties.

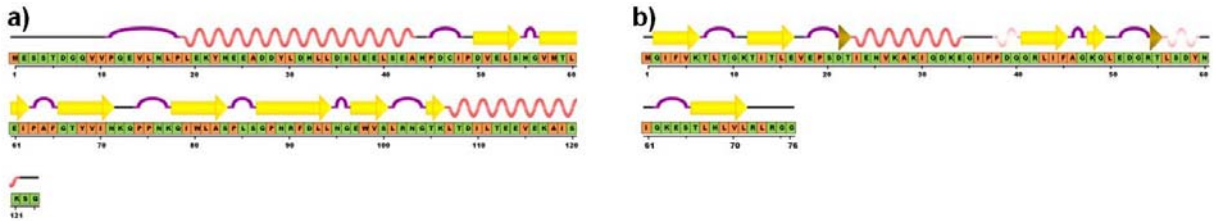


Figure 5.1 Sequence and secondary structure obtained from the PDB structures of a) Yfh1 (2GA5), and b) Ubq (1UBQ). Hydrophobic and non-hydrophobic amino acids are marked with green and orange, respectively. Alpha helical domains are indicates in red; 3_{10} helix in pink; beta sheets in yellow; beta bridges in tan; turns in purple; and coils in black.

5.2.2 Analyses protocol(s)

Structural persistence

We have calculated the parameter P,[93,94] corresponding to the secondary structural persistence, to determine the extent of secondary structural instability with respect to a reference structure, as discussed in previous chapters. The value of P will vary from ~0 to 1, with P values closer to 1 denoting a greater persistence and vice versa.

Water density fluctuation

We have estimated the water density fluctuation,[95] ($P(n^*)$), from the respective surfaces of the secondary structural domains, within a thin probe volume as defined in our earlier study.[69] $P(n^*)$ is obtained by calculating the density of water, divided by the mean density and then maximized by the maximum probability, which is further fitted to the Gaussian distribution

$$P(n^*) = e^{-\frac{(n^*-1)^2}{\alpha}} \tag{5.1}$$

The corresponding α values obtained from the fitted curves gives us the extent of width of the Gaussian distribution and its increase is proportional to the increase in surface hydrophobicity.

Velocity autocorrelation function and vibrational density of states

We have computationally determined the low frequency vibrational density of states (VDOS) procured from Fourier transformation of the corresponding velocity auto-correlation function (VACF),[96-98] for the oxygen atoms of the bulk water as well as hydration water molecules around the different domains (whole protein, beta, helix, hydrophobic beta or $beta_H$, hydrophobic helix or $helix_H$) at the aforesaid temperatures, for both Yfh1 and UBQ.

The VACF of the hydration layer has been calculated as,

$$C_v^O(t) = \frac{\langle \vec{v}_i(0).\vec{v}_i(t) \rangle}{\langle \vec{v}_i(0).\vec{v}_i(0) \rangle} \tag{5.2}$$

where $\vec{v}_i(t)$ is the velocity of the oxygen atom of a water molecule at time *t*. The dynamic exchange of the water molecules between the bulk and the hydration layer has been taken care of, which means only water molecules that reside continuously within the hydration layer (within 4 Å of the protein heavy atoms) over the time *t* are considered. As in several previously reported studies,[96-98] Fourier transformation of the VACF is carried out to obtain the vibrational density of states (VDOS) in the frequency domain; appropriate unit conversions yield VDOS as functions of frequencies in cm^{-1}.

5.3 Results and discussions

Structural propensities and surface hydration

We begin by comparing the temperature dependent structural unfolding characteristics of Yfh1 and Ubq. We first calculated the radius of gyration, R_g (mean values provided in Table 5.1 and 5.2), and the structural persistence, P, of each protein using the last 50 ns of the simulations at each temperature. The mean values obtained for each protein at the temperature of maximum stability (T_s), were used to normalize the R_g and P values, and were designated as $R^{/}$ and $P^{/}$, respectively. Thus, for each protein, $R^{/}$ and $P^{/}$ were 1.0 at T_s. In Figure 5.2, we have plotted $R^{/}$ and $P^{/}$ as functions of temperature for Yfh1 and Ubq, and depicted representative snapshots of both proteins at T_c, T_s and T_h. For Yfh1, the value of $R^{/}$ varies from a maximum of 1.51 (at T_{c1}, commensurate with a 50% increase in R_g) to 0.96 (at T_h), indicating high variability in the degree of compactness with temperature. In Ubq, the variability of $R^{/}$ is distinctly lower in comparison, varying from a maximum value of 1.12 to a minimum value of 0.94. High temperature dependence further characterizes the $P^{/}$ *vs.* temperature plot of Yfh1, indicating the pronounced vulnerability of this protein to thermal denaturation. In comparison, $P^{/}$ is stable for Ubq, and remains close to 1.0 within temperature range in question. Interestingly, however, there is a slight increase in structural stability and compactness observed in both proteins as the temperature increases from T_{h2} to T_h. The representative snapshots of the proteins reflect the structural trends brought forward by the $R^{/}$ and $P^{/}$ analyses. While both proteins are compact and

possess the secondary structural elements of the crystal structures at T_s, there is a distinctly greater structural loss for Yfh1 as the temperatures deviate from T_s, while the structural loss for Ubq is nominal. As discussed in our previous study,[69] there is near total loss of the beta sheet domains of Yfh1 at T_c while the helical domains are largely retained. At T_h, Yfh1 retains a large fraction of the beta sheet domains, while there is a smaller extent of helical unfolding associated with high temperature melting.

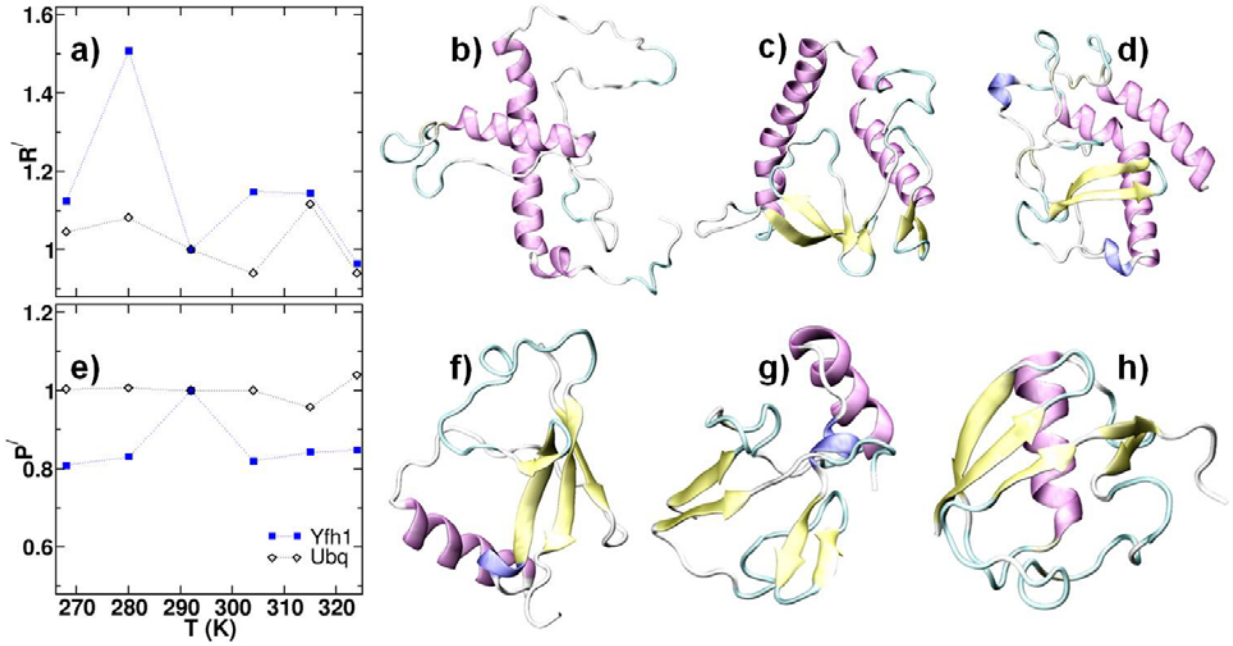


Figure 5.2 Plots of $R^{/}$ and $P^{/}$ for Yfh1 and Ubq as a function of temperature, as well as representative snapshots of Yfh1 and Ubq at selected temperatures. a) evolution of $R^{/}$ of Yfh1 at all the simulated temperatures; snapshots of Yfh1 at b) T_c, c) T_s, d) T_h; e) evolution of $P^{/}$ of Yfh1 at all the simulated temperatures; snapshots of Ubq at f) T_c, g) T_s, f) T_h.

We point out here that despite the unfolding of Yfh1, reflected in reduced $P^{/}$ values at temperatures below and above T_s, the overall decrease in compactness is more distinct at cold denaturing conditions. This observation is in qualitative agreement with a recently reported FRET study of Yfh1 performed within a slightly different temperature range, that showed increasing compaction in the unfolded states of this protein with rise in temperature, and overall higher compaction of the heat denatured states compared to the cold denatured states.[64] The protein was found to be most compact at 320 K, with marginal re-expansion in protein size with further increase in temperature. A similar trend in compactness with temperature has also been noted in an intrinsically disordered protein,[56] and we note that both studies connect these

observations to the alterations in the hydrophobic effect with changes in temperature. Interestingly, our analyses of the simulations on the 100 ns timescales depict that Yfh1 has the minimum R_g at T_h, with an overall decrease in compactness up to T_{c1}, which is the experimentally observed temperature of onset of its cold denaturation. Sampling on longer timescales may reveal if this is a transient phenomenon and possible underlying mechanisms for such observed loss in compactness at the onset of cold denaturation.

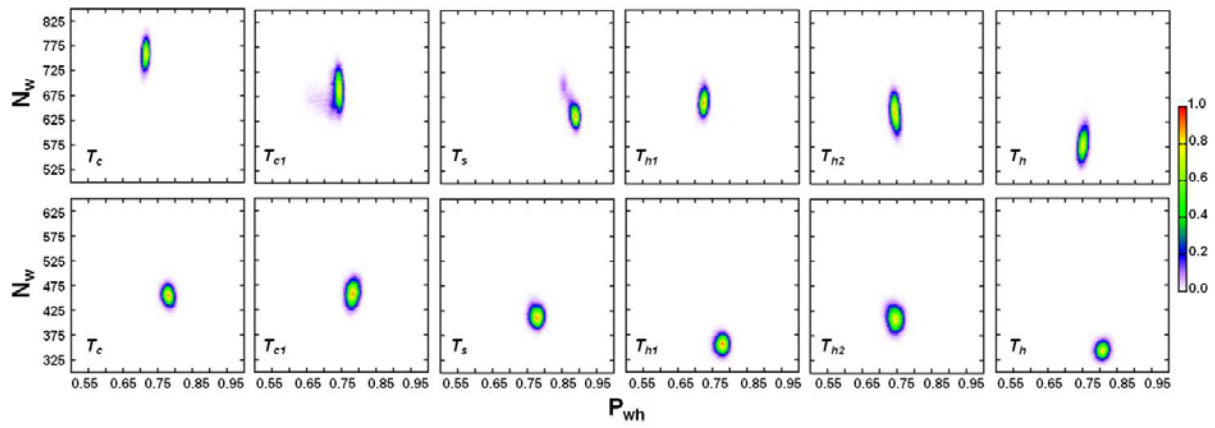


Figure 5.3 Probability distributions of P_{wh} *vs.* N_w (obtained from the last 50 ns of simulations) for Yfh1 (upper row) and Ubq (lower row), at all the simulated temperatures.

Several previous reports have demonstrated a direct association of protein unfolding with enhanced surface hydration at low temperatures.[15,49,69] However, it remains unclear whether the folded states of all globular proteins respond in the same manner to rapid changes in temperature, or whether individual thermal responses are specific to a given protein sequence. For greater insights, we systematically inspect the hydration characteristics of Yfh1 and Ubq systems over the range of simulated temperatures as a function of the structural persistence. In Figure 5.3, we have depicted normalized distributions of (N_w, P_{wh}), where N_w are number of water molecules within the hydration layer (4 Å from the protein heavy atoms) and P_{wh} are structural persistence of the whole protein molecules, with data obtained from the last 50 ns of the simulation trajectories. For Yfh1 at T_s, the distribution is centered at (0.892, 636). There is a distinct lowering in P_{wh} at temperatures greater than T_s, with the centers of distributions at T_{h1}, T_{h2} and T_h being (0.727, 665), (0.742, 643) and (0.747, 578), respectively. Thus, with increasing temperature, there is a marginal decrease in hydration commensurate with the loss in structural stability and enhanced compactness (discussed in further detail in the next section). However,

unlike the distributions at the higher temperatures, the distribution at T_{c1}, centered at (0.742, 703), is characterized by enhanced surface hydration, which increases dramatically at T_c, when the distribution centers at (0.717, 761). Thus, while both increase and decrease in temperature bring about protein unfolding within the simulated timescales, the extent of unfolding is relatively higher, and accompanied by a commensurate increase in surface hydration, as the temperature is systematically lowered to T_c. In comparison, the thermal response of Ubq follows a distinctly different pattern. As the temperature is increased from T_s where the distribution centers at (0.777, 414), to T_{h1}, T_{h2} and T_h, the distribution centers shift to (0.782, 358), (0.747, 410) and (0.812, 344). Thus, for the thermally stable Ubq, at temperatures sufficiently below its melting temperature, the systematic heating is accompanied not just by a minor compactification, but also by a marginal structural stabilization and decrease in surface hydration. Further, in sharp contrast with Yfh1, upon lowering the temperature, Ubq is found to maintain both the structural integrity and level of hydration observed at T_s, with the distributions centered at (0.782, 458) and (0.782, 455) at T_{c1} and T_c, respectively.

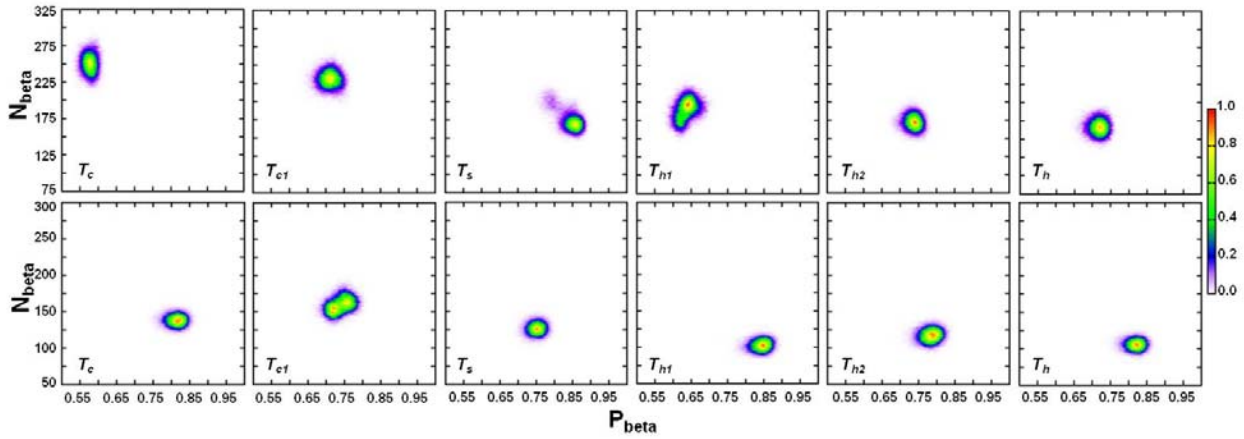


Figure 5.4 Probability distributions of P_{beta} *vs.* N_{beta} (obtained from the last 50 ns of simulations) for Yfh1 (upper row) and Ubq (lower row), at all the simulated temperatures.

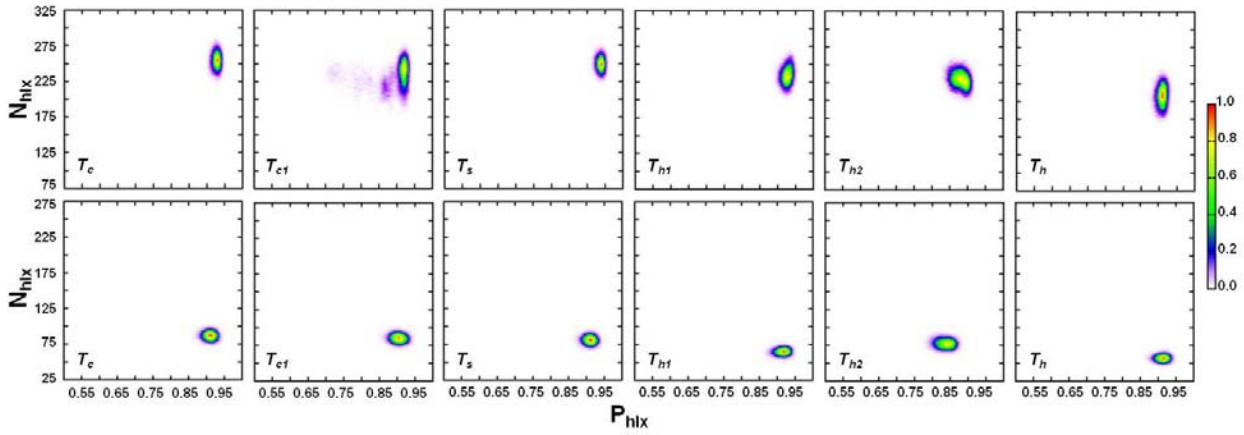


Figure 5.5 Probability distributions of P_{hlx} *vs.* N_{hlx} (obtained from the last 50 ns of simulations) for Yfh1 (upper row) and Ubq (lower row), at all the simulated temperatures.

Table 5.1 Tabulation of the mean R_g (Å), maximum values of P_{wh}, N_w, P_{beta}, N_{beta}, P_{helix}, N_{helix}, as well as α values (10^{-2}) corresponding to the $beta_H$ (α_b) and $helix_H$ (α_h) residues, at simulated temperatures (T); for Yfh1.

T (K)	R_g	P_{wh}	N_w	P_{beta}	N_{beta}	P_{hlx}	N_{hlx}	α_b	α_h
T_c	18.13	0.717	761	0.577	250	0.927	256	1.28	2.25
T_{c1}	24.31	0.742	703	0.712	232	0.922	245	1.61	2.73
T_s	16.11	0.892	636	0.857	167	0.942	249	2.31	2.20
T_{h1}	18.50	0.727	665	0.642	197	0.927	236	2.61	3.14
T_{h2}	18.44	0.742	643	0.742	171	0.877	232	2.46	3.38
T_h	15.54	0.747	578	0.717	166	0.907	206	2.52	5.45

Table 5.2 Tabulation of the mean R_g (Å), maximum values of P_{wh}, N_w, P_{beta}, N_{beta}, P_{helix}, N_{helix}, as well as α values (10^{-2}) corresponding to the $beta_H$ (α_b) and $helix_H$ (α_h) residues, at simulated temperatures (T); for Ubq.

T (K)	R_g	P_{wh}	N_w	P_{beta}	N_{beta}	P_{hlx}	N_{hlx}	α_b	α_h
T_c	13.35	0.782	455	0.817	137	0.907	87	3.33	7.17
T_{c1}	13.82	0.782	458	0.722	156	0.902	87	3.28	7.25
T_s	12.78	0.777	414	0.752	126	0.912	82	3.88	8.20
T_{h1}	12.01	0.782	358	0.847	104	0.922	66	8.30	9.48
T_{h2}	14.26	0.747	410	0.792	120	0.842	78	4.91	9.22
T_h	12.02	0.812	344	0.822	106	0.912	58	6.47	24.99

As reported in our previous study, the enhanced hydration of Yfh1 observed at T_c is primarily associated with the unfolded beta sheet domains. In Figure 5.4, we compare the distributions (N_{beta}, P_{beta}), where N_{beta} and P_{beta} are the hydration number and structural persistence of the beta sheet domains, over all the simulated temperatures, for both the proteins. For Yfh1, the centers of the distribution shift from (0.857, 167) at T_s, to (0.642, 197) at T_{h1}, which corresponds to the melting temperature of the protein. Interestingly, the distribution shifts to (0.742, 171) at T_{h2}, and to (0.717, 166) at T_h. Thus, changes in structural stability of the beta sheet domains with increase in temperature are not accompanied by distinct changes in hydration upon elevation of temperature. On the other hand, upon lowering of the temperature from T_{c1} to T_c, we observe a correlated increase in structural loss and increased hydration, with the center of the distribution shifting from (0.712, 232) at T_{c1} to (0.577, 250) at T_c. In comparison, the beta sheet domains of Ubq stabilize with lowering of temperature, with the distribution centered at (0.817, 137) at T_c. Upon heating, the structural integrity of the beta sheet domains are largely preserved and accompanied by small fluctuations in the hydration level. The (N_{hlx}, P_{hlx}) distributions for the helical domains, where N_{hlx} and P_{hlx} are the hydration number and structural persistence of the helical domains, are presented in Figure 5.5. The data show that for Yfh1, a small extent of helical melting, but no enhancement in surface hydration, accompanies the increase in temperature. In addition, the lowering of temperature is not accompanied with any noticeable loss of structural integrity of the helical domains. The helical domains of Ubq behave in a manner similar to the protein's beta sheet domains, with no major unfolding or changes in surface hydration with either an increase of decrease in temperature. The (N_w, P_{wh}), (N_{beta}, P_{beta}) and (N_{hlx}, P_{hlx}) values for Yfh1 and Ubq at all the simulated temperatures are provided in Table 5.1 and 5.2 respectively.

Local characteristics of surface hydration

It has emerged recently that the extent of fluctuation in water density in the vicinity of a surface is a measure of its effective hydrophobicity.[95,99] In our previous work,[69] we have shown that at the cold denaturation temperature, the normalized density fluctuations near the hydrophobic residues[100] of the beta sheet domains in Yfh1 reduced sharply at T_c. In Figure 5.6 (a, b, c, d), we present the normalized water density fluctuations (Eqn. 5.1) corresponding to the hydrophobic residues of the beta sheet ($beta_H$), as well as the helical domains ($helix_H$) at the simulated temperatures for both Yfh1 and Ubq, the 'α' values are tabulated in Table 5.1 and 5.2. For both

domains of Ubq, we find significantly broader $P(n^*)$ distributions at T_h compared to those obtained at T_c and T_s. While the distribution at T_c is narrowest, it does not deviate significantly from the distribution at T_s. This analysis clearly shows that for Ubq, there is a marked enhancement in effective hydrophobicity at T_h compared to T_s, and a very marginal reduction at T_c. Interestingly, we find the same trend of $P(n^*)$ with temperature in the helical domain of Yfh1, but not in this protein's beta sheet domains. In contrast, for this domain in Yfh1, the $P(n^*)$ distribution obtained at T_h is only negligibly broader than that obtained at T_c, while the distribution obtained at T_c is markedly narrower. In Figure 5.6 e, we have compared the water density fluctuations at T_h, T_s and T_c around the non-hydrophobic residues of the beta sheet domains ($beta_{NH}$) of Yfh1; the corresponding 'α' values are provided in Table 5.3. In contrast to the sharp narrowing of the density distribution in the $beta_H$ domain, the density fluctuations around the $beta_{NH}$ domains at T_c are only marginally narrower than those obtained at T_s and T_h. Noting that the beta sheet domains are uniformly unfolded, and with nearly equal solvent accessible surface area per residue in the $beta_H$ and $beta_{NH}$ domains, as reported in our earlier work (*in Suppor. Info.*),[69] this data suggests that the reduced fluctuation at T_c around $beta_H$ arises mainly due to a reduction in the effective hydrophobicity at this temperature.

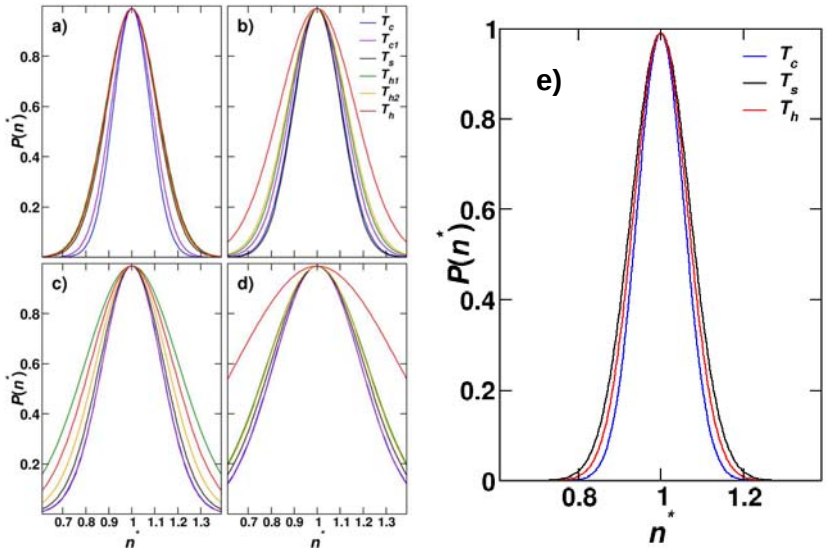


Figure 5.6 Distributions of $P(n^*)$ (obtained from the last 50 ns of simulations) for a) $beta_H$, and b) $helix_H$ of Yfh1, at the simulated temperatures; and c) $beta_H$, and d) $helix_H$ of Ubq, at the simulated temperatures; e) $beta_{NH}$ domains of Yfh1, at temperatures T_c, T_s, T_h.

Table 5.3 α values (10^{-2}) corresponding to the beta$_{NH}$ (α_{bN}) residues, at temperatures T_c, T_s, T_h; for Yfh1.

T (K)	α_{bN}
T_c	0.65
T_s	1.03
T_h	0.85

A significant body of work over the last two decades has already established that the introduction of macromolecular solutes dramatically alters the structural and dynamical behavior of bulk water.[13,40,49,51,71,73,74,76-78,97,98,101-103] Computer simulations have been especially useful in providing fundamental insights into how interfacial water on biomolecular surfaces deviates from its bulk characteristics as a function of local physico-chemical factors.[15,99,104] Furthermore, very recent studies have started to elucidate the unique characteristics of hydrophobic hydration.[57,76,77] Water molecules in the vicinity of hydrophobic domains are characterized by slower (translational and orientational) dynamics.[76,77] Spectroscopic studies indicate corresponding red shift (with decrease in temperature) in the solute-correlation spectra of the OH bonds of the hydrogen bonded hydration water molecules, due to increased water ordering surrounding the hydrophobic solutes.[57] Thus changes in temperature have a strong influence in the intermolecular hydrogen bonding dynamics of water.

It is believed that transient hydrogen bonds formed by water molecules with their neighbors lead to caging effects within the picoseconds time regimes, which inhibits their translational and rotational motions and results in restricted (or 'rattling') motions.[105] These effects, further enhanced in sub-freezing temperature domains and in the presence of flexible macromolecules, can be captured in computer simulations by calculating the velocity autocorrelation functions (VACF) and the corresponding vibrational density of states (VDOS) of hydration water molecules.[96,98] Herein, we have performed these calculations to investigate how the temperature dependent unfolding and effective hydrophobicity of key domains influence the dynamical behavior of the local hydration water molecules.

In Figure 5.7, we first compare the VACF for the oxygen atoms (Eqn. 3) of bulk water with those in the hydration layer of Yfh1 and Ubq at T_c, T_s and T_h; corresponding data for the other temperatures are provided in Figure 5.8. Similar calculations for the hydrogen atoms of the bulk and hydration water that reflect librational modes of water molecules have been reported to effectively remain unchanged with temperature.[96] The decay curves of $C_v^O(t)$ for the bulk water are associated with well-defined sinks at about 0.1 ps, whose depths correspond to the extent of restriction and enhancement in rigid dynamics of the water molecules. For the hydration layer water molecules, the sinks in the decay curves are deeper than bulk water irrespective of temperature, due to substantial back-scattering of water molecules arising from the influence of protein surface. With decrease in temperature, the signature of stronger back scattering is reflected in further deepening of the corresponding minima (Figures 5.7 and 5.8). Interestingly, the decay curves corresponding to the hydration layer of the whole proteins are found to behave similarly relative to bulk water at all the temperatures.

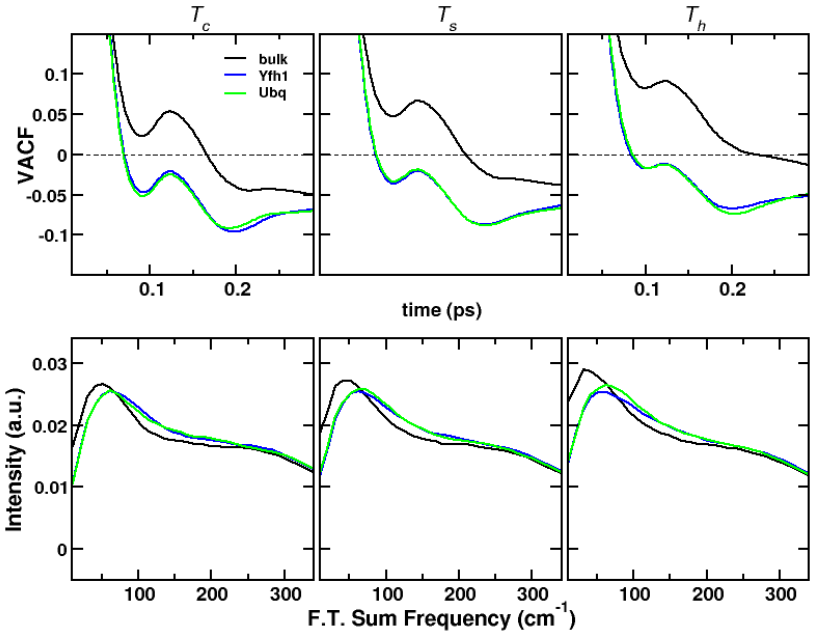


Figure 5.7 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the whole protein domains of Yfh1 and Ubq, compared with the corresponding bulk behavior, at the selected temperatures (T_c, T_s, T_h).

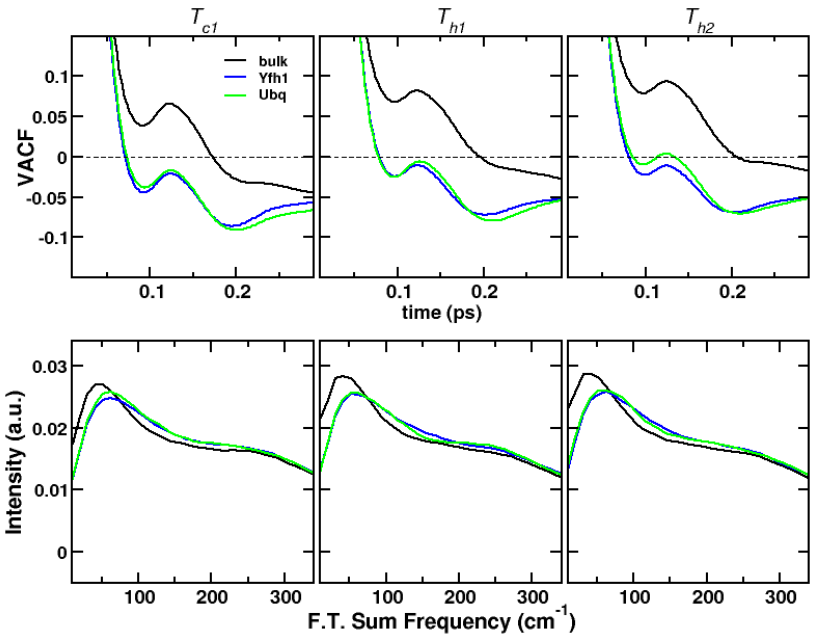


Figure 5.8 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the whole protein domains of Yfh1 and Ubq, compared with the corresponding bulk behavior, at the selected temperatures (T_{c1}, T_{h1}, T_{h2}).

To further elucidate the effect of local structural perturbation on the dynamical behavior of hydration water with changes in temperature, we obtained the plots of $C_v^O(t)$ separately for the different secondary structural domains (beta, beta$_H$, helix, helix$_H$) of Yfh1 (Figure 5.9) and Ubq (Figure 5.11) at T_c, T_s, T_h. The data corresponding to the other temperatures for the two proteins are provided in Figure 5.10 and 5.12 respectively. The peak minima of beta, beta$_H$, helix, helix$_H$ for the two proteins are provided in Table 5.4 and 5.5, respectively. The data for Yfh1 suggests that the corresponding dips of hydration water surrounding beta$_H$ and helix$_H$ are mostly greater than the beta and helical domains. This suggests that water molecules around the selected hydrophobic residues are more rigidified as compared to the corresponding domains in a temperature independent fashion. But while the relative sinks of $C_v^O(t)$ for beta$_H$ of Yfh1 are greater than other domains at temperatures T_s, T_{h1}, T_{h2}, at T_c and T_{c1} they approach the trends of the overall beta, in comparison to that of helix$_H$ (Figures 5.9, 5.10). This is further reflected from the differences in minima of the sinks between beta and beta$_H$ of Yfh1 (Table 5.4), indicative of reduced rattling motions of local hydration water commensurate with the reduced water density at T_c and T_{c1}. The differences in the minima of the corresponding sinks between beta$_H$ and beta

are found to further increase with increase in temperature. Together, these establish reduced restrictions on motions of water molecules due to local unfolding and flexibility of the beta domains at T_c and T_{c1}, commensurate with the decrease in the effective hydrophobicity at these temperatures. Interestingly, the dip of $C_v^O(t)$ for helix_H of Yfh1 is greatest at T_h (Figure 5.9, Table 5.4), showing a greater rigidification of the helices in comparison to the beta sheets, and in accordance with the trends found earlier in the $P(n^*)$ distributions. Also, the distinction between the two minima for helix_H of Yfh1 is nearly lost at T_h. In contrast, while the dip of beta_H for Ubq is slightly greater for most of the temperatures, it is much greater at T_c (Figures 5.11 and 5.12, Table 5.5). Further, in this case, we do not observe any trends in the corresponding sinks for beta_H and helix_H of Ubq with temperature.

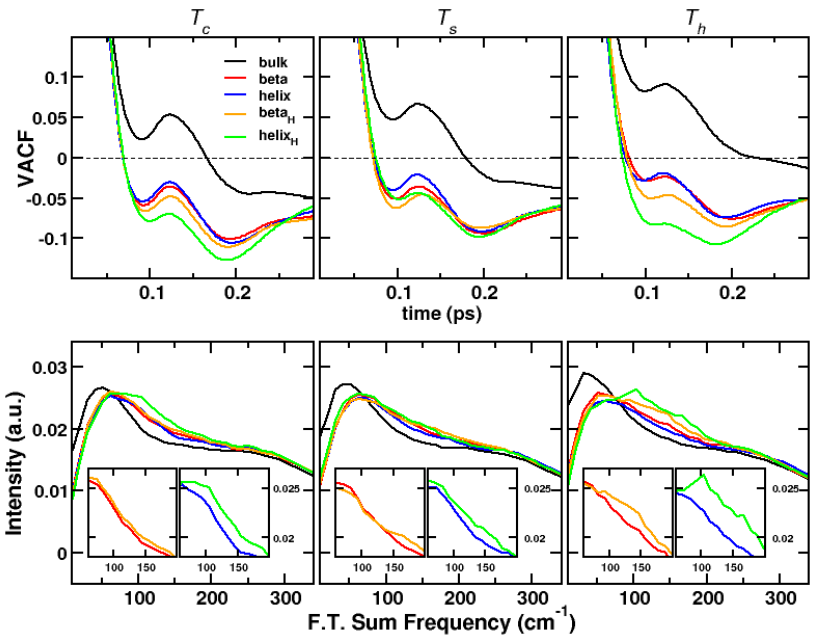


Figure 5.9 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, beta_H, and helix_H domains of Yfh1, compared with the corresponding bulk behavior, at the selected temperatures (T_c, T_s, T_h); differences in VDOS for beta & beta_H and helix & helix_H are represented by the magnified plots in the respective insets.

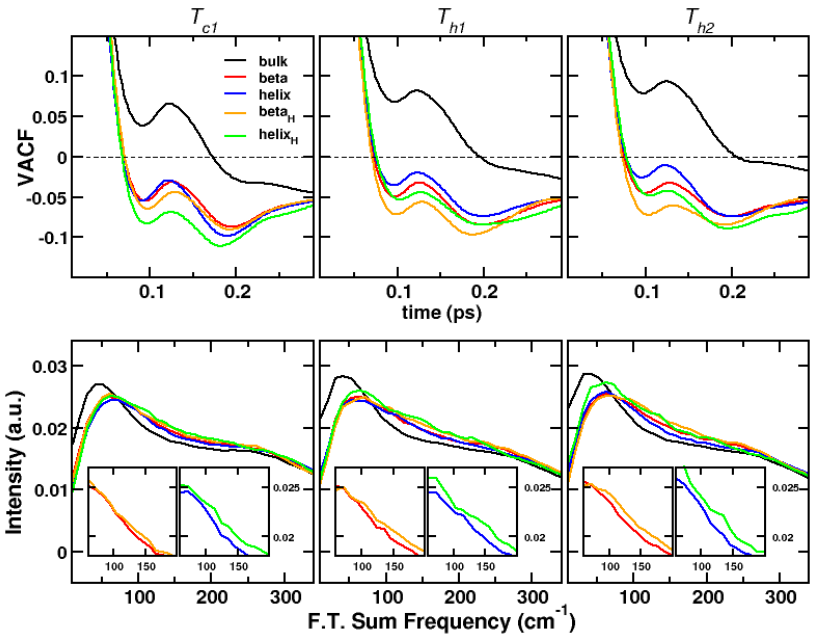


Figure 5.10 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, beta$_H$, and helix$_H$ domains of Yfh1, compared with the corresponding bulk behavior, at the selected temperatures (T_{c1}, T_{h1}, T_{h2}); differences in VDOS for beta & beta$_H$ and helix & helix$_H$ are represented by the magnified plots in the respective insets.

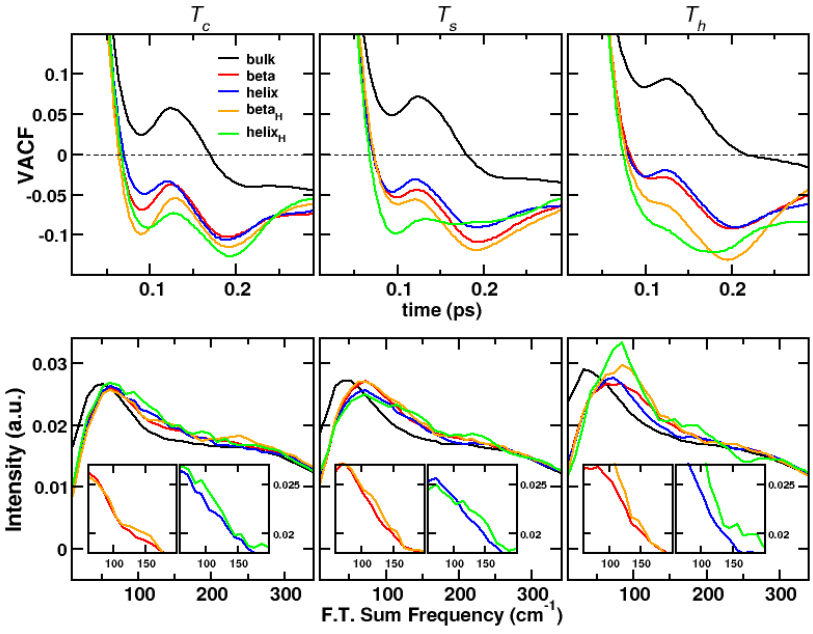


Figure 5.11 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, beta$_H$, and helix$_H$ domains of Ubq, compared with the corresponding bulk behavior, at the selected temperatures (T_c, T_s, T_h); differences in VDOS for beta & beta$_H$ and helix & helix$_H$ are represented by the magnified plots in the respective insets.

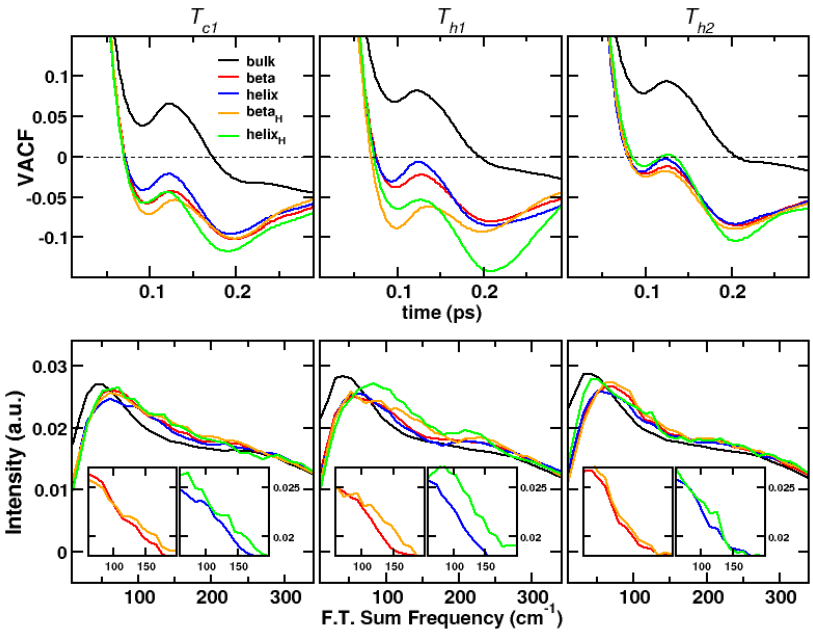


Figure 5.12 Plots of VACF (upper row), and VDOS (lower row) for the hydration water layer of the beta, helix, beta$_H$, and helix$_H$ domains of Ubq, compared with the corresponding bulk behavior, at the selected temperatures (T_{c1}, T_{h1}, T_{h2}); differences in VDOS for beta & beta$_H$ and helix & helix$_H$ are represented by the magnified plots in the respective insets.

Table 5.4 First minima (ps, 10^{-2}) of VACF for the secondary structural domains (beta, beta$_H$, helix, helix$_H$) of Yfh1; as well as respective differences between the ordinates (depth of the minima) of beta & beta$_H$ (Δ_1), and helix & helix$_H$ (Δ_2).

T (K)	beta	beta$_H$	Δ_1	helix	helix$_H$	Δ_2
T_c	(0.094, –5.90)	(0.096, –6.63)	0.73	(0.092, –5.47)	(0.1, –7.87)	2.40
T_{c1}	(0.094, –5.45)	(0.098, –6.46)	1.01	(0.09, –5.46)	(0.096, –8.31)	2.85
T_s	(0.098, –5.10)	(0.098, –6.26)	1.16	(0.094, –4.07)	(0.102, –5.15)	1.08
T_{h1}	(0.096, –4.94)	(0.098, –7.17)	2.23	(0.096, –3.54)	(0.102, –5.30)	1.76
T_{h2}	(0.1, –4.52)	(0.104, –7.22)	2.7	(0.096, –2.61)	(0.106, –4.81)	2.2
T_h	(0.104, –2.83)	(0.108, –5.05)	2.22	(0.098, –2.84)	(0.11, –8.30)	5.46

Table 5.5 First minima (ps, 10^{-2}) of VACF for the secondary structural domains (beta, $beta_H$, helix, $helix_H$) of Ubq; as well as respective differences between the ordinates (depth of the minima) of beta & $beta_H$ (Δ_1), and helix & $helix_H$ (Δ_2).

T (K)	beta	$beta_H$	Δ_1	helix	$helix_H$	Δ_2
T_c	(0.09, –6.89)	(0.09, –9.89)	3.0	(0.094, –4.91)	(0.098, –9.11)	4.20
T_{c1}	(0.098, –5.79)	(0.98, –7.17)	1.38	(0.092, –4.15)	(0.094, –5.75)	1.60
T_s	(0.1, –5.35)	(0.102, –6.16)	0.81	(0.094, –4.65)	(0.098, –9.85)	5.20
T_{h1}	(0.098, –3.81)	(0.098, –8.87)	5.06	(0.094, –3.09)	(0.1, –6.48)	3.39
T_{h2}	(0.1, –2.14)	(0.1, –2.47)	0.33	(0.096, –1.86)	(0.098, –1.18)	–0.68
T_h	(0.108, –2.98)	(0.1, -5.32)	2.34	(0.102, –2.74)	(0.1, –7.74)	–5.0

There are two representative bands obtained from the low-frequency vibrational density of states (VDOS): one at ~50 cm^{-1}, corresponding to the O-O-O bending mode or the transverse oscillations, and the other at ~200 cm^{-1}, representative of the O-O stretching mode or the longitudinal oscillations; of the hydrogen bonded water molecules.[106-110] An increase (or decrease) in the frequency of oscillations will correspond to a blue (or red) shift of the corresponding spectra. It has been argued that the bands at ~50 cm^{-1} do not necessarily occur due to hydrogen bonded water molecules, but mostly due to inhibited transverse oscillations of the water molecules under confinement.[111,112]

In Figures 5.7 and 5.8, we probe the effect of temperature on the bulk and hydration layer water molecules surrounding the whole proteins (Yfh1 and Ubq) and their different domains, by comparing the VDOS from the Fourier transformation of the respective $C_v^O(t)$. While the O-O stretching mode in bulk does not demonstrate marked changes with temperature, the O-O-O bending mode has a gradual blue shift with decrease in temperature due to enhanced rigidity of the water molecules. This is in accordance with studies which show that longitudinal and transverse oscillations of the water molecules are influenced asymmetrically with changes in temperature.[96] Further, the transverse oscillations of water molecules are influenced to a greater extent with decreasing temperature. The O-O-O spectra corresponding to the whole protein hydration layers undergo blue shifts with respect to bulk water, indicating that the strong interaction of hydration water molecules with the protein surface further restricts the motion of water molecules. Notably, however, the extent of blue shift relative to bulk water diminishes with decrease in temperature, plausibly due to the thermal rigidification of both bulk and

hydration layer water. As observed in the $C_v^O(t)$ plots, the VDOS for the hydration layers of the two proteins are nearly indistinguishable from each other at each simulated temperature. However, noticeable differences are found when comparing behavior of the hydration water layer surrounding the secondary structural domains of the two proteins (Figures 5.9, 5.10, 5.11, 5.12). At T_c and T_{c1} (Figures 5.9, 5.10), the bending mode corresponding to hydration waters around beta and helical domains of Yfh1 peak at similar values (~80 cm^{-1}) with comparable intensities. However, closer inspection reveals that at T_c and T_{c1}, the corresponding spectra for $beta_H$ is only slightly blue shifted over the entire beta domain (see *insets*). On the other hand, the spectra for $helix_H$ have broader distribution (peaked at ~100 cm^{-1}) with higher intensity, and they are distinctively blue shifted compared to the helical domains. Thus the differences in the VDOS spectra between beta and $beta_H$ are marginal at T_c and T_{c1}, when compared to that between helix and $helix_H$. These results show in conjunction that the influence of the surface in strengthening the bending mode of water molecules around the apolar residues of the beta sheets decreases noticeably at cold denaturing temperatures (T_c and T_{c1}), in contrast to those of the helices. As a result, the corresponding spectra of $beta_H$ approach those of the beta domains, whereas the spectra of $helix_H$ are blue shifted relative to those of the helical domains. We point out that these changes are concurrent with the sharp decrease in water density fluctuations in the hydration layer surrounding the $beta_H$ domains at the cold denaturing conditions. At most temperatures above T_{c1}, the representative spectra corresponding to the $beta_H$ and $helix_H$ are distinctly blue-shifted when compared to the spectra of the corresponding full domains, signifying rigidification of the transverse oscillations within corresponding hydration layers. Moreover, for the $helix_H$ at at T_h, while the individual peaks at ~80 cm^{-1} and ~200 cm^{-1} are observed to broaden distinctly, we find that overall this spectrum is maximally blue shifted relative to the full helical domain (see Figure 5.9 *inset*). This is commensurate with the distinct broadening of the water density fluctuations and thereby with the increased hydrophobicity. However, for both proteins, the peak corresponding to the trends in the O-O stretching mode of the hydration layers do not have any clear correlation with temperature.

In contrast, the VDOS obtained for the hydration layer water molecules of the domains of Ubq (Figures 5.11, 5.12) demonstrate no distinctive temperature dependent pattern in the O-O-O bending mode. The differences between VDOS corresponding to the beta and $beta_H$ domains,

and to the helix and $helix_H$ domains, are similar across all temperatures. Also at T_h, the peaks of the spectra corresponding to the $beta_H$ and $helix_H$ have higher intensity, and they are considerably blue-shifted over the other domains.

5.4 Conclusion

In this study, using fully atomistic MD simulations, we have correlated changes in the local hydration behavior with temperature in the mesophilic Yfh1 protein, benchmarking the results against those obtained from the thermally stable Ubq protein. The studies spanned the experimentally determined cold (T_c) and heat (T_h) denaturation temperatures of Yfh1. Our results indicate that there is significant variation in the globular compactness and structural propensity of Yfh1 from T_c to T_h, whereas Ubq exhibits a nominal variation. While relative disruption of the structural propensity of the beta sheets of Yfh1 takes place above the temperature of maximal stability (T_s), it is completely lost at cold denaturing temperatures and accompanied by a corresponding increase in hydration propensity; this is not observed for Ubq. The helical domains of both the proteins remain nearly unperturbed within the simulated timescales across all the simulated temperatures. Upon investigating the water density fluctuations around the apolar residues of the beta and helical domains of the proteins, we find evidence of disruption of the overall hydrophobicity in the vicinity of the beta domains of Yfh1 at cold denaturing conditions, and a corresponding strengthening at higher temperatures. In contrast, the beta domains of Ubq retain their effective hydrophobicity at cold denaturing conditions, while the helical domains of both the proteins undergo a strengthening at T_h.

We have probed local dynamical behavior of the hydration water molecules surrounding the protein domains as a function of temperature by calculating the velocity auto-correlation function (VACF) of the water oxygen atoms and the corresponding vibrational density of states (VDOS). The results highlight how the restricted caging effects are altered with temperature, and further, their association with the extent of local hydrophobicity. We find that with decrease in temperature and due to influence of the protein surface, motions of the hydration water molecules surrounding the whole protein domains of Yfh1 and Ubq are eventually restricted compared to the bulk. This is reflected in the deepening of minima in the VACF plots and in the corresponding blue-shifted VDOS spectra for the hydration layer water oxygens over bulk water due to controlled O-O-O bending mode of the water molecules. While these differences between

Yfh1 and Ubq cannot be discerned from the whole protein hydration trends, significant differences in the local movements of water molecules surrounding the different structural domains of these proteins have been established on the basis of the decay curves of VACF and the VDOS spectra. Further, their trends with temperature reflect those found in the water density fluctuations. These analyses show that signatures of protein heat and cold denaturation and associated domain hydrophobicity can be found in the local dynamical behavior of the hydration layer.

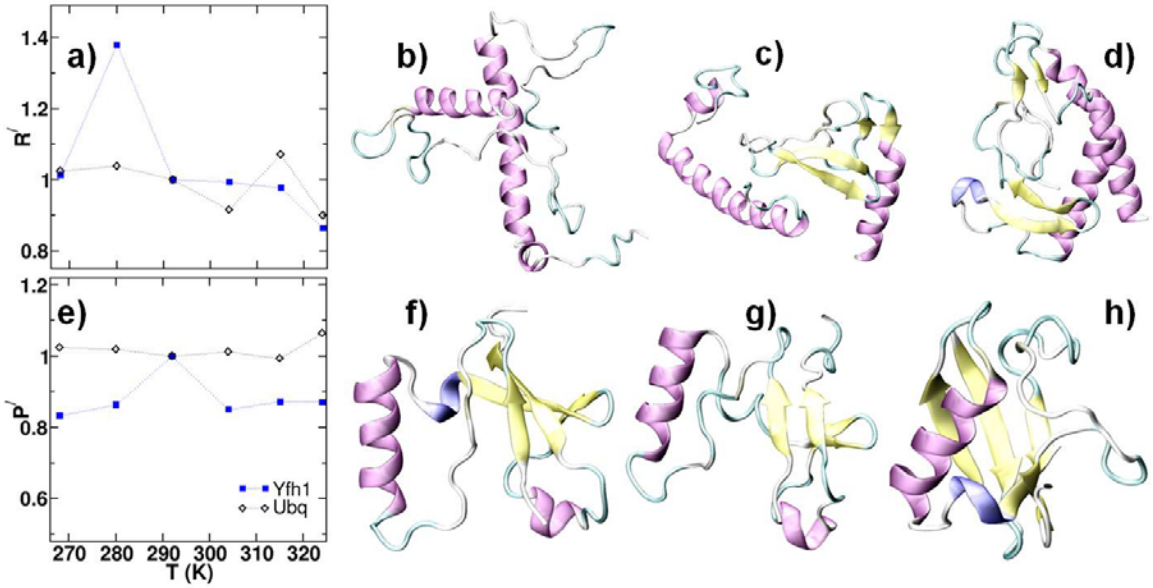


Figure 5.13 Plots of $R^{/}$ and $P^{/}$ for Yfh1 and Ubq as a function of temperature, as well as representative snapshots of Yfh1 and Ubq at selected temperatures, with TIP4P/2005 water model. a) evolution of $R^{/}$ of Yfh1 at all the simulated temperatures; snapshots of Yfh1 at b) T_c, c) T_s, d) T_h; e) evolution of $P^{/}$ of Yfh1 at all the simulated temperatures; snapshots of Ubq at f) T_c, g) T_s, f) T_h.

Before concluding, we point out that temperature dependent protein stabilities could subtly dependent on the force field parameters used. Yet, the thermal denaturation of Yfh1, the qualitative agreement with observations from FRET experiments,[64] and the thermostability of Ubq observed in this study underscore the efficacy of empirical force fields in capturing the essential features of protein thermal stability at regimes not extremely removed from physiological temperatures. However, while three-site water models such as SPC/E and TIP3P have been used in recent studies to study temperature dependent biomolecular phenomena and

associated solvation behavior,[96,113] it is possible that more advanced or polarizable water models may be more effective in capturing these complex phenomena. In this light, we report here that the conformations of Yfh1 and Ubq obtained at the end of our trajectories, when solvated in the TIP4P/2005 water model[92] and simulated for a few ns, yield very similar trends in protein compactness ($R^{/}$) and structure ($P^{/}$); these are reported in Figure 5.13 and Table 5.6. However, it is possible that longer simulations with water models incorporating improved charge and dispersion interactions may yield insightful explanations into phenomena such as the low compaction at the onset of cold denaturation.

Table 5.6 Mean values of R_g obtained from the TIP4P/2005 simulations, for Yfh1 (R_g^{Y}) and Ubq (R_g^{U}).

T (K)	R_g^{Y}	R_g^{U}
T_c	18.21	13.52
T_{c1}	24.78	13.71
T_s	17.96	13.20
T_{h1}	17.85	12.09
T_{h2}	17.56	14.14
T_h	15.52	11.88

5.5 Bibliography

(1) Dill, K. A.; MacCallum, J. L. The Protein-Folding Problem, 50 Years On. *Science*. **2012**, *338*, 1042-1046.

(2) Privalov, P. L. Cold Denaturation of Protein. *Crit. Rev. Biochem. Mol. Biol.* **1990**, *25*, 281-306.

(3) Dobson, C. M. Protein folding and misfolding. *Nature*. **2003**, *426*, 884-890.

(4) Jahn, T. R.; Radford, S. E. The Yin and Yang of protein folding. *FEBS J.* **2005**, *272*, 5962-5970.

(5) Jha, S. K.; Udgaonkar, J. B. Direct evidence for a dry molten globule intermediate during the unfolding of a small protein. *Proc. Natl. Sci. U.S.A.* **2009**, *106*, 12289-12294.

(6) Adinolfi, S.; Nair, M.; Politou, A.; Bayer, E.; Martin, S.; Temussi, P.; Pastore, A. The Factors Governing the Thermal Stability of Frataxin Orthologues: How To Increase a Protein's Stability. *Biochemistry*. **2004**, *43*, 6511-6518.

(7) Falvo, C.; Zhuang, W.; Kim, Y. S.; Axelsen, P. H.; Hochstrasser, R. M.; Mukamel, S. Frequency Distribution of the Amide-I Vibration Sorted by Residues in Amyloid Fibrils Revealed by 2D-IR Measurements and Simulations. *J. Phys. Chem. B.* **2012**, *116*, 3322-3330.

(8) Bagchi, S.; Nebgen, B. T.; Loring, R. F.; Fayer, M. D. Dynamics of a Myoglobin Mutant Enzyme: 2D IR Vibrational Echo Experiments and Simulations. *J. Am. Chem. Soc.* **2010**, *132*, 18367-18376.

(9) Michalet, X.; Weiss, S.; Jäger, M. Single-Molecule Fluorescence Studies of Protein Folding and Conformational Dynamics. *Chem. Rev.* **2006**, *106*, 1785-1813.

(10) Neudecker, P.; Robustelli, P.; Cavalli, A.; Walsh, P.; Lundström, P.; Zarrine-Afsar, A.; Sharpe, S.; Vendruscolo, M.; Kay, L. E. Structure of an Intermediate State in Protein Folding and Aggregation. *Science*. **2012**, *336*, 362-366.

(11) Baiz, Carlos R.; Lin, Y.-S.; Peng, Chunte S.; Beauchamp, Kyle A.; Voelz, Vincent A.; Pande, Vijay S.; Tokmakoff, A. A Molecular Interpretation of 2D IR Protein Folding Experiments with Markov State Models. *Biophys. J.* **2014**, *106*, 1359-1370.

(12) Carrion-Vazquez, M.; Oberhauser, A. F.; Fowler, S. B.; Marszalek, P. E.; Broedel, S. E.; Clarke, J.; Fernandez, J. M. Mechanical and chemical unfolding of a single protein: A comparison. *Proc. Natl. Sci. U.S.A.* **1999**, *96*, 3694-3699.

(13) Martinez, A. V.; Małolepsza, E.; Rivera, E.; Lu, Q.; Straub, J. E. Exploring the role of hydration and confinement in the aggregation of amyloidogenic peptides Aβ16-22 and Sup357-13 in AOT reverse micelles. *J. Chem. Phys.* **2014**, *141*, 22D530.

(14) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. Molecular Simulation of ab Initio Protein Folding for a Millisecond Folder NTL9(1-39). *J. Am. Chem. Soc.* **2010**, *132*, 1526-1528.

(15) Sengupta, N.; Jaud, S.; Tobias, D. J. Hydration Dynamics in a Partially Denatured Ensemble of the Globular Protein Human α-Lactalbumin Investigated with Molecular Dynamics Simulations. *Biophys. J.* **2008**, *95*, 5257-5267.

(16) Sterpone, F.; Bertonati, C.; Briganti, G.; Melchionna, S. Key Role of Proximal Water in Regulating Thermostable Proteins. *J. Phys. Chem. B.* **2009**, *113*, 131-137.

(17) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Accounts Chem. Res.* **2011**, *45*, 83-92.

(18) Liu, F.; Du, D.; Fuller, A. A.; Davoren, J. E.; Wipf, P.; Kelly, J. W.; Gruebele, M. An experimental survey of the transition between two-state and downhill protein folding scenarios. *Proc. Natl. Sci. U.S.A.* **2008**, *105*, 2369-2374.

(19) Cooper, A. Protein Heat Capacity: An Anomaly that Maybe Never Was. *J. Phys. Chem. Lett.* **2010**, *1*, 3298-3304.

(20) Ravłindra, R.; Winter, R. On the Temperature–Pressure Free-Energy Landscape of Proteins. *ChemPhysChem*. **2003**, *4*, 359-365.

(21) Pace, C. N.; Tanford, C. Thermodynamics of the unfolding of *β*-lactoglobulin A in aqueous urea solutions between 5 and 55^{o}. *Biochemistry*. **1968**, *7*, 198-208.

(22) Chen, B. L.; Baase, W. A.; Schellman, J. A. Low-temperature unfolding of a mutant of phage T4 lysozyme. 2. Kinetic investigations. *Biochemistry*. **1989**, *28*, 691-699.

(23) Chen, B. L.; Schellman, J. A. Low-temperature unfolding of a mutant of phage T4 lysozyme. 1. Equilibrium studies. *Biochemistry*. **1989**, *28*, 685-691.

(24) Agashe, V. R.; Udgaonkar, J. B. Thermodynamics of denaturation of barstar: evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry*. **1995**, *34*, 3286-3299.

(25) Anderson, D. E.; Becktel, W. J.; Dahlquist, F. W. pH-Induced denaturation of proteins: a single salt bridge contributes 3-5 kcal/mol to the free energy of folding of T4 lysozyme. *Biochemistry*. **1990**, *29*, 2403-2408.

(26) Martin, S. R.; Esposito, V.; De Los Rios, P.; Pastore, A.; Temussi, P. A. Cold Denaturation of Yeast Frataxin Offers the Clue to Understand the Effect of Alcohols on Protein Stability. *J. Am. Chem. Soc.* **2008**, *130*, 9963-9970.

(27) Davidovic, M.; Mattea, C.; Qvist, J.; Halle, B. Protein Cold Denaturation as Seen From the Solvent. *J. Am. Chem. Soc.* **2008**, *131*, 1025-1036.

(28) Li, Y.; Shan, B.; Raleigh, D. P. The Cold Denatured State Is Compact but Expands at Low Temperatures: Hydrodynamic Properties of the Cold Denatured State of the C-terminal Domain of L9. *J. Mol. Biol.* **2007**, *368*, 256-262.

(29) Hayakawa, I.; Linko, Y.-Y.; Linko, P. Mechanism of High Pressure Denaturation of Proteins. *LWT - Food Sci. Technol*. **1996**, *29*, 756-762.

(30) Bennion, B. J.; Daggett, V. The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Sci. U.S.A.* **2003**, *100*, 5142-5147.

(31) Hawe, A.; Wiggenhorn, M.; van de Weert, M.; Garbe, J. H. O.; Mahler, H.-c.; Jiskoot, W. Forced degradation of therapeutic proteins. *J. Pharm. Sci*. **2012**, *101*, 895-913.

(32) Lavinder, J. J.; Hari, S. B.; Sullivan, B. J.; Magliery, T. J. High-Throughput Thermal Scanning: A General, Rapid Dye-Binding Thermal Shift Screen for Protein Engineering. *J. Am. Chem. Soc.* **2009**, *131*, 3794-3795.

(33) Cauchy, M.; D'Aoust, S.; Dawson, B.; Rode, H.; Hefford, M. A. Thermal Stability: A Means to Assure Tertiary Structure in Therapeutic Proteins. *Biologicals*. **2002**, *30*, 175-185.

(34) Fitzsimons, S. M.; Mulvihill, D. M.; Morris, E. R. Denaturation and aggregation processes in thermal gelation of whey proteins resolved by differential scanning calorimetry. *Food Hydrocoll*. **2007**, *21*, 638-644.

(35) Kolhe, P.; Badkar, A. Protein and solute distribution in drug substance containers during frozen storage and post-thawing: A tool to understand and define freezing–thawing parameters in biotechnology process development. *Biotechnol. Prog*. **2011**, *27*, 494-504.

(36) Tang, X.; Pikal, M. The Effect of Stabilizers and Denaturants on the Cold Denaturation Temperatures of Proteins and Implications for Freeze-Drying. *Pharm. Res*. **2005**, *22*, 1167-1175.

(37) Caldarelli, G.; De Los Rios, P. Cold and Warm Denaturation of Proteins. *J. Biol. Phys*. **2001**, *27*, 229-241.

(38) Griko Yu, V.; Kutyshenko, V. P. Differences in the processes of beta-lactoglobulin cold and heat denaturations. *Biophys. J.* 1994, *67*, 356-363.

(39) Tamura, A.; Kimura, K.; Takahara, H.; Akasaka, K. Cold denaturation and heat denaturation of Streptomyces subtilisin inhibitor: (I). CD and DSC studies. *Biochemistry*. **1991**, *30*, 11307-11313.

(40) Matysiak, S.; Debenedetti, P. G.; Rossky, P. J. Role of Hydrophobic Hydration in Protein Stability: A 3D Water-Explicit Protein Model Exhibiting Cold and Heat Denaturation. *J. Phys. Chem. B*. **2012**, *116*, 8095-8104.

(41) Dias, C. L.; Ala-Nissila, T.; Karttunen, M.; Vattulainen, I.; Grant, M. Microscopic Mechanism for Cold Denaturation. *Phys. Rev. Lett.* **2008**, *100*, 118101.

(42) Romero-Romero, M. L.; Inglés-Prieto, A.; Ibarra-Molero, B.; Sanchez-Ruiz, J. M. Highly Anomalous Energetics of Protein Cold Denaturation Linked to Folding-Unfolding Kinetics. *PLoS ONE*. **2011**, *6*, e23050.

(43) Pometun, M. S.; Peterson, R. W.; Babu, C. R.; Wand, A. J. Cold Denaturation of Encapsulated Ubiquitin. *J. Am. Chem. Soc.* **2006**, *128*, 10652-10653.

(44) Babu, C. R.; Hilser, V. J.; Wand, A. J. Direct access to the cooperative substructure of proteins and the protein ensemble via cold denaturation. *Nat Struct Mol Biol*. **2004**, *11*, 352-357.

(45) Adrover, M.; Esposito, V.; Martorell, G.; Pastore, A.; Temussi, P. A. Understanding Cold Denaturation: The Case Study of Yfh1. *J. Am. Chem. Soc.* **2010**, *132*, 16240-16246.

(46) Luan, B.; Shan, B.; Baiz, C.; Tokmakoff, A.; Raleigh, D. P. Cooperative Cold Denaturation: The Case of the C-Terminal Domain of Ribosomal Protein L9. *Biochemistry*. **2013**, *52*, 2402-2409.

(47) Griko, Y. V.; Privalov, P. L.; Sturtevant, J. M.; Venyaminov, S. Cold denaturation of staphylococcal nuclease. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 3343-3347.

(48) Griko, Y. V.; Privalov, P. L. Calorimetric study of the heat and cold denaturation of *β*-lactoglobulin. *Biochemistry*. **1992**, *31*, 8810-8815.

(49) Adrover, M.; Martorell, G.; Martin, S. R.; Urosev, D.; Konarev, P. V.; Svergun, D. I.; Daura, X.; Temussi, P.; Pastore, A. The Role of Hydration in Protein Stability: Comparison of the Cold and Heat Unfolded States of Yfh1. *J. Mol. Biol.* **2012**, *417*, 413-424.

(50) Lopez, C. F.; Darst, R. K.; Rossky, P. J. Mechanistic Elements of Protein Cold Denaturation *J. Phys. Chem. B*. **2008**, *112*, 5961-5967.

(51) Chandler, D. Interfaces and the driving force of hydrophobic assembly. *Nature*. **2005**, *437*, 640-647.

(52) Agashe, V. R.; Shastry, M. C. R.; Udgaonkar, J. B. Initial hydrophobic collapse in the folding of barstar. *Nature*. **1995**, *377*, 754-757.

(53) Baldwin, R. L. Temperature dependence of the hydrophobic interaction in protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **1986**, *83*, 8069-8072.

(54) Huang, D. M.; Chandler, D. Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 8324-8327.

(55) Southall, N. T.; Dill, K. A.; Haymet, A. D. J. A View of the Hydrophobic Effect. *J. Phys. Chem. B*. **2002**, *106*, 521-533.

(56) Anderson, V. L.; Webb, W. W.; Eliezer, D. Interplay between desolvation and secondary structure in mediating cosolvent and temperature induced alpha-synuclein aggregation. *Phys. Biol.* **2012**, *9*, 056005.

(57) Davis, J. G.; Gierszal, K. P.; Wang, P.; Ben-Amotz, D. Water structural transformation at molecular hydrophobic interfaces. *Nature*. **2012**, *491*, 582-585.

(58) Tsai, C.-J.; Maizel, J. V.; Nussinov, R. The Hydrophobic Effect: A New Insight from Cold Denaturation and a Two-State Water Structure. *Crit. Rev. Biochem. Mol. Biol.* **2002**, *37*, 55-69.

(59) Whitten, S. T.; Kurtz, A. J.; Pometun, M. S.; Wand, A. J.; Hilser, V. J. Revealing the Nature of the Native State Ensemble through Cold Denaturation. *Biochemistry*. **2006**, *45*, 10163-10174.

(60) Ibarra-Molero, B.; Makhatadze, G. I.; Sanchez-Ruiz, J. M. Cold denaturation of ubiquitin. *BBA - Protein Struct. M*. **1999**, *1429*, 384-390.

(61) Kitahara, R.; Okuno, A.; Kato, M.; Taniguchi, Y.; Yokoyama, S.; Akasaka, K. Cold denaturation of ubiquitin at high pressure. *Magnetic Reson. Chem*. **2006**, *44*, S108-S113.

(62) Pastore, A.; Martin, S. R.; Politou, A.; Kondapalli, K. C.; Stemmler, T.; Temussi, P. A. Unbiased Cold Denaturation: Low- and High-Temperature Unfolding of Yeast Frataxin under Physiological Conditions. *J. Am. Chem. Soc.* **2007**, *129*, 5374-5375.

(63) Sanfelice, D.; Puglisi, R.; Martin, S. R.; Di Bari, L.; Pastore, A.; Temussi, P. A. Yeast Frataxin Is Stabilized by Low Salt Concentrations: Cold Denaturation Disentangles Ionic Strength Effects from Specific Interactions. *PLoS ONE*. **2014**, *9*, e95801.

(64) Aznauryan, M.; Nettels, D.; Holla, A.; Hofmann, H.; Schuler, B. Single-Molecule Spectroscopy of Cold Denaturation and the Temperature-Induced Collapse of Unfolded Proteins. *J. Am. Chem. Soc.* **2013**, *135*, 14040-14043.

(65) Ascolese, E.; Graziano, G. On the cold denaturation of globular proteins. *Chem. Phys. Lett.* **2008**, *467*, 150-153.

(66) Amano, K.-i.; Yoshidome, T.; Harano, Y.; Oda, K.; Kinoshita, M. Theoretical analysis on thermal stability of a protein focused on the water entropy. *Chem. Phys. Lett.* **2009**, *474*, 190-194.

(67) Oshima, H.; Yoshidome, T.; Amano, K.-i.; Kinoshita, M. A theoretical analysis on characteristics of protein structures induced by cold denaturation. *J. Chem. Phys.* **2009**, *131*, 205102-205111.

(68) Sanfelice, D.; Temussi, P. A. Cold denaturation as a tool to measure protein stability. *Biophys. Chem*. **2015**. DOI: 10.1016/j.bpc.2015.05.007.

(69) Chatterjee, P.; Bagchi, S.; Sengupta, N. The non-uniform early structural response of globular proteins to cold denaturing conditions: A case study with Yfh1. *J. Chem. Phys.* **2014**, *141*, 205103.

(70) Graziano, G. On the molecular origin of cold denaturation of globular proteins. *Phys. Chem. Chem. Phys.* **2010**, *12*, 14245-14252.

(71) Nandi, N.; Bhattacharyya, K.; Bagchi, B. Dielectric Relaxation and Solvation Dynamics of Water in Complex Chemical and Biological Systems. *Chem. Rev.* **2000**, *100*, 2013-2046.

(72) Pethig, R. Protein-Water Interactions Determined by Dielectric Methods. *Annu. Rev. Phys. Chem*. **1992**, *43*, 177-205.

(73) Marchi, M.; Sterpone, F.; Ceccarelli, M. Water Rotational Relaxation and Diffusion in Hydrated Lysozyme. *J. Am. Chem. Soc.* **2002**, *124*, 6787-6791.

(74) Pal, S.; Bandyopadhyay, S. Importance of Protein Conformational Motions and Electrostatic Anchoring Sites on the Dynamics and Hydrogen Bond Properties of Hydration Water. *Langmuir*. **2013**, *29*, 1162-1173.

(75) Pal, S. K.; Peon, J.; Bagchi, B.; Zewail, A. H. Biological Water: Femtosecond Dynamics of Macromolecular Hydration. *J. Phys. Chem. B.* **2002**, *106*, 12376-12395.

(76) Galamba, N. Water's Structure around Hydrophobic Solutes and the Iceberg Model. *J. Phys. Chem. B.* **2013**, *117*, 2153-2159.

(77) Galamba, N. Water Tetrahedrons, Hydrogen-Bond Dynamics, and the Orientational Mobility of Water around Hydrophobic Solutes. *J. Phys. Chem. B.* **2014**, *118*, 4169-4176.

(78) Das, P.; Matysiak, S. Direct Characterization of Hydrophobic Hydration during Cold and Pressure Denaturation. *J. Phys. Chem. B*. **2012**, *116*, 5342-5348.

(79) He, Y.; Alam, S. L.; Proteasa, S. V.; Zhang, Y.; Lesuisse, E.; Dancis, A.; Stemmler, T. L. Yeast Frataxin Solution Structure, Iron Binding, and Ferrochelatase Interaction. *Biochemistry*. **2004**, *43*, 16254-16262.

(80) Heinig, M.; Frishman, D. STRIDE: a web server for secondary structure assignment from known atomic coordinates of proteins. *Nucleic Acids Res*. **2004**, *32*, W500-W502.

(81) Frishman, D. A., Patrick Knowledge-Based Protein Secondary Structure Assignment. *Proteins.* **1995**, *23*, 566-579.

(82) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics*. **1996**, *14*, 33-38.

(83) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235-242.

(84) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(85) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau,

F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* **1998**, *102*, 3586-3616.

(86) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem*. **2004**, *25*, 1400-1415.

(87) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys*. **1999**, *151*, 283-312.

(88) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.

(89) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613-4621.

(90) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys*. **1977**, *23*, 327-341.

(91) Vijay-Kumar, S.; Bugg, C. E.; Cook, W. J. Structure of ubiquitin refined at 1.8Ã…resolution. *J. Mol. Biol.* **1987**, *194*, 531-544.

(92) Abascal, J. L. F.; Vega, C. A general purpose model for the condensed phases of water: TIP4P/2005. *J. Chem. Phys.* **2005**, *123*, 234505.

(93) Chatterjee, P.; Sengupta, N. Effect of the A30P mutation on the structural dynamics of micelle-bound αSynuclein released in water: a molecular dynamics study. *Eur. Biophys. J.* **2012**, *41*, 483-489.

(94) Jose, J.; Sengupta, N. Molecular dynamics simulation studies of the structural response of an isolated $A\beta_{1\text{-}42}$ monomer localized in the vicinity of the hydrophilic TiO2 surface. *Eur. Biophys. J.* **2013**, 1-8.

(95) Patel, A. J.; Varilly, P.; Jamadagni, S. N.; Hagan, M. F.; Chandler, D.; Garde, S. Sitting at the Edge: How Biomolecules use Hydrophobicity to Tune Their Interactions and Function. *J. Phys. Chem. B.* **2012**, *116*, 2498-2503.

(96) Chakraborty, K.; Bandyopadhyay, S. Effect of temperature on the low-frequency vibrational spectrum and relative structuring of hydration water around a single-stranded DNA. *J. Chem. Phys.* **2015**, *142*, 015101.

(97) Choudhury, N.; Pettitt, B. M. Dynamics of Water Trapped between Hydrophobic Solutes. *J. Phys. Chem. B.* **2005**, *109*, 6422-6429.

(98) Pal, S.; Balasubramanian, S.; Bagchi, B. Dynamics of bound and free water in an aqueous micellar solution: Analysis of the lifetime and vibrational frequencies of hydrogen bonds at a complex interface. *Physical Rev. E.* **2003**, *67*, 061502.

(99) Jose, J. C.; Khatua, P.; Bansal, N.; Sengupta, N.; Bandyopadhyay, S. Microscopic Hydration Properties of the $A\beta_{1-42}$ Peptide Monomer and the Globular Protein Ubiquitin: A Comparative Molecular Dynamics Study. *J. Phys. Chem. B.* **2014**, *118*, 11591-11604.

(100) Moon, C. P.; Fleming, K. G. Side-chain hydrophobicity scale derived from transmembrane protein folding into lipid bilayers. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 10174-10177.

(101) Bizzarri, A. R.; Cannistraro, S. Molecular Dynamics of Water at the Protein-Solvent Interface. *J. Phys. Chem. B.* **2002**, *106*, 6617-6633.

(102) Tarek, M.; Tobias, D. J. Role of Protein-Water Hydrogen Bond Dynamics in the Protein Dynamical Transition. *Phys. Rev. Lett.* **2002**, *88*, 138101.

(103) Tarek, M.; Tobias, D. J. Single-Particle and Collective Dynamics of Protein Hydration Water: A Molecular Dynamics Study. *Phys. Rev. Lett.* **2002**, *89*, 275501.

(104) Sinha, S. K.; Chakraborty, S.; Bandyopadhyay, S. Thickness of the Hydration Layer of a Protein from Molecular Dynamics Simulation. *J. Phys. Chem. B.* **2008**, *112*, 8203-8209.

(105) Larini, L.; Ottochian, A.; De Michele, C.; Leporini, D. Universal scaling between structural relaxation and vibrational dynamics in glass-forming liquids and polymers. *Nat Phys*. **2008**, *4*, 42-45.

(106) Walrafen, G. E.; Chu, Y. C. Linearity between Structural Correlation Length and Correlated-Proton Raman Intensity from Amorphous Ice and Supercooled Water up to Dense Supercritical Steam. *J. Phys. Chem.* **1995**, *99*, 11225-11229.

(107) Walrafen, G. E.; Chu, Y. C.; Piermarini, G. J. Low-Frequency Raman Scattering from Water at High Pressures and High Temperatures. *J. Phys. Chem.* **1996**, *100*, 10363-10372.

(108) Brubach, J. B.; Mermet, A.; Filabozzi, A.; Gerschel, A.; Roy, P. Signatures of the hydrogen bonding in the infrared bands of water. *J. Chem. Phys.* **2005**, *122*, 184509.

(109) Crupi, V.; Dianoux, A. J.; Majolino, D.; Migliardo, P.; Venuti, V. Dynamical response of liquid water in confined geometry by laser and neutron spectroscopies. *Phys. Chem. Chem. Phys.* **2002**, *4*, 2768-2773.

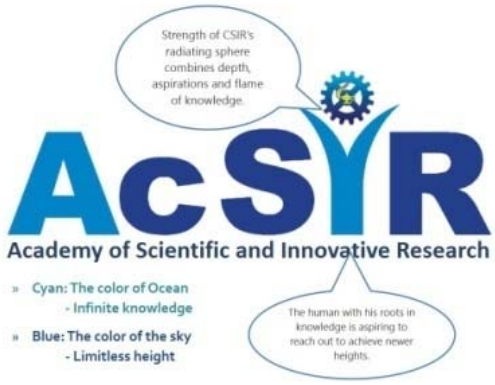
(110) Li, J. Inelastic neutron scattering studies of hydrogen bonding in ices. *J. Chem. Phys.* **1996**, *105*, 6733-6755.

(111) Martí, J.; Padro, J. A.; Guàrdia, E. Molecular dynamics simulation of liquid water along the coexistence curve: Hydrogen bonds and vibrational spectra. *J. Chem. Phys.* **1996**, *105*, 639-649.

(112) Padró, J. À.; Martí , J. An interpretation of the low-frequency spectrum of liquid water. *J. Chem. Phys.* **2003**, *118*, 452-453.

(113) Kuffel, A.; Czapiewski, D.; Zielkiewicz, J. Unusual dynamic properties of water near the ice-binding plane of hyperactive antifreeze protein. *J. Chem. Phys.* **2015**, *143*, 135102.

Chapter 6
Future perspectives



The important thing is not to stop questioning. Curiosity has its own reason for existing.
- Albert Einstein

6.1 Overview

Investigations on the nucleation mechanism of amyloidogenic proteins are critically important in elucidating and controlling the oligomeric tendencies responsible for cytotoxicity in physiological systems.[1,2] Although it has been argued that both pre-fibrillar oligomeric species and insoluble fibrillar aggregates are responsible for disruption of cellular membranes,[2] recent studies suggest that the former plays a major role in introducing cell toxic species and disease propensity.[3,4] Several studies, experimental as well as computational, have been employed in order to study and control the oligomer formation in various amyloidogenic systems.[5-7] Recent studies from our group have focused on the regulation of early mechanistic pathways to oligomer formation in the proximity of non-biological surfaces.[8,9] Additionally, retardation of the nucleation mechanism in hybrid nanochannel-oligomeric systems and further changes due to applying electric field have been carried out.[10,11] However, most of these studies were undertaken with oligomeric systems of amyloidβ (Aβ) peptides, the protein responsible for Alzheimer's that dimerizes rapidly in physiological conditions.[12-14] On the contrary, αSynuclein (αS), the protein responsible for Parkinson's disease, Lewy body dementia and other neurodegenerative diseases, is found to oligomerize with a longer lag phase with a greater number of monomeric units.[5,15-20] Interestingly, the domains of Aβ are more hydrophobic in nature in contrast to αS, which has non-conserved polar amino acids in its sequence. In ongoing studies, we attempt to elucidate the early aspects of oligomer formation of αS protein, and further address the basic differences in the nucleation mechanism for the two proteins.

As discussed previously in *Chapter 1*, hydrophobic channels designed with carbon nanomaterials, are relevant vis-à-vis investigations of solvent transport properties inside transmembrane ion-channels, chemical separators, and nanomembranes.[21-23] Although carbon nanotubes (CNTs) are cytotoxic in nature, those functionalized by introducing special chemical groups (like hydroxyl, ester or amino linkages) may have wide ranging applications in nanotechnology-assisted modern medicinal techniques.[24-28] Studies in our group show that CNTs are capable of regulating the conformational propensities of amyloids.[8,10] Interestingly, water also plays an important role in protein-CNT interactions, with dewetting of apolar groups during adsorption on CNT surfaces, a process similar in nature to hydrophobic collapse and assembly during early protein folding and aggregation.[29,30] In the present context, as well as owing to diverse applications of nano-confined aqueous properties, we attempt to look into the interesting

dynamical and structural properties of water in presence of nano-channels and with the introduction of co-solvents of lower polarity. Insights from these studies are expected to be useful in understanding the structural and dynamical behavior of water under nano-confinement.

6.2 Investigations of the early oligomerization propensity of the wild type and a point mutant αSynuclein: a molecular dynamics study

As previously discussed in *Chapter 3*, Parkinson's Disease (PD) is a neurodegenerative disorder characterized by aggregation and subsequent amyloid-fibril formation of misfolded forms of the αSynuclein (αS) protein.[20,31] The N-terminus region of αS is helix rich with repetition of the non-conserved residue sequence, KTKEGV. The central hydrophobic region (residues 61-95) includes the non-amyloidβ component (NAC) of the peptide aggregates associated with Alzheimer's Disease (AD).[32] The tail region of αS protein is proline and lysine rich and thus acidic in nature, and with no definite secondary structure. The N-terminal region of the protein binds to cell membranes while the C-terminal ends remain free and disordered.[33] The micelle bound αS shows greater spontaneous aggregation compared to the free disordered form.[34] It was also observed that changes in structural conformation from unfolded to partially folded states can be induced with either decrease in pH or increase in temperature.[35] The aggregation of αS is nucleation dependent.[36] The mechanism involves formation of water soluble oligomers that are intermediates for fibril formation and are reported to be cytotoxic.[37-39] Several studies revealed that prefibrillar oligomers and protofibrils of αS are the main pathogenic agents in these diseases.[15,40,41] These oligomers have been marked for the disruption of lipids with strong affinity for αS, while insoluble aggregates lead to complete lipid damage.[42]

Naturally occurring point mutant variants of αS, namely the A30P, A53T, E46K, and the G51D variants, alter the aggregation and membrane-binding propensity.[43-46] While the A53T and E46K forms increased fibrillation propensity, the A30P form enhances the formation of prefibrillar species.[47-49] Further, there is reduction in helical propensity as well as elongation in the fibrillation lag phase for A30P.[50,51] Our previous MD studies have showed differences in the early conformational propensity of the wild type (WT) and A30P variant of αS in explicit solvent.[52]

Recent studies report that endogenous αS are helically folded tetramers under non-denaturing conditions, and in the absence of lipid bilayers or micelles.[17,53] Further, computational studies predict that pre-formed oligomeric αS units with prevalent helical conformation are

capable of higher assemblies.[6,7] However, as of yet, the mechanisms via which αS nucleates into the smallest possible oligomeric units remain unknown. Using atomistic MD simulations, we have attempted to study the underlining factors triggering the formation of small oligomers, for the WT and A30P variant of the αS monomer. Our preliminary results suggest that irrespective of the mutation variant, the smallest αS oligomer is a trimer.

System setup and MD simulations

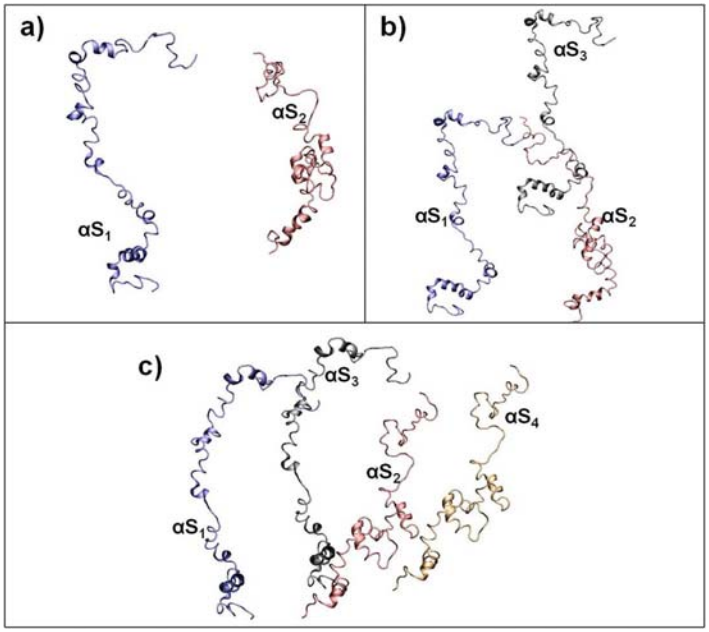


Figure 6.1 Initial setup, in implicit solvent, of the a) dimer (d_N), b) trimer (tr_N), and c) tetramer (tt_N) systems of WT αS. Partially helical monomer conformations are generated from Accelerated MD (AMD).

The solution NMR micellar pdb structure 2KKW[54] was used as the starting structure for our simulation. It was energy minimized with conjugate gradient optimization protocol in the gaseous phase, followed by accelerated MD (AMD) simulations[55] for fifteen nanoseconds. With AMD simulations, partially folded conformers were created and two of the most suitable conformations (based on ^{13}C J-Coupling parameters and their partial helical propensity)[6,7,56] were selected for setting up of the oligomeric systems (dimer, trimer, and tetramer). In each of the systems, the monomers were initially placed in a way that each monomer is at least 70 Å apart from each other (Figure 6.1). Thus monomers in the trimeric system were at the apex of an equilateral triangle; monomeric units of tetramers were on each corners of a square. The same

procedure is followed for both WT (i.e. the native) and A30P (i.e. the mutant) variants. Four independent systems of dimer, trimer and tetrameric systems of WT (dm_N, tr_N, tt_N) and the mutant A30P (dm_M, tr_M, tt_M) were setup by tweaking the relative orientation of the monomers. For convenience, the monomeric units are named as αS_1 and αS_2 (for dimers); αS_1, αS_2 and αS_3 (for trimers); and αS_1, αS_2, αS_3 and αS_4 (for tetramers) respectively. Each of these systems were run for 300 ns with atomistic MD simulations using the CHARMM force field (with CMAP corrections)[57,58] for the protein, in Generalized Born implicit solvent (GBIS) model.[59,60] The simulations were performed at 310 K, with 2 fs time step and configuration data saved every 2 ps leading to a total production run of 1.2 μs for each of these oligomeric systems. The dielectric constant of the continuum water model was kept to be 74.2, the dielectric constant of water at the simulated temperature. The cut-off for non-bonded interactions was kept till 16 Å, with the corresponding smoothing starting at 15 Å. Then for each of these systems, Cartesian principal component analysis (or PCA; see *Chapter 2*),[61,62] were performed on the cumulative data sets, to differentiate the most significant clusters based on their population, from the heterogeneous conformational ensemble. The projections of the two ensembles on the first two principal components are shown in Figure 6.2. The projection yielded 2 clusters (C1 and C2) in d_N, and 3 clusters (C1, C2, C3) in each of d_M, tr_M, tr_N, tt_N, and tt_M.

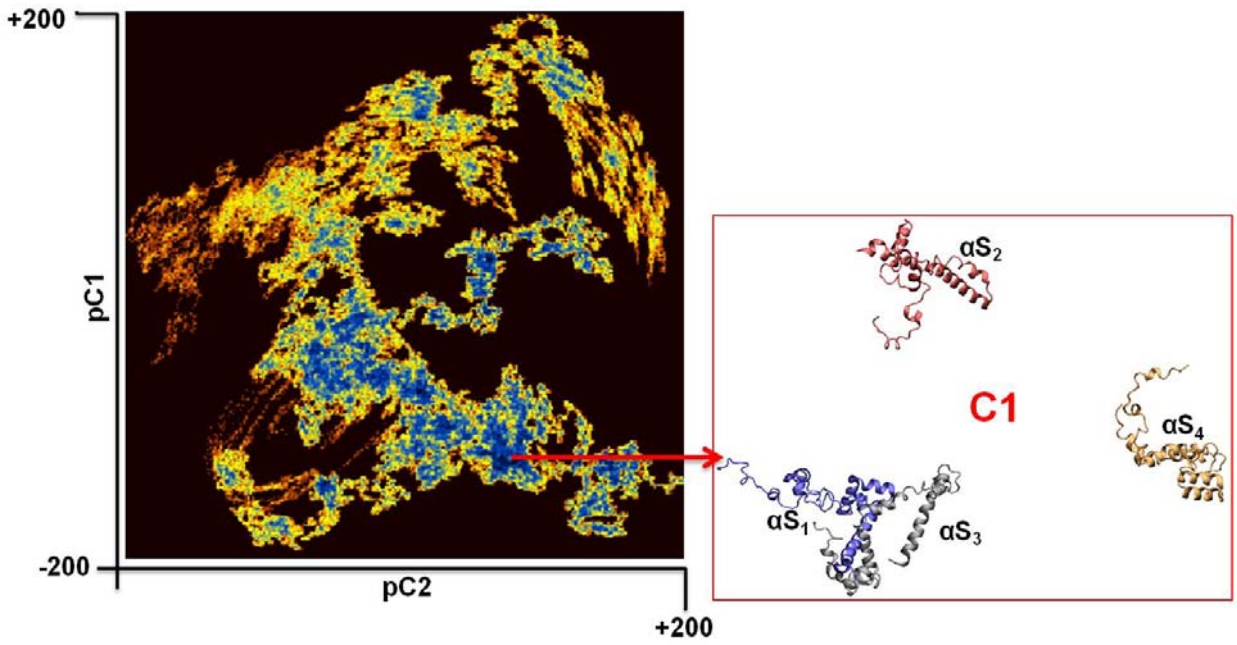


Figure 6.2 Evolution of the clusters from the cartesian PCA for the tt_N system. Two-dimensional distribution of the density function ΔG, on the first two principal components pC1 and pC2, has been illustrated; the cluster corresponding to the greatest population (C1) is additionally presented.

Separation between the monomers

Table 6.1 Mean center of mass distances (comd, in Å) between the monomeric units of the prominent clusters of tr_N oligomeric systems.

Clusters	monomers (tr_N)	αS_2	αS_3
C1	αS_1	268.81	274.57
	αS_2	-	37.39
C2	αS_1	120.45	127.25
	αS_2	-	61.32
C3	αS_1	167.03	222.06
	αS_2	-	90.33

Table 6.2 Mean center of mass distances (comd, in Å) between the monomeric units of the prominent clusters of tr_M oligomeric systems.

Clusters	monomers (tr_M)	αS_2	αS_3
C1	αS_1	73.12	70.90
	αS_2	-	33.20
C2	αS_1	187.08	93.13
	αS_2	-	116.21
C3	αS_1	149.07	27.55
	αS_2	-	131.43

Table 6.3 Mean center of mass distances (comd, in Å) between the monomeric units of the prominent clusters of tt_N oligomeric systems.

Clusters	monomers (tt_N)	αS_2	αS_3	αS_4
C1	αS_1	114.42	41.31	200.58
	αS_2	-	122.20	140.11
	αS_3	-	-	202.33
C2	αS_1	89.40	91.22	131.27
	αS_2	-	106.11	123.91
	αS_3	-	-	124.0
C3	αS_1	109.14	76.13	109.14
	αS_2	-	64.97	89.19
	αS_3	-	-	40.82

Table 6.4 Mean center of mass distances (comd, in Å) between the monomeric units of the prominent clusters of tt_M oligomeric systems.

Clusters	monomers (tt_M)	αS_2	αS_3	αS_4
C1	αS_1	155.01	188.08	172.99
	αS_2	-	162.40	106.17
	αS_3	-	-	156.89
C2	αS_1	122.72	89.95	100.50
	αS_2	-	150.61	72.23
	αS_3	-	-	106.04
C3	αS_1	198.64	56.77	73.99
	αS_2	-	169.88	146.14
	αS_3	-	-	32.15

As simulations were started with at least 70 Å inter-monomer distances, the most primary analysis to detect the oligomerization is by calculating the inter-protein centre of mass distance (comd). The comd was calculated for each frame of the clusters. For the d_N systems, the average comd is 104 and 220 Å for C1 and C2 respectively depicting repulsive interaction between the monomers. Similarly comd was calculated for the three clusters of d_M. The average distances are 125, 170 and 128 Å, emphasizing that monomers moved away from each other in dimeric clusters. On the other hand, when the number of monomers exceeded 2 in the native state, we observed propensity for oligomerization, as signified by an overall reduction in the inter-monomer distance. For tr_N, 2 of the 3 clusters formed oligomers. For C1, the average distance between αS_2 and αS_3 was 37 Å (Table 6.1). For the remaining pairs (αS_1-αS_2 and αS_1-αS_3), the average comd distances were 268 and 274 Å respectively, indicating they have negligible propensity of dimer formation. A similar observation prevailed in the case of C2, where αS_2 and αS_3 formed dimers with an average distance of 61 Å; αS_1 propagated to a larger distance from αS_2 and αS_3, the mean distances of αS_1 from αS_2 and αS_3 are respectively 120 and 127 Å. Similarly, two of the three clusters (C1 and C3) of tr_M exhibited significant probability of dimer formation (Table 6.2). C1 displayed the dimer formation of αS_2 and αS_3, with an average comd distance of 33 Å. C2 didn't show any oligomerization propensity. C3 exhibited dimerization between αS_2 and αS_3 (mean comd of 28 Å). In the C1 of tt_N (Table 6.3), αS_1 and αS_3 formed dimers, the average distance between them to be 41 Å. This indicates that both the monomers have approached each other from an initial separation distance of ~100 Å, as they were originally diagonally opposite to each other (put in each corners of a square with the length of each side to be 70 Å). The αS_2 and αS_4, and αS_1 and αS_3 remained apart from each other. While C2 did not show any oligomer formation, C3 again exhibited oligomerization propensity between αS_2, αS_3, and αS_4 with average distances between αS_3-αS_4 and αS_2-αS_3 to be 40 and 65 Å respectively. For the tt_M system (Table 6.4), only C3 shows oligomerization between αS_1, αS_3 and αS_4. Interestingly, both the variants of tetrameric systems have maximum propensity of forming oligomers. On the contrary trimeric systems are limited to only form dimers. Thus, propensity of formation of oligomers comprising 'n' monomers is only possible with an initial system of discrete 'n+1' monomers.

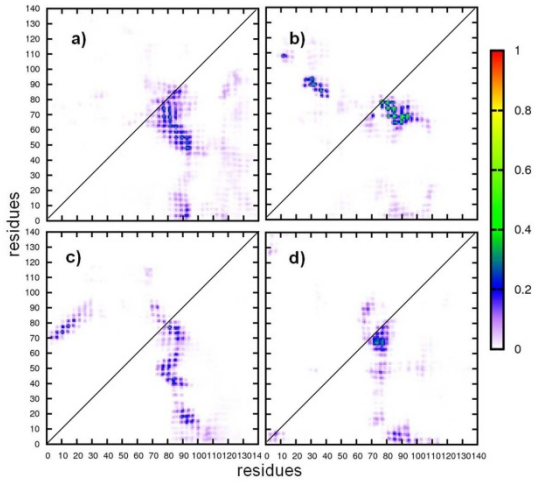


Figure 6.3 Selected plots of contact maps between the monomers for a) C1 of tr_M, b) C3 of tr_M, c) C1 of tt_N, d) C3 of tt_N.

In addition, the contact number was calculated to determine the proximity of each pair of oligomers (Figure 6.3). Two residues were considered to be in contact if the centers of mass of their sidechain or backbone were found within 8 Å between two monomers in the cluster. The upper triangle in the figures depicts backbone-backbone contacts while the lower triangle shows the sidechain-sidechain contacts. The clusters C1 of tr_N, C1 and C3 of tr_M, C1 and C3 of tt_N, and C3 of tt_M, displays significant contacts between the pair of monomers prone to form oligomers. Interestingly, it is observed that they have a common patch with high probability of contacts in the region consisting of residues 60 to 100, the region corresponding to the location of central hydrophobic core (NAC). This shows that NAC of αS plays an important role in the oligomerization pathway of the αS in both the variants.

Structural propensities: preliminary insights

We have verified that the interactions that bring monomeric units close to each other in the oligomeric systems discussed earlier are dominated by electrostatics (data not shown). This is plausibly due to the presence of abundant polar amino acid groups in the sequence of αS. In the same light, it would be interesting to find whether the extent of structural disorder for the interacting monomers is altered due to oligomer formation. As a preliminary analysis, we have estimated the cumulative configurational entropy (Figure 6.4), for each αS monomer present in the respective clusters using Schlitter's method (*Chapter 4*). To calculate the configurational entropy, we have performed 4 ns simulations of a representative conformer of each cluster in

implicit solvent. For the clusters having interacting monomers, marked decrease in configurational entropy is expected. From the comd as well as contact maps, it is evident that clusters C1 and C2 of tr_N, C1 and C3 of tr_M and tt_N, and C3 of tt_M have high propensity of oligomerization.

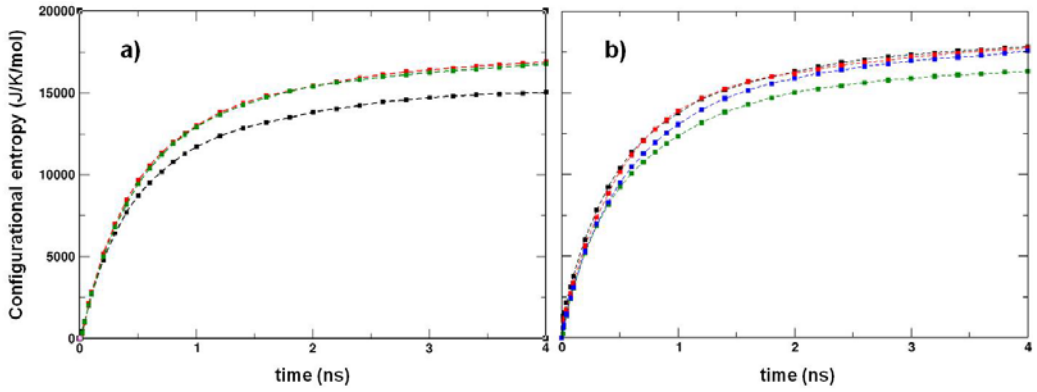


Figure 6.4 Selected plots of cumulative configurational entropy of the monomeric units (αS_1, αS_2, αS_3, and αS_4; denoted in *black*, *red*, *green* and *blue*) in a) C1 of tr_M, and b) C3 of tt_N.

In C1 of tr_N, there is a marked decrease in configurational entropy of αS_1, while entropy of αS_3 is similar to that of the free monomer. The C3 of tt_N shows a decrease in entropy for αS_3 while interacting with αS_4, promoting dimer formation. Similarly, monomeric units of the clusters prone to form oligomers exhibit reduced entropy, in qualitative agreement with previous results.

These results discussed here provide insights into the early propensity for oligomerization of the full-length αS. The structural propensity of monomers in each of these clusters shall be investigated further in order to elucidate the influence of secondary structure in oligomer formation. In addition, it is interesting to note that helical propensity is reinstated in the monomers with the progress of the simulations, although their starting conformations resembled partially folded helical conformations (Figure 6.1, 6.2). Further, the reason behind dimer formation with more than 2 monomeric units, as well as the role of the unbound monomer in the early nucleation will also be investigated. Amongst other things, we are currently investigating plausibility of monomeric exchange between free and oligomer-bound units, and the detailed thermodynamics of αS nucleation; both these aspects involve atomistic MD sampling on timescales exceeding several microseconds.

6.3 Effect of hexafluoro-2-propanol on the dynamics and structure of water within nano-channels: a molecular dynamics study

As discussed previously in *Chapter 1*, water, the key component in biological systems, is a prerequisite for life,[63] and transport of water through nano-channels has an essential role in biological activities.[64-68] Further nano-confined aqueous transport have relevant applications in mimicking biological channels, on-off gates, nano-scale electronic devices, chemical separators, etc. The flux and occupancy of water molecules have been found to depend on dimensions and deformations of the nanometer pores as well as movable external charges, showing an on-off gating mechanism and a wave-like water distribution similar to biological water channels.[69-71] Gating mechanisms, associated with the presence of dipolar molecules, has been found to further vary with separation between the corresponding molecules.[72] Moreover, transitions from empty to filled states are dependent on the mobility of single charges within the symmetrical single-walled carbon nanotube (SWNT).[73] Water permeation further depends on corresponding structure of the outer surface of nano-channels.[74] Also, the mobility of water-associated biomolecules within SWNT can be manipulated by the movement of external charge(s) along the channel length.[75] Additionally, it has been shown that double-walled carbon nanotubes (DWNT) can act like artificial trans-membrane water channels,[76] and increased hydrophilicity of the nano-channel membranes drastically decreases the flow due to disruption of smoothness of the wall as well as its favorable interaction with water.[77] Apart from water molecules, selective flow of ions through nano-channels has also gained attention. The energetics of Na^+, K^+, Cl^- ions while entering inside SWNT of smaller diameters has been found to vary as a function of nano-channel radii, underscoring the significance of nano-channel dimensions in ion-selectivity.[78] Charge-functionalized multi-walled carbon nanotubes (MWNT) exhibited voltage-gated control of ionic flow through the nano-channel core.[79] Also, the mass transport of ions, polar liquids (water, alcohols, etc.) as well as non-polar gases was found to alter significantly within the MWNT pores.[80]

Interestingly, it has been suggested that solvents with altered polarity (like urea, ethanol, glycerol, etc.) help tune the aqueous flow within channels in nano-scale electronic devices.[81] A molecule having size, structure and properties comparable with water (namely methanol) has been recently studied, characterizing the bi-wire like nano-fluidic properties.[82] The spontaneous

mono-wired filling of urea resulted in expulsion of water molecules from the narrow SWNT interior.[83] However, there have been relatively few detailed investigations of the dynamics and structural aspects of water within nano-channels in presence of a foreign molecule. Hexafluoro-2-propanol (HFIP), an alcohol with lower polarity than water, but higher polarity than other alcohols (methanol, ethanol, propanol, etc.) has earlier been useful investigating dynamical aspects of several biological systems.[84-88] In our present study we have selected SWNTs having 1.36 nm diameter, with lengths (L) varying from 11 Å to 50 Å, and calculated the resultant flow as well as the dynamics of water within the nano-channel when exactly identical water boxes are kept on either sides of the SWNTs (Nil_{HFP}). Each Nil_{HFP} system was simulated for 25 ns. Hereafter, we have modified one of the solvent boxes introducing 27 HFIP molecules and removing an appropriate number of water molecules (Water-SWNT-Mix).[89] Five trajectories were generated for each of the Water-SWNT-Mix systems for up to 30ns. We have implemented unbiased atomistic MD simulations protocols with TIP3P[90] water model (as depicted in our earlier studies), to elucidate the altered water properties defined in the present investigation. The flow of water within nano-channels has been compared for the respective systems. Subtle differences in translational as well as rotational dynamics were observed between the Nil_{HFP} and Water-SWNT-Mix systems, with differences increasing gradually with increase in N_{HFP}. Our preliminary studies show how the introduction of cosolutes of different polarity can alter the flow and dynamical behavior of water in nanochannels, opening the scope of investigation for further alterations due to electric field, external point charges, functionalized nanochannels, etc.

System setup and MD simulations

Table 6.5 Abbreviations and their corresponding descriptions, used in describing the model SWNT systems.

Abbreviations	Descriptions
SWNT	Single-walled carbon nanotube
L (Å)	Length of an SWNT
HFIP	Hexa-fluoro-2-propanol
Nil_{HFP}	SWNT systems with pure water (no HFIP) in either of the solvent boxes
Water-SWNT-Mix	SWNT systems with 27 HFIPs in one solvent box
N_{HFP}	Number of HFIP present within an SWNT
t_{Meta}	Time window for constant N_{HFP}

Nine armchair ((10,10); diameter 1.36 nm[91]) SWNTs, the lengths (L) of which varying from 11 Å – 50 Å (11 Å, 15 Å, 20 Å, 25 Å, 30 Å, 35 Å, 40 Å, 45 Å, 50 Å) are generated using Visual Molecular Dynamics (VMD).[92] Each SWNT have two well-equilibrated water boxes on either sides of nanotube containing 2322 TIP3P[90] water molecules in each of them. The nanotubes were aligned along the z-axis with its center almost coinciding with the center of the overall system. The edges of SWNTs were combined with two graphene sheets, hence separating the water-boxes apart. Each such system with identical solvent boxes on either side (Nil_{HFP}) was simulated for 25 ns. In addition, a water box of similar dimension, with a few water molecules being replaced by 27 HFIP[84] molecules (separated by 7.5 Å along z-axis and 10 Å along x and y axis) was simulated for 5 ns, and the end-point coordinates were considered for the final water/HFIP box to be placed in positive z-axis of the SWNTs (Figure 6.5). Five trajectories for a constant SWNT length of the Water-SWNT-Mix systems were generated for 18-30 ns, depending on the metastable time (t_{Meta}), where the number of HFIP (N_{HFP}) remains constant. For all the SWNT systems, the nanotubes and graphene sheets were restrained with a harmonic force of 20 kcal mol^{-1} $Å^{-2}$. The graphene and SWNT carbon atoms were modeled as the sp^2 hybridized carbon atoms of the CHARMM22 force field,[93,94] and the HFIP force field parameters were those reported by Fioroni *et al*.[84] The NAMD 2.9 simulation package was used.[95] Trajectories were generated in the isothermal-isobaric (NPT) ensemble at constant temperature (300 K) and constant pressure (1 bar). Energy minimizations of 15000 steps were performed for each system using the conjugate gradient method. Orthorhombic periodic boundary conditions were applied along the x and y directions, while it was applied to the double of the z-dimension, so that the water/HFIP solvent box does not interact with the pure-water box of the mirror image. A 12 Å cut-off was used for non-bonded interactions with the smoothing starting at 11 Å. Particle mesh Ewald (PME)[96] was used for the long-range electrostatics calculation. Constant temperature was maintained by Langevin Dynamics with a collision frequency of 1 ps^{-1}, and constant pressure was maintained by Nose´–Hoover Langevin piston method.[97] The lengths of the bonds involving hydrogen atoms were constrained using the SHAKE algorithm.[98] A timestep of 2 fs with snapshots saved every 1 ps was used to generate the trajectories. For the dynamical analyses, a 400 ps NPT (1 bar, 300 K; to maintain constant pressure along the z-axis) simulation with 2 fs timestep and snapshots saved every 100 fs were generated, after N_{HFP} is found to be constant for a metastable time domain (t_{Meta}). To compare the structure and dynamics of the nano-confined

water with the bulk water behavior; 4 ns NPT (1 bar, 300K) simulation with 1 fs timestep and snapshots saved every 1 ps, and 400 ps NVT (300K) simulation (after NPT simulation was performed) with 2 fs timestep and snapshots being saved every 100fs were performed respectively. VMD was used for visualization.[92]

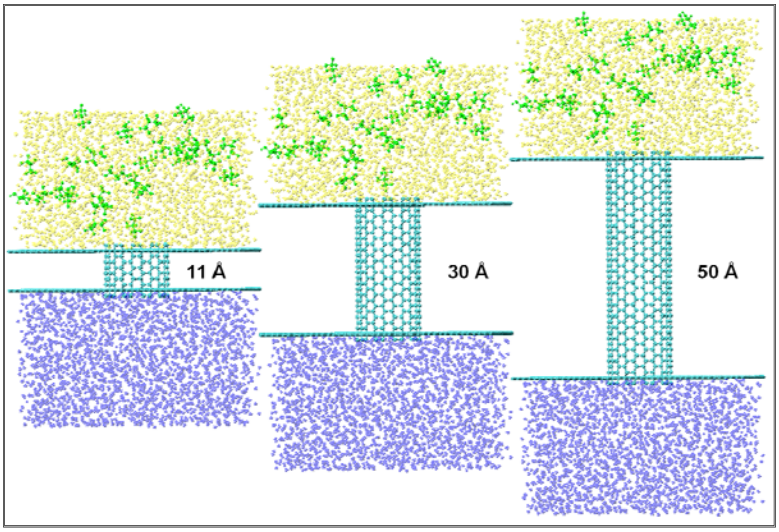


Figure 6.5 Representative snapshots of the starting structures for the water-SWNT-Mix systems. Upper solvent-reservoir contains water (yellow) and HFIP molecules (green). Only pure water molecules (blue) are incorporated in the lower reservoir.

Trajectory analyses

We have estimated the diffusion coefficient from the displacement of water molecules present within the nano-channels. The diffusion coefficient (D) can be computed from the center of mass of the mean square displacement (MSD) of the corresponding molecules as,

$$\left\langle \left| \left(r(t) - r(0) \right) \right|^2 \right\rangle = ADt^{\alpha} \qquad (6.1)$$

Here, r is the center of mass coordinate of water molecule, the angle bracket denotes averaging over all water molecules, t is the time interval, A is the dimensional factor (2, 4 and 6 for 1, 2 and 3-dimensional systems respectively) and α corresponds to the type of diffusion mechanism (single file when α = 0.5, Fickian when α = 1 and ballistic when α = 2).[99] Generally the axial diffusion coefficient (along the z-axis) is determined for diffusion within the nano-channels, where A = 2 and ΔZ^2 (MSD along z-axis) determines the corresponding squared displacement, in units of $Å^2$.

Results and discussions

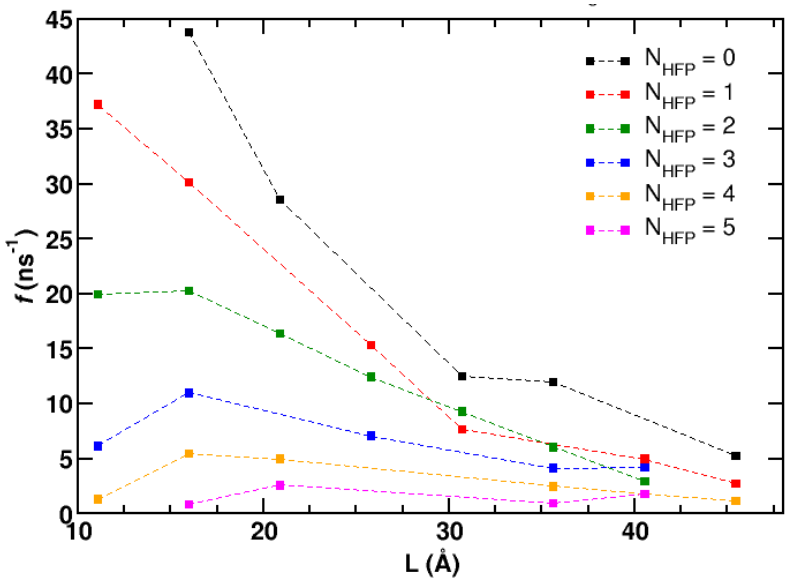


Figure 6.6 *f* as a function of nanotube length for systems with different N_{HFP}.

Earlier studies have correlated the change in flow of water inside channels with that of nanotube length.[100] Flow (*f*) has been defined as the summation of up-flux and down-flux; up-flux and down-flux are the total number of water molecules per nanosecond conducted through SWNT from bottom to top (−z/+z) and from top to bottom (+z/−z), respectively. The flow (*f*) through nanotubes had been fitted to the normal exponential function

$$f = A_0 e^{-L/L_0} + f_0 \qquad (6.2)$$

A_0, L_0 and f_0 are constants. Although the data is not fitted to the exponential function, we find that flow obtained from Nil_{HFP} systems and from systems with N_{HFP}=0 are almost identical with each other (data not shown), having a near exponential pattern in the decrease in flow rate with increase in L, and similar to previous studies. Further, as a preliminary investigation, we have obtained the flow as a function of L for different N_{HFP} systems (Figure 6.6). We again find a near exponential decrease in flow of water within channels for systems with less N_{HFP}, but the exponential pattern in the decrease in flow is lost as N_{HFP} increase within the channels. Further with greater N_{HFP} values, the flow remains almost unchanged as a function of SWNT lengths. Continuing investigations include determination of up-flux, down-flux and flux of water within

nanochannels as functions of L for different N_{HFP} systems. In addition, whether the exponential decrease in flow also exists as a function of N_{HFP} for systems with different L can be elucidated.

Structure and thermodynamic properties of water within nanochannels are essentially different from bulk water. This behavior of water has been observed in nanochannels with non-polar cavities,[101-106] or exclusively polar or apolar walls.[107-109] Several studies have elucidated translational dynamics of water with their corresponding diffusion coefficients.[99,106,110-115] Herein, we have plotted the mean square displacement ΔZ^2 for the water molecules of selected systems (Figure 6.7, 6.8), corresponding D and α values are tabulated below (Table 6.6, 6.7).

In Figure 6.7, we initially notice that there are subtle differences in the diffusion patterns corresponding to representative SWNT systems (of lengths 20 Å, 35 Å, 45 Å) with Nil_{HFP} and N_{HFP} = 0. Although, the slope of ΔZ^2 moderately decreases with increase in L for both Nil_{HFP} and N_{HFP} = 0 systems, an alteration in the diffusive trends for these systems has been observed, with increase in L. ΔZ^2 corresponding to the systems with N_{HFP} = 0 gradually increases with increase in L, although the diffusive trends of Nil_{HFP} systems are similar for the selected systems of interest. This is also validated from the tabulated values of α in Tables 6.6, which proves that introduction of foreign molecules in water (even when not present within the nanochannels) can alter the dynamics of nano-confined water molecules.

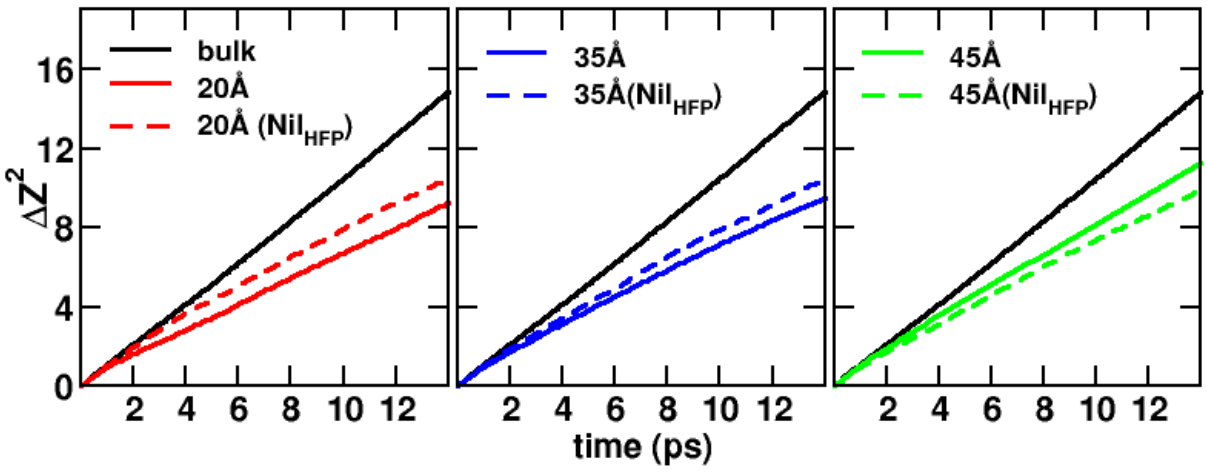


Figure 6.7 Plots of mean square displacement along z-axis (ΔZ^2) for systems with N_{HFP} =0 and Nil_{HFP} systems, having 20 Å, 35 Å and 45 Å lengths of SWNTs.

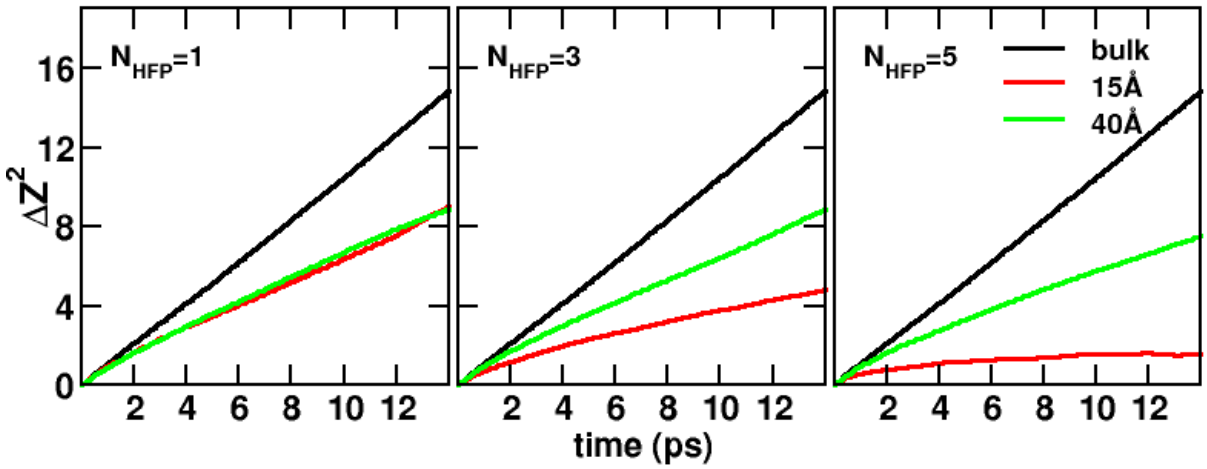


Figure 6.8 Plots of mean square displacement along z-axis (ΔZ^2) for water molecules within nanochannels of different lengths for a constant N_{HFP} (1, 3, 5)

As an extension to these primary investigations, we have additionally taken two extreme cases from the simulated SWNT systems (the corresponding nanochannel lengths to be 15 Å and 40 Å), to compare the translational trends in different systems. We find that with increase in N_{HFP}, the slopes of ΔZ^2 decrease gradually for SWNTs of smaller L. Also, the slopes of ΔZ^2 increase with increase in L for a particular N_{HFP} (Figure 6.8). But while the increase in diffusive pattern with increase in L is marginally reflected for the systems with N_{HFP} = 1, noticeable differences between the diffusion properties of the selected systems (40 and 15 Å) are observed with increase in N_{HFP} (Figure 6.8). The increase in translational properties of water in systems with greater N_{HFP} resembles earlier findings, where water within infinite SWNTs with narrower diameter resembles the Fickian diffusion pattern instead of single file diffusion.[99] The flattening of the slope of ΔZ^2 for water molecules within the SWNT of smaller length (15Å) with greater number of N_{HFP} (= 5), and the corresponding near zero α values for these systems (Table 6.7), indicate highly sub-diffusive nature in the translational behavior of nanoconfined water for these systems.

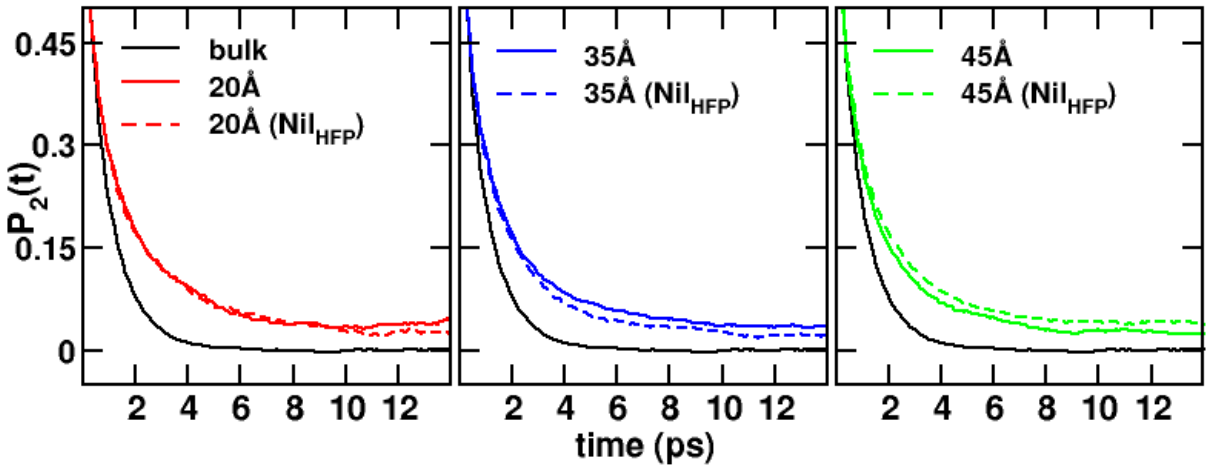


Figure 6.9 Plots of $P_2(t)$ for selected systems with N_{HFP} =0 and Nil_{HFP} systems, having 20 Å, 35 Å and 45 Å lengths of SWNTs.

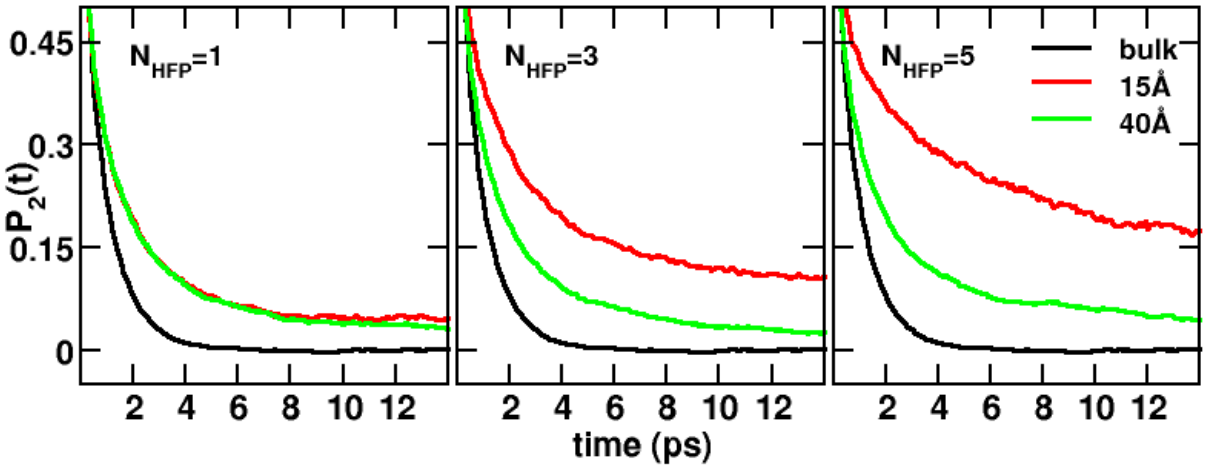


Figure 6.10 Plots of $P_2(t)$ corresponding to water molecules within nanochannels of different lengths for a constant N_{HFP} (1, 3, 5).

Earlier studies have calculated the translational as well as rotational dynamics, to elucidate the properties like ice-like mobility of nano-confined water[106] or electric field induced unidirectional water flow with rotation-translational coupling.[116] To determine the rotational dynamics of water and further illuminate on the correlation between the rotational and translational dynamics, we have obtained the orientational auto-correlation analyses ($P_2(t)$) for the dipole moment unit vector of bulk as well as water within the selected nano-channels, as discussed in *Chapter 4*. The corresponding τ, β and $\langle\tau\rangle$ values are reported in Tables 6.8, 6.9. Resembling ΔZ^2, $P_2(t)$ for bulk also has a faster decay compared to any of the systems (Figure

6.9). Further, from Figure 6.9, we see that although there are small differences in the trends of Nil_{HFP} systems and systems with $N_{HFP} = 0$, differences in trends of $P_2(t)$ for systems with $N_{HFP} = 0$ is also observed with increase in channel length, in qualitative agreement with previously determined ΔZ^2 trends. This is also reflected from the tabulated $<\tau>$ values in Table 6.8.

$P_2(t)$ plots (Figure 6.10) corresponding to water molecules of the selected systems with constant N_{HFP} clearly shows that slowering of the rotational relaxation is directly proportional to increase in N_{HFP}. Further, increased magnitudes of $<\tau>$, for the SWNT of smaller length (15 Å) but with greater N_{HFP} (= 5), matches with the trends of near zero α value (Table 6.9). It can be thus concluded that foreign molecule influenced translational and rotational properties of water molecules are similar in nature (considering the current case study with HFIP), with minute alteration in their trends in the SWNT systems of interest.

Table 6.6 Tabulated D and α values from the mean square displacement (ΔZ^2) of bulk water, Nil_{HFP} and for water within channels where $N_{HFP} = 0$.

Systems	N_{HFP}	L (Å)	*D*	*α*
Bulk	-	-	0.47	1.04
Nil_{HFP}	-	20	0.67	0.78
		35	0.59	0.82
		45	0.52	0.86
Water-SWNT-Mix	0	20	0.33	1.00
		35	0.55	0.82
		45	0.43	0.98

Table 6.7 Tabulated D and α from the mean square displacement (ΔZ^2) for water molecules within nanochannels of different lengths with non-zero N_{HFP} (1, 3, 5).

Systems	N_{HFP}	L (Å)	*D*	*α*
Water-SWNT-Mix	1	15	0.26	1.07
		40	0.56	0.78
Water-SWNT-Mix	3	15	0.39	0.68
		40	0.35	0.96
Water-SWNT-Mix	5	15	0.94	-0.07
		40	0.40	0.84

Table 6.8 Tabulated τ, β and <*τ*> values from orientational autocorrelation analyses for bulk water behavior, water molecules present within nano-channels of the Nil_{HFP} systems, and water within channels where $N_{HFP} = 0$.

Systems	N_{HFP}	L (Å)	*τ*	*β*	<*τ*>
Bulk	-	-	0.56	0.7	0.71
Nil_{HFP}	-	20	0.54	0.43	1.49
		35	0.55	0.46	1.30
		45	0.50	0.39	1.78
Water-SWNT-Mix	0	20	0.51	0.39	1.82
		35	0.52	0.41	1.62
		45	0.52	0.43	1.43

Table 6.9 Tabulated τ, β and <τ> values from orientational autocorrelation analyses for water within channels, with non-zero N_{HFP} (1, 3, and 5).

Systems	N_{HFP}	L (Å)	τ	β	<τ>
Water-SWNT-Mix	1	15	0.53	0.38	2.04
		40	0.58	0.41	1.80
Water-SWNT-Mix	3	15	0.98	0.32	6.89
		40	0.59	0.42	1.72
Water-SWNT-Mix	5	15	1.63	0.24	50.43
		40	0.52	0.36	2.38

Continuing prospects

Presently we are elucidating the tetrahedral order parameter (Q)[117,118] (*Chapter 4*) of water molecules (considering all possible atoms as potential hydrogen bonding sites) for bulk water, water within nano-channels. We see that compared to bulk, Q is considerably reduced within the nano-channel owing to the reduced availability of potential donors for hydrogen binding sites (data not shown). Earlier studies have indicated a decrease in water-water hydrogen bonds from 3.7 to 2.9 at the water-nanotube interface.[119] Also, it has been proposed that water molecules entering the nano-channel lose on an average of 2 hydrogen bonds,[21] commensurate with our findings of decreased Q for nano-confined water compared to bulk.

Our preliminary results provide insights on the flow as well as dynamical and structural aspects of water molecules within the channels in presence of HFIP. The work is undertaken in continuation from the elucidated findings, in determining the properties of remaining systems. Based on the results to be obtained from the presently undertaken investigations, protocols in altering these trends like applying applying electric field, introducing external point charges, or creating functionalized nano-channels, may be further implemented. Theoretical calculations have already predicted the generation of hydroelectric voltage on water-occupied uncapped SWNT.[120] Our results thus open the possibilities of future works of studying the properties of

water in presence of a foreign molecule due to the application of electric field, hence enlightening the mechanistic aspects of nano-electronic devices, chemical separators, and so on. We attempt to further calculate the velocity autocorrelation function and integrate it to obtain the diffusion coefficients, and compare those with the obtained values of D from the non-linear mean square displacement data.

6.4 Bibliography

(1) Selkoe, D. J. Folding proteins in fatal ways. *Nature*. **2003**, *426*, 900-904.

(2) Dobson, C. M. Protein folding and misfolding. *Nature*. **2003**, *426*, 884-890.

(3) Das, A. K.; Rawat, A.; Bhowmik, D.; Pandit, R.; Huster, D.; Maiti, S. An Early Folding Contact between Phe19 and Leu34 is Critical for Amyloid-*β* Oligomer Toxicity. *ACS Chem. Neurosci.* **2015**, *6*, 1290-1295.

(4) Mannini, B.; Mulvihill, E.; Sgromo, C.; Cascella, R.; Khodarahmi, R.; Ramazzotti, M.; Dobson, C. M.; Cecchi, C.; Chiti, F. Toxicity of Protein Oligomers Is Rationalized by a Function Combining Size and Surface Hydrophobicity. *ACS Chem. Biol.* **2014**, *9*, 2309-2317.

(5) Bartels, T.; Choi, J. G.; Selkoe, D. J. α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature*. **2011**, *477*, 107-110.

(6) Gurry, T.; Ullman, O.; Fisher, C. K.; Perovic, I.; Pochapsky, T.; Stultz, C. M. The Dynamic Structure of α-Synuclein Multimers. *J. Am. Chem. Soc.* **2013**, *135*, 3865-3872.

(7) Ullman, O.; Fisher, C. K.; Stultz, C. M. Explaining the Structural Plasticity of *a*-Synuclein. *J. Am. Chem. Soc.* **2011**, *133*, 19536-19546.

(8) Jana, Asis K.; Sengupta, N. Adsorption Mechanism and Collapse Propensities of the Full-Length, Monomeric $A\beta_{1-42}$ on the Surface of a Single-Walled Carbon Nanotube: A Molecular Dynamics Simulation Study. *Biophys. J.* **2012**, *102*, 1889-1896.

(9) Jose, J.; Sengupta, N. Molecular dynamics simulation studies of the structural response of an isolated $A\beta_{1-42}$ monomer localized in the vicinity of the hydrophilic TiO_2 surface. *Eur. Biophys. J.* **2013**, *42*, 487-494.

(10) Jana, A. K.; Tiwari, M. K.; Vanka, K.; Sengupta, N. Unraveling origins of the heterogeneous curvature dependence of polypeptide interactions with carbon nanostructures. *Phys. Chem. Chem. Phys.* **2015**.

(11) Jana, A. K.; Sengupta, N. Aβ self-association and adsorption on a hydrophobic nanosurface: competitive effects and the detection of small oligomers via electrical response. *Soft Matter*. **2015**, *11*, 269-279.

(12) Chong, S.-H.; Ham, S. Atomic-level investigations on the amyloid-β dimerization process and its driving forces in water. *Phys. Chem. Chem. Phys.* **2012**, *14*, 1573-1575.

(13) Chong, S.-H.; Ham, S. Impact of chemical heterogeneity on protein self-assembly in water. *Proc. Natl. Sci. U.S.A.* **2012**, *109*, 7636-7641.

(14) Lopez del Amo, J. M.; Schmidt, M.; Fink, U.; Dasari, M.; Fändrich, M.; Reif, B. An Asymmetric Dimer as the Basic Subunit in Alzheimer's Disease Amyloid β Fibrils. *Angew. Chem. Int. Ed.* **2012**, *51*, 6136-6139.

(15) Conway, K. A.; Lee, S.-J.; Rochet, J.-C.; Ding, T. T.; Williamson, R. E.; Lansbury, P. T. Acceleration of oligomerization, not fibrillization, is a shared property of both α-synuclein mutations linked to early-onset Parkinson's disease: Implications for pathogenesis and therapy. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 571-576.

(16) Iwai, A.; Masliah, E.; Yoshimoto, M.; Ge, N.; Flanagan, L.; Rohan de Silva, H. A.; Kittel, A.; Saitoh, T. The precursor protein of non-A*β* component of Alzheimer's disease amyloid is a presynaptic protein of the central nervous system. *Neuron*. **1995**, *14*, 467-475.

(17) Wang, W.; Perovic, I.; Chittuluru, J.; Kaganovich, A.; Nguyen, L. T.; Liao, J.; Auclair, J. R.; Johnson, D.; Landeru, A.; Simorellis, A. K.; Ju, S.; Cookson, M. R.; Asturias, F. J.; Agar, J. N.; Webb, B. N.; Kang, C.; Ringe, D.; Petsko, G. A.; Pochapsky, T. C.; Hoang, Q. Q. A soluble α-synuclein construct forms a dynamic tetramer. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 17797-17802.

(18) Spillantini, M. G.; Schmidt, M. L.; Lee, V. M. Y.; Trojanowski, J. Q.; Jakes, R.; Goedert, M. *α*-Synuclein in Lewy bodies. *Nature*. **1997**, *388*, 839-840.

(19) Resende, R.; Marques, S. F.; Ferreiro, E.; Simões, I.; Oliveira, C.; Pereira, C. F. Effect of α-Synuclein on Amyloid β-Induced Toxicity: Relevance to Lewy Body Variant of Alzheimer Disease. *Neurochem. Res.* **2013**, *38*, 797-806.

(20) Hardy, J.; Selkoe, D. J. The Amyloid Hypothesis of Alzheimer's Disease: Progress and Problems on the Road to Therapeutics. *Science*. **2002**, *297*, 353-356.

(21) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature*. **2001**, *414*, 188-190.

(22) Joseph, S.; Aluru, N. R. Pumping of Confined Water in Carbon Nanotubes by Rotation-Translation Coupling. *Phys. Rev. Lett.* **2008**, *101*, 064502.

(23) Choudhury, N. Effect of surface hydrophobicity on the dynamics of water at the nanoscale confinement: A molecular dynamics simulation study. *Chem. Phys.* **2013**, *421*, 68-76.

(24) Sun, Y.-P.; Fu, K.; Lin, Y.; Huang, W. Functionalized Carbon Nanotubes: Properties and Applications. *Accounts Chem. Res.* **2002**, *35*, 1096-1104.

(25) Pantarotto, D.; Singh, R.; McCarthy, D.; Erhardt, M.; Briand, J.-P.; Prato, M.; Kostarelos, K.; Bianco, A. Functionalized Carbon Nanotubes for Plasmid DNA Gene Delivery. *Angew. Chem. Int. Ed*. **2004**, *116*, 5354-5358.

(26) Paul, S.; Abi, T. G.; Taraphder, S. Structure and dynamics of water inside endohedrally functionalized carbon nanotubes. *J. Chem. Phys.* **2014**, *140*, 184511.

(27) Abi, T. G.; Taraphder, S. Free energies of proton transfer by polar amino acid sidechain analogues anchored to the outer wall of single walled carbon nanotubes. *Comp. Theor. Chem.* **2014**, *1027*, 19-25.

(28) Khabashesku, V. N.; Margrave, J. L.; Barrera, E. V. Functionalized carbon nanotubes and nanodiamonds for engineering and biomedical applications. *Diam. Relat. Mater.* **2005**, *14*, 859-866.

(29) Rasaiah, J. C.; Garde, S.; Hummer, G. Water in Nonpolar Confinement: From Nanotubes to Proteins and Beyond. *Annu. Rev. Phys. Chem.* **2008**, *59*, 713-740.

(30) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Accounts Chem. Res.* **2012**, *45*, 83-92.

(31) Hardy, J.; Higgins, G. Alzheimer's disease: the amyloid cascade hypothesis. *Science*. **1992**, *256*, 184-185.

(32) Uéda, K.; Fukushima, H.; Masliah, E.; Xia, Y.; Iwai, A.; Yoshimoto, M.; Otero, D. A.; Kondo, J.; Ihara, Y.; Saitoh, T. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. *Proc. Natl. Sci. U.S.A.* **1993**, *90*, 11282-11286.

(33) Eliezer, D.; Kutluay, E.; Bussell Jr, R.; Browne, G. Conformational properties of α-synuclein in its free and lipid-associated states1. *J. Mol. Biol.* **2001**, *307*, 1061-1073.

(34) Lee, H.-J.; Choi, C.; Lee, S.-J. Membrane-bound α-Synuclein Has a High Aggregation Propensity and the Ability to Seed the Aggregation of the Cytosolic Form. *J. Biol. Chem.* **2002**, *277*, 671-678.

(35) Uversky, V. N.; Li, J.; Fink, A. L. Evidence for a Partially Folded Intermediate in α-Synuclein Fibril Formation. *J. Biol. Chem.* **2001**, *276*, 10737-10744.

(36) Wood, S. J.; Wypych, J.; Steavenson, S.; Louis, J. C.; Citron, M.; Biere, A. L. alpha-synuclein fibrillogenesis is nucleation-dependent. Implications for the pathogenesis of Parkinson's disease. *J. Biol. Chem*. **1999**, *274*, 19509-19512.

(37) Winner, B.; Jappelli, R.; Maji, S. K.; Desplats, P. A.; Boyer, L.; Aigner, S.; Hetzer, C.; Loher, T.; Vilar, M.; Campioni, S.; Tzitzilonis, C.; Soragni, A.; Jessberger, S.; Mira, H.; Consiglio, A.; Pham, E.; Masliah, E.; Gage, F. H.; Riek, R. In vivo demonstration that α-synuclein oligomers are toxic. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 4194-4199.

(38) Karpinar, D. P.; Balija, M. B.; Kugler, S.; Opazo, F.; Rezaei-Ghaleh, N.; Wender, N.; Kim, H. Y.; Taschenberger, G.; Falkenburger, B. H.; Heise, H.; Kumar, A.; Riedel, D.; Fichtner, L.; Voigt, A.; Braus, G. H.; Giller, K.; Becker, S.; Herzig, A.; Baldus, M.; Jackle, H.; Eimer, S.; Schulz, J. B.; Griesinger, C.; Zweckstetter, M. Pre-fibrillar α-synuclein variants with impaired beta-structure increase neurotoxicity in Parkinson's disease models. *EMBO J*. **2009**, *28*, 3256-3268.

(39) Meratan, A. A.; Ghasemi, A.; Nemat-Gorgani, M. Membrane Integrity and Amyloid Cytotoxicity: A Model Study Involving Mitochondria and Lysozyme Fibrillation Products. *J. Mol. Biol.* **2011**, *409*, 826-838.

(40) Bucciantini, M.; Calloni, G.; Chiti, F.; Formigli, L.; Nosi, D.; Dobson, C. M.; Stefani, M. Prefibrillar amyloid protein aggregates share common features of cytotoxicity. *J. Biol. Chem*. **2004**, *279*, 31374-31382.

(41) Gosavi, N.; Lee, H. J.; Lee, J. S.; Patel, S.; Lee, S. J. Golgi fragmentation occurs in the cells with prefibrillar α-synuclein aggregates and precedes the formation of fibrillar inclusion. *J. Biol. Chem*. **2002**, *277*, 48984-48992.

(42) Chaudhary, H.; Stefanovic, A. N. D.; Subramaniam, V.; Claessens, M. M. A. E. Membrane interactions and fibrillization of α-synuclein play an essential role in membrane disruption. *FEBS Lett*. **2014**, *588*, 4457-4463.

(43) Zarranz, J. J.; Alegre, J.; Gómez-Esteban, J. C.; Lezcano, E.; Ros, R.; Ampuero, I.; Vidal, L.; Hoenicka, J.; Rodriguez, O.; Atarés, B.; Llorens, V. n.; Gomez Tortosa, E.; del Ser, T.; Muñoz, D. G.; de Yebenes, J. G. The new mutation, E46K, of α-synuclein causes Parkinson and Lewy body dementia. *Ann. neurol.* **2004**, *55*, 164-173.

(44) Kruger, R.; Kuhn, W.; Muller, T.; Woitalla, D.; Graeber, M.; Kosel, S.; Przuntek, H.; Epplen, J. T.; Schols, L.; Riess, O. Ala30Pro mutation in the gene encoding α-synuclein in Parkinson's disease. *Nat. Genet*. **1998**, *18*, 106-108.

(45) Fares, M.-B.; Bouziad, N. A.; Dikiy, I.; Mbefo, M. K.; Jovičić, A.; Kiely, A.; Holton, J. L.; Lee, S.-J.; Gitler, A. D.; Eliezer, D.; Lashuel, H. A. The Novel Parkinson's Disease Linked Mutation G51D Attenuates In Vitro Aggregation and Membrane Binding of α-Synuclein, and Enhances its Secretion and Nuclear Localization in Cells. *Human Mol. Genet.* **2014**.

(46) Polymeropoulos, M. H.; Lavedan, C.; Leroy, E.; Ide, S. E.; Dehejia, A.; Dutra, A.; Pike, B.; Root, H.; Rubenstein, J.; Boyer, R.; Stenroos, E. S.; Chandrasekharappa, S.; Athanassiadou, A.; Papapetropoulos, T.; Johnson, W. G.; Lazzarini, A. M.; Duvoisin, R. C.; Di Iorio, G.; Golbe, L. I.; Nussbaum, R. L. Mutation in the a-Synuclein Gene Identified in Families with Parkinson's Disease. *Science*. **1997**, *276*, 2045-2047.

(47) Conway, K. A.; Harper, J. D.; Lansbury, P. T. Accelerated in vitro fibril formation by a mutant alpha-synuclein linked to early-onset Parkinson disease. *Nat. Med*. **1998**, *4*, 1318-1320.

(48) Conway, K. A.; Lee, S. J.; Rochet, J. C.; Ding, T. T.; Williamson, R. E.; Lansbury, P. T., Jr. Acceleration of oligomerization, not fibrillization, is a shared property of both α-synuclein mutations linked to early-onset Parkinson's disease: implications for pathogenesis and therapy. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 571-576.

(49) Greenbaum, E. A.; Graves, C. L.; Mishizen-Eberz, A. J.; Lupoli, M. A.; Lynch, D. R.; Englander, S. W.; Axelsen, P. H.; Giasson, B. I. The E46K mutation in α-synuclein increases amyloid fibril formation. *J. Biol. Chem*. **2005**, *280*, 7800-7807.

(50) Bussell, R., Jr.; Eliezer, D. Effects of Parkinson's disease-linked mutations on the structure of lipid-associated alpha-synuclein. *Biochemistry*. **2004**, *43*, 4810-4818.

(51) Bussell, R., Jr.; Eliezer, D. Residual structure and dynamics in Parkinson's disease-associated mutants of alpha-synuclein. *J. Biol. Chem*. **2001**, *276*, 45996-46003.

(52) Chatterjee, P.; Sengupta, N. Effect of the A30P mutation on the structural dynamics of micelle-bound αSynuclein released in water: a molecular dynamics study. *Eur. Biophys. J.* **2012**, *41*, 483-489.

(53) Bartels, T.; Choi, J. G.; Selkoe, D. J. α-Synuclein occurs physiologically as a helically folded tetramer that resists aggregation. *Nature*. **2011**, *477*, 107-110.

(54) Rao, J. N.; Jao, C. C.; Hegde, B. G.; Langen, R.; Ulmer, T. S. A Combinatorial NMR and EPR Approach for Evaluating the Structural Ensemble of Partially Folded Proteins. *J. Am. Chem. Soc.* **2010**, *132*, 8657-8668.

(55) Markwick, P. R. L.; McCammon, J. A. Studying functional dynamics in bio-molecules using accelerated molecular dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20053-20065.

(56) Jose, J. C.; Chatterjee, P.; Sengupta, N. Cross Dimerization of Amyloid-β and αSynuclein Proteins in Aqueous Environment: A Molecular Dynamics Simulations Study. *PLoS ONE*. **2014**, *9*, e106883.

(57) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B*. **1998**, *102*, 3586-3616.

(58) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400-1415.

(59) Onufriev, A.; Bashford, D.; Case, D. A. Modification of the Generalized Born Model Suitable for Macromolecules. *J. Phys. Chem. B*. **2000**, *104*, 3712-3720.

(60) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Struct. Funct. Bioinf.* **2004**, *55*, 383-394.

(61) Lin, Y.-S.; Bowman, Gregory R.; Beauchamp, Kyle A.; Pande, Vijay S. Investigating How Peptide Length and a Pathogenic Mutation Modify the Structural Ensemble of Amyloid Beta Monomer. *Biophys. J.* **2012**, *102*, 315-324.

(62) Nguyen, P. H.; Li, M. S.; Derreumaux, P. Amyloid oligomer structure characterization from simulations: A general method. *J. Chem. Phys.* **2014**, *140*, 094105.

(63) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2007**, *108*, 74-108.

(64) Ball, P. Water as an Active Constituent in Cell Biology. *Chem. Rev.* **2007**, *108*, 74-108.

(65) Murata, K.; Mitsuoka, K.; Hirai, T.; Walz, T.; Agre, P.; Heymann, J. B.; Engel, A.; Fujiyoshi, Y. Structural determinants of water permeation through aquaporin-1. *Nature*. **2000**, *407*, 599-605.

(66) Agre, P. Aquaporin Water Channels (Nobel Lecture). *Angew. Chem. Int. Ed.* **2004**, *43*, 4278-4290.

(67) Miyazawa, A.; Fujiyoshi, Y.; Unwin, N. Structure and gating mechanism of the acetylcholine receptor pore. *Nature*. **2003**, *423*, 949-955.

(68) Sui, H.; Han, B.-G.; Lee, J. K.; Walian, P.; Jap, B. K. Structural basis of water-specific transport through the AQP1 water channel. *Nature*. **2001**, *414*, 872-878.

(69) Wan, R.; Li, J.; Lu, H.; Fang, H. Controllable Water Channel Gating of Nanometer Dimensions. *J. Am. Chem. Soc.* **2005**, *127*, 7166-7170.

(70) Li, J.; Gong, X.; Lu, H.; Li, D.; Fang, H.; Zhou, R. Electrostatic gating of a nanometer water channel. *Proc. Natl. Sci. U.S.A.* **2007**, *104*, 3687-3692.

(71) Zhou, X.; Wu, F.; Kou, J.; Nie, X.; Liu, Y.; Lu, H. Vibrating-Charge-Driven Water Pump Controlled by the Deformation of the Carbon Nanotube. *J. Phys. Chem. B.* **2013**, *117*, 11681-11686.

(72) Meng, X. W.; Wang, Y.; Zhao, Y. J.; Huang, J. P. Gating of a Water Nanochannel Driven by Dipolar Molecules. *J. Phys. Chem. B.* **2011**, *115*, 4768-4773.

(73) Lu, H.; Zhou, X.; Wu, F.; Xu, Y. Effect of Charge on Water Filling/Emptying Transitions of Nanochannel. *J. Phys. Chem. B.* **2008**, *112*, 16777-16781.

(74) Gong, X.; Li, J.; Zhang, H.; Wan, R.; Lu, H.; Wang, S.; Fang, H. Enhancement of Water Permeation across a Nanochannel by the Structure outside the Channel. *Phys. Rev. Lett.* **2008**, *101*, 257801.

(75) Xiu, P.; Zhou, B.; Qi, W.; Lu, H.; Tu, Y.; Fang, H. Manipulating Biomolecules with Aqueous Liquids Confined within Single-Walled Nanotubes. *J. Am. Chem. Soc.* **2009**, *131*, 2840-2845.

(76) Liu, B.; Li, X.; Li, B.; Xu, B.; Zhao, Y. Carbon Nanotube Based Artificial Water Channel Protein: Membrane Perturbation and Water Transportation. *Nano Lett*. **2009**, *9*, 1386-1394.

(77) Majumder, M.; Corry, B. Anomalous decline of water transport in covalently modified carbon nanotube membranes. *Chem. Comm*. **2011**, *47*, 7683-7685.

(78) Song, C.; Corry, B. Intrinsic Ion Selectivity of Narrow Hydrophobic Pores. *J. Phys. Chem. B*. **2009**, *113*, 7642-7649.

(79) Majumder, M.; Zhan, X.; Andrews, R.; Hinds, B. J. Voltage Gated Carbon Nanotube Membranes. *Langmuir*. **2007**, *23*, 8624-8631.

(80) Majumder, M.; Chopra, N.; Hinds, B. J. Mass Transport through Carbon Nanotube Membranes in Three Different Regimes: Ionic Diffusion and Gas and Liquid Flow. *ACS Nano*. *5*, **2011**, 3867-3877.

(81) Tu, Y.; Zhou, R.; Fang, H. Signal transmission, conversion and multiplication by polar molecules confined in nanochannels. *Nanoscale*. **2010**, *2*, 1976-1983.

(82) Nakamura, Y.; Ohno, T. Biwire structure of methanol inside carbon nanotubes. *Chem. Phys. Lett.* **2012**, *539–540*, 123-127.

(83) Xiu, P.; Tu, Y.; Tian, X.; Fang, H.; Zhou, R. Molecular wire of urea in carbon nanotube: a molecular dynamics study. *Nanoscale*. **2012**, *4*, 652-658.

(84) Fioroni, M.; Burger, K.; Mark, A. E.; Roccatano, D. Model of 1,1,1,3,3,3-Hexafluoro-propan-2-ol for Molecular Dynamics Simulations. *J. Phys. Chem. B*. **2001**, *105*, 10967-10975.

(85) Ennaceur, S. M.; Sanderson, J. M. Micellar Aggregates Formed Following the Addition of Hexafluoroisopropanol to Phospholipid Membranes. *Langmuir*. **2004**, *21*, 552-561.

(86) Roccatano, D.; Fioroni, M.; Zacharias, M.; Colombo, G. Effect of hexafluoroisopropanol alcohol on the structure of melittin: A molecular dynamics simulation study. *Protein Sci*. **2005**, *14*, 2582-2589.

(87) Yoshida, K.; Yamaguchi, T.; Adachi, T.; Otomo, T.; Matsuo, D.; Takamuku, T.; Nishi, N. Structure and dynamics of hexafluoroisopropanol-water mixtures by x-ray diffraction, small-angle neutron scattering, NMR spectroscopy, and mass spectrometry. *J. Chem. Phys.* **2003**, *119*, 6132-6142.

(88) Wei, G.; Shea, J.-E. Effects of Solvent on the Structure of the Alzheimer Amyloid-β(25–35) Peptide. *Biophys. J.* **2006**, *91*, 1638-1647.

(89) Fioroni, M.; Burger, K.; Mark, A. E.; Roccatano, D. Model of 1,1,1,3,3,3-Hexafluoro-propan-2-ol for Molecular Dynamics Simulations. *J. Phys. Chem. B*. **2001**, *105*, 10967-10975.

(90) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926-935.

(91) Cao, D.; Wu, J. Self-Diffusion of Methane in Single-Walled Carbon Nanotubes at Sub- and Supercritical Conditions. *Langmuir*. **2004**, *20*, 3759-3765.

(92) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics*. **1996**, *14*, 33-38.

(93) MacKerell, A. D.; Bashford, D.; Bellott; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B.* **1998**, *102*, 3586-3616.

(94) Mackerell, A. D.; Feig, M.; Brooks, C. L. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem*. **2004**, *25*, 1400-1415.

(95) Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K. NAMD2: Greater Scalability for Parallel Molecular Dynamics. *J. Comput. Phys.* **1999**, *151*, 283-312.

(96) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577-8593.

(97) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant pressure molecular dynamics simulation: The Langevin piston method. *J. Chem. Phys.* **1995**, *103*, 4613-4621.

(98) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **1977**, *23*, 327-341.

(99) Striolo, A. The Mechanism of Water Diffusion in Narrow Carbon Nanotubes. *Nano Lett*. **2006**, *6*, 633-639.

(100) Su, J.; Guo, H. Effect of nanotube-length on the transport properties of single-file water molecules: Transition from bidirectional to unidirectional. *J. Chem. Phys.* **2011**, *134*, 244513.

(101) Yin, H.; Hummer, G.; Rasaiah, J. C. Metastable Water Clusters in the Nonpolar Cavities of the Thermostable Protein Tetrabrachion. *J. Am. Chem. Soc.* **2007**, *129*, 7369-7377.

(102) Collins, M. D.; Hummer, G.; Quillin, M. L.; Matthews, B. W.; Gruner, S. M. Cooperative water filling of a nonpolar protein cavity observed by high-pressure crystallography and simulation. *Proc. Natl. Sci. U.S.A.* **2005**, *102*, 16668-16671.

(103) Iwamatsu, S.-i.; Uozaki, T.; Kobayashi, K.; Re, S.; Nagase, S.; Murata, S. A Bowl-Shaped Fullerene Encapsulates a Water into the Cage. *J. Am. Chem. Soc.* **2004**, *126*, 2668-2669.

(104) Vaitheeswaran, S.; Yin, H.; Rasaiah, J. C.; Hummer, G. Water clusters in nonpolar cavities. *Proc. Natl. Sci. U.S.A.* **2004**, *101*, 17002-17005.

(105) Ramachandran, C. N.; Sathyamurthy, N. Water clusters in a confined nonpolar environment. *Chem. Phys. Lett.* **2005**, *410*, 348-351.

(106) Mashl, R. J.; Joseph, S.; Aluru, N. R.; Jakobsson, E. Anomalously Immobilized Water: A New Water Phase Induced by Confinement in Nanotubes. *Nano Lett*. **2003**, *3*, 589-592.

(107) Giovambattista, N.; Debenedetti, P. G.; Rossky, P. J. Effect of Surface Polarity on Water Contact Angle and Interfacial Hydration Structure. *J. Phys. Chem. B.* **2007**, *111*, 9581-9587.

(108) Park, J. H.; Aluru, N. R. Diffusion of water submonolayers on hydrophilic surfaces. *Applied Phys. Lett*. **2008**, *93*, 253104.

(109) Wang, C.; Lu, H.; Wang, Z.; Xiu, P.; Zhou, B.; Zuo, G.; Wan, R.; Hu, J.; Fang, H. Stable Liquid Water Droplet on a Water Monolayer Formed at Room Temperature on Ionic Model Substrates. *Phys. Rev. Lett*. **2009**, *103*, 137801.

(110) Liu, Y.; Wang, Q.; Wu, T.; Zhang, L. Fluid structure and transport properties of water inside carbon nanotubes. *J. Chem. Phys.* **2005**, *123*, 234701.

(111) Liu, Y.; Wang, Q.; Zhang, L.; Wu, T. Dynamics and Density Profile of Water in Nanotubes as One-Dimensional Fluid. *Langmuir*. **2005**, *21*, 12025-12030.

(112) Liu, Y. W., Qi. Transport behavior of water confined in carbon nanotubes. *Phys. Rev. B*. **2005**, *72*, 4.

(113) Joseph, S.; Aluru, N. R. Why Are Carbon Nanotubes Fast Transporters of Water? *Nano Lett*. **2008**, *8*, 452-458.

(114) Zheng, Y.-g.; Ye, H.-f.; Zhang, Z.-q.; Zhang, H.-w. Water diffusion inside carbon nanotubes: mutual effects of surface and confinement. *Phys. Chem. Chem. Phys*. **2012**, *14*, 964-971.

(115) Barati Farimani, A.; Aluru, N. R. Spatial Diffusion of Water in Carbon Nanotubes: From Fickian to Ballistic Motion. *J. Phys. Chem. B*. **2011**, *115*, 12145-12149.

(116) Joseph, S.; Aluru, N. R. Pumping of Confined Water in Carbon Nanotubes by Rotation-Translation Coupling. *Phys. Rev. Lett*. **2008**, *101*, 064502.

(117) Chau, P. L.; Hardwick, A. J. A new order parameter for tetrahedral configurations. *Mol. Phys*. **1998**, *93*, 511-518.

(118) Errington, J. R.; Debenedetti, P. G. Relationship between structural order and the anomalies of liquid water. *Nature*. **2001**, *409*, 318-321.

(119) Walther, J. H.; Jaffe, R.; Halicioglu, T.; Koumoutsakos, P. Carbon Nanotubes in Water: Structural Characteristics and Energetics. *J. Phys. Chem. B.* **2001**, *105*, 9980-9987.

(120) Yuan, Q.; Zhao, Y.-P. Hydroelectric Voltage Generation Based on Water-Filled Single-Walled Carbon Nanotubes. *J. Am. Chem. Soc.* **2009**, *131*, 6374-6376.

Copyrights and permissions

Rightslink Printable License

NATURE PUBLISHING GROUP LICENSE
Aug 21, 2015

This is a License Agreement between Prathit Chatterjee ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3693540029226
License date	Aug 21, 2015
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Protein folding and misfolding
Licensed content author	Christopher M. Dobson
Licensed content date	Dec 18, 2003
Volume number	426
Issue number	6968
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print and electronic
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	2
High-res required	no
Figures	Figure 2 (**Figure 1.1 in the dissertation**)
Author of this NPG article	no
Your reference number	None
Title of your thesis / dissertation	Protein thermal stability, conformational dynamics and solvent properties: insights with atomistic Molecular Dynamics simulations
Expected completion date	Sep 2015

Reproducing AAAS Material in your Thesis or Dissertation
NOTE: If you are the Original Author of the AAAS article being reproduced, please refer to your License to Publish for rules on reproducing your paper in a dissertation or thesis. AAAS permits the use of content published in its journals *Science*, *Science Technical Medicine*, and *Science Signaling*, but only provided the following criteria are met.

1. If you are using figure(s)/table(s), permission is granted for use in print and electronic versions of your dissertation or thesis (**Figure 1.2 in the dissertation**).
2. A full text article may be used only in print versions of a dissertation or thesis. AAAS does not permit the reproduction of full text articles in electronic versions of theses or dissertations.
3. The following credit line must be printed along with the AAAS material: "From [Full Reference Citation]. Reprinted with permission from AAAS."
4. All required credit lines and notices must be visible any time a user accesses any part of the AAAS material and must appear on any printed copies that an authorized user might make.
5. The AAAS material may not be modified or altered except that figures and tables may be modified with permission from the author. Author permission for any such changes must be secured prior to your use.
6. AAAS must publish the full paper prior to your use of any of its text or figures.
7. If the AAAS material covered by this permission was published in *Science* during the years 1974–1994, you must also obtain permission from the author, who may grant or withhold permission, and who may or may not charge a fee if permission is granted. See original article for author's address. This condition does not apply to news articles.
8. If you are an Original Author on the AAAS article being reproduced, please refer to your License to Publish for rules on reproducing your paper in a dissertation or thesis.

Permission covers the distribution of your dissertation or thesis on demand by a third party distributor (e.g. ProQuest / UMI), provided the AAAS material covered by this permission remains in situ and is not distributed by that third party outside of the context of your Thesis/Dissertation.

Permission does not apply to figures/photos/artwork or any other content or materials included in your work that are credited to non-AAAS sources. If the requested material is sourced to or references non-AAAS sources, you must obtain authorization from that source as well before using that material. You agree to hold harmless and indemnify AAAS against any claims arising from your use of any content in your work that is credited to non-AAAS sources. By using the AAAS Material identified in your request, you agree to abide by all the terms and conditions herein.

AAAS makes no representations or warranties as to the accuracy of any information contained in the AAAS material covered by this permission, including any warranties of merchantability or fitness for a particular purpose.

Rightslink Printable License

ROYAL SOCIETY OF CHEMISTRY LICENSE
Aug 27, 2015

This is a License Agreement between Prathit Chatterjee ("You") and Royal Society of Chemistry ("Royal Society of Chemistry") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Royal Society of Chemistry, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3697120109306
License date	Aug 27, 2015
Licensed content publisher	Royal Society of Chemistry
Licensed content publication	Physical Chemistry Chemical Physics
Licensed content title	On the molecular origin of cold denaturation of globular proteins
Licensed content author	Giuseppe Graziano
Licensed content date	Sep 30, 2010
Volume number	12
Issue number	42
Type of Use	Thesis/Dissertation
Requestor type	academic/educational
Portion	figures/tables/images
Number of figures/tables/images	1 (**Figure 1.3 in the dissertation**)
Format	print and electronic
Distribution quantity	3
Will you be translating?	no
Order reference number	None

Title of the thesis/dissertation: Protein thermal stability, conformational dynamics and solvent properties, insights with atomistic molecular Dynamics simulations

Expected completion date Sep 2015

Estimated size	200
Total	0.00 USD

RightsLink Printable License
ELSEVIER LICENSE
Dec 04, 2015

This is a License Agreement between Prathit Chatterjee ("You") and Elsevier ("Elsevier") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Elsevier, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

Supplier	Elsevier Limited The Boulevard,Langford Lane Kidlington,Oxford,OX5 1GB,UK
Registered Company Number	1982084
Customer name	Prathit Chatterjee
Customer address	Lab 1733, CMC Pune, Maharashtra 411008
License number	3761851112905
License date	Dec 04, 2015
Licensed content publisher	Elsevier
Licensed content publication	Journal of Molecular Biology
Licensed content title	The Occupancy of Ions in the K+ Selectivity Filter: Charge Balance and Coupling of Ion Binding to a Protein Conformational Change Underlie High Conduction Rates
Licensed content author	Yufeng Zhou,Roderick MacKinnon
Licensed content date	7 November 2003
Licensed content volume	333
Licensed content issue	5
Number of pages	11
Start Page	965
End Page	975
Type of Use	reuse in a thesis/dissertation
Portion	figures/tables/illustrations
Number of figures/tables/illustrations	1 (**Figure 1.4 a in the dissertation**)
Format	print
Are you the author of this Elsevier article?	No

Will you be translating?	No
Original figure numbers	Figure 1
Title of your thesis/dissertation:	Protein thermal stability, conformational dynamics and solvent properties, insights with atomistic Molecular Dynamics simulations
Expected completion date	Sep, 2015
Estimated size (number of pages)	200
Elsevier VAT number	GB 494 6272 12
Permissions price	0.00 USD

RightsLink Printable License
NATURE PUBLISHING GROUP LICENSE
Dec 03, 2015

This is a License Agreement between Prathit Chatterjee ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3761161303124
License date	Dec 03, 2015
Licensed content publisher	Nature Publishing Group
Licensed content publication	Nature
Licensed content title	Structural mechanism of plant aquaporin gating
Licensed content author	Susanna TörnrothHorsefield, Yi Wang,Kristina Hedfalk,Urban Johanson,Maria Karlsson et al.
Licensed content date	Feb 9, 2006
Volume number	439
Issue number	7077
Type of Use	reuse in a dissertation / thesis
Requestor type	academic/educational
Format	print
Portion	figures/tables/illustrations

Number of figures/tables/illustrations	1
Highres required	no
Figures	Figure 4 (**Figure 1.4 c in the dissertation**)
Author of this NPG article	no
Your reference number	None
Title of your thesis /dissertation:	Protein thermal stability, conformational dynamics and solvent properties: insights with atomistic Molecular Dynamics simulations.
Expected completion date	Sep 2015
Estimated size (number of pages)	200

Total 0.00 USD

ॐ पूर्णमदः पूर्णमिदम् पूर्णात् पूर्णमुदच्यते ।
पूर्णस्य पूर्णमादाय पूर्णमेवावशिष्यते ॥

Om,
That is Infinite, This is Infinite;
Infinite originates from Infinite.
If Infinite is removed from Infinite;
what remains is again Infinite.